

**Discovery of *Pseudomonas* Natural Products
Involved in the Biological Control of Potato
Pathogens**

Alaster David Moffat

John Innes Centre

A thesis submitted September 2021 to the University of East Anglia for the degree of Doctor of Philosophy. This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

Potato common scab and late blight, caused by *Streptomyces scabies* and *Phytophthora infestans*, respectively, are serious diseases affecting one of the world's largest and most important food crops. The lack of stable interventions has shifted recent focus toward biological control (biocontrol) methods. *Pseudomonas* isolates have shown significant promise as bacterial biocontrol agents, occurring in soils worldwide with high inter-strain diversity and potential for natural product biosynthesis. This thesis details investigations into the biosynthetic potential of environmental *Pseudomonas* strains isolated from a potato field, with a focus on discovering novel natural products active against plant pathogens. Investigations focused on a strain showing strong biocontrol phenotypes, Ps652. Initially, this strain showed strong inhibition of phytopathogens but with few biosynthetic gene clusters (BGCs) identified by common methods. A variety of methods were used to identify the determinants of the strong biocontrol phenotype shown by this strain, including activity-guided isolation of natural products and transposon mutagenesis. 3,7-dihydroxytropolone (3,7-HT) is reported here as being produced by a *Pseudomonas* isolate for the first time. 3,7-HT shows improved activity towards *Streptomyces scabies* compared to 7-hydroxytropolone, but does not fully explain activity of Ps652 against *P. infestans*. Additionally, investigations were made into putative RiPP BGCs containing DUF692 proteins in environmental *Pseudomonas* isolates Ps706 and Ps708. These BGCs appeared associated with phytopathogen inhibition in previous work, and were studied here using bioinformatics, gene deletions, and heterologous expression approaches.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Acknowledgements

I would like to extend my sincere gratitude to all those who have helped make my time at the John Innes Centre Molecular Microbiology department as productive as it has been. Firstly, my thanks go to Dr Thomas Scott, without whom I would have been lost in the first 6 months of my work. Secondly, I would like to thank Dr David Widdick for his infinite patience in explaining the nuances and tricks of common molecular biology techniques. Thirdly, my supervisor Dr Andrew Truman for giving me the freedom to explore the possibilities of my work, while providing the expertise and input necessary to ensure its success. Fourth, my unending gratitude goes to the laboratory support team at the John Innes Centre, without whom I would have needed to have been much more organised, and would have been much less productive. Fifth, my appreciation goes to the platform staff who made performic specialist analyses so much more accessible, and in particular Dr Martin Rejzek who acted as my chemistry guru through the final years and months of my work. I would also like to acknowledge my collaborators at the Earlham Institute, Dr Adam Elliston and Dr Jose Carrasco Lopez, for their help with the automation work which saved a great deal of time and repetitive strain injury. My gratitude goes to Dr Javier Santos-Aberturas and Dr Natalia Miguel-Vior who frequently engaged in stimulating discussion that helped guide my work. Finally, I thank my mother, Debbie, whose funding of my undergraduate days allowed me to pursue my interest in biology; my sister, Emily, whose own experience of doing a Ph.D. was useful more times than I can count; and Devonne, without whom none of this would have been possible.

This work was funded by the UKRI-BBSRC Norwich Research Park Biosciences Doctoral Training Partnership.

Author's Declaration

The research described in this thesis was conducted at the John Innes Centre between October 2017 and September 2021. All data described here are original and were obtained by the author unless otherwise attributed in the text. No part of this thesis has previously been submitted for a degree at this or any other academic institution. Initial isolation of environmental *Pseudomonas* strains was performed, and selected mutants were obtained, by Dr Francesca Stefanato. The pGME6032 expression vector was produced by Dr Natalia Miguel-Vior. Programming of automation equipment for the transposon mutagenesis screen was performed by Dr Adam Elliston at the Earlham Institute (Norwich, UK). Field trials for assessing Ps652 and Ps925 efficacy against potato diseases were part-designed, and the field work carried out, by Mr Graham Tomlinson of VCS Potatoes.

The transposon mutagenesis workflow and results described in **chapter 2** was the subject of the following publication and is presented in **Appendix 1**:

Moffat, A. D., *et al.* (2021) A biofoundry workflow for the identification of genetic determinants of microbial growth inhibition. *Synth. Biol. (Oxf)*, 6, ysab004.

The updated RiPPER genome mining workflow described in **chapter 4** was the subject of the following publication and is presented in **Appendix 2**:

Moffat, A. D., *et al.* (2021) A user guide for the identification of new RiPP biosynthetic gene clusters using a RiPPER-based workflow. *Methods Mol. Biol.*, 2296, 227-247.

List of Abbreviations

ACP	Acyl-Carrier Protein
ADIC	2-amino-2-desoxyisochorismic acid
AHL(s)	Acyl-Homoserine Lactone(s)
ALDH	Aldehyde Dehydrogenase
AMP	Adenosine Monophosphate
AOCHC	(1 R,6 S)-6-amino-5-oxo-2-cyclohexene-1-carboxylic acid
BGC	Biosynthetic Gene Cluster
BLAST	Basic Local Alignment Search Tool
BLS	Bacterial Leaf Spot
BP	Boiling Point
bp	Base Pairs
Mbp	Megabase pairs
BPC	Base Peak Chromatogram
CAD	Charged Aerosol Detection
CAS	Chrome Azurol S
CDS	Coding Sequence(s)
(CC-)NLR	(Coiled-coiled Domain-) Nucleotide-binding Domain Leucine-rich Repeat
CFU(s)	Colony Forming Unit(s)
(C)LP(s)	(Cyclic) Lipopeptide(s)
DAPG	Diacetyl Phloroglucinol
DHHA	(5S,6S)-6-amino-5-hydroxy-1,3-cyclohexadiene-1-carboxylic acid
DNA	Deoxyribonucleic Acid

DUF	Domain of Unknown Function
ECH	Enoyl-CoA Hydratase
EIC	Extracted Ion Chromatogram
EtOAc	Ethyl Acetate
EPS	Extracellular Polymeric Substance
FAD	Flavin Adenine Dinucleotide
FRAC	Fungicide Resistance Action Committee
GC-MS	Gas Chromatography-Mass Spectrometry
GFP	Green Fluorescent Protein
GMP	Guanosine Monophosphate
GWAS	Genome-Wide Association Studies
HCN	Hydrogen Cyanide
HDTMA	Hexadecyltrimethylammonium
HHPDC	Hexahydro-phenazine-1,6-dicarboxylic acid
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HSQC	Heteronuclear Single Quantum Coherence
IAA	Indole Acetic Acid
IMA	Instant Mash Agar
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ISR	Induced Systemic Resistance
JA	Jasmonic Acid
LB	Lysogeny Broth

LC-MS	Liquid Chromatography-Mass Spectrometry
MeCN	Acetonitrile
MeOH	Methanol
MHz	Megahertz
MKB	Modified King's Broth
MLST	Multilocus Sequence Typing
(m)RNA	(Messenger) Ribonucleic Acid
MS²	Mass Spectrometry-Mass Spectrometry (MS/MS)
MT	Methyltransferase
MYM	Maltose Yeast Extract Medium
NAD	Nicotinamide Adenine Dinucleotide
NADH	Reduced Nicotinamide Adenine Dinucleotide
NMR	Nuclear Magnetic Resonance Spectroscopy
NRP(S)	Non-Ribosomal Peptide (Synthetase)
OD	Optical Density
PAA	Phenylacetic Acid / Phenylacetate
PAMP(s)	Pathogen-Associated Molecular Pattern(s)
PCA	Phenazine Carboxylic Acid
PCP	Peptidyl Carrier Protein
PDC	Phenazine Dicarboxylic Acid
PDA	Potato Dextrose Agar
PGP	Plant Growth Promotion
PGPF	Plant Growth Promoting Fungi
PGPR	Plant Growth Promoting Rhizobacteria

PK(S)	Polyketide (Synthase)
PLP	Phospholipoprotein
PTI	PAMP-Triggered Immunity
QS	Quorum Sensing
QTL	Quantitative Trait Locus / Quantitative Trait Loci
RiPP	Ribosomally Synthesised-Post Translationally Modified Peptide
ROS	Reactive Oxygen Species
rpm	Revolutions Per Minute
RSA	Rye Sucrose Agar
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
RCBD	Randomised Complete Block Design
SA	Salicylic Acid
SAR	Systemic Acquired Resistance
SFM	Soy Flour Mannitol Medium
SIFT-MS	Selected-Ion Flow-Tube Mass Spectrometry
SNP	Single Nucleotide Polymorphism
TAD	Take-all Decline
TDA	Tropoditheitic Acid
TE	Thioesterase
TOFMS	Time-of-flight Mass Spectrometry
T(x)SS	Type (x) Secretion System
VOC(s)	Volatile Organic Compound(s)
UK	United Kingdom
US(A)	United States (of America)

UV	Ultraviolet
XU(s)	Exchange Unit(s)
YPD	Yeast Extract Potato Dextrose Medium

List of Figures

Figure 1.1. Factors in the soil environment contributing to the health of plants	22
Figure 1.2. The disease triangle illustrating the factors governing plant disease	28
Figure 1.3. Notable <i>Pseudomonas</i> natural products with biocontrol activities	44
Figure 1.4. The biosynthesis of phenazine natural products in <i>Pseudomonas</i>	52
Figure 1.5. <i>Phytophthora</i> lesion on a potato leaf	60
Figure 1.6. Rapid progression of potato late blight disease over a two week period	61
Figure 1.7. Breeding for late blight resistance in potatoes	65
Figure 1.8. Potato common scab caused by <i>Streptomyces scabies</i>	70
Figure 1.9. Thaxtomin phytotoxins produced by scab-causing <i>Streptomyces</i> species	71
Figure 2.1. MLST phylogenetic tree for Ps652 created using autoMLST	77
Figure 2.2. Cross-streak assays of Ps652 with <i>Streptomyces scabies</i> 87-22	79
Figure 2.3. <i>Phytophthora infestans</i> inhibition by Ps652	81
Figure 2.4. Schematic of split-plates and possible molecular interactions	83
Figure 2.5. Split-plate assays of Ps652 against <i>Streptomyces scabies</i> 87-22	84
Figure 2.6. Split plate assays of Ps652 against <i>Phytophthora infestans</i>	85
Figure 2.7. Bioactive extracts from Ps652 grown on RSA medium	87
Figure 2.8. Cross-streak assays for Ps652 , Ps652 Δ Pyo, and Ps652 Δ NRPS5 against <i>Streptomyces scabies</i> 87-22	89
Figure 2.9. Hydrogen cyanide production by Ps652	92
Figure 2.10. Hydrogen cyanide production by Ps652 and 652 Δ HCN	93
Figure 2.11. Bioassays against <i>S. scabies</i> 87-22 and <i>P. infestans</i> for Ps652 and Ps652 Δ HCN	94
Figure 2.12. Split-plate assays for Ps652 and Ps652 Δ HCN against <i>Streptomyces scabies</i> 87-22 and <i>Phytophthora infestans</i>	96
Figure 2.13. Extraction of a water soluble biologically active molecule from Ps652 Δ HCN	98
Figure 2.14. Purification of a tetrapeptide unrelated to the biological activity of Ps652 Δ HCN ...	99
Figure 2.15. Purification of pyoverdine from Ps652 Δ HCN by bioactivity-guided fractionation .	101
Figure 2.16. Fluorescence and MS ² spectrum for the purified molecule from Ps652 Δ HCN	103
Figure 2.17. LC-MS of Ps652 Δ HCN and Ps652 Δ HCN Δ Pyo cultures	104
Figure 2.18. Biological activity of Ps652 pyoverdine against <i>Streptomyces scabies</i> 87-22	105
Figure 2.19. Design of the automated transposon mutagenesis workflow for Ps652 Δ HCN	109
Figure 2.20. Representative hits from automated transposon mutant screening	111
Figure 2.21. Cross-streak assays for 23 high confidence mutants identified through automated screening of the transposon mutant library	112

Figure 2.22. Growth curves for selected transposon mutants	114
Figure 2.23. Putative biosynthetic gene cluster identified in Ps652 through amplicon sequencing of transposon mutants	116
Figure 2.24. Cross-streak assay for Ps652 Δ HCN Δ TE against <i>S. scabies</i> 87-22	118
Figure 2.25. Complementation of the Ps652 tropolone cluster thioesterase	120
Figure 3.1. Tropolone natural products.....	127
Figure 3.2. Formation of iron-chelates by 7-hydroxytropolone	131
Figure 3.3. Branch point between phenylacetate degradation and tropolone biosynthesis	134
Figure 3.4. Mechanisms of producing tropolone natural products	135
Figure 3.5. Tropolone and derivatives biosynthetic pathway in <i>Streptomyces</i>	136
Figure 3.6. Comparison of Biosynthetic Gene Clusters in Tropolone-producing bacteria	141
Figure 3.7. Proposed biosynthetic pathway for tropolone production in Ps652	144
Figure 3.8. Detection of 7-hydroxytropolone and related iron chelates in mass spectral data from Ps652 fermentations	146
Figure 3.9. Detection of a dihydroxylated tropolone in mass spectral data from Ps652 Δ HCN fermentations	147
Figure 3.10. Extraction of bioactive molecules from fermentations of Ps652 Δ HCN.....	149
Figure 3.11. LC-MS chromatograms for Sephadex LH-20 fractions in tropolone purification	151
Figure 3.12. Separation of 7-hydroxytropolone and 3,7-hydroxytropolone by reversed phase HPLC using a polar C18 stationary phase.....	152
Figure 3.13. ^1H NMR in CD_3OD for the purified 3,7-dihydroxytropolone	154
Figure 3.14. ^{13}C NMR in CD_3OD for the purified 3,7-dihydroxytropolone	155
Figure 3.15. Resonance forms of the purified dihydroxytropolone	156
Figure 3.16. Binding of 7-hydroxytropolone and 3,7-dihydroxytropolone to iron on CAS agar plates	158
Figure 3.17. Biological activity of 7-hydroxytropolone and 3,7-dihydroxytropolone against <i>Streptomyces scabies</i> 87-22 in vitro.....	160
Figure 3.18. Biological activity of 7-hydroxytropolone and 3,7-dihydroxytropolone against <i>Phytophthora infestans</i> 88069 in vitro.	161
Figure 3.19. In-frame gene deletions made in the Ps652 tropolone biosynthetic gene cluster.	163
Figure 3.20. Cross-streak assays for Ps652 tropolone cluster mutants against <i>S. scabies</i> 87-22	166
Figure 3.21. Iron-binding assay for Ps652 Δ HCN and mutant derivatives using CAS agar plates	167
Figure 3.22. Loss of tropolone production in tropolone cluster mutants	169
Figure 3.23. Accumulation of a shunt metabolite in tropolone cluster mutants	170

Figure 3.24. Performance of Ps652 as a biocontrol agent in field trials for protection of potato crops against <i>Phytophthora</i>	172
Figure 3.25. Updated proposed biosynthetic pathway for biosynthesis of tropolones in <i>Pseudomonas</i>	177
Figure 4.1. Comparative genomics for field isolated Pseudomonads with biocontrol potential	181
Figure 4.2. Pep1, Pep2, and Pep4 gene clusters as identified by antiSMASH	182
Figure 4.3. Alignment of MbnB from <i>Methylosinus trichosporium</i> OB3b with DUF692 protein from the Ps706 Pep1 cluster	185
Figure 4.4. Overview of the biological logic of the RiPPER genome mining tool	186
Figure 4.5. The steps and tools in the RiPPER genome mining workflow	187
Figure 4.6. Networks of Short Peptides Retrieved from RiPPER-EGN and Aligned Sequences ..	189
Figure 4.7. Pep1 Gene Cluster with Short Peptides Belonging to Networks 1 and 2	190
Figure 4.8. 85% Identity Unrooted Maximum Likelihood Tree of DUF692 Proteins in <i>Pseudomonas</i>	192
Figure 4.9. Evolutionary Relationship Between DUF692 Proteins and Networks of RiPPER-identified Short Peptides	194
Figure 4.10. Conservation of genes from the Ps706 Pep1 gene cluster by Multigeneblast analysis	196
Figure 4.11. Serial dilution spot assay on LB plates with 0.05 mM sodium Chlorite for Ps706 and Pep cluster mutants	199
Figure 4.12. Cross streaks of environmental isolate Ps706 and Pep cluster mutants against <i>Streptomyces scabies</i> 87-22	201
Figure 4.13. Cross streaks of environmental isolate Ps708 and Pep cluster mutants against <i>Streptomyces scabies</i> 87-22	202
Figure 4.14. Bioassays of Ps706 and Pep gene cluster mutants against <i>Phytophthora infestans</i>	204
Figure 4.15. Bioassays of Ps706 and Pep gene cluster mutants against <i>Gaeumannomyces graminis</i> var. <i>tritici</i>	205
Figure 4.16. Schematic of the Pep1 gene cluster	208
Figure 4.17. Effect of heterologous expression of Pep1 gene cluster in strain Ps723 on the inhibition of <i>Streptomyces scabies</i> 87-22 and <i>Gaeumannomyces graminis</i> var. <i>tritici</i>	209
Figure 4.18. Chromatograms from LC-MS of Ps723 heterologous expression strains in YPD medium.....	212
Figure 4.19. Expression of Pep1 gene cluster in <i>Pseudomonas putida</i> EM383	214

Figure 4.20. Schematic of Pep1 and Pep1Δ2282 gene clusters heterologously expressed in <i>Pseudomonas putida</i> EM383.....	216
Figure 4.21. Profiles of m/z 583.25 in <i>Pseudomonas putida</i> EM383 heterologous expression cultures	218
Figure 4.22. XCMS Analysis of Pep1 Expression Cultures	220
Figure 4.23. UV Absorption spectrum at 405nm for Pep1 expression cultures	221
Figure 5.1. Randomised complete block design (RCBD) for late blight field trials.....	266
Figure 5.2. Field trial study design for control of potato common scab using Ps652 and Ps925 as biological treatments.	269

Table of Contents

Chapter 1: Introduction	21
1.1 Natural product discovery for biological control of plant pathogens: a brief overview	21
1.2 The Soil Environment	21
1.2.1 Competition in the Soil	23
1.2.2 The Role of the Plant	25
1.2.3 Environmental Conditions and plant disease	26
1.2.4 Implications	29
1.3 The <i>Pseudomonas</i> genus and biological control	29
1.3.1 Bacterial Pathogen Inhibition by <i>Pseudomonas</i> species	32
1.3.2 Fungal Pathogen Inhibition	34
1.3.3 Oomycete Pathogen Inhibition	36
1.3.4 Arthropod Pathogen Inhibition	37
1.3.5 Limitations	39
1.4 Natural Product Discovery & Relevance to Biocontrol	39
1.4.1 Advances in Genomics for Discovery of Natural Product Gene Clusters.....	40
1.4.2 Activation of Silent BGCs	41
1.4.3 Prospects for Natural Product Discovery with Biocontrol Applications	42
1.5 <i>Pseudomonas</i> Natural Products	42
1.5.1 Overview	42
1.5.2 Polyketides: 2,4-diacetylphloroglucinol (2,4,-DAPG).....	45
1.5.3 NRPS: Viscosin	46
1.5.4 PKS-NRPS hybrids: Pyoluteorin	48
1.5.5 Phenazines	49
1.5.6 Volatile Organic Compounds: 1-undecene.....	53
1.5.7 Ribosomally-synthesised Post-translationally Modified Peptides (RiPPs).....	55
1.6 <i>Phytophthora infestans</i> & <i>Streptomyces scabies</i> as important potato pathogens	57
1.6.1 <i>Phytophthora infestans</i>	58
1.6.1.1 Overview	58
1.6.1.2 Plant Response.....	62
1.6.1.3 Fungicides	62
1.6.1.4 Breeding for Resistance	63
1.6.1.5 Forecasting.....	66
1.6.2 <i>Streptomyces scabies</i>	66
1.6.2.1 Overview	66

1.6.2.2	Breeding Scab-resistant Varieties	72
1.6.2.3	Cultural Practices	72
1.6.2.4	Chemical & Biological Control Methods	73
1.7	Thesis Objectives	75
1.7.1	Phenotype-guided Discovery of Natural Products	75
1.7.2	Investigation into Bioinformatically Identified Gene Clusters.....	75
Chapter 2:	Identifying the Natural Products Determining Biocontrol Activity in Environmental	
	<i>Pseudomonas</i> Isolate Ps652	76
2.1	Introduction and chapter aims	76
2.2	Initial Bioassays	76
2.2.1	<i>Streptomyces scabies</i>	78
2.2.2	<i>Phytophthora infestans</i>	80
2.2.3	Split-plate assays	82
2.2.4	Preparation of a Biologically Active Extract from Ps652 Cultures	86
2.3	Searching the Genome	88
2.3.1	Improved Illumina & Nanopore genome.....	88
2.3.2	AntiSMASH Analysis	88
2.3.3	Manual BLAST searches.....	90
2.3.4	Identification & Deletion of HCN Cluster.....	90
2.4	Isolation of Ps652 Pyoverdine by Activity-guided Purification.....	97
2.4.1.1	Extraction from 5 Litre Culture Extract	100
2.4.1.2	8-step Bioactivity-guided Purification of Ps652 Molecule	100
2.4.1.3	Identification of Ps652 Pyoverdine-related compound by HRMS and MS ²	102
2.4.1.4	Anti- <i>scabies</i> Activity of the Purified Pyoverdine	102
2.5	Transposon Mutagenesis.....	106
2.5.1	Screen Design & Considerations	106
2.5.2	Screen Output and Mutant Selection.....	108
2.5.3	Cross-streak Assays with Selected Mutants	110
2.5.4	Growth Curves & Further Mutant Selection	113
2.5.5	Transposon Sequencing & Tropolone Cluster	115
2.6	Validation of the Phenotype Using Clean Knockouts and Complementation	117
2.6.1	Deletion of the Tropolone Cluster Thioesterase	117
2.6.2	Complementation of Acyl-Coa Thioesterase in Ps652ΔHCNΔTE	119
2.7	Discussion & Conclusions.....	121
2.7.1	Discussion.....	121

2.7.2	Conclusions	123
Chapter 3:	Production of Tropolones by Ps652.....	125
3.1	Biosynthesis of Tropolone	125
3.1.1	Chapter Aims	125
3.1.2	Overview of Tropolone Natural Products in Bacteria.....	125
3.1.3	Tropolone Examples in Bacteria	128
3.1.3.1	Isatropolones	128
3.1.3.2	Roseobacticide and Tropodithietic acid.....	128
3.1.3.3	7-hydroxytropolone.....	129
3.1.4	Biosynthesis of Tropolones in Bacteria	132
3.1.5	Tropolones in <i>Pseudomonas</i>	137
3.2	Predicted Tropolone Biosynthesis in Ps652.....	138
3.2.1	Pathway Comparison by MultiGeneBlast.....	138
3.2.2	Proposed Biosynthetic Pathway for 7-Hydroxytropolone in <i>Pseudomonas</i>	142
3.3	Isolation of a Dihydroxylated Tropolone from Ps652	145
3.3.1	Detection of Tropolones in Ps652 Culture Extracts.....	145
3.3.2	Initial Extraction	148
3.3.3	Size-exclusion Chromatography	150
3.3.4	Semi-preparative HPLC.....	150
3.3.5	Characterisation of 3,7-dihydroxytropolone by ¹ H and ¹³ C NMR.....	153
3.3.6	Iron Binding of Hydroxytropolones	157
3.3.7	Biological Activity against <i>Streptomyces scabies</i> 87-22.....	159
3.3.8	Biological Activity against <i>Phytophthora infestans</i>	159
3.4	Effect of Gene Deletions in the 3,7-dihydroxytropolone Biosynthetic Gene Cluster	162
3.4.1.1	Cross Streak Assays.....	164
3.4.1.2	Iron-binding Assays on CAS agar.....	165
3.4.1.3	Production of Tropolones	168
3.4.1.4	Searching for Tropolone Biosynthetic Intermediates by LC-MS	168
3.5	Field Trails for Ps652 against potato diseases	171
3.5.1	<i>Phytophthora infestans</i>	171
3.5.2	<i>Streptomyces scabies</i>	173
3.6	Discussion & Conclusions.....	173
3.6.1	Summary / Discussion of Results.....	173
3.6.2	Future Work	178
Chapter 4:	Bioinformatically Identified RiPP-like Clusters.....	180

4.1	Introduction & Aims	180
4.2	Bioinformatic Analysis of Pep Gene Clusters	183
4.2.1	Overview of Pep gene clusters	183
4.2.2	Searching for conserved DUF692-associated short peptides using RiPPER & EGN	188
4.2.3	Phylogenetics of DUF692 Proteins	191
4.2.4	Establishing BGC boundaries using MultiGeneBlast.....	195
4.3	Gene Deletion Experiments	197
4.3.1	Deletion of Pep Gene Clusters.....	197
4.3.2	Investigating the link between DUF2282-DUF692-DUF2063 operon & oxidative stress	198
4.3.3	Bioassays with Deletion Strains.....	200
4.3.3.1	<i>Streptomyces scabies</i>	200
4.3.3.2	<i>Phytophthora infestans</i>	203
4.3.3.3	<i>Gaeumannomyces graminis</i>	203
4.4	Heterologous Expression of Pep1 Gene Cluster.....	206
4.4.1	Cloning the Pep1 Gene Cluster	206
4.4.2	Expression in Strain Ps723 (Bioassays)	206
4.4.3	Expression in Strain Ps723 (Metabolomics)	210
4.4.4	Heterologous Expression in Strain EM383	213
4.4.4.1	Overview of Strain EM383 as a Heterologous Host	213
4.4.4.2	Expression of Pep1 in YPD medium (Metabolomics).....	213
4.4.4.3	Cloning of Pep1 Δ DUF2282.....	215
4.4.4.4	Expression of Pep1 Δ DUF2282 in <i>Pseudomonas putida</i> EM383	217
4.4.4.5	XCMS Analysis	219
4.5	Discussion & Conclusions.....	222
4.5.1	Summary / Discussion of Results.....	222
4.5.2	Conclusions and Future Perspectives.....	225
Chapter 5:	Materials and Methods.....	226
5.1	Materials.....	226
5.1.1	Strains.....	226
5.1.2	Plasmids.....	232
5.1.3	Chemicals, Reagents, and Media.....	234
5.1.3.1	Chemicals & Reagents.....	234
5.1.3.2	Growth Media.....	234
5.1.4	Oligonucleotides.....	237
5.2	General methods.....	242

5.2.1	<i>E. coli</i> methods	242
5.2.2	<i>Pseudomonas</i> strain methods	243
5.2.2.1	Growth conditions	243
5.2.2.2	DNA Extraction.....	243
5.2.3	Transformations	243
5.2.3.1	<i>Pseudomonas</i> transformations	243
5.2.3.2	<i>E. coli</i> DH5 α	244
5.2.4	Cloning and sequencing	244
5.2.4.1	PCR.....	244
5.2.4.2	Agarose Gel Electrophoresis	245
5.2.4.3	Purification of DNA from Agarose Gels and Reaction Mixtures.....	245
5.2.4.4	DNA digestions.....	246
5.2.4.5	DNA Ligations.....	247
5.2.4.6	Gibson Assembly.....	247
5.2.4.7	Sequencing.....	247
5.2.5	Making in-frame deletion strains	247
5.3	Metabolomics and LC-MS	248
5.3.1	Sample preparation (general)	248
5.3.1.1	Preparation of tropolone BGC mutant extracts.....	248
5.3.2	LC-MS conditions.....	249
5.3.3	Accurate mass	249
5.3.4	Analysis of results.....	249
5.3.4.1	XCMS online pairwise analysis.....	250
5.4	Biological Assays	250
5.4.1	Cross streaks.....	250
5.4.2	Spot assays against <i>Phytophthora infestans</i> and <i>Gaeumannomyces graminis</i> var. <i>tritici</i>	250
5.4.3	Split-plate assays	251
5.4.4	Testing extracts of <i>Pseudomonas</i> cultures & purification fractions.....	251
5.4.5	Iron binding assays	252
5.5	Purification and Partial characterisation of Pyoverdine from Ps652	252
5.5.1	Scaled up production cultures.....	252
5.5.2	Initial ethyl acetate extraction	252
5.5.3	Sample debulking by flash chromatography	253
5.5.3.1	500g XAD-4	253
5.5.3.2	120g C18	253
5.5.3.3	Further C18 flash chromatography.....	254
5.5.4	Preparative C18 HPLC.....	254

5.5.5	Semi-preparative HPLC.....	255
5.5.5.1	C18	255
5.5.5.2	Pentafluorophenyl (PFP).....	256
5.5.5.3	Polar C18.....	257
5.5.6	High-resolution mass spectrometry (HRMS).....	258
5.6	Purification of Tropolones from Ps652.....	258
5.6.1	Production cultures	258
5.6.2	Extraction of Tropolones from 2L Culture	258
5.6.3	Size Exclusion Chromatography using Sephadex LH-20	259
5.6.4	Semi-preparative HPLC.....	259
5.6.5	NMR analysis of Purified Tropolone.....	260
5.6.5.1	3,7-dihydroxytropolone	260
5.7	Bioinformatic analysis of Pep gene clusters	260
5.7.1	AntiSMASH	260
5.7.2	Alignment of MbnB and DUF692.....	260
5.7.3	RiPPER workflow	260
5.8	Bioinformatic analysis of Ps652	261
5.8.1	Genome assembly	261
5.8.2	AntiSMASH	262
5.8.3	Manual BLAST.....	262
5.9	Automated Transposon Mutagenesis screen	262
5.9.1	Creation of a transposon mutant library of Ps652ΔHCN.....	262
5.9.2	Automated screening of mutant library against <i>S. coelicolor</i> M145	263
5.9.3	Selection of mutants for further verification	264
5.9.4	Growth curves	264
5.9.5	Amplicon sequencing of mutants.....	264
5.10	Field Trials	265
5.10.1	Field trials for <i>Phytophthora infestans</i>	265
5.10.1.1	Late blight field trial study design.....	265
5.10.1.2	Preparation of biocontrol strain inocula.....	267
5.10.2	Field trials for potato common scab.....	267
5.10.2.1	Common scab field trial study design	267
5.10.2.2	Preparation of inocula for dipping of seed tubers.....	270
5.10.2.3	Preparation of inocula for drip irrigation.....	270
	References	271

Appendix 1317
Appendix 2325

Chapter 1: Introduction

1.1 Natural Product Discovery for Biological Control of Plant Pathogens: a brief overview

It is well established that the soil microbiome is extensive, complex, and heterogeneous. Only fractions of the true biodiversity of these environments have been cultured, but advances are being made in this area (1, 2). Simultaneously, metagenomics and systems biology approaches are showing us just how much remains to be understood with regards to bacterial communities in nature (3). The variety of interactions in microbial ecology are diverse, but can involve bacteria and fungi, oomycetes, insects, nematodes, plants, or others. Key factors in these interactions include direct or indirect competition for nutrients, antibiosis via antimicrobial natural products, or contact-mediated inhibition via specific secretion systems that may be coupled to toxin systems (4). Biological control (biocontrol) is the use of select bacterial species, communities, or their metabolites to control plant pathogens. These biocontrol strains may function either as endophytic colonisers of the plant itself, or simply occupy the rhizospheric portion of soil (5, 6), acting through many different competitive mechanisms, often simultaneously and inextricably (7).

1.2 The Soil Environment

In order to understand how biocontrol solutions might be applied, it is important to first grasp the nuances of the soil environment and the range of interactions that take place. These will be discussed in the following section, and an overview is provided in Figure 1.1.

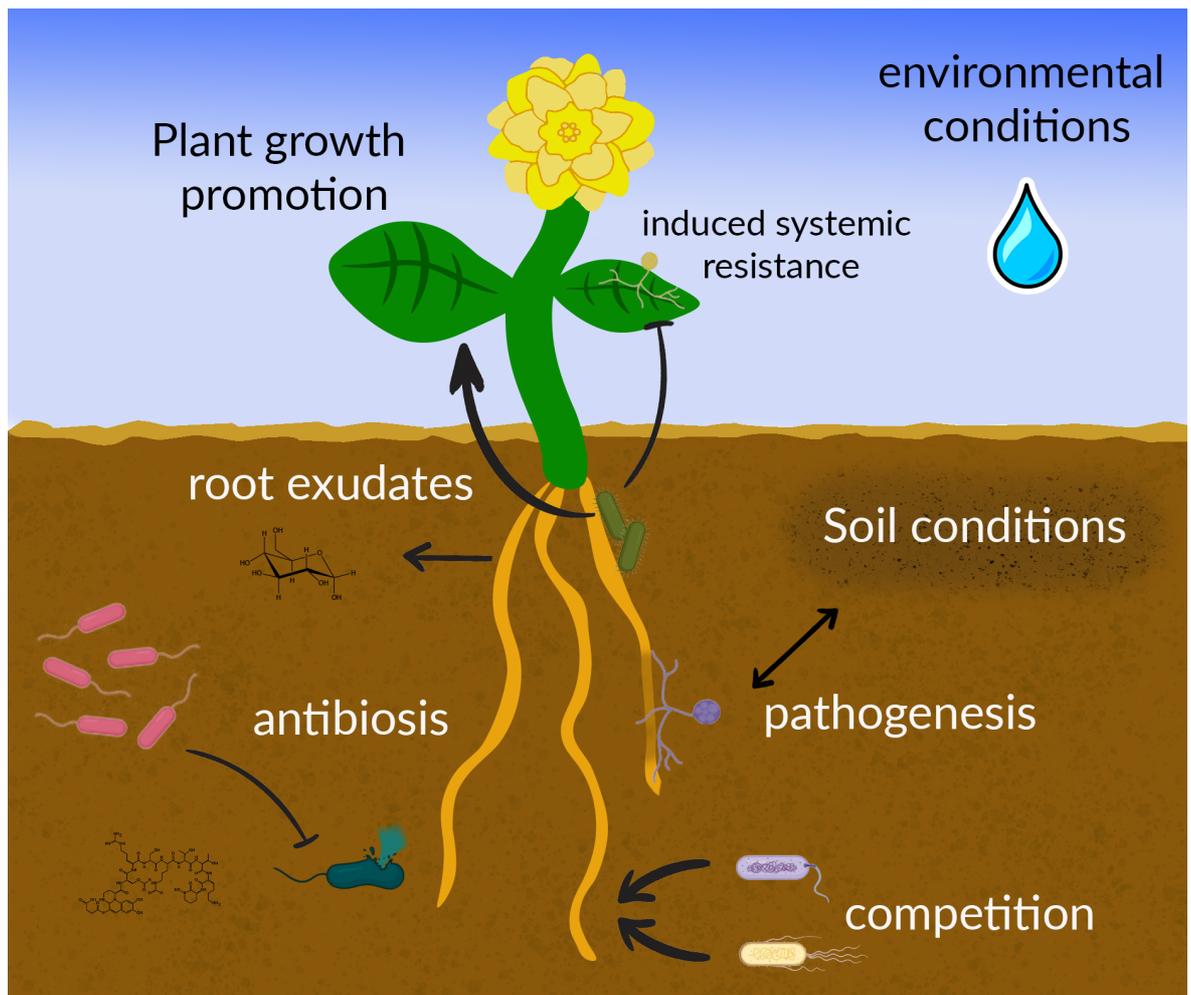


Figure 1.1. Factors in the soil environment contributing to the health of plants. This schematic shows the main factors in the soil resulting in plant health or disease, including interactions with microorganisms. Environmental conditions like rain, humidity, sunlight, and temperature affect both growth of the plant directly, and the ability of different microorganisms to grow in the rhizosphere (soil environment) or phyllosphere (leaf environment). Root exudates, like sugars, produced by the plant change the soil environment directly attached to the roots, and may promote the growth of specific microbial taxa. Some of these may promote plant growth by producing plant hormones or aiding the sequestration of key nutrients. Microorganisms in the soil may compete for resources like root exudates or space on the root, by a number of mechanisms, excluding competitors. In some cases, this is due to direct antibiosis through the production of antimicrobial compounds. Some microbial taxa are obligate or opportunistic plant pathogens, and may affect plant structures beneath ground, above it, or both. The plant's immune system plays a role in resisting disease and may be primed by other microbes. Finally, soil conditions such as pH, richness in organic matter, or moisture content can affect all of these interactions, and ultimately the health of the plant.

1.2.1 Competition in the Soil

Application of specific microbial populations in soil for the control of a plant pest or pathogen centres around the interaction between the competing organisms, as well as the host plant. Competition for space and resources is the driving force of this interaction, as the vast majority of soil microflora are lacking nutrients, and in constant battle to selectively accumulate plant-derived resources (8, 9). Nicholson originally characterised the nature of this competition as ‘scramble’ and ‘contest’ (10). In the dynamic world of rhizospheric and phyllospheric environments, an interplay of both is likely at work and needed to fully understand the nature of microbial warfare. For a review of mechanisms of competition in and between microbial populations, the reader is directed to these specialised reviews (11, 12).

Nicholson’s ‘scramble’ is the competitive strategy of rapidly assimilating the available resource(s) so as to constitute a competitive exclusion of your rival. This may include the essential nutrients for microorganisms to grow: carbon, nitrogen, phosphorus, sulfur, hydrogen, calcium, iron, and other metals (12). In the context considered here, the resource in question is likely to be either the valuable and highly nutritious root exudates, including bioavailable iron, or the space on the root / plant itself, one of the most nutrient-rich habitats on earth for a microorganism with plants depositing up to 40% of their photosynthetically stored carbon into the rhizosphere (13). Jacques Monod’s investigations into bacterial growth dynamics show us that, where a single nutrient required for growth is scarce, ability to assimilate this at the expense of your competitor will essentially preclude their proliferation (14). Siderophores, secondary metabolites that allow accumulation of iron, offer a competitive advantage in this way. Space is also a valuable resource in the context of soil, and many bacteria scramble to monopolise this too. Natural products like lipopeptide surfactants aid motility on plant surfaces, and enhance ability to colonise plant surfaces (15). Where there is a specific nutrient source, for example a plant root, phenotypes allowing preferential colonisation of this interface can exclude the growth of a rival unable to reach the nutrient source (16). The formation of a biofilm composed of extracellular polymeric substance (EPS) can serve as a means of retaining nutrients and preventing their dispersal to non-kin in such cases.

On the other hand, Nicholson's 'contest' is the direct competition model, that we often term antibiosis. This is also regulated by the production of specific molecules. It is an intuitive competitive model to understand, whereby a metabolite toxic to your competitor either kills them or restricts their growth, excluding them from the niche in question (e.g. root surface). An example may be found in the R-tailocin bacteriocins produced by *Pseudomonas chlororaphis*, where tailocin production is required for competitiveness in rhizosphere communities (17). This mode of interaction is by far the most extensively investigated in attempts to derive new biocontrol solutions, and is still considered the most potent form of antagonism (18). However, some suggest that repeated investigations into this mode of action has biased our perspective on how relevant it is to real-world antagonism (7). Typically, researchers will target a specific plant pathogen, and attempt to selectively isolate bacteria from specific soil environments that display potential to inhibit growth of the pathogen. However, there are multiple complications of this approach. Firstly, one cannot guarantee that the full biosynthetic potential of all isolates are expressed under laboratory conditions, leading to innovative approaches to explore this biosynthetic space. Secondly, the issue of rediscovery plagues these studies as much as the wider field of natural product discovery, although this will be discussed later in this review. Thirdly, there can be no surety that any interaction observed *in vitro* will be replicated *in planta* (19, 20), particularly given the lability of these metabolites in soils (21). Finally, we may view the utility of bacterial natural products in a disparate manner to the producer (22), as can be shown in the story of phenazine (23), or the signal-like nature of sub-inhibitory levels of many natural products (24). The role of natural product antibiotics in the environment is therefore still hotly debated (12). In order for antibiosis to be a viable strategy in controlling plant pests and diseases, a detailed understanding of the ecology is useful.

A third mode of competition, a blend of direct and indirect, is that which we shall call here 'interference'. A diverse range of microbial strategies and adaptations are regulated by population-level responses, or Quorum Sensing (QS). In many Proteobacteria the signal in question is the production of Acyl Homoserine Lactones (AHLs) which, through two-component systems, regulate responses on a large scale once a sufficient population size

is achieved - or in a low diffusivity environment (25). A number of enzymes have been characterised which display the ability to degrade these signals, interfering with the population synchronised response of the original producer (26, 27). It has yet to be extensively investigated how such signalling disruptions apply in the realms of microbial ecology of the soil, with most studies limited to *in vitro* interactions at present. However, it has been established that synthetic AHLs are mineralised upon addition to soil; a function lost upon sterilisation of the soil (28). This interference may preclude your competitor from forming a biofilm and limiting your access to nutrients (29).

Overall, we can predict that competition is likely a strong factor in rhizospheric or phyllospheric interactions given at least one of the three following factors: (i) there is overlap between the niche or required resources, (ii) cells are spatially mixed on a scale such that resources become shared, or (iii) available resources are scarce relative to cell densities (12). Such conditions prevail in the nutrient-limited, highly heterogenous, and microbially diverse environments of agricultural fields.

1.2.2 The Role of the Plant

The overall outcome of any biocontrol efforts ultimately must also consider the third component in the system, the plant itself. Of particular relevance are the effects of Plant Growth Promotion (PGP) and Induced Systemic Resistance (ISR) by microorganisms. Plant growth promotion is a property of many free-living soil microbes, including those investigated as biocontrol agents, whereby growth of the plant is enhanced through a number of mechanisms, without forming a direct symbiosis with the organisms in question. Plant Growth Promoting Rhizobacteria (PGPR) or Fungi (PGPF) may either do so directly through the production of phytohormones like auxins (30, 31) altering root architecture and boosting growth (32, 33), or indirectly through the solubilisation of minerals like Fe, N, and P (34). In some cases, PGP activity may even protect the plant against disease indirectly (35). Induced systemic resistance was first described in 1991, and relates to the ability of certain rhizobacteria, mainly *Pseudomonas*, *Serratia*, and *Bacillus* species as well as fungi like *Fusarium* and *Trichoderma* species to enhance whole plant resistance to a broad range of pathogens upon local colonisation by PGPR or PGPF (36). This process is at least partially

influenced by the production of natural products (37). Independent of the salicylic acid (SA) signalling network that characterises the acquisition of systemic resistance after induction by pathogen attack, beneficial rhizobacteria-ISR is coordinated by jasmonic acid (JA) and ethylene pathways that are effective against necrotrophic pathogens (38). However, PGPR have been identified that do activate the SA-dependent pathways and can enhance resistance against biotrophic pathogens (39). Given the observation that coactivation of both SA and JA enhances plant immune response, PGPR or PGPF capable of activating both signalling pathways could be particularly potent primers of plant immunity (40). Such a biocontrol agent may be preferable where no specific target pathogen exists, to bolster the plant's defences against a wide range of pathogens. Perhaps counterintuitively, these bacteria and fungi must suppress the plant immune response to some extent to permit their own colonisation of these niches (41), and natural products were recently implicated in this process too, as *B. subtilis* BSn5 reduces flg22-mediated immune response through the production of the lantibiotic subtilomycin (42).

1.2.3 Environmental Conditions and Plant Disease

The 'Disease Triangle' illustrates the need for three factors to result in plant disease: plant susceptibility, pathogen virulence, and environmental conditions (43) (Figure 1.2). Environmental conditions encompass a plethora of factors including but not necessarily limited to soil moisture, pH, relative humidity, nutrient availability, temperature, soil aeration, C/N balance, micronutrient ratios, soil composition, and atmospheric CO₂ levels. Many of these are difficult, if not impossible, to manipulate and are likely to undergo extensive change in future in the context of our shifting climate (44, 45). Beyond simple reduction in yield, the effects of environmental change can also impact plant product quality, like increased mycotoxin concentrations in grains with rising CO₂ (46). As such, it is crucial we understand how a plant's environment can predispose to disease, and make or break biocontrol efforts (7). Such knowledge can inform disease risk and inform application of mitigation strategies. For example, risk of bacterial leaf spot of pepper (BLS) caused by *Xanthomonas euvesicatoria* was recently found to relate to macro and micronutrient concentrations and their ratios, through the plant SAR pathways (47). Likewise, *Arabidopsis* susceptibility to *Pseudomonas syringae* infection was found to be dependent on soil

composition and soil amendment (48). Interestingly, even post-harvest diseases are influenced by environmental conditions at time of harvest, as was demonstrated with *Salmonella* in tomatoes (49).

The environment is frequently a major factor limiting the success of biocontrol strategies. *Ralstonia pickettii* control of *R. solanacearum* infection of tomato in China was correlated with mean temperature during the growing season, where higher mean temperatures resulted in poorer control of the pathogen (50). *R. pickettii* grew best below 20 °C, whereas the pathogen grew best closer to 30 °C, meaning the ratio of biocontrol agent to pathogen was lowest at temperatures conducive to *R. solanacearum* pathogenicity (51). Soil pH and moisture content have likewise been demonstrated to affect the ability of biocontrol fungus *Heteroconium chaetospora* to protect against clubroot disease in Chinese cabbage (52). In greenhouse tests the biocontrol fungus reduced disease incidence by up to 100% in some conditions, but soil moisture of 80% or low pH around 5.5 could completely eliminate the biocontrol effect. Perhaps of greatest concern for the prospects of biological control of plant disease is the finding of Cray *et al.*, who showed a biocontrol strain of *Bacillus* that near-completely eliminated growth of potato-pathogenic *Phytophthora* and *Fusarium* strains *in vitro*, yet on potato tubers actually promoted growth of Fusaria under chaotropic conditions by up to 570% (53). They propose the bacterium supplied the fungus with kosmotropic metabolites, allowing growth under unfavourable conditions for the pathogen. Whether this is likely to occur under more natural conditions (field conditions) than their assay is yet to be seen, but highlights the complications environmental conditions pose for development of biocontrol strategies. These are factors we must keep in mind when attempting to exploit environmental bacteria for purposes of phytopathogen control.

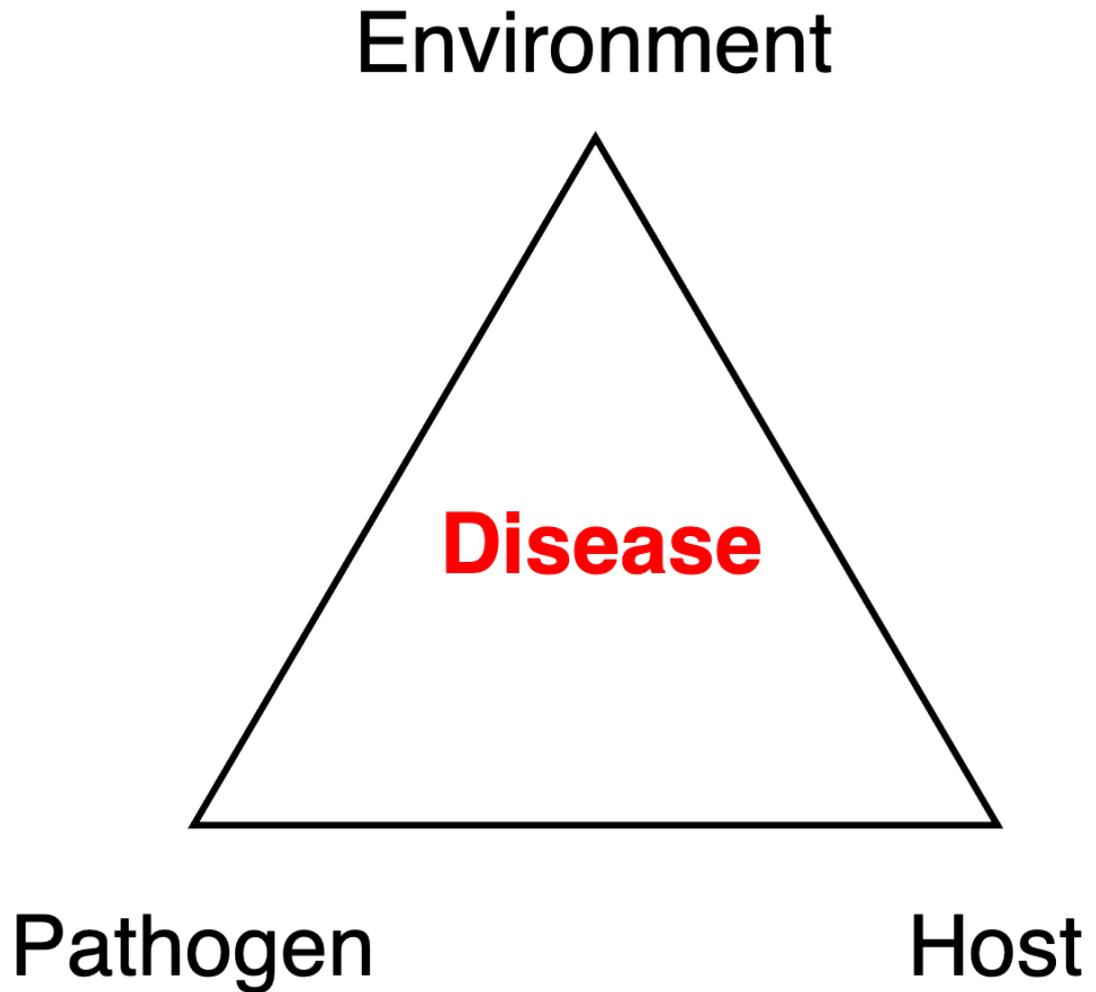


Figure 1.2. The disease triangle illustrating the factors governing plant disease. All three elements are required to be suitable for plant disease to develop. Environmental factors may include soil moisture, pH, relative humidity, nutrient availability, temperature, soil aeration, C/N balance, micronutrient ratios, soil composition, and atmospheric CO₂ levels. The pathogen includes the particular organism causing disease and the adaptations it possess that allow it to be virulent. The host plant refers to the susceptibility of the plant in question to disease, including the cultivar, the health of the individual plant, its immune system, and any other factors enabling or preventing it from dealing with infection.

1.2.4 Implications

Ultimately, biocontrol outcomes can be difficult to predict, and the ecological and evolutionary stability over time troublesome to anticipate. All the above factors have implications for the development of biocontrol solutions. Emerging methods are focusing on harnessing the power of entire communities of microbes for human benefit, through enrichment of specific populations within the rhizosphere or plant-associated soil environment, which can lead to measurable decreases in disease incidence (54) – this has been particularly demonstrated with Take-All of wheat (55) but also in pea (56) and tobacco (57). A comprehensive review of biological control of soil-borne pathogens was published in 2005 by Haas and Défago (58), which this review aims to augment. Many of these biocontrol strains have been commercialised, particularly for *Pseudomonas* species (59), a genus with particular promise.

1.3 The *Pseudomonas* Genus and Biological Control

Proteobacteria have been identified as the most prevalent taxa in many disease-suppressive soils, particularly *Gammaproteobacteria* (including *Pseudomonas*) and *Betaproteobacteria* (60). *Pseudomonas* is a genus of *Gammaproteobacteria* that is largely heterogenous, demonstrating both large interspecies and intraspecies diversity, with a relatively small proportion of genes comprising the core genome (61). Silby and colleagues (6) reviewed in 2011 the variety of environments adapted to by members of this group, and the range of metabolic capabilities adopted as a consequence. Crucially, pseudomonads are found ubiquitously in soils around the world, with unique populations in each region, thereby living and growing alongside possibly every known plant pathogen today (62). In some cases, the pseudomonad is the pathogen itself, with the genus containing both plant growth promoting (PGP) and plant pathogenic species, for example in *Pseudomonas fluorescens* and *Pseudomonas syringae* respectively. Many strains of this genus have already been targeted for use in crop protection programmes and as producers of compounds noted to decrease incidence of disease through inhibition of plant pathogens. Some of these have been applied commercially (Table 1.1).

Commercial biocontrol agents fall broadly into three categories: those applied directly to site of infection when and where needed (including seeds pre-germination), those applied in a singular region of the plant with the view that they will colonise and spread to the rest of the plant, and those that may have persistent benefits after a singular or regular release (63). Others are displaying the ability to control post-harvest spoilage in food crops (64-66), contamination of stored food products by human pathogens (67, 68), and even speed up the growth of commercial food crops (69). However, existing selective isolation practices have traditionally been targeted towards other genera, like the well-known members of the *Actinobacteria*, and it is unlikely that the full scope of phenotypic diversity with regards to phytopathogen inhibition within this genus has been captured. Accordingly, it stands to reason that a more comprehensive bioprospecting approach with a focus on *Pseudomonas* could recover isolates with high potential to improve current agricultural practices and reduce the use of pesticides and other harmful agrochemicals worldwide. It has been demonstrated that many agrochemicals and their additives are retained for long periods under certain application practices (70, 71), and such chemicals have been linked to reductions in biodiversity (72) and pollinator efficiency (73, 74), increases in incidence of late-onset diseases like Parkinson’s (75) and global pollution. For an overview of pesticide usage and risks or benefits, see ref. (76). Examples of the diverse biological control phenotypes noted in this genus to date will be explored in the following sections.

Table 1.1 Notable commercial biocontrol strains / formulations derived from *Pseudomonas* species and related Proteobacteria. Non-exhaustive list of biocontrol strains of *Pseudomonas* that have been licensed and sold for control of phytopathogens, the product names they were marketed under, and which pathogens they were targeted at controlling in which species.

Product Name	Strain	Applied to	Active Against	Reference(s)
Tx-1	<i>P. aurofaciens</i>	Turf Grass	Various fungal pathogens	(77-79)
BioSave (Various)	<i>P. syringae</i>	Citrus and Pome fruit	Various fungal pathogens	(63, 80)

Victus or Conquer	<i>P. fluorescens</i> <i>NUB 12089</i>	Mushrooms	<i>Pseudomonas tolassii</i>	(81)
Deny*	<i>Burkholderia cepacia</i> †	Alfalfa, barley, beans, clover, cotton, peas, grain sorghum, vegetables crops, wheat	<i>Rhizoctonia</i> , <i>Pythium</i> , <i>Fusarium</i> , nematodes	(82)
Intercept*	<i>Burkholderia cepacia</i> †	Maize, vegetables, cotton	<i>Rhizoctonia solani</i> , <i>Fusarium spp.</i> , <i>Pythium sp.</i>	(82)
PSSOL	<i>P. solanacearum</i> (non-pathogenic)	Vegetables	<i>Pseudomonas solanacearum</i>	(80)
Blightban A506	<i>P. fluorescens</i> A506	Almond, apple, cherry, peach, pear, potato, strawberry, tomato, various vegetables	<i>Erwinia amylovora</i>	(80)
Cedomon, Cerall, Cedress	<i>P. chlororaphis</i> MA342	Barley and Oats (Cedomon) Wheat (Cerall) Pea and Carrot (Cedress)	Various Seed-borne diseases	(83-85)
Blue Circle*	<i>Burkholderia cepacia</i> †	Vegetables	<i>Fusarium</i>	(86)
Howler	<i>P. chlororaphis</i> strain AFS009	Berries / fruits, various vegetables, beans & peas, tuber / root crops, tobacco & specialty crops	<i>Botrytis</i> , <i>Pythium</i> , <i>Phytophthora</i> , <i>Rhizoctonia</i> , <i>Colletotrichum</i> , <i>Sclerotinia</i>	(87)
D7	<i>P. fluorescens</i> strain D7	Wheat, barley, triticale, oats, rangeland	Downy Brome, Medusahead, Japanese Brome, Jointed goatgrass	(88)
Proradix	<i>Pseudomonas sp.</i> strain DSMZ 13134	Seed potatoes, flowers, tomatoes, cucumbers, peppers, eggplant, lettuce, cabbage	Soil-borne fungi	(89)

*, Withdrawn from market; †, previously *Pseudomonas cepacia*.

1.3.1 Bacterial Pathogen Inhibition by *Pseudomonas* Species

Pseudomonas antibacterial compounds have been well studied, and include various classes of molecule from lipopeptides, volatile compounds, small-molecule antibiotics, as well as larger antimicrobials like peptides and proteins (61). Many of these compounds are involved in the biocontrol phenotypes exhibited by members of this genus and are covered in more depth in section 1.5. Curiously however, abolition of many determinants of *in vitro* inhibition through deletion mutants is not always predictive of *in planta* disease prevention, as was demonstrated for protection of apple flowers against *Erwinia amylovora* by a *Pseudomonas orientalis* isolate (90). Safracin, pyoverdine, and phenazine mutants retained efficacy on flowers, demonstrating that we must be cautious in overinterpreting *in vitro* experiments with well-determined inhibitory molecules, although such knowledge can still be of value. In line with the ‘interference’ model discussed above in section 1.2.1, production of an acylase quorum-quenching enzyme by *P. segetis* P6 mediates soft-rot diseases of root vegetables, reducing *Dickeya solani*, *Pectobacterium atrosepticum* and *Pectobacterium carotovorum* symptoms on potatoes and carrots (91). Likewise, the strain was able to protect tomato plants against *P. syringae* pv. tomato *in vivo*. Volatile compounds produced by *P. fluorescens* and *P. chlororaphis* isolates have been implicated in the inhibition of quorum sensing and formation of biofilms by the crown-gall pathogen *Agrobacterium tumefaciens*; it was established and verified experimentally that ketones produced by these strains not only inhibited biofilm formation but killed *Agrobacterium* in existing biofilms (26, 92). Volatile organic compounds (VOCs) produced by *P. fluorescens* strains also inhibit another important plant pathogen, *Ralstonia solanacaerum*, the causative agent of bacterial wilt disease that affects a wide range of plant species. In both cases the interactions were determined to be bacteriostatic, but the former studies implicated ketones like 2-nonanone, 2-heptanone, and 2-undecanone, whereas the latter focused on compounds such as long chain alcohols like 2-decanol or 2-tridecanol, as well as aromatic volatiles like benzaldehyde or ethyl benzene and benzothiazole (93).

P. fluorescens P142 has also shown *in planta* efficacy against *R. solanacearum* in tomato plants, which the authors speculated could be due to niche exclusion or priming of the plant immune system, highlighting the various ways in which biocontrol strains may work (94). *Streptomyces scabies*, an economically important potato pathogen for which no consistently effective control measure is documented, was inhibited by *Pseudomonas* LBUM223, and a biocontrol *in planta* assay revealed an ability to remediate disease severity (95). The isolate displayed the ability to reduce induction of expression of the key pathogenicity factor thaxtomin, and much of the inhibitory activity was attributed to production of phenazine-1-carboxylic acid. However, *in vitro* suppression of *S. scabies* growth was not completely abolished with the loss of phenazine production, indicating multifactorial inhibition – a persistent issue in this field. Antibiotics pseudopyronine A and B, produced by an isolate of *P. putida* (BM11M1), were discovered by Gross and colleagues to display significant activity towards a pathogenic pseudomonad, *P. savastanoi* (96). BW11M1 was also found to produce a previously undiscovered group of cyclic lipopeptides, the xantholysins, which effected inhibition of *Xanthomonas* isolates and also were active against fungi and some Gram-positive bacteria (97). The authors stressed the importance of the findings for control of bacterial plant pathogens, citing a review which placed *Xanthomonas axonopodis*, *Xanthomonas campestris*, and *Xanthomonas oryzae* as 3 of the top 10 bacterial phytopathogens (98).

In an attempt to increase specificity of biological treatments, bacteriocin-producing strains are being investigated to control crop diseases caused by bacteria. These molecules are ribosomal peptides, produced to inhibit other bacteria often with a narrow spectrum of activity, and may function to competitively exclude other taxa from a niche (99, 100). Tailocins produced by *Pseudomonas fluorescens* SF4c have shown effectiveness in reducing bacterial-spot of tomato, caused by *Xanthomonas vesicatoria* Xcv Bv5-4a, reducing both disease incidence and severity even when applied up to 12 hours after infection in pot trials (101). Similar results have been shown for prevention of infection of *Nicotiana benthamiana* by *P. syringae* pv. *syringae* B728a, where a number of strains of *P. syringae* were able to produce tailocins that prophylactically prevent colonisation and infection of the plant host by the phytopathogen (102). Astonishingly, no cells of B728a could be recovered from some of the plants, even though they had been dipped in a 10^7 CFU/mL

inoculum, highlighting the potent protective effect that bacteriocin-based treatments can offer. Included in the list of top 10 bacterial plant pathogens by Mansfield *et al.* (98) are many of the genera mentioned in this section, highlighting the role *Pseudomonas* isolates could play in future biocontrol efforts.

1.3.2 Fungal Pathogen Inhibition

Inhibition of fungal pathogens is well known for this genus, with volatile molecules like hydrogen cyanide and 1-undecene being regularly implicated in activity (103, 104). Other volatiles like dimethyl disulphide from *P. stutzeri* (E25) have been linked to antagonism of *Botrytis cinerea* (105), and volatiles from *P. chlororaphis* subsp. *aureofaciens* SPS-41 demonstrated broad spectrum inhibition of many fungal isolates, including crucial plant pathogen *Alternaria alternata* (106). Likewise, the production of chitinase enzymes that are able to degrade the cell walls of fungi by targeting chitin, potentially causing osmotic lysis of these phytopathogens. Some have found chitinase-producing pseudomonads to inhibit *Rhizoctonia solani* *in vitro*, but were unable to completely correlate chitinase production with inhibition, suggesting the production of additional antifungals in this case (107).

A *Pseudomonas putida* isolate, B2017, is suggested as a potential biocontrol candidate that also inhibits *R. solani* on potato, as well as *Fusarium oxysporum* f.sp. *radices-lycopersici* in tomato, and *Sclerotinia sclerotiorum* on lettuce (108). Interestingly, the strain appears to lack almost all well-characterised determinants of such activity, which was ultimately linked to its production of the pyoverdine siderophore, and that the precise pyoverdine produced by this strain made it more potent than that isolated from *P. putida* KT2440 (109). The lipopeptide antimicrobials ecomycins B and C, isolated from *Pseudomonas viridiflava*, were tested against a panel of 12 plant-pathogenic fungi, including *F. oxysporum*, *R. solani*, and *S. sclerotiorum*, demonstrating effectiveness at 100 µg/mL or less in all cases (110). Pafungin, isolated from *P. aeruginosa* K-187, is a thermostable high molecular weight non-enzymatic antifungal that resembles a complex carbohydrate, and inhibited *Fusarium oxysporum* damping off of alfalfa seedlings, showing 80% survival rate (111). *Fusarium solani* on the other hand was inhibited by a *P. aeruginosa* protease, which prevented spore

germination and hyphal elongation (112). Müller *et al.* were able to demonstrate *in vitro* inhibition of fusaria and alternaria on moistened wheat grains, as well as demonstrating reduced production of mycotoxins like deoxynivalenol and zearalenones by *Fusarium culmorum* and *Fusarium graminearum* when co-inoculated with fluorescent pseudomonads; similar results were also reported for mycotoxin production by *Alternaria tenuissima* (113). Furthermore, mycotoxin degraders have also shown biocontrol potential (114). This provides promise not only for inhibition of pathogens in the field, but also improvement of grain quality in cereals, where mycotoxin contamination is a constant concern.

Müller *et al.* found that key antibiotic gene clusters like pyrrolnitrin, phenazines, pyoluteorin, and phloroglucinols were not sole predictors of antagonistic activity (113). However, in some cases individual molecules can have a profound effect, as was observed in *P. synxantha* strain 2-79, which demonstrated improved control of fungal pathogens of wheat when transformed with genes for the production of the antibiotic pyrrolnitrin, even *in planta* (115). Others have demonstrated *in planta* that strains of *P. fluorescens* (B1) can significantly ameliorate disease phenotype in potato and lettuce plants in field conditions (116). *Pseudomonas* sp. RU47 has also proven effective against *R. solani* (strains Ben3 & AJ868459) in both potato and lettuce models, even with variation of soil type, which can regularly be a confounding factor in biocontrol efforts (117). 2,4-DAPG is a well-known antifungal molecule produced exclusively by pseudomonads that has been demonstrated to inhibit *R. solani* (118) as well as Take-all disease of wheat caused by *Gaeumannomyces graminis var. tritici* (119-121), and has been linked to soil suppressivity to Tobacco root rot (57) and antagonism to *Verticillium* fungi (122). Suppression of *Erwinia amylovora*, the causative agent of Fire Blight, has also been observed in field trials by *Pseudomonas fluorescens* A506, but was only observed in a single trial (123).

Unfortunately, in many of these cases the definitive molecular interactions between fungus and pseudomonad are either not known or not investigated, especially for small molecules, and are usually limited to phenazines, HCN, and phloroglucinol compounds (122). Additionally, many determinants of inhibition co-occur in a single isolate and are often not simply additive, complicating matters further (7). As an example, the

Pseudomonas strain ST–TJ4 was found to produce cellulase, protease, 1-undecene, and harbour genes for production of phenazines, pyrrolnitrin, and hydrogen cyanide, all contributing to the inhibition of 11 phytopathogens tested, including *Fusarium* sp. *Phytophthora*, and *Rhizoctonia* (124). While multiple co-acting mechanisms of inhibition are critical for effective and stable inhibition, almost all studies indicate further uncharacterised molecules involved in suppression, leaving scope for discovery of novel antifungal natural products from *Pseudomonas*.

1.3.3 Oomycete Pathogen Inhibition

Pseudomonas isolates have well-documented abilities to inhibit oomycete pathogens of plants. *Phytophthora infestans* is a devastating pathogen of potatoes, causing late blight and tremendous loss of yield even with the use of fungicides (125). Many cyanogenic *Pseudomonas* isolates are able to completely inhibit the growth of *Phytophthora*, with 1-undecene producers also demonstrated to be effective (103, 126). Pyrrolnitrin has also been shown to have a role in combating late blight of tomato caused by *Phytophthora* (127). Phenazine-1-carboxylic acid produced by pseudomonads alters the transcriptome of *P. infestans*, crucially affecting the secondary metabolome. However, potentially problematic for biocontrol efforts was the observation that phenazine caused the upregulation of effector proteins in the pathogen (128) – demonstrating the need for *in planta* experiments when selecting and prioritising biocontrol candidates. Other oomycete pathogens such as the closely related *Phytophthora capsici*, causal agent of blight and fruit rot in various commercial species, is inhibited by *P. chlororaphis* volatiles (106), and *Phytophthora parasitica*, a pathogen of many plant genera ranging from vegetables to forest trees (129) is inhibited by *Pseudomonas* sp. EA6 by means of an excreted extracellular glucanase enzyme that is able to break down the cell walls of the oomycete pathogen. The conserved nature of the target makes resistance emergence unlikely in this case, and demonstrates the necessity for investigations in antimicrobial proteins produced by biocontrol strains, which have traditionally been overlooked in favour of small molecules or short peptides (100).

Chitinases are often identified in screens for bioactive compounds from *Pseudomonas*, but predicted to have low efficacy against oomycetes owing to the canonical absence of chitin from their cell walls. However, recent evidence indicates at least some water moulds temporarily display chitin in their walls, before conversion to cellulose, which may restore chitin as a potential target for biocontrol protein discovery against these pathogens (130). Recent reports of anti-*Phytophthora palmivora* proteins purified from a *P. aeruginosa* isolate were highly similar to a catalase, protease, protease IV, and chitin binding domain from the Global Proteome Machine Database, giving further indications that cell-free protein preparations may be of use when the producing organism is closely related to a human pathogenic strain, as in this instance. A mode of action was not proposed, but growth defects including swollen hyphal tips and short branched mycelium were demonstrated (131).

The promising finding of El-Sayed *et al.* (129) is of particular note, as their *Pseudomonas* isolate displayed activity against a further four *Phytophthora* species, and *Phytophthora* comprise six of the top 10 oomycete pathogens of plants (132). Eighth on the list is *Pythium ultimum*, the causative agent of root rot of various plants of economic importance, has also been demonstrated to be vulnerable to a beta-1,6-glucanase enzyme, albeit from a biocontrol fungus (133). Given these findings, future investigations on the biocontrol of oomycete pathogens via extracellular hydrolases, other enzymes, and natural products are warranted.

1.3.4 Arthropod Pathogen Inhibition

Interactions between members of the *Pseudomonas* genus and insects are extensive, as both insects and bacteria are amongst the most widely distributed taxonomic groups on earth. The range of interactions extends all the way from symbiotic to pathogenic, with countless examples between the two extremes (134).

Many pseudomonads are producers of hydrogen cyanide by glycine oxidation, as well as cyclic lipopeptides, and are in these cases lethal to insects upon injection, as demonstrated with *Galleria mellonella* (135). Plant-beneficial pseudomonads can also produce

insecticidal toxins, as both *P. protogens* CHA0 and Pf-5 were demonstrated to kill *Manduca sexta* and *Galleria mellonella* by production of a large protein toxin (Fit) (136). This was found to be related to the Mcf toxin from *Photobacterium luminescens*, being transferred to pseudomonads through horizontal gene transfer, and now restricted to the *protegens* and *chlororaphis* groups (137). Olcott and colleagues (138) were able to demonstrate not only lethality, but developmental and morphological defects of *Drosophila melanogaster* after being fed either CHA0, Pf-5, or SBW25. While the causative agents were originally not elucidated, it was concluded that the GacAS system crucial to the lethality of Pf-5. It was later established through a mutagenesis study that rhizoxin, a macrocyclic lactone synthesised by an NRPS-PKS hybrid (139), as well as orfamide A and chitinase contributes to oral toxicity of Pf-5 in *Drosophila* (140, 141). Ruffner *et al* further extended our understanding of the lethal effects of insect-ingested pseudomonads in various arthropod models (142). They were able to show that ingested CHA0 is able to kill larvae of *Plutella xylostella*, *Heliothis virescens*, and *Spodoptera littoralis* through Fit toxin production, under the regulation of GacA. However, for *P. chlororaphis* O6, Fit toxin production was deemed dispensable for killing of the root knot nematode *Meloidogyne hapla in planta* (143). Against green peach aphid nymphs, the same phenomenon was observed, with a role for HCN in aphicidal activity but no effect of Fit toxin, phenazine, or pyrrolnitrin (144).

Beyond natural products, secretion systems like the T6SS have also been demonstrated crucial to insect gut invasion and infection for some biocontrol strains (145). Plant growth-promoting rhizobacteria can even protect plants against arthropod pests indirectly, through activation of ISR pathways (146), as has been demonstrated for *P. fluorescens* in rice (147), okra (148), and *Arabidopsis* (149), and *P. putida* isolates against cucumber beetles in cucumbers (150). An overview of the distribution and phylogeny of toxicity to arthropods within the genus is presented in ref (151). So far, evidence for control of moths, flies and midges, and true bugs has been reported but the link to effects of individual compounds is complicated by multifactorial inhibition (152). If a more comprehensive understanding of molecular warfare in rhizosphere and phyllosphere interactions can be obtained in many of these cases, this may translate to the development of novel, environmentally safe, biopesticides. Furthermore, the ubiquity of *Pseudomonas* isolates in

the rhizosphere and phyllosphere (153) highlights the promise of bioprospecting for such compounds.

1.3.5 Limitations

Multiple limitations exist to developing *Pseudomonas* biocontrol formulations, from their lack of sporulation hindering long-term storage, regulatory hurdles and cost of development, to poor understanding of soil ecology in the context of agricultural systems and addition of biocontrol bacteria. Many preliminary assessments of potential biocontrol agents may take place in small-scale controlled environments, with sterilised soils or soil substitutes employed. However, there are multiple lines of evidence demonstrating that promising activity *in vitro* does not correlate perfectly with success *in planta* (19, 20). Such complications highlight the need for improved screening techniques or understanding of soil ecology in the agricultural system being investigated, including how the entire microbial community shifts in response to treatments (154). An understanding of molecular players in microbial interactions in the soil, and the factors that trigger their production, could vastly enhance efficacy of crop protection efforts. In instances where direct application of bacterial preparation to the plant is not efficacious, applying the key inhibitory natural product directly may prove preferable. In such a case, the compound must have suitable physical properties regarding stability, toxicity, solubility in appropriate solvents, and ease of production. These constraints mean very few bacterial natural products are commercialised for crop protection purposes.

1.4 Natural Product Discovery & Relevance to Biocontrol

The production of antimicrobial secondary metabolites was a key focus for novel antibiotic discovery in the 50s, 60s, and 70s, focusing around selective isolation procedures and novel molecule discovery. However, we may have drawn misleading conclusions about the functions of these molecules in the natural environment (155), which has the potential to lead us astray in biocontrol efforts for agricultural applications, where bacteria behave differently to a laboratory environment. Others have more recently developed our understanding of the role of these molecules in the natural environment and extended it

beyond simple antibiosis and antagonism (22), leading the way to a microbial ecology-based array of biocontrol strategies.

However, the future prospects of the traditional novel species-novel molecule approach are not as dim as they appeared just a decade ago, with the iChip enhancing biodiversity recovery (1) and leading to novel drug lead compounds (156) while our understanding of microbial secondary metabolism regulation deepens (157, 158). The potential of coupling this understanding to genome-mining approaches may bring about a second wave of antimicrobial discovery from environmental isolates (159). Applying “old fashioned” approaches to new environments like exotic soils, marine niches, or unusual environments has already yielded new compounds of interest (160-165). Other advances are similarly revolutionising the prospecting for useful microbial natural products with desirable properties.

1.4.1 Advances in Genomics for Discovery of Natural Product Gene Clusters

Perhaps some of the most notable advances have been *in silico*. Despite increased recovery of microbes from the environment, more clusters are detected bioinformatically than we can assign products to through fermentation and chemical experiments. The rapid expansion of genomic datasets provided the endless opportunity to mine genomes for natural product biosynthetic gene clusters, but such analysis was often difficult and tedious. Bioinformatic tools like antiSMASH (166) have revolutionised the way natural product researchers find new compounds, and show many strains have untapped biosynthetic potential in the laboratory (167). Clusters that specify the production of non-ribosomal peptides (NRPs) or polyketides (PKs), are easily identified through such analyses, and their modular, assembly line production system allows products to be predicted with high accuracy (168-170). Other tools are more specialised towards specific classes of natural product that may not be detected by antiSMASH or traditional genome mining, like RiPPER (171), RODEO (172), and DeepRiPP (173) for the detection of Ribosomally-synthesised Post-translationally modified Peptides (RiPPs). However, many still rely on similarity to known clusters, potentially reducing their ability to pick up true novelty in clusters and chemistry (174). The widespread uptake of these new tools across the field

has turned the field's prominent issue on its head, as microorganisms are now recognised to be far more gifted in secondary metabolite production than previously thought. As a result, there are now many candidate clusters with no known products (orphan clusters), requiring new approaches to activate the expression of silent biosynthetic gene clusters (BGCs) *in vitro* and link them to their molecular products (174).

1.4.2 Activation of Silent BGCs

The commonly accepted mantra in natural product discovery for explaining the gap between clusters identified *in silico* and products observed in bioactivity or chemical screening, is that such clusters are expressed poorly or not at all under laboratory growth conditions. A common approach to effect production of molecules is to try a range of different media in fermentation efforts, and analysing these cultures with metabolomic tools to probe for new molecules. Additionally, co-cultivation approaches may provide natural elicitors to stimulate natural product biosynthesis and enhance recovery of biosynthetic potential (175). Other strategies include using molecules from the strain's native environment to activate gene cluster expression, such a soil extract as in the case of closthioamide (165). Addition of soil extract to *Clostridium cellulolyticum* fermentations drastically changed the metabolite profile, with closthioamide simply being the most abundant in the culture. Rare earth metals scandium and lanthanum proved useful in the activation of cryptic gene clusters in *Streptomyces coelicolor* A3(2) (176). However, the concentrations of metals used in this case likely do not represent natural values in the environment, but highlights it is productive to be adventurous when examining the biosynthetic potential of a strain, even for one so well studied as *S. coelicolor* A3(2). Random mutagenesis approaches when combined with reporter systems like LacZ or GFP fusions show promise in uncovering the identity of factors involved in natural product biosynthesis that it would be difficult, if not impossible, to predict beforehand. Such an approach was used by Mao *et al.* to activate three BGCs in *Burkholderia thailandensis*, ultimately linking pyrimidine nucleotide limitation to production of malleilactone (177).

1.4.3 Prospects for Natural Product Discovery with Biocontrol Applications

Many studies in the field of biocontrol still employ low-throughput and so-called “antiquated” methods, often failing to elucidate the mechanism of *in vitro* antagonism or the molecular players involved. With this considered, the search for novel compounds of relevance to biocontrol efforts can learn from the wider field of natural product discovery, where significant strides are being made towards efficient and effective pipelines for finding novel chemistry with desirable activity. The contribution of this strategy to agricultural improvement and biocontrol is not well-surveyed, but genome mining strategies themselves have been employed to better understand the plant-microbe relationship (178). Doubtless, utilising a similar strategy to discovery of antimicrobials for human pathogens may at least contribute to the discovery of novel molecules to aid crop yields while decreasing agrochemical use; while employing high-throughput sequencing and computational approaches (54, 179) to understand microbial communities in plant disease may enhance biocontrol strain discovery and application.

1.5 *Pseudomonas* Natural Products

1.5.1 Overview

A 2016 analysis concluded that the *P. fluorescens* species complex contained a pan-genome of 30,848 coding sequences (CDSs), with only 1,344 core genome CDSs, demonstrating enormous inter-strain diversity, a figure which has likely grown since (180). As less than a fifth of the genes in *Pseudomonas aeruginosa* have been assigned functional biological roles (181), it is likely such a gulf between number of genes identified and number of genes assigned function also exists in the *P. fluorescens* complex. Significant potential for biocontrol, compounds regulating PGP, and biofertilisation lurks within this sea of uncharacterised protein coding sequences. A 2016 evaluation puts the known *Pseudomonas* secondary metabolites at 119 molecules from 30 different molecule families (182). Additionally, numerous unique features of biosynthetic clusters within this genus have been identified, exemplifying the untapped molecular and biosynthetic prospects within (Table 2 in ref (61)). The diversity of compounds produced by this group mirrors the

diversity of strains within the species. Key classes of biosynthetic gene clusters and example products of each will be discussed from the polyketides, non-ribosomal peptides, lipopeptides and cyclic lipopeptides, volatile organic compounds, ribosomally synthesised post-translationally modified peptides (RiPPs), as well as those products of unusual biosynthetic mechanism. Functional importance of example products will also be dealt with in short. Some examples are shown in Figure 1.3.

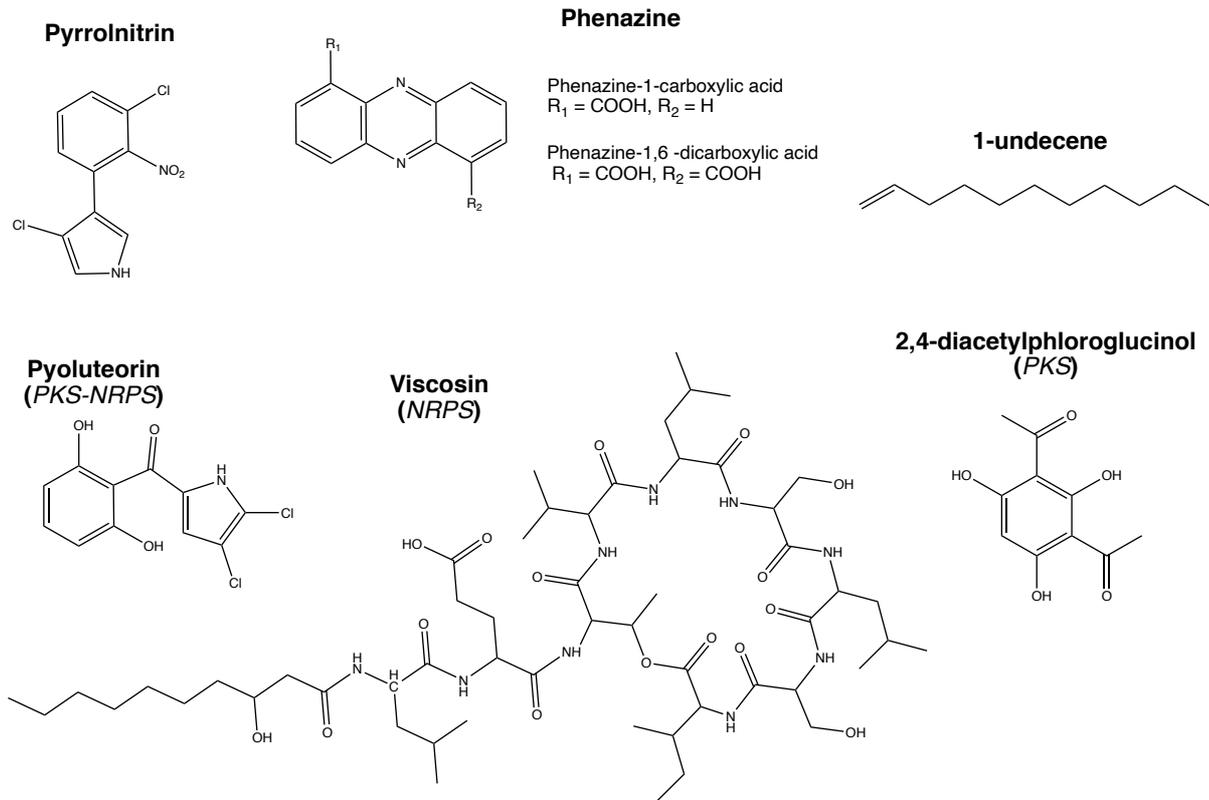


Figure 1.3. Notable *Pseudomonas* natural products with biocontrol activities. Representative molecules from the various molecular classes and biosynthetic pathways are shown. All have been shown to have antimicrobial activities relevant to biocontrol.

1.5.2 Polyketides: 2,4-diacetylphloroglucinol (2,4,-DAPG)

The polyketide 2,4-diacetylphloroglucinol (2,4-DAPG) (Figure 1.3) is a well-known phenolic antifungal produced by the *P. protegens* and *P. corrugata* groups (180), known to be important in suppressive soils to Take-All Decline (TAD) (183). The *P. protegens* species was proposed after observations that 2,4-DAPG-producing strains clustered separately from non-producing strains in the *P. fluorescens* group (184). In Take-all suppressive soils, certain agricultural practices allow the preferential expansion of populations of pseudomonads that produce this phenol to 10^5 CFU/g root tissue. 2,4-Diacetylphloroglucinol targets the mitochondrial proton gradient, disrupting the proton motive force in eukaryotes. Reported activities are antibacterial, antifungal, antihelminthic, antitumour, and even phytotoxic at high enough concentrations (185-187). Its role in biocontrol is already well-established, with activity against fungal pathogens of plants, and an ability to induce systemic resistance in the plant (188). In contrast, recent evidence suggests that 2,4-DAPG can act as a virulence factor in wheat and tomato models for pathogenic members of the *P. fluorescens* complex like *P. brassicacearum* Q8r1-96 (189). Relevant to biocontrol applications, there is evidence that host plant can affect transcription of this gene cluster, possibly hindering or helping pathogen inhibition depending on the host plant used (190).

Polyketides are produced by either Type I, II, or III Polyketide Synthases (PKS). These assembly line biosynthetic pathways use simple starting materials to produce complex and varied final products through decarboxylative condensation. Chains often originate from thioesters of monoacyl-CoAs, and extend via addition of malonyl-CoA or methylmalonyl-CoA building blocks (191). Type I PKS systems are highly modular, and type II PKS systems are aggregates of individual monofunctional proteins. Type III Polyketide Synthases differ from type I and II PKS systems in that they have no acyl carrier protein (ACP) domain, and are self-contained enzymes that act iteratively (192). 2,4-DAPG is biosynthesised by the latter, and this is encoded by a 6.5kb DNA fragment containing a four gene operon, *phlACBD*, for its production, which can be transferred to non-producing strains and allow them to produce 2,4-diacetylphloroglucinol (187). *phIE* is found downstream, under the influence of a σ^{54} -dependent promoter, and encodes a permease for export of 2,4-DAPG (120). *phIA* encodes a 3-oxoacyl-[ACP] synthase, *phIB* encodes a small protein showing

homology to DNA-binding proteins, and *phlC* has the hallmarks of an acetyl-CoA acetyltransferase (120). Together, PhIABC catalyse the conversion of the phloroglucinol precursor to diacetylphloroglucinol (193). PhID, the polyketide synthase in the cluster, is responsible for synthesising the precursor for the subsequent reaction, and shows homology to plant PKSs (187). By sequential condensation of three malonyl-CoA units, 3,4-diketoheptanedioate is produced (a polyketide), which is then cyclised to form phloroglucinol (194). PhIA, PhIB, PhIC are then required to convert this to the diacetyl form. Regulatory proteins are also found flanking the cluster. PhIF can block the biosynthesis of DAPG, by binding to *phlO*, a sequence upstream of *phIABCD*, and preventing transcription as part of a complex (195). Two additional genes, *phIG* and *phIH* are found alongside *phIF*. Where PhIG is thought to contribute to the degradation of DAPG (196), and PhIH demonstrates a form of transcriptional regulation of *phIA* (197). Other regulatory inputs to the system include the GacA/GacS two-component system (198), a global regulator of antibiotic production in *Pseudomonas*; RpoD, a housekeeping sigma factor whose overexpression increases 2,4-DAPG production (199); and RpoS, the stationary phase sigma factor, where deletion mutants overproduce this phenolic compound (200).

1.5.3 NRPS: Viscosin

Non-ribosomal peptides are produced by non-ribosomal peptide synthetases (NRPS), and are abundant within the pseudomonads, found in all species complexes (185). As the name suggests, they are peptide natural products produced non-ribosomally, and so are independent of mRNA. Accordingly, they may contain one or more non-proteinogenic amino acids, α -hydroxy acids, β -amino acids, and fatty acids (201). This facilitates enormous functional diversity, and reported activities are extensive – likely enhanced by further modification of the peptide post-synthesis (202). Antibacterial, antifungal, anti-oomycete, phytotoxic, antitumour, haemolytic, biosurfactant, anti-protozoan, antiviral, and cytotoxic capacities of non-ribosomal peptides have been observed, with some demonstrating three or more of these concomitantly (185). Notable products of these cluster types that are inextricably involved in biocontrol are the cyclic lipopeptides (CLPs) – many of which act against fungal plant pathogens (203). Rhizospheric pseudomonads are prolific producers of these biosurfactants, so have been the focus of much of the research

in this area. CLPs consist of an oligopeptide lactone ring with an aliphatic tail – making them amphiphilic molecules capable of disrupting cell membranes (203). The cyclic lipopeptide viscosin (Figure 1.3), produced by some environmental *Pseudomonas fluorescens* isolates, displays antibacterial, anti-oomycete, antiviral, antiprotozoal and biosurfactant properties (204). In addition, its expression aids PGP qualities in producing strains by enabling spreading motility, while also protecting against plant pathogens (205). Effects on biofilm dispersal have been noted through *viscA* mutant studies (206), which may further impact microbial competition in the soil. Viscosin-group lipopeptide producers also have potential in direct biocontrol efforts, having been demonstrated to lyse the zoospores of *Phytophthora infestans* (203, 207). The GacA/GacS two-component system in *Pseudomonas* regulates the production of LPs or CLPs, but the environmental conditions that trigger production are unknown. The viscosin class of NRPS CLPs contain 9 amino acid residues, but considerable diversity in length of cyclic moieties are observed (203).

The enzyme complexes present in these clusters lend themselves to genome-mining-based discovery efforts, and commonalities between different NRPS pathways have allowed for the development of bioinformatics tools to annotate these clusters, by looking for homology to known domains in sequence data (208). Core domains in these pathways are the Adenylation (A), Peptidyl-carrier Protein (PCP), and Condensation domains (C). A-domains catalyse the addition of a specific amino acid to the growing chain by activating it as an amino acyl adenylate, then transferring it to the PCP-domain. Specific conserved sequences within A-domains allow bioinformaticians to determine the amino acid it will adenylate and add to the chain, meaning the amino acid composition of the entire natural product can be loosely determined from sequence data alone (61). As the number and order of NRPS modules correspond to the number and sequence of amino acids in the final product, the term ‘colinearity rule’ was coined and relates to this phenomenon. The C-domains catalyse peptide bond formation between peptidyl intermediates and amino acids bound to the PCP-domains. The entire process is modular, successively adding amino acids to the growing chain, with molecule flux through sequential modules of A, PCP, and C-domains. Other domains may also be present, which can modify amino acids before they are added to the nascent molecule. Cyclisation (Cy) domains catalyse oxazole or thiazole ring formation, Reductase (R) domains reduce these to oxazolidine or thiazolidine

equivalents, Epimerisation (E) domains convert L amino acids to their D form, and Methyltransferase (MT) domains add a methyl group to nitrogen. Such additions to NRPS modules drastically enhance the diversity of end-products. Furthermore, loading modules and thioesterase-containing termination modules are required for a functional pathway (61).

Emerging methods for the modification of these pathways via the use of exchange units (XUs) (209), facilitated by their highly modular nature, to generate novel compounds with new / directed activity. Bozhüyük and colleagues' method (209) of using A-T-C tridomains fused in the C-terminal end of the C-A domain linker region to generate novel NRPSs without substantial loss of yield has dramatically enhanced the prospects within this category of natural products. This will no doubt be of interest to those working in the field of *Pseudomonas* natural products and development of biocontrol strains or compounds. With how prevalent these NRPS clusters are in *Gammaproteobacteria* (210, 211), there is substantial value in such an approach.

1.5.4 PKS-NRPS hybrids: Pyoluteorin

Some bioactive metabolites produced by *Pseudomonas* fail to follow the well-defined systems above, and are instead produced by hybrid systems. Pyoluteorin, a chlorinated aromatic compound (Figure 1.3), is one such molecule that also has long established bioactivity and potential in biocontrol with it and its derivatives having noted antibacterial and anti-protozoal (212), and anti-oomycete activities (213). Reference (213) details the observation of a producing strain inhibiting the important plant pathogen *Pythium ultimum*, which causes large economic losses in agriculture (61). There is evidence that production of pyoluteorin in laboratory culture represses the biosynthesis of the phenolic antimicrobial 2,4-DAPG (197). Analysis also demonstrated that it cannot be directly added to soil, as rapid adsorption and deactivation occurs - such issues may have precluded its commercialisation. Furthermore, observed toxicity to plants may limit applications (214).

As mentioned, a convergence of two biosynthetic systems occurs to produce pyoluteorin. The biosynthetic gene cluster (*plt*) consists of 17 genes spanning 30kb. A PKS system

generates a resorcinol ring, while NRPS machinery is used to form a dichloropyrrolyl residue. The type I PKSs PltB and PltC produce the resorcinol ring, but have a number of curious features including the lack of a loading module, and a number of defective domains (61). L-proline is the precursor of the dichloropyrrole, and is amino-adenylated by PltF to form L-prolyl-AMP, and then transferred to the phosphopantetheinyl arm of the PltL PCP. Subsequent prolyl desaturation occurs by the action of the PltE dehydrogenase, and the molecule is dichlorinated at positions 4 and 5 by the PltA halogenase to yield dichloropyrrolyl-S-PltL. The type I PKS PltB and PltC then add a further three malonyl-CoA units, and PltG cyclises the product to resorcinol to produce bioactive pyoluteorin (61).

1.5.5 Phenazines

This well-studied class of *Pseudomonas* natural products has broad antibacterial, cytotoxic, and antifungal activities, and function via the generation of reactive oxygen species (ROS) to cause oxidative stress (185). The basic skeleton is shown in Figure 1.3. With over 150 representative compounds (215) including the pigment pyocyanin, significant understanding of the genetics, biosynthesis, and natural roles of these molecules have been elucidated. The ability to produce phenazines has also been shown to be enriched in rhizospheric environments, highlighting their potential role in plant health (216). Due to their biosynthesis being distinct from common classes of natural products like NRPs and PKSs, an overview of the biosynthesis is shown in Figure 1.4.

All phenazine molecules contain nitrogen atoms and derive from a branching of the shikimate biosynthetic pathway, which usually produces the aromatic amino acids from phosphoenolpyruvate and erythrose-4-phosphate (217). In this process, two molecules of 2,3-dihydro-3-oxo-anthranilic acid are paired, producing phenazine-1-carboxylic acid (PCA). Genetic approaches have deduced a seven gene BGC with a core set of five enzymes, the *phz* operon, required for the production of either phenazine-1,6-dicarboxylic acid (PDC) or PCA from chorismic acid, but no genetic distinction has been elucidated to date to predict which of these two compounds a strain will produce. Further tailoring enzymes are required to produce downstream derivatives of these molecules, like pyocyanin. Blankenfheldt & Parsons reviewed the biosynthetic pathways for these compounds, and

readers are directed to their review for a more detailed description (215). The *phz* operon consists of the *phzABCDEFG* genes, which act sequentially to form PDC or PCA from chorismic acid. PhzC is a 3-deoxy-D-arabinoheptulo-7-phosphate synthase, involved in the biosynthesis of chorismic acid, the precursor for phenazine production (218). PhzE acts as a homodimer, where the glutamine amidotransferase of one subunit transfers ammonia to the menaquinone, siderophore, tryptophan domain of the partner. The reaction catalysed is the conversion of chorismate to 2-amino-2-desoxyisochorismic acid (ADIC). Unlike the related anthranilate synthases, PhzE is unable to convert ADIC to anthranilate, for reasons yet to be uncovered. ADIC is then acted upon by the hydrolytic enzyme PhzD, via an acid/base catalytic mechanism involving aspartic acid and lysine residues, to generate (5S,6S)-6-amino-5-hydroxy-1,3-cyclohexadiene-1-carboxylic acid (DHHA). PhzF subsequently isomerises DHHA to form a highly reactive aminoketone, (1R,6S)-6-amino-5-oxo-2-cyclohexene-1-carboxylic acid (AOCHC). AOCHC is toxic to cells, and is hypothesized to react with amines in proteins, so is rapidly removed via condensation by PhzB. Here, two molecules of AOCHC are converted to hexahydrophenazine-1,6,-dicarboxylic acid (HHPDC), which can then act as a substrate for the final reaction that produces PDC or PCA. It is noted that *Pseudomonas* species contain a duplicated *phzB*, *phzA*, that cannot catalyse this reaction *in vitro*, but plays a role nonetheless as mutants display a 75% reduction in production of HHPDC. Finally, the homodimer PhzG acts via either of two pathways to produce PDC or PCA via two oxidations or an oxidative carboxylation and an oxidation respectively. Others have demonstrated that, as this enzyme can catalyse either, it is the relative concentrations of enzymes in the pathway (PhzF, PhzB, PhzG) as well as the oxygen availability that affect which is produced or the balance between the production of both. However, recent evidence suggests PDC can actually form spontaneously in the absence of PhzG (219).

Evidence from the laboratory of Lars Dietrich has deepened our understanding of the natural function of these redox active antibiotics, demonstrating roles in colony morphology, coupling phenazine production to cyclic-di-GMP levels and alginate biosynthesis (23). Their work demonstrates that phenazines function as extracellular electron shuttles, able to oxidise NADH to NAD⁺, and sustain respiration in the anoxic layers of biofilms. These advances demonstrate the need to re-evaluate how natural products

researchers approach secondary metabolism, and further focus on how these molecules may function in natural, non-laboratory, systems. For further insight on re-examining secondary metabolism in natural microbial systems, readers are directed to this older but outstanding review (22).

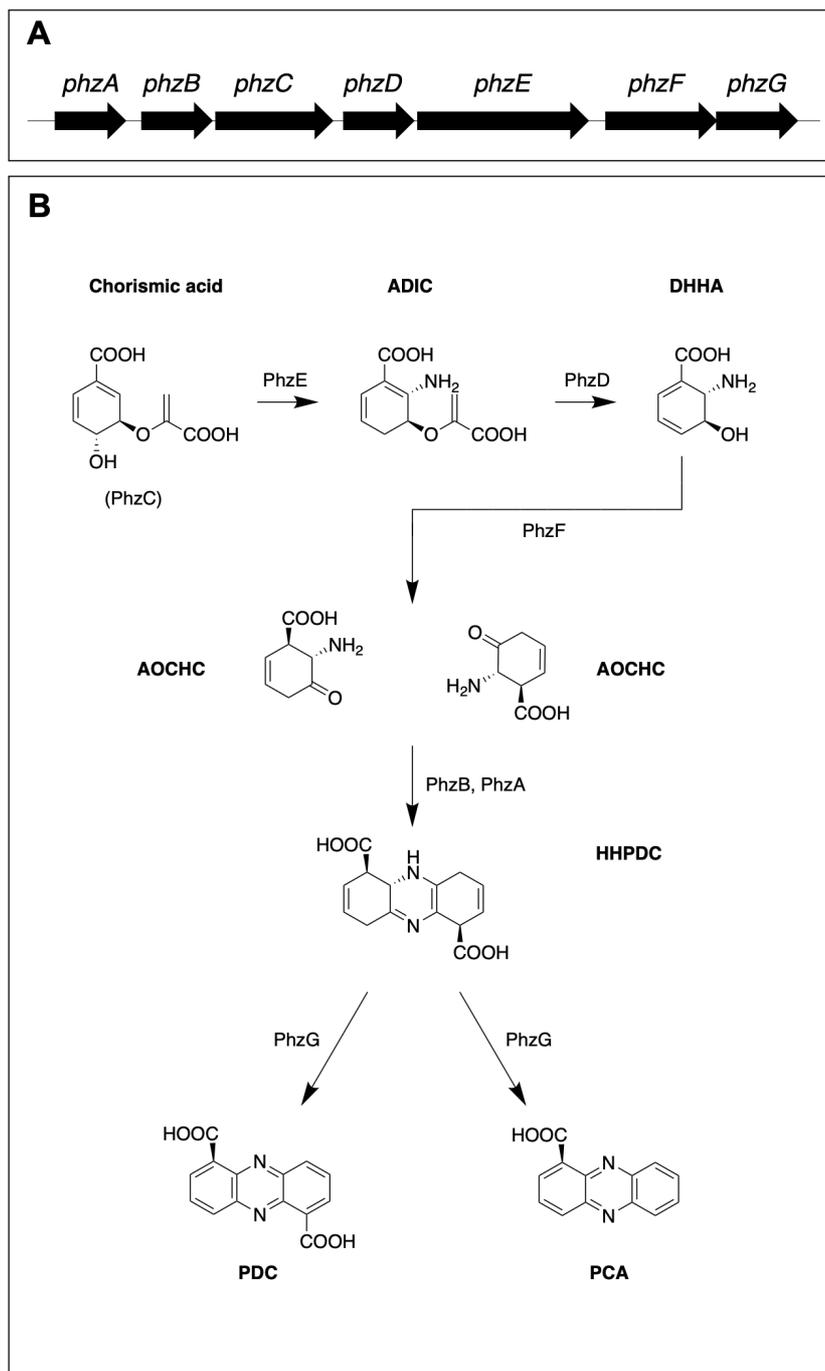


Figure 1.4. The biosynthesis of phenazine natural products in *Pseudomonas*. (A) A seven-gene operon, *phzABCDEFG* is responsible for the production of phenazines. (B) PhzC acts to provide chorismic acid, the precursor for this pathway. PhzE, PhzD, and PhzF then act sequentially to produce (1 R,6 S)-6-amino-5-oxo-2-cyclohexene-1-carboxylic acid (AOCHC). Two molecules of AOCHC are condensed by PhzB, aided by PhzA through an unknown mechanism, yielding a single molecule of hexahydrophenazine-1,6,-dicarboxylic acid (HHPDC). PhzG then acts to produce either phenazine dicarboxylic acid (PDC) or phenazine carboxylic acid (PCA), although there is evidence PDC may form spontaneously. The relative concentrations of PhzF, PhzB, and PhzG, as well as oxygen availability affects the balance of PDC and PCA.

1.5.6 Volatile Organic Compounds: 1-undecene

Volatile organic compounds (VOCs) like acids, sulphur compounds, ketones or alcohols, hydrocarbons, terpenes, and nitrogen-containing molecules or inorganics like hydrogen cyanide, ammonia, and nitrous oxide have all been noted to be produced by bacteria (220). Antifungal, anti-oomycete, and anti-arthropod activities have been noted, but there are few reports of an antibacterial volatile produced by another bacterium in the literature.

Recently however, a novel class of long chain aliphatic nitrile volatiles was reported, with activity against staphylococci (221). There are however plentiful accounts of antifungal volatiles, but their modes of action or bacterial producers are mostly uncharacterised (104). The volatile organic hydrocarbon 1-undecene (Figure 1.3), is a well-known inhibitor of many plant pathogens, including *Phytophthora infestans*. Volatiles, particularly HCN, but including sulphur compounds and long-chain ketones, have been demonstrated to be particularly effective against *Phytophthora*, and will likely play a large role in future attempts at biocontrol of this devastating pathogen (103). Biocontrol prospects are enhanced by observations that volatiles produced by rhizospheric bacteria can also both promote plant growth and induce system resistance to pathogens (222-224). Other volatiles produced by pseudomonads, like hydrogen cyanide, have also been demonstrated crucial to phytopathogen inhibition including arthropods (135) and cyanogenic *Pseudomonas* strains have the strongest inhibitory effect on *P. infestans* (103). Here we shall focus on 1-undecene, which Hunziker and colleagues observed to be a potent inhibitor of this potato pathogen (103), and an analysis by Rui *et al.* (225) demonstrated it to be ubiquitously distributed throughout the entire *Pseudomonas* genus – curiously including phytopathogenic species, where it likely functions to inhibit competing fungi that may be faster growing (226). Its production is regulated via the GacAS two-component system, coupling environmental conditions to production of anti-phytopathogen compounds (227) and explaining heterogenous production on different media (226).

The biosynthetic mechanism remained elusive for a long time, possibly because few were focusing on the role of bacterial volatiles in biocontrol. Its biosynthesis by bacteria was eventually elucidated by Rui *et al.* in 2014 (225), aiming at discovering a method of

converting fatty acids to hydrocarbons as a source of renewable fuels. Through a number of convergent techniques, they discovered a single gene encoded a member of a novel family of non-heme Fe(II) enzymes that act via oxidative decarboxylation of fatty acids to produce hydrocarbons. ¹³C feeding experiments using potential substrates found lauric Acid, a 12-carbon fatty acid, to be the precursor of 1-undecene (C:11) biosynthesis, with the terminal CO₂ being removed to form the product. *In vitro* reconstitution of the reaction discovered the need for Fe(II) as a cofactor, and that the reaction consumed O₂ and stalled when not available. The gene was termed *undA*, and was found to be capable of transferring the capacity to produce 1-undecene to *E. coli* when expressed on a fosmid. It encodes a 261 amino acid protein, UndA, with some substrate promiscuity, capable of converting fatty acids no longer than 14 carbons or not shorter than 10 carbons were to their [M-1]-carbon alkenes. The proposed biosynthetic mechanism involves binding of lauric acid to Fe(II), subsequent O₂ binding to produce Fe(III)-superoxide, then abstraction of the Lauric acid β-hydrogen. A single-electron transfer then generates 1-undecene along with CO₂ and water (225).

While Bacterial VOCs may display promise as potent inhibitors of various phytopathogens (104, 228), and benefit from an improved diffusibility in the soil microenvironment when compared to soluble molecules (220), methods for the discovery and characterisation of novel VOCs are limited and laborious (229). Such metabolites generally fall into ketones, alcohols, hydrocarbons, acids, terpenes, and sulphur or nitrogen-containing compounds (229). While distinguishing between volatile or soluble compound mediated inhibition can be achieved by simple use of barrier plates, identification of VOCs is troublesome. Gas chromatography coupled to mass spectrometry is the go-to in this area, and a workflow is provided by Audrain and colleagues (229). Headspace analysis GC-MS (229) is widely applied to monitor production of volatiles and is the gold-standard at present. Other methods are emerging that allow real-time analysis of volatile production, like selected-ion flow-tube mass spectrometry (SIFT-MS) (230). It is, however, not without issue. Analysis of the volatilome of 24 *P. aeruginosa* clinical isolates revealed 391 non-redundant compounds through GCxGC-TOFMS analysis (231). With such a variety of volatiles, not only is it extremely difficult to disentangle the effects of each individual molecule, but obtaining a pure molecule (if the compound is novel) for characterisation and further testing proves

complex. A headspace analysis approach to biocontrol molecule discovery was taken by Fernando and colleagues (228). In their method, nitrogen gas was passed continuously over shaking cultures for 48 hours, and the VOCs trapped in activated charcoal. Compounds were then eluted from charcoal with methylene chloride and analysed by GC-MS. The authors demonstrated the antifungal effects of pure compounds identified in their analysis. Headspace analysis was recently advanced through the implementation of semisolid growth surfaces in headspace vials, which can be sampled directly by an autosampler for GC-MS (232). This seemingly small step forward will allow for higher throughput, automated screening of samples, and was demonstrated to enhance VOC recovery over plate-based techniques.

1.5.7 Ribosomally-synthesised Post-translationally Modified Peptides (RiPPs)

RiPPs are genetically encoded peptide natural products that undergo post-ribosomal modifications. Biosynthesis occurs via the ribosomal route, translating coding mRNA into the precursor peptide product. The nascent peptide then undergoes further post-translational modification of specific amino acids in the sequence, generating significant diversity in the resultant molecules. Isomerisation, methylation, and cyclisation are some such elaborations to the linear peptide sequence. Precursor peptides contain both leader peptide and core peptides, with C-terminal extensions also found in some cases. The core peptide is that which is subject to modification and will be released as the final bioactive metabolite. The leader peptide is found N-terminal to the core peptide, and functions to recruit modifying enzymes to the core peptide. This system allows staggering diversity of final molecules, permitting mutation events to occur with the core peptide sequence without limiting the ability to recruit modifying enzymes (233). The leader peptide is finally cleaved from the mature core peptide by a protease and the bioactive RiPP transported to its site of action. Precursor peptide sequences are often followed by a suite of modifying enzymes in an operon, flanked by transporter genes. A detailed description of this class of natural products may be found in ref. (233).

Many ribosomal peptide antimicrobials are found in the *Pseudomonas* genus, but are frequently larger, unmodified peptides called bacteriocins. The pyocins from *P. aeruginosa*

are one such example, and comprise the F-Type pyocins, the R-Type pyocins, pyocins S1-S7, and the M1 pyocins. All have been reported as having only antibacterial activity but have diverse targets and molecular size ranging from 289 to 777 amino acids, demonstrating their potential. Targets of pyocins include the cell membrane, transfer and ribosomal RNA, DNA, and Lipid II (185). At these sites they usually act via degradation of DNA, RNA, Lipid II, or pore formation (234). *Pseudomonas* pyocins have mostly been associated with type V and Type VI secretion systems, where they act as the toxin effectors (234).

Many of these typically have a narrow target strain range affecting only a certain bacterial species or close relatives of the producer itself (234) in the case of the pyocins. This offers considerable prospects for biocontrol, where off-target effects would constitute barriers to application of novel strain or compound formulations for use in agriculture. The modular nature of their targeting and toxicity domains has also facilitated hybrid molecule development with novel activities (235), and it is easy to envisage the creation of such hybrids to target specific plant pathogens. Accompanying these gene clusters are co-synthesised immunity proteins that bind and prevent toxicity to the producing organism.

In contrast with the relatively abundant bacteriocins, which are unmodified ribosomally encoded peptides, few RiPPs have been characterised as being produced by the *Pseudomonas* genus. Microcin B-like compounds have also been found in genomes of *Pseudomonas syringae*, and produced heterologously in *E. coli*, indicating RiPPs can be found within the genus and more likely await discovery (236). A RiPP-like antimetabolite was recently characterised (237) encoded in the genome of a plant pathogen, *Pseudomonas syringae* pv. *maculicola* ES4326. This example not only extended the known post-translational modifications, but a new method of peptide small molecule biosynthesis was elucidated. The leader peptide, PmaA, but not the final product, 3-thiaglutamate, was genetically encoded. A short LanB-like protein (PmaB), adds a C-terminal peptide bonded cysteine residue to PmaA, which is then modified by other genes in the cluster, before its final release as 3-thiaglutamate, reforming PmaA for a subsequent round of biosynthesis. No activity has yet been reported for this metabolite. However, such discovery highlights

the extensive potential of this genus for displaying novel biochemistry currently unknown in natural products research.

Characterised production and diversity of these molecules is well known in prolific antimicrobial producers like *Streptomyces*, but is understudied in the *Pseudomonas* genus, and represents a potential reservoir of bioactive metabolites from a group of microorganisms found to be a major component of the rhizosphere. With the advent of bioinformatics tools like RiPPER (238, 171) and RODEO (172) to detect these gene clusters, efforts can be made to use genome mining approaches for the discovery of novel biocontrol candidate molecules in this class. Furthermore, the plasticity and tolerance of tailoring enzymes (239) with regards to the core region of the precursor peptide is predicted to enhance the synthetic biology potential of these biosynthetic pathways, allowing tailor-made biocontrol molecules to be designed with suitable biochemical properties and specificities. Modifying and reconstructing biosynthetic gene clusters in *Pseudomonas* has been aided by the identification and improvement of *P. putida* KT2440 as a versatile heterologous host (240), and the development of one-step recombinatorial cloning in yeast followed by transposase-mediated chromosomal integration (241).

1.6 *Phytophthora infestans* & *Streptomyces scabies* as Important Potato Pathogens

With potato crops constituting the fifth largest food crop world-wide (242), the world's population projected to continue increasing until 2100 (243), and shifting attitudes towards chemical pesticides, there is a requirement for development of novel strategies to mitigate food crop loss through phytopathogen infection. *Phytophthora infestans* and *Streptomyces scabies* are two critical potato plant pathogens, with limited mitigation strategies. Current understanding of their biology will be reviewed, along with an overlook of strategies being developed to deal with them.

1.6.1 *Phytophthora infestans*

1.6.1.1 Overview

P. infestans is one of various species in the *Phytophthora* genus, which encompasses many devastating plant pathogens. As mentioned previously, this genus accounts for 6 of the top 10 oomycete pathogens of plants (132), causing many plant diseases (244). Oomycetes are described as being spore-forming fungus-like eukaryotes that can reproduce sexually or asexually, a factor that dramatically increases their pathogenicity. Oomycetes fall within the kingdom *Stramenopila*, being closely related to diatoms, rather than true fungi. Unlike fungi, oomycetes are diploid, lacking the true haploid life stage (245). A heterothallic organism, it exists as two known mating types, A1 and A2, and sexually produced oospores occur when mycelia of the two mating types interact (246). Like bacterial spores, these oospores are thick-walled and promote survival outside the host in soil or waste plant matter (247), for up to several years (248). *Phytophthora* is known to differentiate into 11 different cell types, with different specialisations (249). This finding highlights the potential complexity of infection, while also providing more potential drug targets. At the genetic level, ~74% of the genome is gene-poor and repeat-rich, and contains hundreds of putative disease effector proteins (250). Analyses have demonstrated this region shows elevated transcription during infection stages, and also is more rapidly evolving than the gene-rich, repeat-poor core region of the genome (251). A comprehensive description of the pathogen's life cycle may be found in ref (252).

P. infestans is infamous as the causal agent of the 19th century Irish Potato Famine, as well as other similar famines in Europe at the time. It is estimated that 1 million deaths occurred in Ireland alone, with a further 1 million individuals emigrating abroad. The disease it causes, known as Late Blight, is not limited to potatoes, but can also affect other members of the *Solanaceae*, like tomatoes. Originating in Mexico (253), it first occurred in the United States in 1843 (254), soon thereafter spreading to Belgium via imported seed potatoes, and later to the rest of Europe and the UK. It is now the most costly potato pathogen to control in the world, costing billions to control annually (255, 256), and it is predicted near-eradication of the disease would increase annual potato yield by almost 80 million tons

(256). Disease is initially difficult to detect, with only minor lesions, sometimes as small as 1-2mm, appearing on stems and leaf tips in 3-4 days, and not all plants are affected simultaneously (Figure 1.5). In later stages of infection, these lesions enlarge, and the white hyphae may be visible on the abaxial surface of leaves. The collapse of the entire plant can occur merely 5 days after infection, providing a prohibitively small window for intervention, even with close monitoring. In some cases, crop losses of 100% can occur, and in only 7-10 days (254) (Figure 1.6). Adding to concerns, emergence of phenylamide-resistant isolates has occurred (257, 258), with additional evidence that environmental conditions may limit the effectiveness of applied fungicides. It is evident that new methods of controlling this pathogen must be developed, to mitigate losses to both economic and food productivity.



Figure 1.5. *Phytophthora* lesion on a potato leaf. Early signs of *Phytophthora* infection of potato plants are small lesions on leaves that are easy to miss, but can progress to crop failure in a matter of weeks. Image courtesy of G. Tomalin of VCS Potatoes.

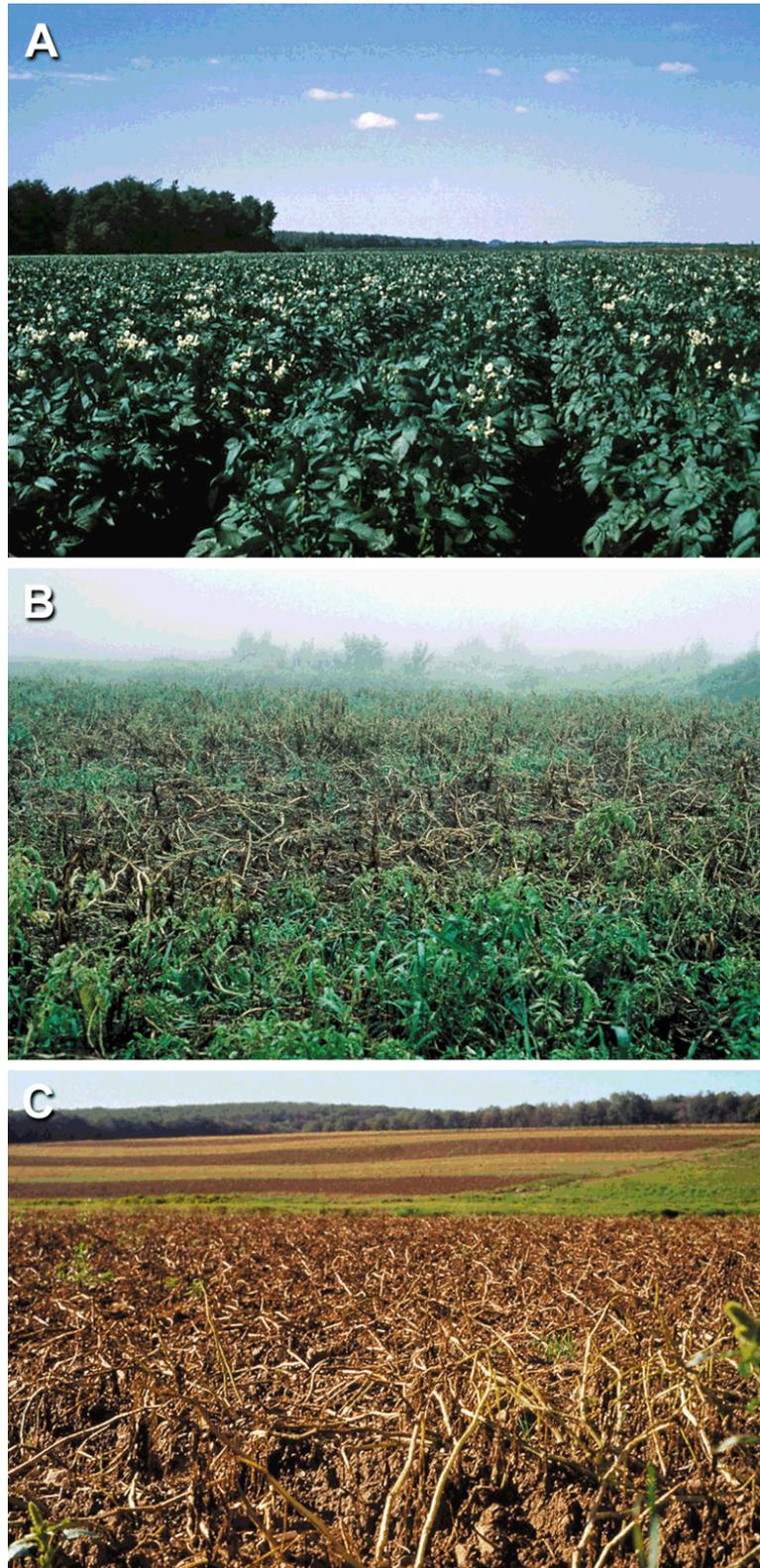


Figure 1.6. Rapid progression of potato late blight disease over a two week period. Plants that appear healthy and free of disease one week (A) can be visibly diseased the next week (B) and totally destroyed just a week after that (C). Reproduced with permission from Fry et al. (2008). *Phytophthora infestans*: the plant (and R gene) destroyer. *Mol. Plant Path.* 9(3). 385-402.

1.6.1.2 Plant Response

Defensive responses in the plant in response to infection occur as a two-stage process. Initially, detecting mechanical pressure at the site of *Phytophthora* penetration peg, plants mobilise actin microfilaments in a reversible manner, forming a radially-organised support structure under developing appressoria. This process is accompanied by callose formation, secretion of enzymes, and a ROS burst. If hyphae succeed in penetrating the plant, a pathogen-specific immune response occurs through recognition of pathogen associated molecular patterns (PAMPs), inducing PAMP-triggered immunity (PTI), with accompanying programmed cell death. A more detailed description of infection from the perspective of both plant and *Phytophthora* is available in ref. (254). For an overview of the cell biology of *Phytophthora* infections in late blight, readers are directed to the comprehensive review of Whisson *et al.* (259).

1.6.1.3 Fungicides

Control of *Phytophthora*-based crop diseases are traditionally managed using fungicides, which may be applied as many as 25 times per growing season in Europe (260). Fry cites the US National Agricultural Statistics Service claiming more than 2000 tonnes of fungicides were used in the US in 2001 for the control of *P. infestans* (252), constituting a great expense to both growers and the environment. Field trials suggest that this usage could be reduced by up 80% if cisgenic crops were used to stack resistance genes (256). The broad-spectrum nature of many commercially applied fungicides means resistance rapidly follows deployment of new fungicides, as was observed with metalaxyl. However, some U.S. isolates are still sensitive to some commonly used oomycete-targeted fungicides (261).

Resistance to FRAC-group (Fungicide Resistance Action Committee) fungicides in A1 and A2 mating type *Phytophthora* isolates from the United Kingdom has been assessed, belonging to the following groups: Pyridinylmethyl-benzamides, CAA-fungicides, Carbamates, Qil fungicides, and uncouplers of oxidative phosphorylation. For isolates relatively new to the region, like 36_A2, resistance to many of these compounds was higher (262). The report cites numerous cases of isolates resistant to many of these compounds on the European

continent and the U.S., indicating resistance to fungicides is an ongoing issue, and in some cases provides isolates with a selective advantage (263).

No fungicide currently available to growers targets a specific and essential process in *Phytophthora* that not only would it bypass environmental harms from using broad spectrum agents, but resistance emergence would be prohibited by deleterious effects of mutations to alter or circumvent the targeted process. This is the 'silver bullet' approach, and it is not clear whether such an agent could exist.

1.6.1.4 Breeding for Resistance

Transfer of resistance (R) genes from wild relatives of the cultivated potato *Solanum tuberosum*, such as *S. bulbocastanum*, *S. okadae*, *S. demissum*, and *S. stoloniferum* has been explored as a means of mitigating late blight (Figure 1.7). However, there is significant difficulty in crossing wild potato species with the cultivated potato, owing to the tetraploidy and extensive heterozygosity of domesticated potato (252), and many demonstrate low efficiency (254). Wild potato varieties may also differ in ploidy from *S. tuberosum*, and may not permit direct hybridisation efforts. Such constraints meant that it took 49 years to develop the Bionica cultivar used commonly in organic agriculture, containing only a single R gene from the wild relative (256). However, there have been recent advances in accelerating transfer of R genes from wild relatives for transfer to cultivated potato (264), leveraging novel methods for R gene identification in poorly annotated sequences (265), combined with modern sequencing techniques. Furthermore, Fry noted that, given the 10-15 year period of developing a cultivar, and the rapidity with which *Phytophthora* circumvents R gene-mediated resistance, no R gene transfer strategy has contributed extensively to mitigating potato late blight in practical terms (252). Some hope that combining various resistance genes in a single cultivar, known as 'pyramiding', will extend the longevity of resistant lines in the field. At least 20 R genes are now available for transformation into *S. tuberosum* (266). Unfortunately, this is considered difficult to achieve with conventional breeding approaches (267), and cisgenesis has been proposed as a means to achieve such durable resistance. In cisgenesis, plants are genetically modified but only with naturally existing genes from the crop plant or sexually compatible relatives,

speeding up cultivar development and eliminating the off-target effects associated with conventional breeding (268). Unfortunately, regulatory constraints still hamper the application of such an approach (267, 269). Additionally, there is no guarantee this strategy will succeed, as historically isolates insensitive to newly discovered R genes have been observed, despite lack of application of these R genes in cultivars (270). Recently, a CC-NLR protein R-gene, Rpi-amr1, was identified in *Solanum americanum* that was able to confer resistance to 19 isolates of *P. infestans* when transferred to Maris Piper potato plants, offering hope for durable resistance through R-gene-based strategies (260).



Figure 1.7. Breeding for late blight resistance in potatoes. Two cultivars are shown side by side in an organic commercial potato field. On the left is a non-resistant cultivar, and on the right is a cultivar with an R gene introduced by natural breeding. Image courtesy of G. Tomalin of VCS Potatoes.

1.6.1.5 Forecasting

As free water availability is crucial to infection, forecasting measures have been developed to assess risks to crops and optimise mitigation strategies. Most commonly cited are the Beaumont Period (271), which improved on the original 'Dutch Rules', and the subsequent Smith period (272). A Beaumont period is defined as 48 hours in which temperature does not drop below 10 °C and relative humidity exceeds 75%. A Smith period is defined as a minimum two consecutive days where minimum temperature is 10 °C and on each day relative humidity exceeds 90% for 11 hours or more. Some suggest given that populations of *Phytophthora* have changed in the half century or more since these periods were proposed for late blight risk prediction, the indices themselves should be updated to better predict risk (273). In 2016, the Hutton criteria were presented at an agronomist's meeting, providing a much needed improved forecasting of disease outbreaks (274). The updated system decreased the minimum period of >90% humidity from 11 to 6 hours, which significantly improved past outbreak prediction, and removed much of the spatial variability in performance of the previous criteria. Still, only 69% of historical outbreaks were predicted by this system, highlighting there is further progress to be made. Other measures have been developed that provide improved forecasting during periods of atypical weather (275). These measures are typically used to direct and inform application of fungicides in a prophylactic manner, as applied fungicides are generally not considered curative. For many of the fungicides, a limited number of applications can be performed per season, due to resistance concerns. Accordingly, timing application to coincide with the period of greatest risk is ideal (276). However, it must be noted that forecasting is not 100% accurate (274), and combined with emerging resistance to applied fungicides, may have limited scope in future control of the disease.

1.6.2 *Streptomyces scabies*

1.6.2.1 Overview

Streptomyces scabies, or *Streptomyces scabiei*, is a pathogenic member of the family *Streptomycetaceae*, a filamentous soil bacterium found in soils worldwide. The type strain

was described in 1989 by Lambert and Loria (277). Unlike other members of this family, it is plant pathogenic rather than saprophytic, infecting a wide range of vegetable tubers and tap root crops (278). Morphologically, it grows as branched filamentous hyphae with relatively few cross walls, and forms aerial hyphae with spiral chains of 20 or more smooth grey spores (Figure 1.8A). These dormant spores are resistant to desiccation allowing long-term survival in soils as an enduring source of inoculum, or may be transported in water or by eukaryotic vectors enabling disease spread (279). With a genome size of 10,148,695bp (Genbank accession FN554889), *S. scabies* 87-22 has the one of the largest genomes of any streptomycete. Genes involved in infection are widespread in the genome but cluster in two notable regions, the toxicogenic and the colonisation regions (278). Occurring wherever potatoes are cultivated, it is perhaps best known as the causative agent of potato common scab. Selective isolation of actinomycetes from scab lesions has also demonstrated an array of distantly related *Streptomyces* species able to cause common scab (278), a phenomenon which was observed in the original description of the species (277). This list may be expanding all the time, as pathogenicity is determined by a mobile 660kb pathogenicity island that may be transferred by mating to non-pathogenic species, rendering them able to infect plants. However, not all that receive the pathogenicity island turn plant-pathogenic, and it is evident there are further determinants of infectivity (280).

A survey of pathogenic and non-pathogenic isolates in the United States demonstrated that a nitrated dipeptide of tryptophan and phenylalanine called thaxtomin (Figure 1.9), is essential for plant infection and was observed in all pathogenic isolates (281). Presence of NRPS genes for its production (*txtAB*) can be used to detect scab-forming pathogens in environmental samples (282). Thaxtomin A is the predominant phytotoxin produced in scab-causing isolates, containing a 4-nitroindole moiety that is essential for activity. The current biosynthetic model begins with a nitrated tryptophan residue, which is condensed to a phenylalanine residue, forming the diketopiperazine ring, resulting in thaxtomin C. Thaxtomin C is then N-methylated to form thaxtomin B, and finally hydroxylated to form thaxtomin A (283). The origin of 4-nitrotryptophan remained a mystery until it was revealed *txtE* encoded a cytochrome P450 enzyme capable of nitrating L-tryptophan using nitric oxide and oxygen (284). Stimulated by cellobiose and suberin (285), this phytotoxin is predicted to function by inhibiting cellulose biosynthesis (286), and ultimately induces

programmed cell death (287). While thaxtomin is the primary virulence determinant, other factors have been demonstrated to impact pathogenicity. Recently, an additional phytotoxin was found to be produced by this pathogen, N-coronafacoyl-L-isoleucine (288), which is thought to contribute towards disease severity, but is not essential for infection and is also found in non-pathogens (289). This class of phytotoxins consist of a polyketide-derived bicyclic ring, linked via an amide bond to an amino acid, and appear to be common in many plant pathogenic bacteria, including various other pathogens of *Solanaceae* (289).

Also in *S. scabies*, *nec1* encodes for a necrogenic factor, that is theorised on the basis of GC content to have been transferred horizontally from another genus (290, 291), and enhances pathogenicity as a secreted virulence protein (292). Additionally, *S. scabies* 87-22 encodes for TomA, a secreted hydrolase of the plant antimicrobial α -tomatine, allowing it to subvert plant defences (293). Some propose the true function of tomatinase enzymes is to downregulate induced plant defences, through a degradation product of α -tomatine hydrolysis (294). Despite Ca^{2+} and H^+ influx during infection, it appears this stems from a hypersensitive response rather than a classical plant defence response. Like many rhizobacteria, *S. scabies* is capable of producing plant hormones like indole-3-acetic acid (IAA) (295); however, it appears this pathway is not negatively regulated by tryptophan availability, as thaxtomin production is. Excess tryptophan therefore induces IAA production, reduces phytotoxin production, and leads to reduced disease symptoms in plants (296). However, other studies suggest the reduced disease symptoms also stem from enhanced thaxtomin tolerance when IAA levels are higher, rather than simply being a function of lower phytotoxin levels (297). Overall, presence of *txtAB* genes is the most commonly used metric for identifying and quantifying scab-causing streptomycetes in soils.

Common Scab (Figure 1.8B) as a disease presents differently depending on severity. Sparse, colourless corky lenticels may be present in less affected tubers, progressing to darker raised or pitted lesions in more heavily affected potatoes. Scab-infected potato tubers are not deemed inedible, nor is their consumption considered a risk to human health. Unlike Late Blight, it is unlikely to put farmers out of business, or vastly reduce food productivity. However, crop quality is considerably reduced and has tangible economic impact (298, 299). The Potato Council estimate that roughly 87% of UK potato crops experienced some

level of common scab, and claim that 1M tonnes of potatoes are wasted each year, to which common scab contributes (300). The United States Department of Agriculture sets objective standards for determining damage to potatoes by scab, deeming those with over 5% surface area coverage to be 'damaged', and those with over 25% to be 'seriously damaged' (301). Environmental conditions are implicated in disease onset, as with potato Late Blight. Soils may be conducive or suppressive to potato common scab, influenced by biotic and abiotic factors. It is widely recognised that a pH above 5.2 is a risk factor for *Streptomyces scabies*, but may not be a limiting factor for *S. acidiscabies*. Additional risk factors include temperatures of 20-22 °C and less than 70% soil moisture (301). Abiotic risk factors include high soil carbon, nitrogen, calcium, and iron, as well as high C/N ratio (179). At present, there is no consistently successful approach to managing common scab, but breeding of resistant potato cultivars is the most reliable approach (302).

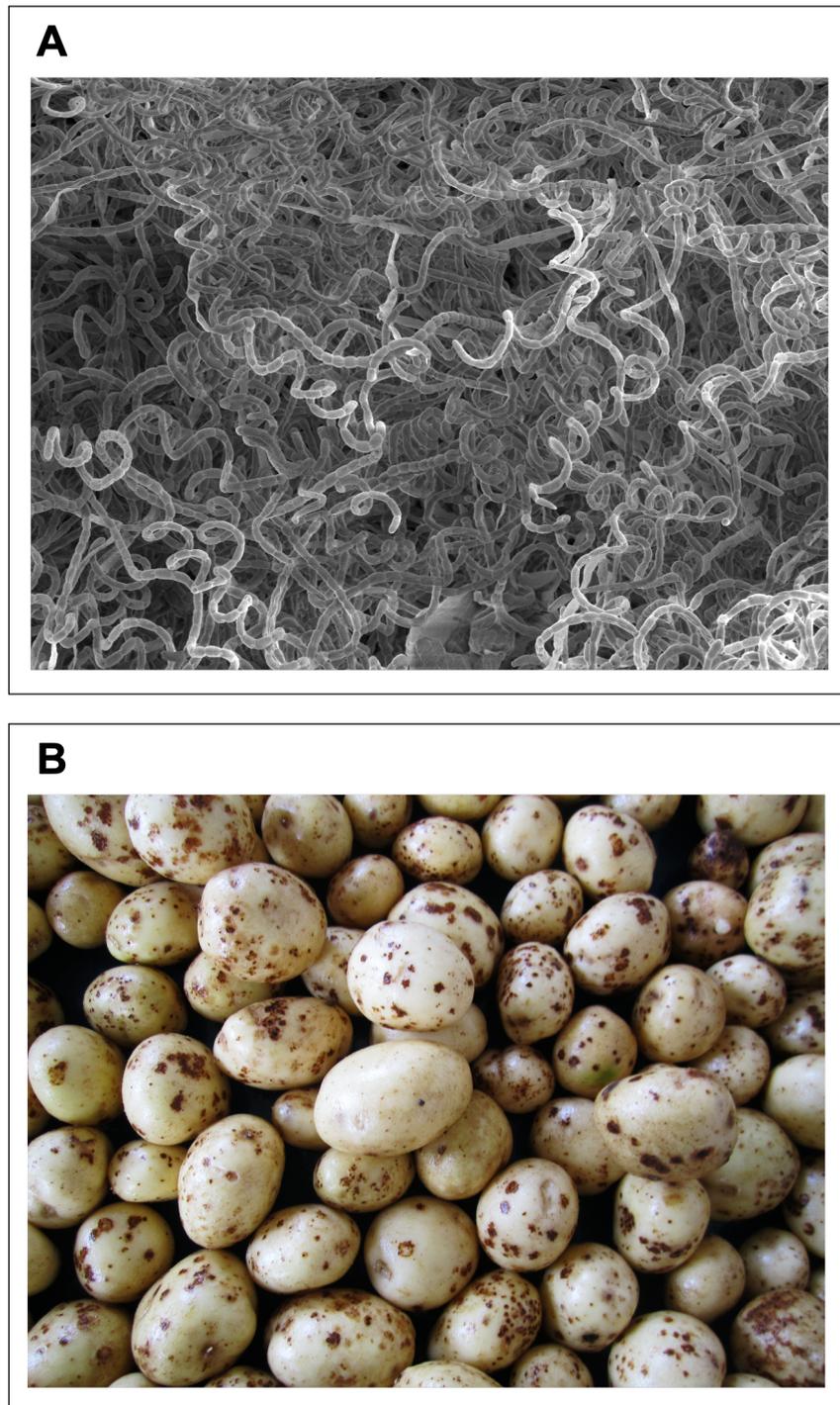
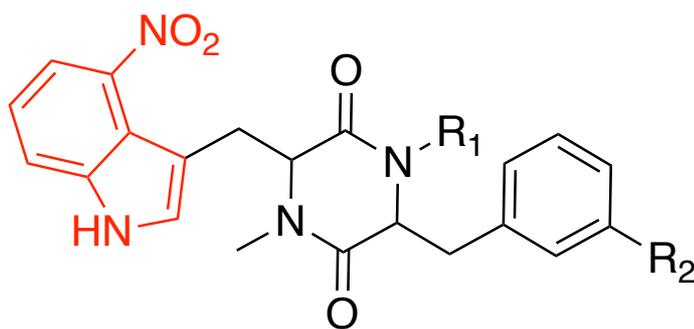


Figure 1.8. Potato common scab caused by *Streptomyces scabies*. (A) Scanning electron micrograph of *Streptomyces scabies* 87-22 showing the spiral chains of spores characteristic of this species. Credit: K. Findlay, John Innes Centre. (B) Potato common scab caused by *Streptomyces* sp. on Maris Piper potatoes. Affected tubers are not inedible, but less palatable and of lower economic value. Scab severity varies depending on environmental conditions and the interventions used. The tubers shown are of intermediate severity, showing significant scab coverage but without significant coverage of deep pitted lesions, Image courtesy of G. Tomalin of VCS Potatoes.



Thaxtomin C
 $R_1 = H, R_2 = H$

Thaxtomin B
 $R_1 = CH_3, R_2 = H$

Thaxtomin B
 $R_1 = CH_3, R_2 = OH$

Figure 1.9. Thaxtomin phytotoxins produced by scab-causing *Streptomyces* species. Highlighted in red is the 4-nitroindole moiety.

1.6.2.2 Breeding Scab-resistant Varieties

It has not yet proven possible to generate potato cultivars completely immune to potato scab, but cultivars exist with different susceptibilities. From within the tetraploid domestic potato (*S. tuberosum*) more tolerant cultivars have been developed, but are limited by the resistance genes already present. Analyses of tuber development and onset of disease symptoms revealed scab-tolerant cultivars to have greater numbers of cell layers and layer thickness in the periderm, as well as improved suberisation (303). Other work on genomic selection combined SNP data with phenotypic data from field trials, and identified a SNP in a WRKY transcription factor, which have roles in activation or repression of R genes (304). Such studies can shed light on how best to develop potato cultivars for common scab tolerance or resistance. Efforts have also been made to introgress genes from wild relatives, as is the case for breeding Late Blight resistance. Notably, a QTL for common scab resistance was recently mapped (305), and introduced into new clones (306). However, despite showing high fertility and improved resistance to scab than the currently available varieties, high glycoalkaloid levels were observed as in previous efforts to introgress genes from *Solanum chacoense*. The authors propose these clones may serve as parents in the generation of new cultivars, ultimately aiming to retain scab resistance but improve marketability. Similar work using protoplast fusion to introgress genes from *Solanum brevidens* was able to introduce a level of scab resistance, but not able to completely restore chromosome number or crossability of the *S. tuberosum* (307). Cisgenic methods described above may help resolve these issues inherent in breeding efforts. However, it must be noted that no clones have yet been developed that show complete resistance to potato common scab caused by *S. scabies* and related organisms, highlighting the need for investigation into alternative strategies, including combinations of existing approaches.

1.6.2.3 Cultural Practices

In some instances, controlling the environmental conditions may alleviate disease symptoms or reduce incidence, but there is a balance between precluding phytopathogen growth and maintaining conditions favourable for plant growth (45). In the control of common scab of potato, a number of cultural practices are in widespread application that

centre around modifying the environmental conditions to curb disease incidence and severity, but are plagued by lack of reliability (308). Commonly applied is irrigation of potato crops during the crucial period of tuber initiation in the first six weeks (309, 310) and is considered to be one of the most reliable means of control, despite conflicting data (311). Reducing soil pH to below 5.2 is also considered to be effective in alleviating common scab, as *S. scabies* is intolerant to such conditions (312) and the disease noted to be largely absent in regions with acidic soils (313). Sadly, this approach often fails depending on the predominant strain in the region, as *Streptomyces acidiscabies* also causes common scab in acidic soils (314). Overall, the incidence and severity of common scab on potato crops is dependent on a range of interacting factors, differing again by region and cultivar (315)

1.6.2.4 Chemical & Biological Control Methods

A variety of chemical and biological applications have been applied to control common scab, with varying degrees of success. A two-year field study conducted in Canada tested biological control methods like seed application of *Bacillus subtilis* and *Enterobacter cloacae*, and soil application of mustard meal, as well as chemical treatments of chloropicrin, fludioxonil, mancozeb, manganese sulphate, and Pic-Plus which is chloropicrin-based fumigant (316). Many reduced disease incidence, severity, or both. Surprisingly, mustard meal was the most effective in both cases, reducing incidence by 36% and severity by 63.1%. The authors recommend combining mustard meal application to soil, with seed application of *B. subtilis*, which itself reduced severity by 56.1%. It is also reported that pentachloronitrobenzene has been used in some regions, but does not appear to provide any added benefit beyond simply increasing soil moisture content (310). An Australian study evaluated mancozeb, benzothiazole, and fluazinam for control of *S. acidiscabies*-caused scab, and observed effectiveness of all three compared to untreated controls in an artificially infested soil (317). Fluazinam was most effective, and disease in these tubers was either low or absent. Soil amendment with rice bran, traditionally used in southwest Japan, has demonstrated effectiveness in field trials, inhibiting *S. scabies* in young tubers by increasing populations of antagonistic actinomycetes. Inoculation of plants in field experiments reduced disease, but was still stronger with concomitant addition of rice bran, suggesting further mechanisms of disease reduction (318). Other soil

amendments and supplements like peat and iron can protect against common scab in pot trials, showing a synergistic effect by shifting microbial communities closer to that of a naturally suppressive soil while also enhancing plant defence (319). While the authors noted a pH shift after addition of peat and iron to the soils, the values remained in the range considered conducive to development of potato common scab caused by *Streptomyces* species. Also of note was the proportional increase of Proteobacteria with reference to decrease of actinobacteria in the soil samples in this study. Other strategies showing some success in bot pot and field trials include foliar treatment of potato plants with 2,4-dichlorophenoxyacetic acid, a synthetic auxin, reducing disease severity on both disease resistant and conducive cultivars (320). Others have explored direct application of specific bacterial isolates to combat the disease. Sarwar and colleagues successfully applied non-pathogenic *Streptomyces* A1RT, which produces isatropalone C, in pot trials against common scab caused by *S. scabies*, *S. turgidiscabies*, and *S. stelliscabiei* (321). However, their approach used sterile vermiculite substrate in place of natural soils, perhaps limiting the direct applicability of their findings to commercial potato growth. Lin *et al.* were able to demonstrate effectiveness of *Bacillus amyloliquefaciens* BA01 in both pot and field trials, linking inhibition to the natural products surfactin, iturin A, and fengycin, although inhibitory activity was greater than the application of these compounds alone or in combination (322). The approach was preventative in the field rather than curative, but benefitted from ease of application, which consisted of pouring bacterial cells in solution directly onto the soil surrounding potato plants. To complicate matters further, the success of a particular antagonist can vary between growing seasons, environmental factors, and geographic location (323). To date there appears no treatment that displays consistent results across different regions, soils, or time.

1.7 Thesis Objectives

The project aims to identify *Pseudomonas* natural product gene clusters associated with inhibition of the important potato pathogens *Streptomyces scabies* and *Phytophthora infestans*. Where such clusters or their products are uncharacterised, the focus will be on isolating and characterising both the chemical nature of any molecules and their biosynthetic pathways.

1.7.1 Phenotype-guided Discovery of Natural Products

Following selective isolation efforts from potato field soil samples in 2015, a *Pseudomonas* isolate, Ps652, was retrieved that shows high level of inhibition against two critical potato pathogens, *Phytophthora infestans* and *Streptomyces scabies*. From the original genome sequence, there was little clue as to the mechanism of inhibition when investigated using bioinformatic tools. The work presented in chapters 2 and 3 reports efforts to (i) identify the genetic determinants of phytopathogen inhibition by Ps652 using bioinformatic and genetic tools, (ii) uncover the natural products involved in biological control by Ps652, characterising any novel molecules, and (iii) investigate the potential for Ps652 to control potato pathogens in the field.

1.7.2 Investigation into Bioinformatically Identified Gene Clusters

Following initial work by Dr F. Stefanato *et al.* (324), gene clusters were identified that correlated with inhibition of potato pathogens and appeared to be previously uncharacterised biosynthetic gene clusters encoding for Ribosomally synthesised Post-translationally modified Peptides (RiPPs). Chapter 4 describes efforts to (i) investigate this hypothesis, (ii) discern their role in phytopathogen inhibition, and (iii) characterise any novel natural products associated with the clusters.

Chapter 2: Identifying the Natural Products Determining Biocontrol

Activity in Environmental *Pseudomonas* Isolate Ps652

2.1 Introduction and chapter aims

Pseudomonas strain Ps652 is an environmental isolate recovered from agricultural soils from a commercial potato field during the work of Stefanato *et al.* (324). In the original phenotypic characterisation of all 240 strains, and the genotypic characterisation of the sequenced 69 strains, Ps652 stood out. Firstly, it is phylogenetically distinct from the other isolates (Figure 4.1), falling into the *putida* group rather than the *fluorescens* group with the closest match being *Pseudomonas* sp. UC 17F4 (GCF_900101695) by Multi-Locus Sequence Typing (MLST) with an estimated average nucleotide identity of 93.5% (325) (Figure 2.1). Accordingly, Ps652 likely represents a novel species in the *Pseudomonas putida* group, falling below the 94% cut-off proposed by Richter and Rosselló-Móra (326) and ratified by Peix *et al.* (327) for *Pseudomonas* taxonomy. Secondly, during high throughput phenotyping of isolated pseudomonads, Ps652 showed considerable inhibition of all tested phytopathogens (Figure 4.1, shaded squares). However, from the initial bioinformatic analysis performed in that study, Ps652 seemed to show relatively little biosynthetic potential, making its biocontrol phenotype quite unlikely (Figure 4.1, shaded and unshaded circles). An updated bioinformatic analysis of Ps652 biosynthetic potential will be provided in section 2.3. Accordingly, the aims of this work were to (i) investigate the biosynthetic potential of Ps652, (ii) find the genetic determinants of phytopathogen inhibition by Ps652 and (iii) link those, where possible and appropriate, to the natural products involved in the phenotype.

2.2 Initial Bioassays

In order to verify the original results outlined in ref (324), initial bioassays against key phytopathogens were repeated, and follow-up experiments carried out to further probe the implications of those phenotypes.

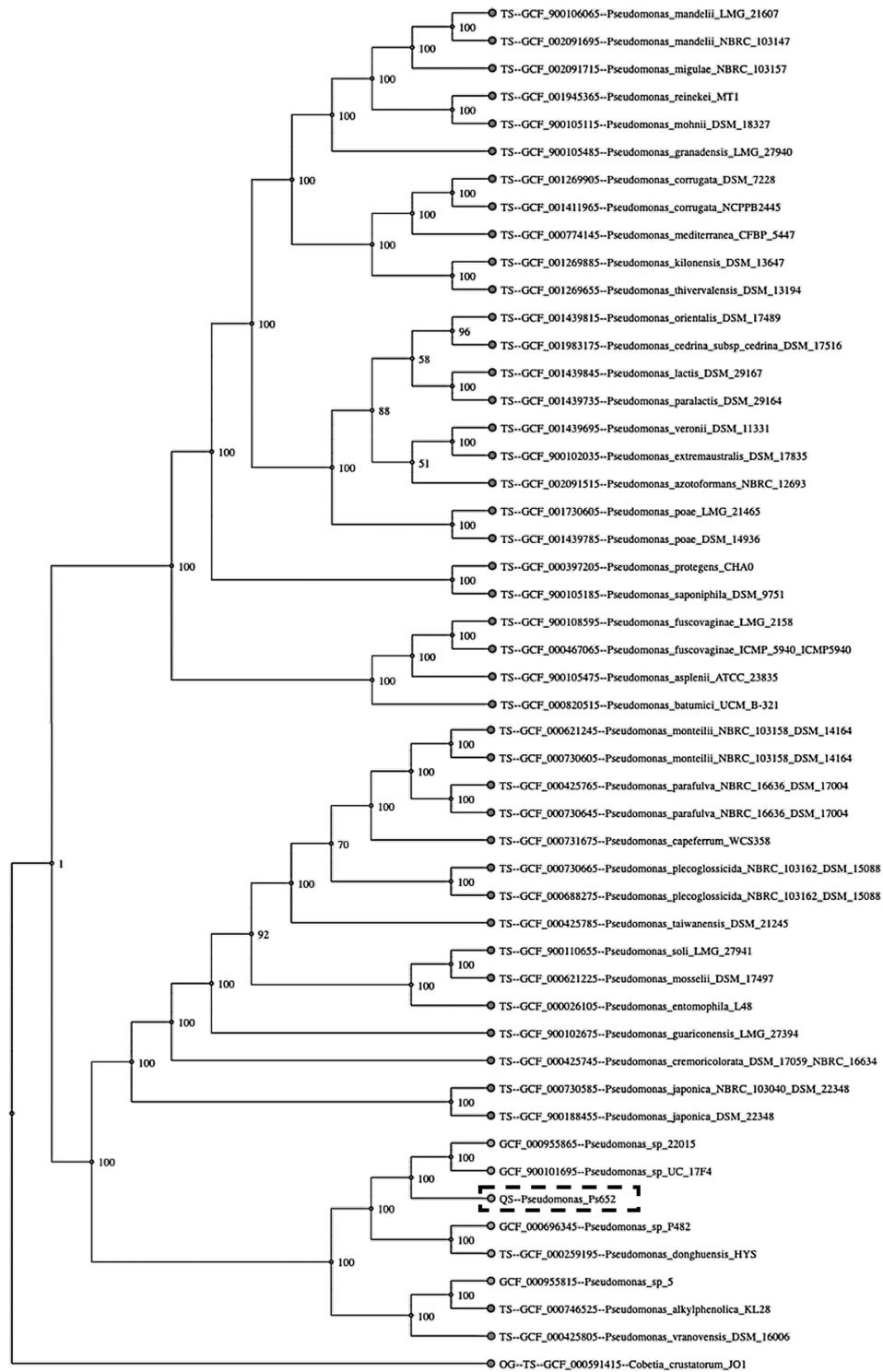


Figure 2.1. MLST phylogenetic tree for Ps652 created using autoMLST. Shown is the tree output from the autoMLST tool, with Ps652 indicated by the hatched box. Ps652 falls within the *putida* group and likely represents a novel species.

2.2.1 *Streptomyces scabies*

As described in section 1.6.2, *Streptomyces scabies* is a significant pathogen of potato crops worldwide, and is the source of great economic loss for growers due to lowering of tuber quality. In this study, the *S. scabies* strain 87-22 is used, which possesses additional virulence factors that render it particularly damaging (293). As such, it seemed an appropriate indicator strain for this study. Cross-streak assays were performed as per section 5.4.1, and imaged either after five days or once interspecies interactions were visible. Cross-streaks were attempted on three different media, as substrate may considerably affect the phenotype in microbial interactions. Ps652 indeed showed strong inhibition of *S. scabies*, but with strongly media dependent effects (Figure 2.2). MYM agar proved to be the most consistent in terms of phenotype, and was therefore used for all further work (Figure 2.2B). The increase in inhibitory activity with the addition of yeast extract to SFM medium (Figure 2.2C) was attributed to increased production of hydrogen cyanide by Ps652, which will be explored further in section 2.3.4. Overall, the cross-streak assay verified that Ps652 is a potent inhibitor of *S. scabies* and should be prioritised for further characterisation.

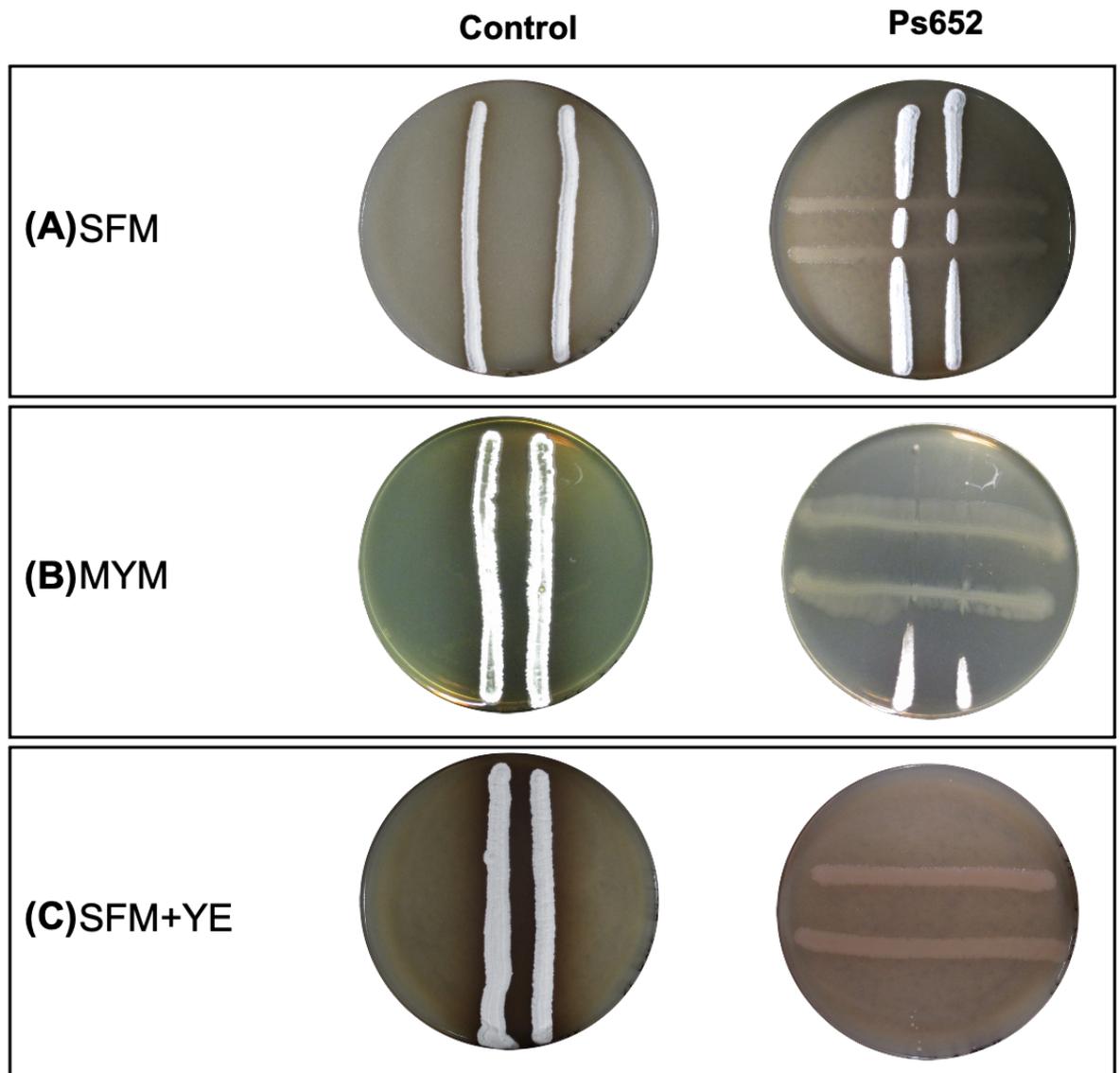


Figure 2.2. Cross-streak assays of Ps652 with *Streptomyces scabies* 87-22. (A) Cross-streak assay performed on SFM agar. (B) Cross-streak assay performed on MYM agar. (C) Cross-streak assay performed on SFM+YE agar. All images are representative of three biological replicates. Control is *S. scabies* 87-22 alone, with no cross-streaking.

2.2.2 *Phytophthora infestans*

As described in section 1.6.1, *P. infestans* is a serious pathogen of potato crops, and can be particularly devastating should environmental conditions favour its growth. Given that Ps652 was isolated from a commercial potato field, and had strong activity against *S. scabies*, screening against the causative agent of late blight seemed an obvious next step. An isolate capable of inhibiting the growth of two critical potato pathogens would be a high priority candidate for development as a biocontrol solution. Preliminary data (324) already suggested Ps652 was a potent inhibitor of *P. infestans*, more so than almost any other member of the strain collection. As such, a spot assay was performed to verify these results on Rye Sucrose Agar (RSA). A plug was taken from the growing edge of a *P. infestans* mycelium, and placed in the centre of the plate, then 15 μ L spots of overnight culture of Ps652 was added to the plate in a triangle around the plug (see methods). This assay demonstrated almost complete growth inhibition of *P. infestans* by Ps652, as the phytopathogen was almost entirely confined to the plug (Figure 2.3). Given the severely limited growth, efforts to quantify the area of mycelium and percentage inhibition were unsuccessful, despite such experiments being reported in the literature for other strains (103).



Figure 2.3. *Phytophthora infestans* inhibition by Ps652. Compared to the control (left, *P. infestans* only), Ps652 (right) was able to almost completely preclude the growth of *P. infestans* over a 10-day period. Quantification of growth inhibition was not attempted due to the extremely low levels of growth by the pathogen precluding accurate measurement.

2.2.3 Split-plate assays

The two assays used to probe for activity against *S. scabies* and *P. infestans* allowed for inhibition due to both soluble, diffusible compounds, and volatile organic compounds, both of which are regularly implicated in biocontrol phenotypes (see sections 1.5 and 1.5.6 for details). The use of split-plates, which have two separate compartments, allows separation of these different modes of compound delivery. In such assays, the strain to be tested and phytopathogen indicator are inoculated on opposite sides of the divide. Volatile compounds will easily diffuse through the headspace of the plate, over the barrier in the split-plate, whereas compounds which are agar-soluble and non-volatile will not be exchanged and so are unable to reach the pathogen (Figure 2.4). Simply, the split-plate assay determines whether inhibitory compounds are volatile or not. Testing Ps652 against *S. scabies* in this assay was performed both on SFM (Figure 2.5 A,B) and MYM agar (Figure 2.5 C,D). *S. scabies* and Ps652 were streaked in the same orientations as in the original cross-streak assay (Figure 2.2) for maximum correspondence of results. Each organism is immediately adjacent to the barrier to mirror the proximity in the original assay. The results from the split-plate assays for Ps652 against *S. scabies* suggest minimal effects of volatile compounds when comparing Ps652 inoculation with the controls both on SFM (Figure 2.5B vs Figure 2.5A) and on MYM (Figure 2.5D vs Figure 2.5C). This indicated the phenotype observed in Figure 2.2 can most likely be attributed to soluble compounds. A similar story can be observed for Ps652 versus *P. infestans* (Figure 2.6), where on RSA there is no observable inhibition compared to the control due to volatiles (Figure 2.6B vs Figure 2.6A) when a barrier is erected between the strains. However, significant inhibition due to volatiles is observed when Ps652 is grown on Lennox agar in its own compartment, due to the overproduction of hydrogen cyanide which is not observed on RSA (Figure 2.6C). Despite this, the original assay was performed on RSA (Figure 2.3), making this result largely irrelevant.

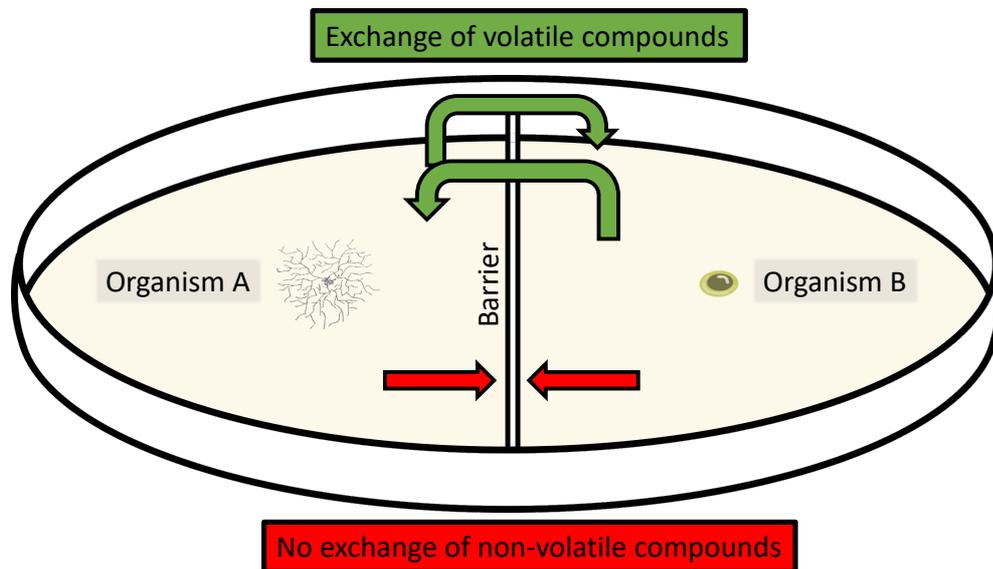


Figure 2.4. Schematic of split-plates and possible molecular interactions. With each organism inoculated on to opposite sides of the central barrier that bisects the plate, only interactions mediated by volatile compounds are possible. Non-volatile compounds can diffuse within each side, but not cross the barrier to affect the other organism. The assay also allows for different media to be used for each organism if required.

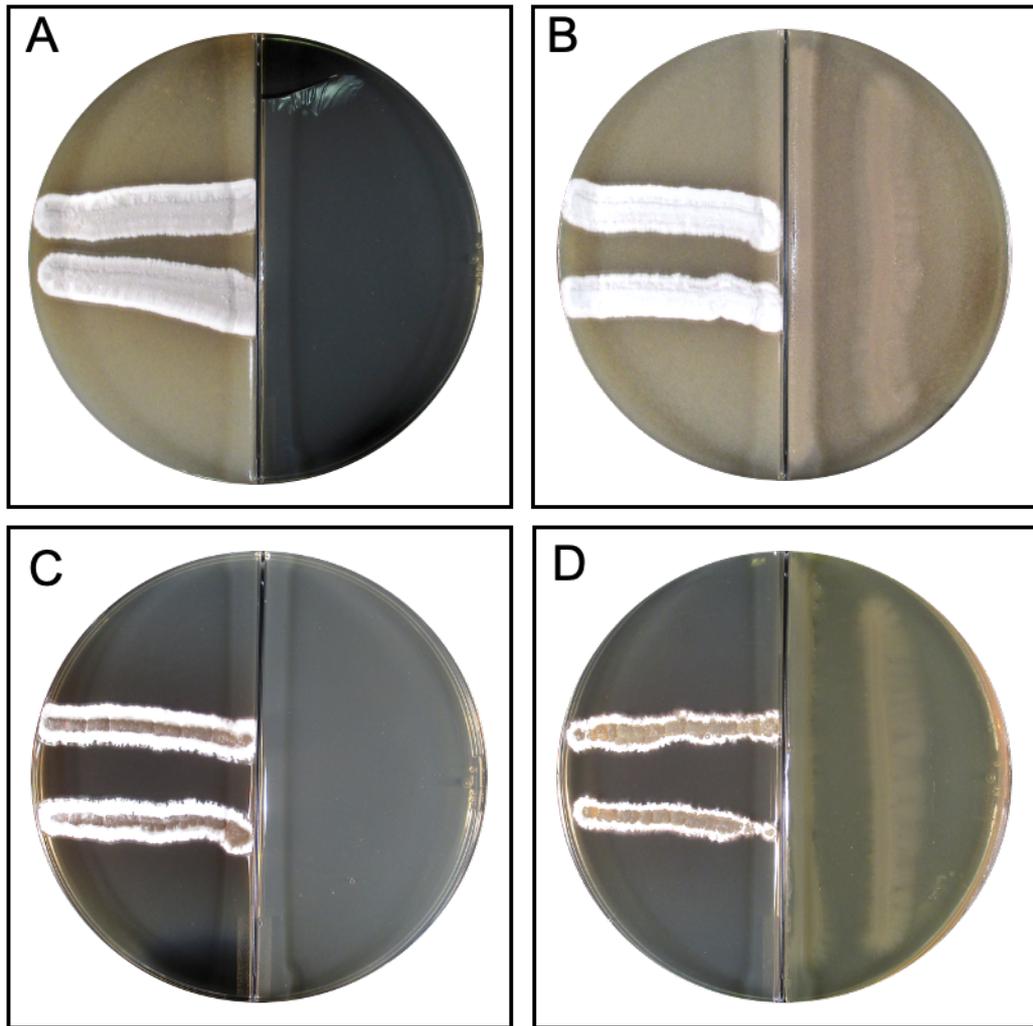


Figure 2.5. Split-plate assays of Ps652 against *Streptomyces scabies* 87-22. (A) *S. scabies* (left) control on SFM agar. (B) *S. scabies* (left) versus Ps652 (right) on SFM agar. (C) *S. scabies* (left) control on MYM agar. (D) *S. scabies* (left) versus Ps652 (right) on MYM agar. No evidence of inhibition by volatile compounds was observed under these conditions. All plates are representative of three biological replicates. Note, Ps652 is streaked immediately adjacent to the barrier in all relevant plates (2 x Ps652 per plate), although this may be difficult to observe in the images.

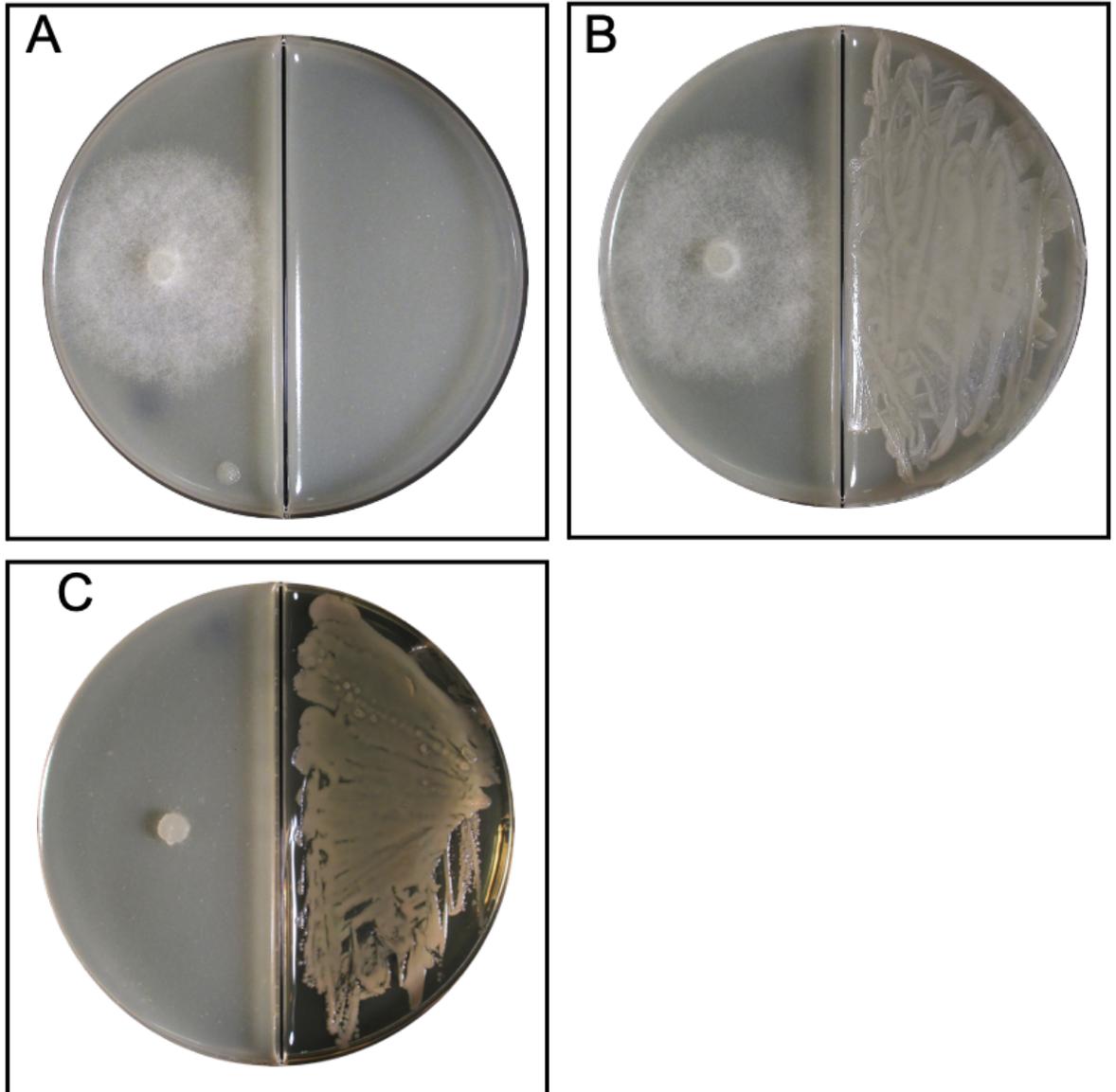


Figure 2.6. Split plate assays of Ps652 against *Phytophthora infestans*. (A) Control plate with *P. infestans* inoculated on the left side of the barrier, with no inoculation on the right (media only, RSA). (B) *P. infestans* inoculated on the left side of the barrier, Ps652 on RSA medium on the right. No inhibition due to volatiles is seen when compared to the control. (C) *P. infestans* inoculated on the left side of the barrier, and Ps652 on Lennox agar on the right side. Complete inhibition of *P. infestans* is seen in this plate due to overproduction of hydrogen cyanide by Ps652 on Lennox media. All plates are representative of three biological replicates.

2.2.4 Preparation of a Biologically Active Extract from Ps652 Cultures

Split-plate assays suggested that the compound responsible for anti-*scabies* and anti-*Phytophthora* activity by Ps652 was due to the presence of a non-volatile compound diffusing in the agar. It was thought quite likely that this represented a novel natural product with activity relevant to commercial agriculture. The next step was to test whether this compound could be extracted from the agar using a solvent, and re-tested for the activity of interest. After initial testing, RSA appeared the ideal medium from which to extract, as significant batch to batch variability was observed with MYM agar, and activity on SFM agar was highly inconsistent. Plates were grown for 24 hours before being extracted with methanol, which should extract both hydrophilic and hydrophobic compounds as it is water-miscible. The extracts were then tested for activity against *S. scabies*, which was chosen for speed of growth and ease of handling compared to *P. infestans*, on instant mash agar (IMA) media. Compared to the control, a zone of inhibition could be observed around disks containing 150 μ L methanol (MeOH) extracts from Ps652 grown on RSA, indicating Ps652 produces a non-volatile natural product that is extractable with solvents (Figure 2.7). While initial bioinformatic analysis had failed to identify any natural product BGCs that may explain the phenotype, a more tailored approach was used in an effort to find the source of the biological activity in Ps652.

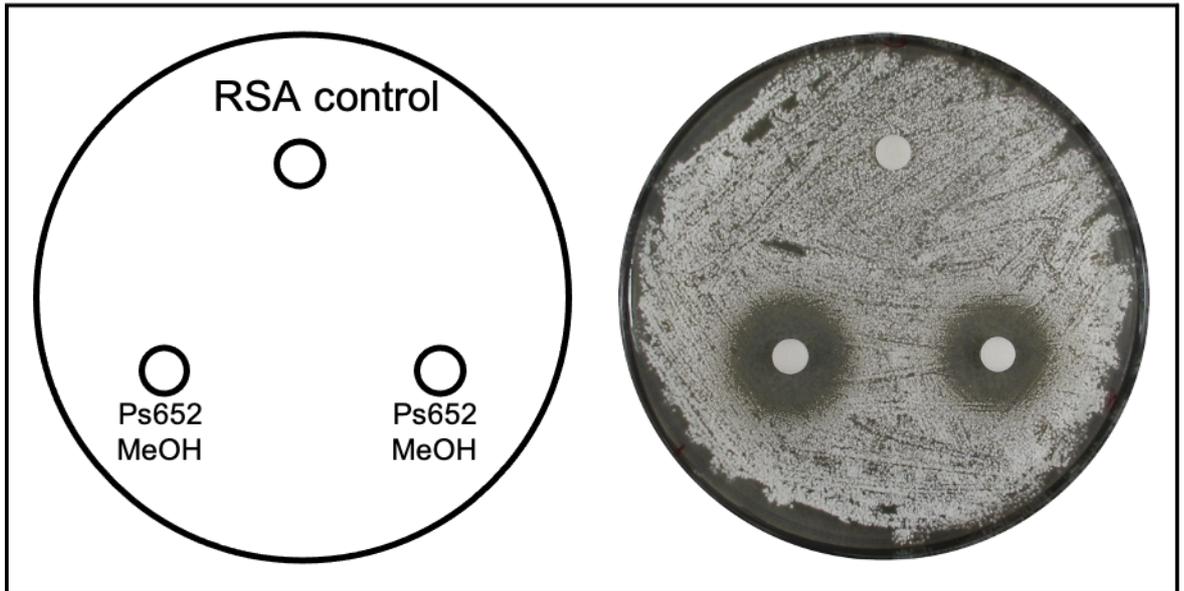


Figure 2.7. Bioactive extracts from Ps652 grown on RSA medium. Methanol extracts of RSA and Ps652 grown on RSA. Each disk is loaded with 150 μ L of MeOH extract and allowed to dry fully before application to the plate. Zones of inhibition are visible around Ps652 MeOH extracts, indicating the production of an antibacterial natural product. The extraction was performed in duplicate, as per the image.

2.3 Searching the Genome

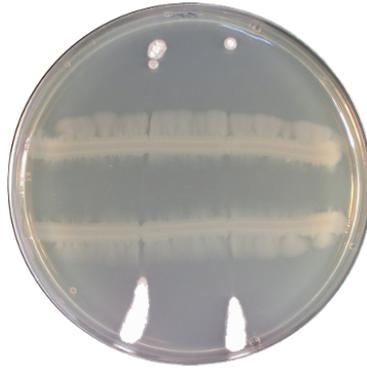
2.3.1 Improved Illumina & Nanopore Genome

The original genome sequence for this strain, obtained as part of (324), was produced through short read sequencing only and, as such, consisted of 82 contigs. This fragmented genome made it difficult for antiSMASH predictions due to cluster proximities to contig edges. There may also be genes present which span two contigs, or genes missing entirely from the genome due to missing sequence. Accordingly, a combined short read (Illumina) and long read (Oxford Nanopore) genome sequence for this strain was obtained from MicrobesNG (Birmingham, UK). It was possible to assemble this genome into a single scaffold (see methods for details). This revealed a genome size of 5.96 Mbp, with 62.19% GC, and 5371 genes. This improved genome was deemed sufficiently complete for bioinformatic investigations into the strain's biocontrol potential.

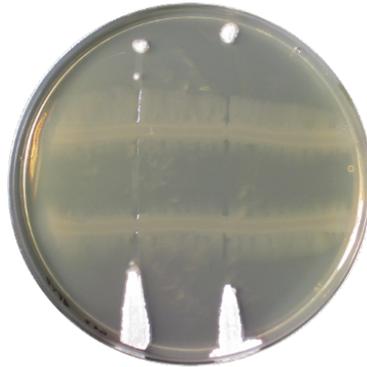
2.3.2 AntiSMASH Analysis

To assess the biosynthetic potential of Ps652, the improved genome was analysed using antiSMASH 5.0 (328). Compared to the original antiSMASH analysis with version 3.0 (324), no significant new results were obtained that explained the strong on-plate activity of this strain when using the improved genome, with only seven clusters identified. Three NRPS regions were identified involved in the production of pyoverdine, one NRPS-like region predicted to be involved in the biosynthesis of a single L-threonine-related metabolite, one arylpolyene, one short type III PKS of unknown product, and one NAGGN region. None of these appeared likely candidates for the inhibitory activity shown by Ps652 towards the tested phytopathogens. However, since pyoverdine had been linked to biocontrol phenotypes previously (109), a deletion strain Ps652 Δ Pyo was made by allelic exchange (181). Additionally, in the event that it produced a small antimetabolite, a mutant strain Ps652 Δ NRPS5 was produced with a deletion in the the NRPS-like region predicted to make a threonine-related metabolite. Neither of these was found to have a role in anti-*scabies* activity in the cross-streak assay (Figure 2.8). Loss of activity against *P. infestans* was not screened for at this time, as the pathogen was unavailable.

Ps652



Ps652 Δ Pyo



Ps652 Δ NRPS5

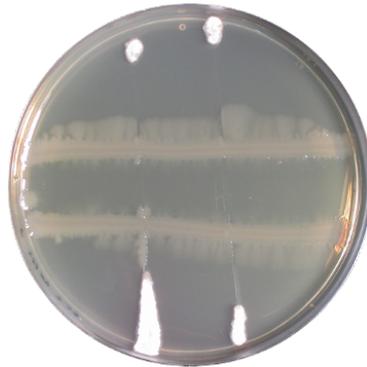


Figure 2.8. Cross-streak assays for Ps652 , Ps652 Δ Pyo, and Ps652 Δ NRPS5 against *Streptomyces scabies* 87-22. Ps652 shows strong inhibition of *S. scabies*, with no loss of inhibitory phenotype for pyoverdine and NRPS5 deletion strains, indicating these are not involved in the biocontrol phenotype under investigation. All images are representative of three biological replicates.

2.3.3 Manual BLAST Searches

Pseudomonas isolates are recorded as producing a wide variety of natural products, many of which show unusual biosynthetic pathways (see section 1.5) that are not currently identified by the automated approach of antiSMASH (328). Accordingly, manual BLAST (329) searches were conducted for key genes in the biosynthesis of molecules linked to bioactivity of *Pseudomonas* isolates in the literature at the time (Table 2.1). The only gene for which a match was obtained was *hcnA*, from the pathway responsible for biosynthesis of hydrogen cyanide (HCN) by glycine oxidation (330).

Table 2.1. Determinants of *Pseudomonas* bioactivity identified from literature, and their presence in Ps652.

Compound	Gene	Presence
1 - undecene	<i>undA</i>	×
Phenazine	<i>phzG</i>	×
Pyoluteorin	<i>pltG</i>	×
2,4-DAPG	<i>phlD</i>	×
Pyrrrolnitrin	<i>prnD</i>	×
2-hexyl,5-propyl resorcinol	<i>darB</i>	×
Hydrogen cyanide	<i>hcnA</i>	✓

2.3.4 Identification & Deletion of HCN Cluster

The *hcnA* gene was the only identified determinant of biological activity from the bioinformatic analyses that might explain the antagonism towards phytopathogens. The genome was checked for the remaining two genes that make up this cluster and allow hydrogen cyanide production by glycine oxidation. Both *hcnB* and *hcnC* were found to be present (Figure 2.9A), suggesting Ps652 could produce hydrogen cyanide, which was

confirmed by using filter paper soaked in Feigl-Anger reagent (see methods), which produces a blue colour when in contact with HCN (Figure 2.9B).

Next it was essential to uncover whether the production of hydrogen cyanide was responsible, at least in part, for the inhibition of *S. scabies* and *P. infestans*. Accordingly, it was decided to produce an in-frame deletion mutant of the entire *hcn* cluster in Ps652 by allelic exchange. Ps652 Δ HCN was produced, and verified by PCR and sequencing as standard practice, but was also verified phenotypically using the Feigl-Anger filter paper disk to confirm visually that production of HCN had been abolished (Figure 2.10).

Once elimination of HCN production had been established in Ps652 Δ HCN, testing for loss of inhibition against the phytopathogens was the obvious next step. Again, the cross-streak assay was used for this purpose. Surprisingly, the loss of HCN production only partially reduced the inhibitory phenotype, with regions of complete growth inhibition of *S. scabies* still visible between the two *Pseudomonas* streaks (Figure 2.11A). Against *P. infestans*, the effect was even more striking, with no loss of inhibition observed for the Δ HCN strain (Figure 2.11B). These were consistent with the earlier results in split-plate assays. This result hinted at the presence of another unknown determinant of bioactivity, which was not evident from bioinformatic analysis of the Ps652 genome using existing secondary metabolite gene cluster prediction tools, or by searching for known biocontrol compounds. Given these facts, it appeared likely that Ps652 might produce a novel biocontrol compound.

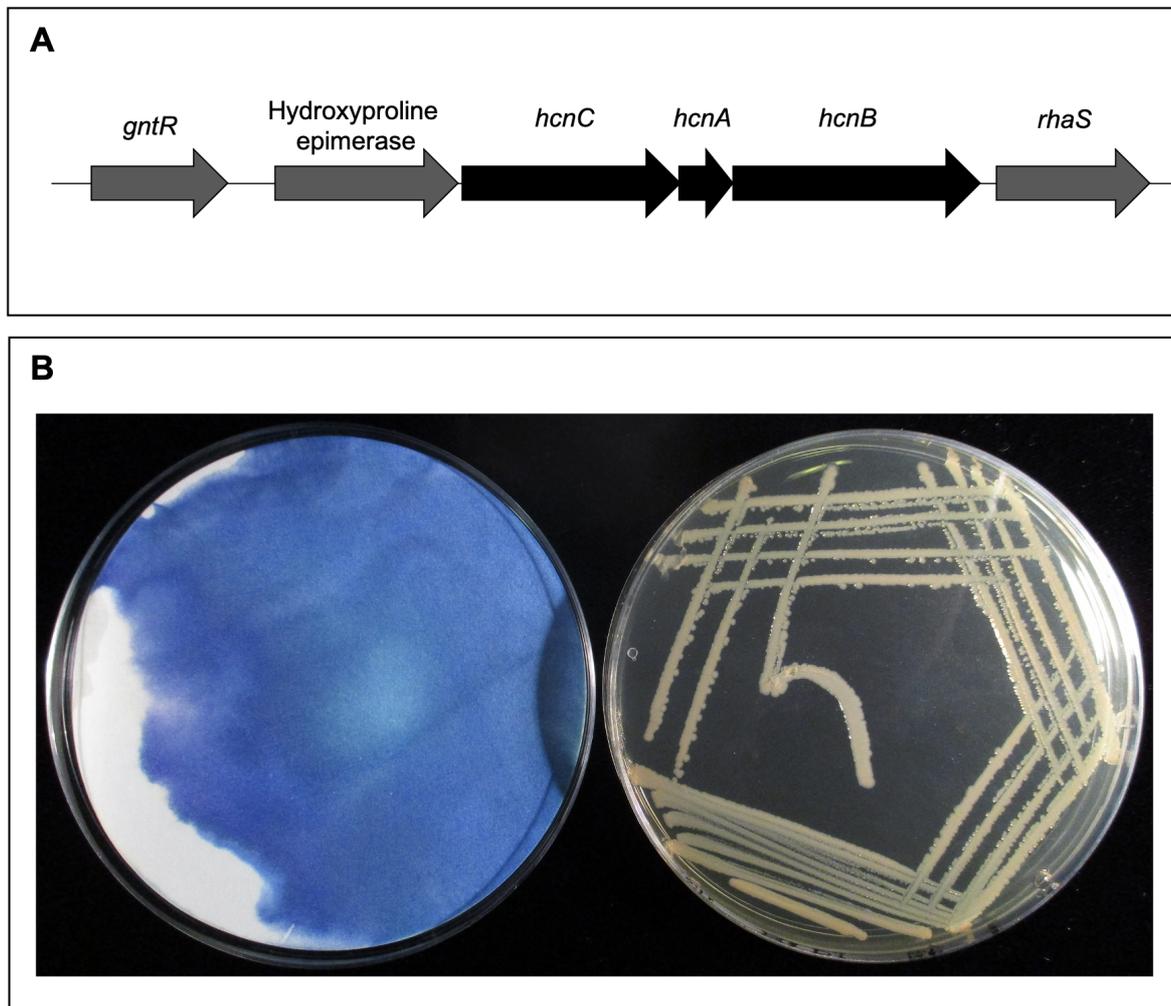


Figure 2.9. Hydrogen cyanide production by Ps652. (A) The *hcnABC* gene cluster in Ps652. (B) Hydrogen cyanide production by Ps652 grown on Lennox agar. On the left is the petri dish lid containing a filter paper disk soaked in Feigl-Anger reagent, which turns blue upon contact with cyanide; the filter paper was left in for the duration of the incubation period.

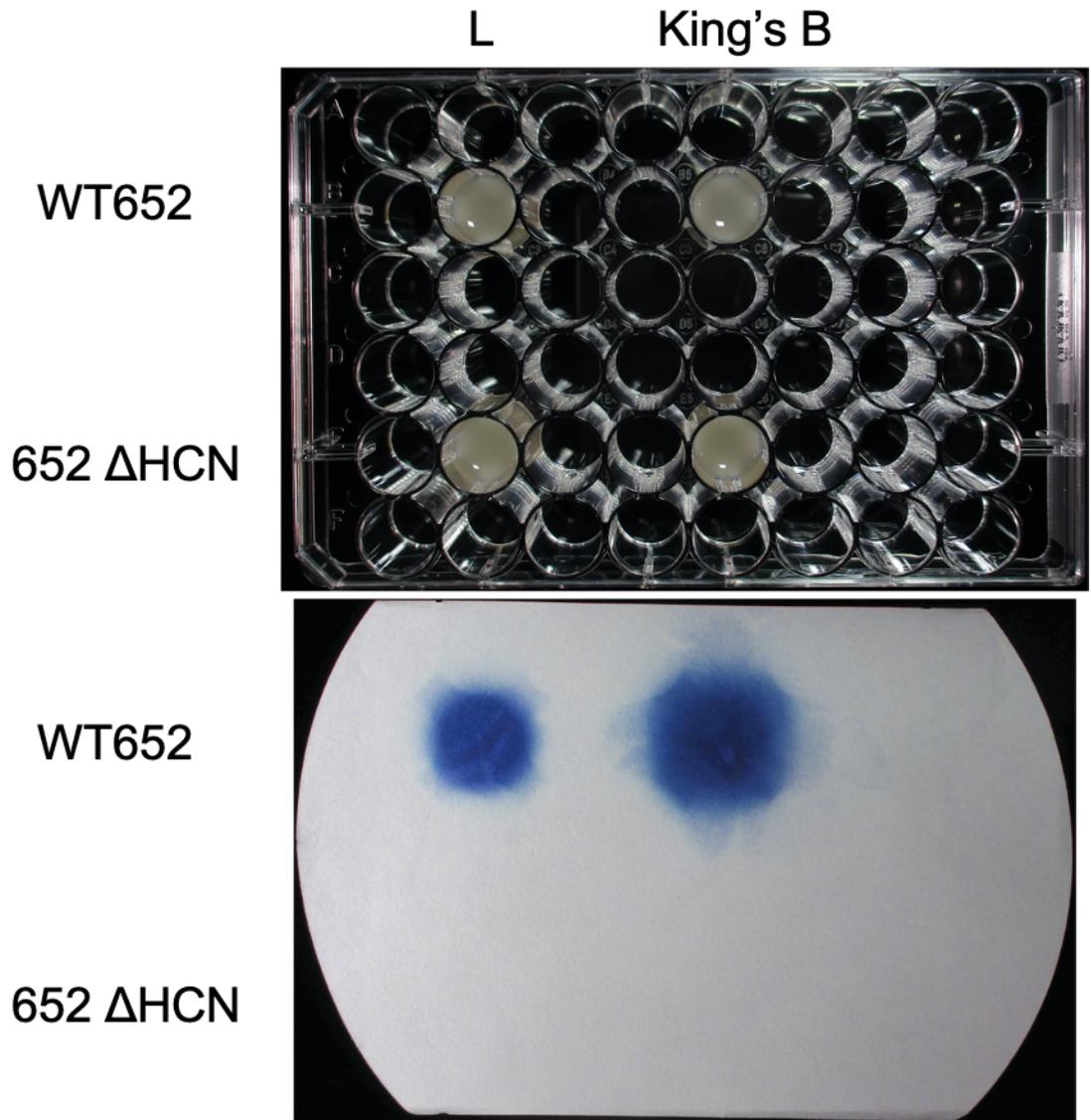


Figure 2.10. Hydrogen cyanide production by Ps652 and 652 Δ HCN. 48-well plate containing overnight culture of Ps652 and Ps652 Δ HCN in Lennox broth or King's broth as indicated (top), and filter paper disk soaked in Feigl-Anger reagent which was present in the lid of the plate during incubation (bottom). Blue spots indicate hydrogen cyanide production by the wild type Ps652, but none by the cyanide mutant Ps652 Δ HCN.

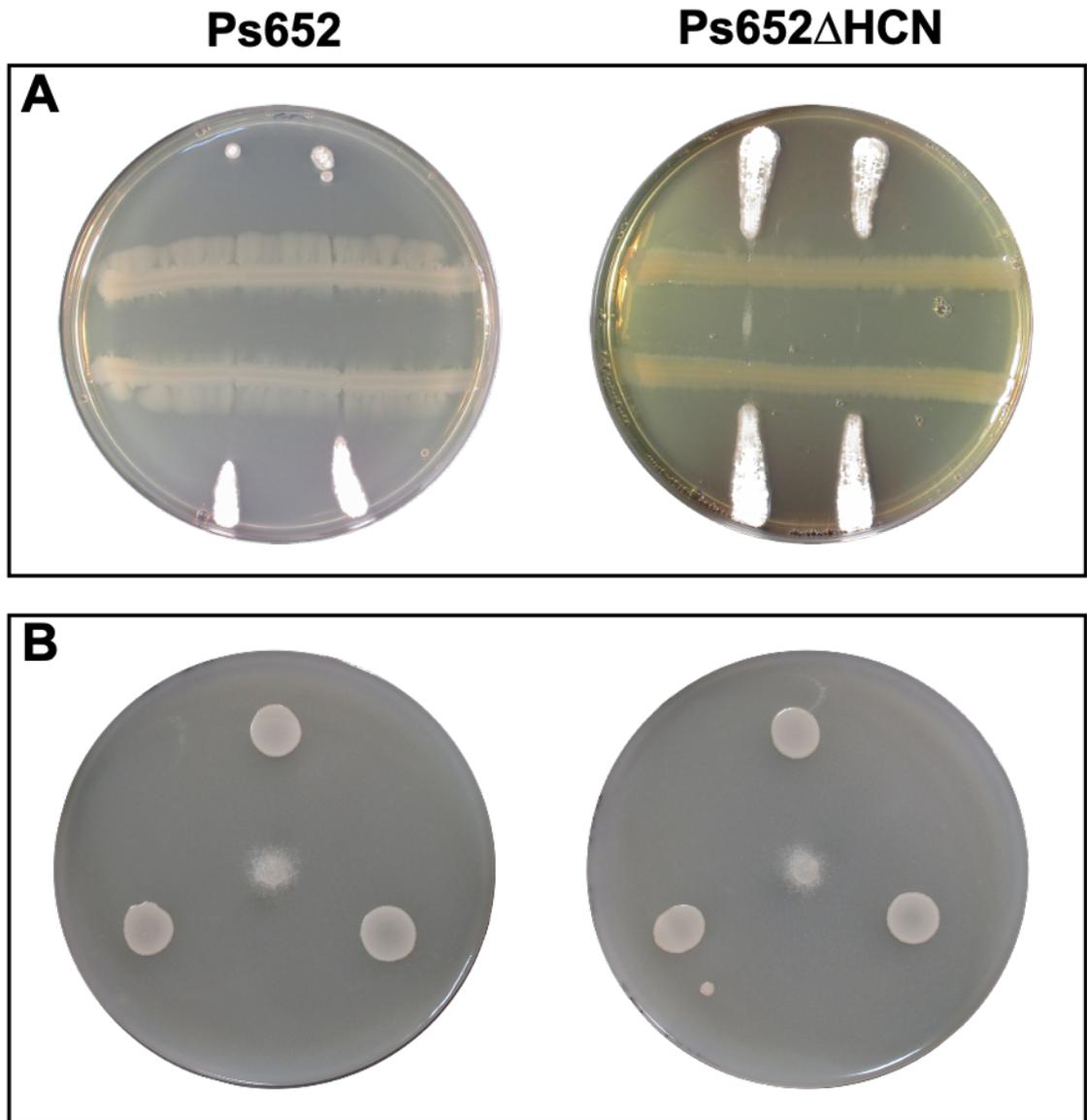


Figure 2.11. Bioassays against *S. scabiei* 87-22 and *P. infestans* for Ps652 and Ps652ΔHCN. (A) Cross-streak assay for Ps652 and Ps652ΔHCN against *S. scabiei*. The inhibitory phenotype is only partially explained by the production of HCN, with zones of complete growth inhibition for *S. scabiei* between the *Pseudomonas* streaks in Ps652ΔHCN plates. (B) Spot assays for Ps652 and Ps652ΔHCN against *P. infestans*. No loss of inhibition is observed with the deletion of the *hcn* cluster from Ps652. All images are representative of three biological replicates.

In order to determine whether this potential new compound was a different volatile organic compound, split-plate assays were repeated as previously with the addition of Ps652 Δ HCN (Figure 2.6). Ps652 or Ps652 Δ HCN was inoculated by spreading onto one side of the split plate, while either *S. scabies* or *P. infestans* was inoculated onto the other side. As before, split plate assays for *S. scabies* were performed by replicating the cross-streak assay but with a barrier between the pseudomonad and the pathogen, and each was streaked immediately adjacent to the barrier. As anticipated, there was no significant effect of abolishing hydrogen cyanide production on the inhibition of *S. scabies*, in agreement with previous data (Figure 2.12A). This assay also indicated that the molecule responsible for inhibition of *S. scabies* by Ps652 was a non-volatile compound, whose ability to affect growth of the streptomycete was hindered by the barrier in the plate; again, matching previous data (Figure 2.7). A similar result is observed for the split plate assays against *P. infestans*, where no difference in inhibition is observed between the wild type Ps652 and the control, or Ps652 and the cyanide mutant Ps652 Δ HCN (Figure 2.12B). Again, this pointed towards the presence of a highly potent non-volatile compound being produced by Ps652 that is responsible for the significant inhibitory activity seen in previous assays. Despite the already documented effects of HCN on *P. infestans* (103), it had been observed in preliminary assays that HCN production on RSA by Ps652 was very low, explaining the lack of effect observed here. However, when *P. infestans* is grown on RSA as normal, but Ps652 is inoculated on Lennox agar, hydrogen cyanide production is high, and completely inhibits the growth of *P. infestans* (Figure 2.12C). Surprisingly, under these conditions, there also appears to be production of an unknown volatile by Ps652 Δ HCN, that enables near complete inhibition of *P. infestans* growth comparable to levels observed in spot assays. Although interesting, this was deemed unlikely to be the cause of the phenotype observed in spot assays (Figure 2.11B), and was not considered further. These assays confirmed Ps652 produces a non-volatile potentially novel compound capable of inhibiting both *S. scabies* and *P. infestans*, that is not due to the *hcnABC* cluster.

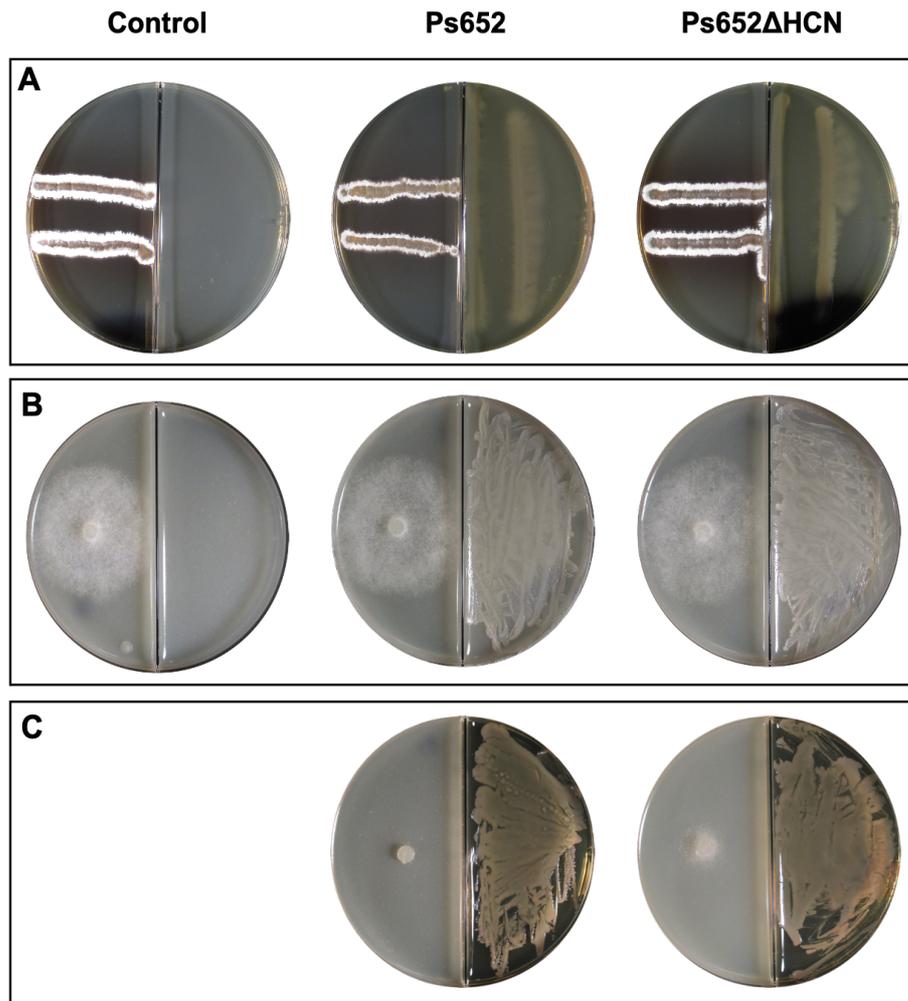


Figure 2.12. Split-plate assays for Ps652 and Ps652ΔHCN against *Streptomyces scabies* 87-22 and *Phytophthora infestans*. (A) Split-plate assays for Ps652 and Ps652ΔHCN versus *S. scabies*. No difference in inhibition is evident between the wild type and cyanide mutant. (B) Split-plate assays for Ps652 and Ps652ΔHCN versus *P. infestans*. Both pathogen and pseudomonad are plated on RSA. No difference in inhibition is evident between the wild type Ps652 and the cyanide mutant Ps652ΔHCN. (C) Split-plate assays for Ps652 and Ps652ΔHCN versus *P. infestans* with *Pseudomonas* strains plated on Lennox agar. Strong cyanide production on Lennox agar allow complete inhibition of *P. infestans* by wild type Ps652. Some activity is lost in the cyanide mutant Ps652ΔHCN, but there is also significant inhibition of *P. infestans*, indicating a secondary volatile produced on Lennox agar. All images are representative of three biological replicates. Control for (A) is *S. scabies* inoculated on the left of the barrier, and uninoculated MYM agar on the right. Control for (B) is *P. infestans* inoculated on the left of the barrier, with uninoculated RSA on the right of the barrier.

2.4 Isolation of Ps652 Pyoverdine by Activity-guided Purification

Given that previous data indicated the non-volatile antimicrobial was solvent-extractable (Figure 2.7), extraction using different solvents was tested to inform further downstream purification to elucidate the identity of the molecule responsible for the inhibitory phenotype. Ps652 Δ HCN was used given the lack of effect of volatiles, and was grown on MKB before being extracted with ethyl acetate (EtOAc) to extract non-polar metabolites. The activity of interest was water-soluble, with very little activity present in the organic fraction (Figure 2.13). These results indicate an unknown water-soluble, polar antibiotic produced by Ps652 that is active against *S. scabies* 87-22. Given that it was possible to extract the biologically active molecule using water, a purification of this molecule was attempted by blind fractionation followed by bioactivity assessment to locate the fraction(s) containing the molecule of interest. A previous attempt had been made to use bioactivity-guided purification to isolate the molecule of interest, but was unsuccessful, ultimately resulting in the purification of a tetrapeptide unrelated to the activity (Figure 2.14). In this following purification, the intention was to follow the biological activity at every stage of the process to avoid loss of the active compound. The full purification is described in detail in the methods.

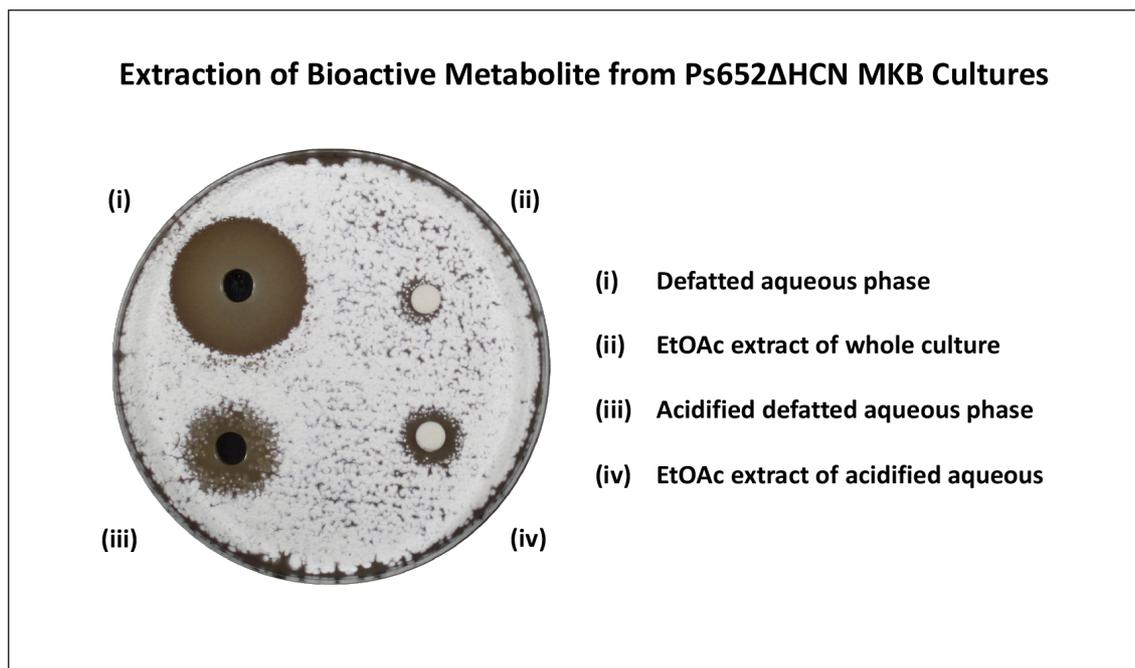


Figure 2.13. Extraction of a water soluble biologically active molecule from Ps652 Δ HCN. Culture was extracted 3x with ethyl acetate, and both aqueous and organic phases tested for biological activity, which remained in the aqueous phase. Acidification of a small sample (10 mL) followed by repeat ethyl acetate extraction failed to provide any significant benefit, with the majority of activity remaining in the aqueous phase, with some possible degradation. All wells or disks were loaded with the equivalent of 1 mL of original culture. The challenge organism used is *S. scabies* 87-22.

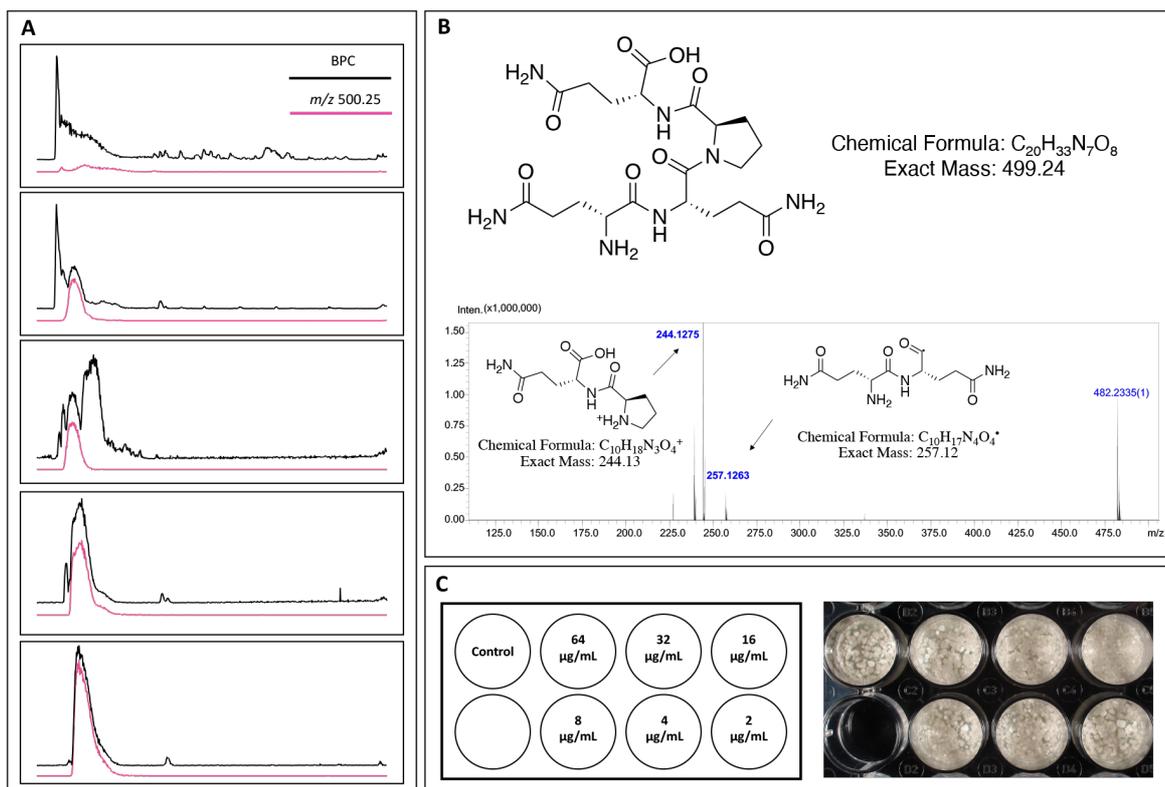


Figure 2.14. Purification of a tetrapeptide unrelated to the biological activity of Ps652 Δ HCN. The molecule was purified only partially guided by biological activity, and was not responsible for the suppression of *S. scabiei* by Ps652 Δ HCN. (A) Chromatograms from stages of the purification, with extracted ion chromatogram for the purified molecule shown at each stage in pink. (B) Predicted structure of the purified tetrapeptide, QQPQ, with MS² data showing fragment ions consistent with the structure. (C) Biological activity against *S. scabiei* 87-22. Pure compound was added to cooled IMA medium immediately before pouring at the concentrations shown. No activity was noted at any concentration.

2.4.1.1 Extraction from 5 Litre Culture Extract

To ensure sufficient material remained at the end of the purification process to allow for characterisation by NMR and mass spectrometry, as well as further biological assays, five litres of starting culture was used for fermentations. Ps652 Δ H₂CN was grown in 5 L of modified King's B medium, concentrated to 1 L by rotary evaporation, and extracted 3x with 500 mL EtOAc. Each fraction was tested for biological activity, and acidification was trialled on a 10 mL aliquot of the residual aqueous phase, followed by repeat EtOAc extraction, but did not provide any benefit or allow the molecule of interest to be extracted wholly into the organic phase (Figure 2.13). The defatted aqueous phase was used for further purification of the molecule.

2.4.1.2 8-step Bioactivity-guided Purification of Ps652 Molecule

With no prior knowledge of the nature or identity of the compound being purified, targeted chromatographic steps could not be designed. Instead, chromatographic methods were chosen based on the available stationary phases, with modifications made in each subsequent step based on the effectiveness of the previous step. As a result, eight steps of purification were needed, comprising three steps of flash chromatography for sample debulking, two steps of preparative-scale HPLC, and three steps of semi-preparative-scale HPLC (Figure 2.15). However, after the first step of semi-preparative HPLC, 20% of the total remaining material had to be used to see a zone of inhibition against *S. scabies*. Accordingly, the decision was taken to purify the largest peak in the sample observed by UV absorbance at 254nm, as it was expected the most dominant component was likely the target molecule. After a further two steps of semi-preparative-scale HPLC, the fraction containing this peak was determined to be sufficiently pure by LC-MS. The sample was observed to be a yellow/brown resin, of which there was 1.82 mg.

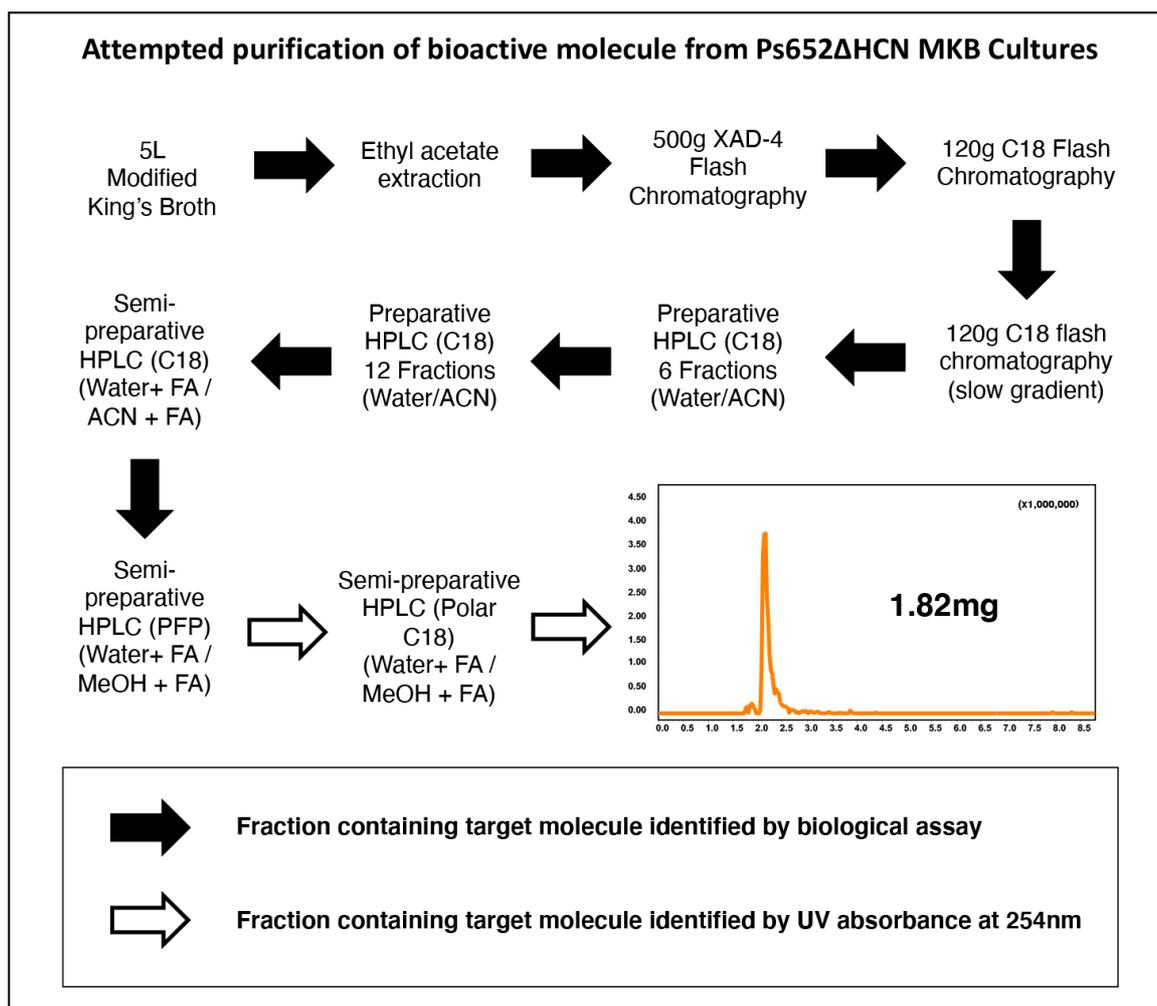


Figure 2.15. Purification of pyoverdine from Ps652ΔHCN by bioactivity-guided fractionation. An eight step purification was performed, following the biological activity by agar-well diffusions assay of all fractions at every possible step (filled arrows), until insufficient material remained to use this method, and the most prominent peak by UV at 254nm was used (hollow arrows). 1.82 mg of pure material was obtained, with a mass of 1328.6 daltons.

2.4.1.3 Identification of Ps652 Pyoverdine-related compound by HRMS and MS²

High-resolution mass spectrometry (HRMS) data was obtained for the molecule of interest, revealing an accurate mass of 1328.6189 daltons, substantially larger than had been anticipated. Given the similarity of this mass to *Pseudomonas* pyoverdine siderophores (331), the sample was redissolved in 1 mL of Milli-Q H₂O and imaged under UV light on a transilluminator. Consistent with pyoverdine, the molecule showed the characteristic fluorescence associated with these siderophores when imaged under UV light (Figure 2.16A). Analysis of MS² data revealed shared fragments with the ferribactin chromophore presented in ref (332) (Figure 2.16B). It was concluded on this basis that the purified compound was either the Ps652 pyoverdine, or a pyoverdine-related metabolite. This was confirmed through comparative metabolomics between Ps652ΔHCN and a newly generated pyoverdine-null derivative, Ps652ΔHCNΔPyo (Figure 2.17). The loss of a doubly charged ion, *m/z* 656.28 was observed, with no other obvious changes to the metabolic profile, and this ion corresponded to a doubly-charged dehydrated derivative of the purified molecule ($[M - H_2O + 2H]^{2+}$). This confirmed the purified molecule was a pyoverdine-related metabolite.

2.4.1.4 Anti-*scabies* Activity of the Purified Pyoverdine

Given that it had not been possible to test the HPLC fractions for biological activity after the sixth step of purification, it was necessary to test the purified pyoverdine-related metabolite for activity against *S. scabies* 87-22. Using the agar-well diffusion method, 100 μg of pure compound was added to an IMA plate inoculated with *Streptomyces* spores, in 100 μL of Milli-Q H₂O. No biological activity was evident (Figure 2.18), consistent with the previous data for a Ps652 pyoverdine deletion strain (Figure 2.8). It was concluded on this basis that the compound did not represent the target molecule. Efforts were therefore directed at uncovering the genetic basis of biosynthesis of the unknown compound, as efforts to purify the compound by bioactivity guided means had proven unsuccessful. In the face of no clues as to its origin, finding the genetic basis of the biocontrol phenotype would require a different approach.

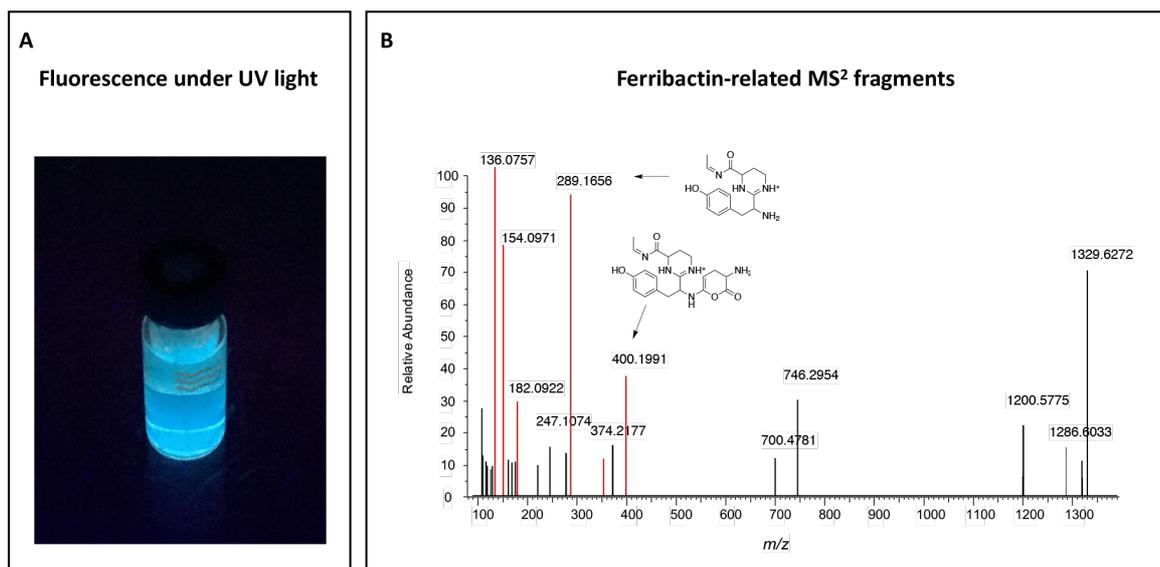


Figure 2.16. Fluorescence and MS² spectrum for the purified molecule from Ps652ΔHCN. (A) Purified molecule redissolved in Milli-Q H₂O and illuminated using a UV transilluminator, showing the characteristic fluorescence of pyoverdines. (B) MS² spectrum of the purified molecule, showing characteristic fragment ions that match with published fragments for the ferribactin chromophore highlighted in red.

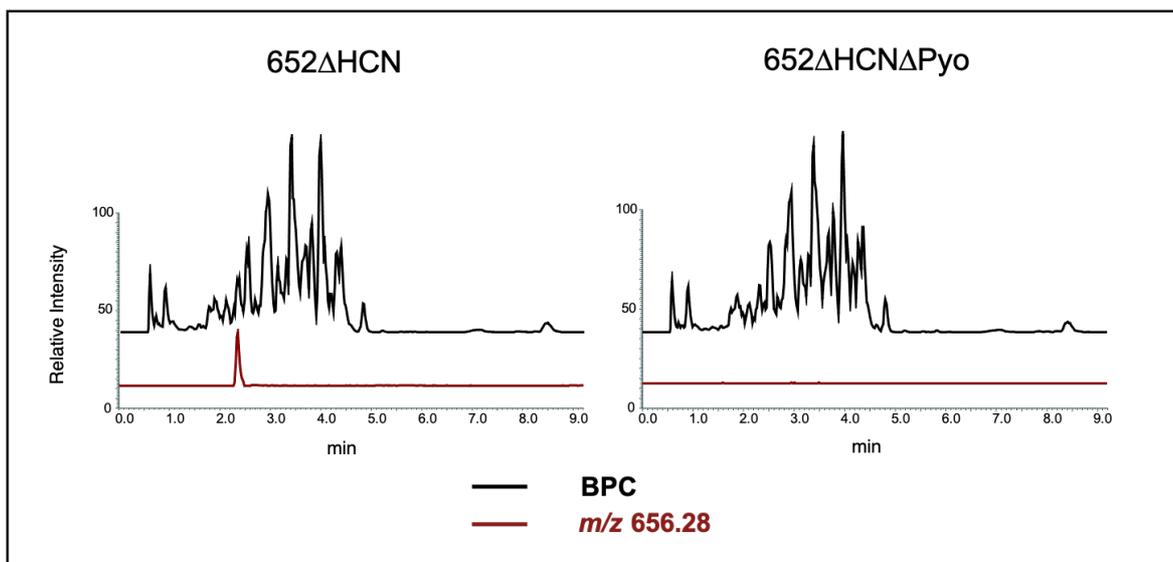


Figure 2.17. LC-MS of Ps652ΔHCN and Ps652ΔHCNΔPyo cultures. Base peak chromatograms (BPC) and extraction ion chromatograms for m/z 656.28(2) (red) are shown, indicating the loss of pyoverdine from culture supernatant in the pyoverdine-null mutant. m/z 656.28 corresponds to a doubly-charged dehydrated congener of the purified pyoverdine from Ps652ΔHCN ($[M - H_2O + 2H]^{2+}$). Data are representative of three biological replicates.

**Biological activity of Ps652
pyoverdine**



Streptomyces scabies 87-22

Figure 2.18. Biological activity of Ps652 pyoverdine against *Streptomyces scabies* 87-22. 100 μg of pure compound was dissolved in 100 μL Milli-Q H_2O and added to an agar well of an IMA plate onto which a lawn of *Streptomyces* spores had been spread. Plates were incubated until *S. scabies* had begun to sporulate. No zone of inhibition was observed.

2.5 Transposon Mutagenesis

Beyond the compound of interest being non-volatile, there were few clues as to its biosynthetic origin beyond it not being derived from the most common classes of biosynthetic pathways (PKS, NRPS, etc.), with the possible exception of the uninvestigated type III PKS identified by antiSMASH 5.0. As such, it was decided that the approach most likely to have success was to create and screen a library of random mutants. Again, given the lack of any known features, such as a characteristic UV absorption wavelength, or a mass identifiable by LC-MS approaches, the library would need to be screened phenotypically.

2.5.1 Screen Design & Considerations

In light of previous results demonstrating that HCN production was partially responsible for the inhibitory phenotype, the cyanide deletion mutant Ps652 Δ HCN was chosen as the parent strain for this work. This ensured that any hits in the screening stage would be due to new genes involved in the phenotype, rather than generating many false positive hits consisting of *hcn* mutants. Furthermore, it was possible that knockouts in important genes may be masked by the presence of HCN.

To be confident that the screen would generate useful data, a library was needed that comprised mutants with only one disrupted gene each, but was also completely random. To this end, it was decided to use the pALMAR3 plasmid, which contains the mariner transposon (333). Mariner-type transposons have inverted tandem repeats flanking a transposase gene, that is randomly inserted into the genome at TA dinucleotide sites (334). The Ps652 genome contains 228,924 such sites, meaning that the transposon was likely to insert more or less randomly across the genome without strong bias for specific genes or regions of the chromosome. Next, it was necessary to ensure enough mutants would be generated to elucidate the genetic basis for inhibition of phytopathogens. 30x96-well plates was deemed sufficient for this purpose, totalling 2,880 mutants. This was based on (i) any essential genes would be screened out at the library creation stage, as mutations in

essential genes would be lethal and so not produce pickable colonies and (ii) the estimated number of accessory genes in environmental *Pseudomonas* strains, which is around ~3000 (335). Accordingly, the screen ensured a library size comparable to the number of non-essential genes. This allows for detection of biocontrol determinants encoded by as little as a single gene. A phenotype controlled by more than one gene, such as traditional secondary metabolite pathways, would likely result in multiple hits from this library size.

Preliminary testing (data not shown), and advice from Dr D. Widdick regarding phenotypic screens of this kind, suggested that there was a significant risk of false negative hits if using a traditional 96-pin replicator approach (336) to inoculate the library from 96-well plates onto a single square plate for screening. Mutants that had lost the ability to produce a diffusible natural product compound would be masked by its production from surrounding colonies on the same plate. This therefore necessitated the inoculation of mutants in individual wells of a multi-well plate. Unfortunately, 96-well plates would prove too small for any loss of inhibitory effect due to diffusible compounds, as the *Pseudomonas* colony would occupy the entire well and may still be capable of contact-dependent inhibition. As such, it was decided to use 24-well plates for the screening stage. This presented significant complications, as each mutant would need to be transferred from 96-well format to 24-well format individually and reliably. With significant risk of human error and repetitive strain injury, it was decided to automate the library generation and screening in collaboration with the Biofoundry at the Earlham Institute (Norwich, UK). This work is published in ref (337) (see appendix 1). An overview of the screen design is presented in Figure 2.19. In short, Ps652ΔHCN was transformed with pALMAR3 and plated onto Lennox agar plates with tetracycline selection at a density that would best suit automated colony picking. Multiple separate transformations were performed to minimise the effects of clonality in the population. Colonies were picked automatically by Hamilton Microlab STARplus (Hamilton, Bonaduz AG, Switzerland), performed in collaboration with Dr A Elliston (Earlham Institute). Using the accompanying EasyPick II software, it was possible to set desired colony parameters to eliminate representation of unwanted clones in the library, and further screen out those with significant growth defects. Only colonies with a circularity factor over 0.05, and an area between 0.4 and 15mm² were picked. These were transferred into individual wells of 96-well plates, and grown overnight. These mutants

were then transferred into individual wells of 24-well plates, into which *S. coelicolor* M145 had been spread uniformly on RSA using the inbuilt shaker, by dipping the pipette tips into the well of 96-well plates without taking up any liquid then simply touching the surface of the agar in 24-well plates. *S. coelicolor* M145 was chosen as a non-pathogenic proxy for *S. scabies* 87-22, which at that time could not be used in the Earlham Institute Biofoundry due to plant health licence restrictions. Preliminary work had shown Ps652ΔHCN was capable of inhibiting the growth of *S. coelicolor* M145, and it was deemed unlikely that the strain would produce multiple different unknown antibiotics active against *Streptomyces* given the previous bioinformatic analysis. After 5 days of incubation, plates were imaged and mutants selected. The remaining steps will be dealt with in the following sections.

2.5.2 Screen Output and Mutant Selection

A library of 2,880 mutants was created as above, and screened using the automated bioassays against *S. coelicolor* M145 in 120x24-well plates. After five days of incubation, interspecies interactions were visible to the naked eye, and potential candidate mutants could be selected. All 120 plates were checked manually, and plates imaged at this stage. Some mutants were labelled 'low confidence' and others 'high confidence' hits. Low confidence mutants were those that had lost inhibition of *Streptomyces*, but appeared to either have compromised growth in RSA or still displayed a small, but visible, zone of inhibition. High confidence mutants were those whose growth appeared normal, and the streptomycete grew immediately adjacent to the pseudomonad with no visible zone of inhibition. Examples of both are presented in Figure 2.20. 23 high confidence hits were obtained from the screen, including 29C1 and 29D2 (Figure 2.20), and taken forward for manual verification. The advantage of the automated approach is that it allowed generation of a small number of mutants that could then be investigated in more detail using existing approaches.

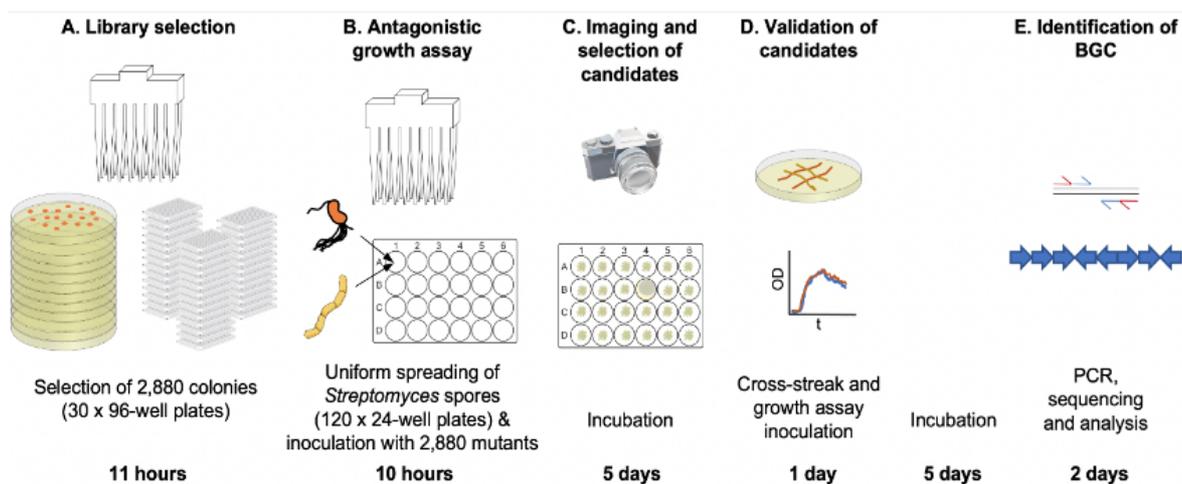


Figure 2.19. Design of the automated transposon mutagenesis workflow for Ps652 Δ HCN. (A) A library of 2,880 random mutants of Ps652 Δ HCN is picked automatically on a Hamilton Microlab Star system. (B) The library is screened using automation for clones having lost inhibitory activity towards streptomycetes in 24-well plates. (C) Candidate mutants are identified visually. (D) High-confidence candidates are screened manually for loss of inhibition against *S. scabies* 87-22, and growth defects. (E) Amplicon sequencing of mutants is performed, and the genetic determinants of the phenotype under investigation are elucidated. Timescales for each step are shown below each step. This approach offers significant advances in error rate, time, and reproducibility. Reproduced with permission from Moffat *et al.* (337) (see appendix 1).

2.5.3 Cross-streak Assays with Selected Mutants

These 23 high confidence hits were restreaked to single colonies, and new stocks made from overnight cultures of single colonies to ensure only a single genotype was present in each. To confirm that mutants unable to inhibit the growth of *S. coelicolor* were also unable to inhibit the growth of *S. scabies* 87-22, the 23 mutants were subsequently tested in the cross-streak assay for activity against the latter. All were tested in biological triplicate, which would not have been feasible with the full set of 2,880 mutants, which would have required 8,840 plates. The assay was performed as previously, with Ps652 Δ HCN as the positive control for comparison. Six of the 23 high confidence mutants were unable to inhibit *S. scabies* as well as *S. coelicolor* (Figure 2.21). These were 14G7, 19C1 24D9, 28E3, 29C1, and 29D2. It is not known why the other 17 failed to inhibit *S. coelicolor* but not *S. scabies*, and this was not investigated further. However, it is possible that the disparity could be accounted for partially by the different media used for each assay, as the automated screening was performed on RSA, and the cross-streak assay on MYM. 19C1 was identified by Sanger sequencing as a *recA* mutant. RecA is known to be involved genetic recombination, and so initiating DNA repair following a number of environmental stresses (338, 272). Accordingly, it is likely that 19C1 may contain multiple other mutations in the genome as a result, that could be effecting the phenotype observed. As such, it was treated as an off-target hit and not addressed further. The remaining five strains were taken forward as those likely to have transposon insertions in regions of the genome genuinely related to the inhibition of *S. scabies*.

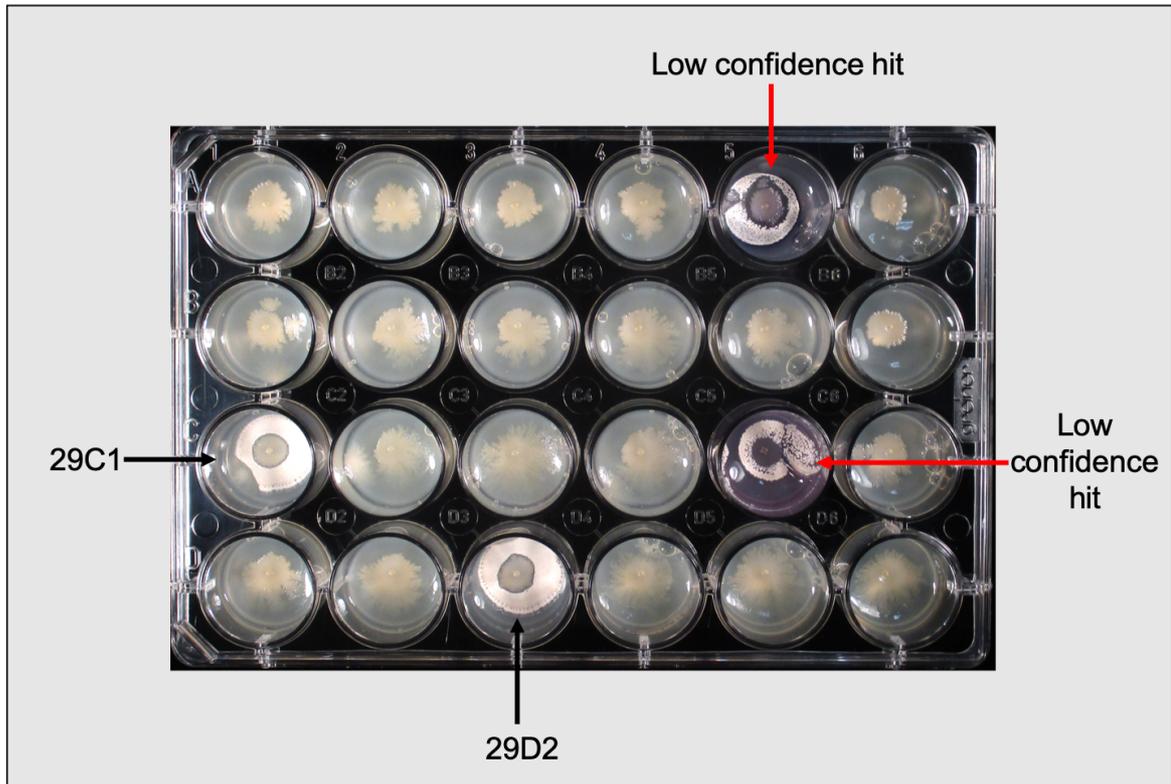


Figure 2.20. Representative hits from automated transposon mutant screening. Shown is a 24-well plate from the screening of transposon mutants. Low confidence hits and two high confidence hits, 29C1 and 29D2, are indicated. All other wells contain mutants growing sufficiently well and without comprised inhibition of streptomycetes, such that no visible growth for *S. coelicolor* M145 can be observed. Reproduced with permission from Moffat *et al.* (337) (see appendix 1).

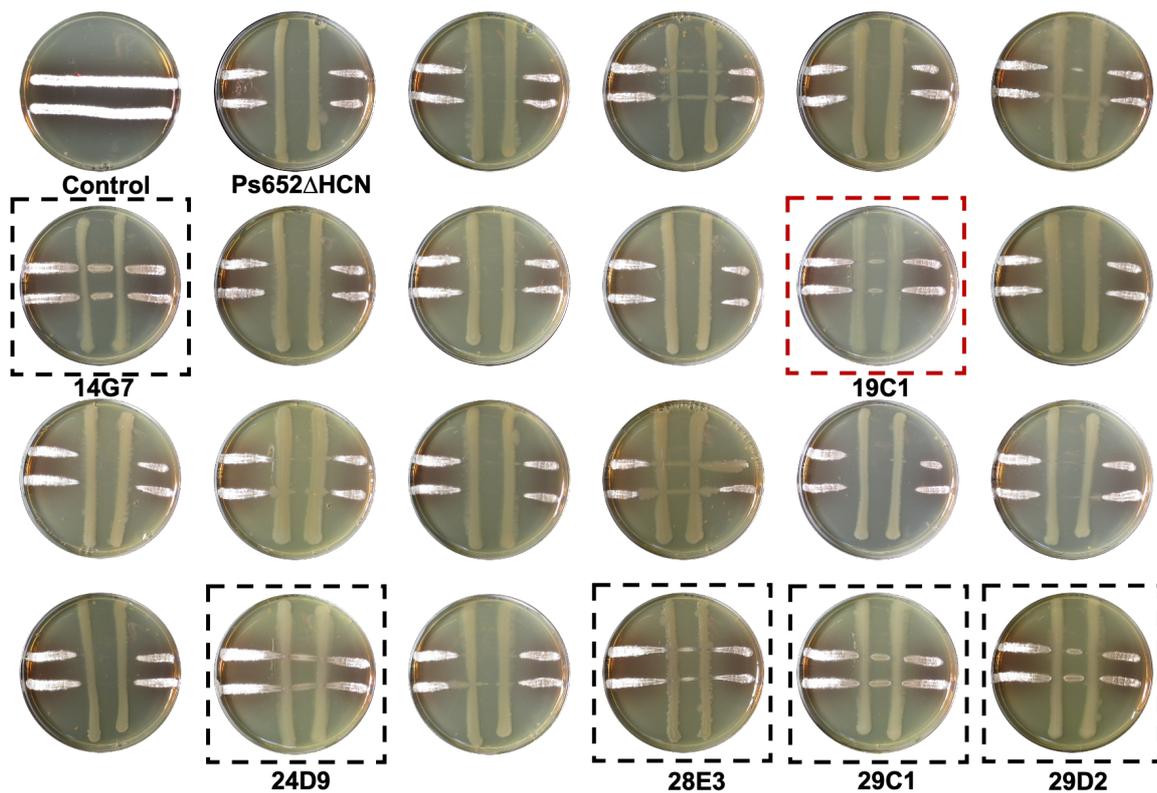


Figure 2.21. Cross-streak assays for 23 high confidence mutants identified through automated screening of the transposon mutant library. Compared to the control, only 14G7, 19C1, 24D9, 28E3, 29C1, and 29D2 showed loss of inhibition of *S. scabies*, evident by the streptomycete growing between the two *Pseudomonas* streaks in the plate. 19C1 (red box) was identified by amplicon sequencing as a *recA* mutant and not considered further.

2.5.4 Growth Curves & Further Mutant Selection

For the remaining five strains, it was essential to confirm that they were absent of any growth defects which could explain the loss of inhibition. While these strains had been observed to grow normally without any notable defects on both RSA and MYM, it was concluded that generating growth curves in liquid media would prove more definitive than visual inspection of colonies on solid agar. Growth curves were generated in triplicate for each strain, as well as Ps652 Δ HCN as a control, over 48 hours in LB medium, in a 96-well plate. Four of the mutants grew as well as, or better than, the parent strain Ps652 Δ HCN (Figure 2.22). One mutant, 28E3, showed an obvious growth defect in this assay with a longer lag time and lower maximal optical density. As such, 28E3 was not carried forward to the next stage. Amplicon sequencing revealed that 28E3 contained a transposon insertion within the RmlA-encoding *rfbA* gene responsible for the production of a thymidyltransferase involved in the biosynthesis of deoxythymidine diphosphate-L-rhamnose, a cell wall component (339). Again, this mutant was deemed an off-target hit, and so disregarded. At this stage, Four mutants remained (14G7, 24D9, 29C1, 29D2) that were both unable to inhibit *S. scabies* 87-22, and showed normal growth characteristics. It was hypothesised that these would contain transposon insertions in a gene or gene cluster responsible for production of an unknown natural product, linked to the biocontrol phenotype observed in earlier experiments.

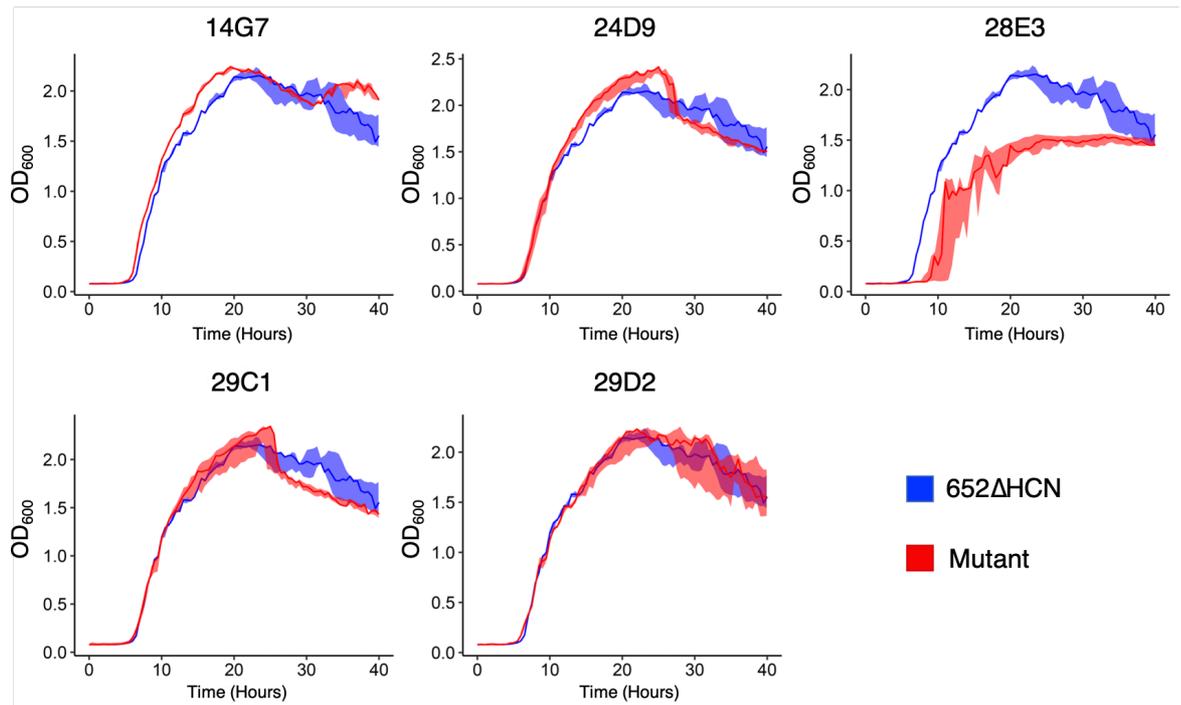


Figure 2.22. Growth curves for selected transposon mutants. Shown are the growth curves for transposon mutants 14G7, 24D9, 28E3, 29C1, and 29D2. All except 28E3 show normal growth characteristics when compared to the parent strain, Ps652ΔHCN. Line graphs represent three separate wells of 96-well plates each. Solid lines represent the median OD₆₀₀ values, and the shaded area represent the minimum and maximum values at each time point. Reproduced with modifications with permission from Moffat *et al.* (337) (see appendix 1)

2.5.5 Transposon Sequencing & Tropolone Cluster

To identify the genetic basis for loss of inhibition towards *S. scabies* in the selected transposon mutants, and hopefully identify the genetic basis for *S. scabies* inhibition in Ps652, amplicon sequencing of transposon insertion sites was performed for these mutants as per the methods. The retrieved sequences were mapped back to the Ps652 genome using a local installation of BLAST, which bypassed the potential issue of inexact sequence matches generated by PCR or sequencing errors. All four mutants contained transposon insertions within the same 8 kbp region of the chromosome, in what resembled a biosynthetic gene cluster (Figure 2.23). This region had not been identified during antiSMASH 5.0 analysis (328). However, during the preparation of a manuscript for publication of these results (337), a study was published that identified this cluster through a similar transposon mutagenesis approach. That study linked the cluster to antifungal activity toward *Macrophomina phaseolina* by a *Pseudomonas donghuensis* isolate, but not to antibacterial activity toward *E. coli* or *B. subtilis* (340). This contrasted with previous evidence that the genomic region was in fact involved in antibacterial activity of *P. donghuensis* P482 (341). It is likely that by conducting their transposon screen in a strain without any antibacterial gene clusters knocked out, the ability to link that cluster to antibacterial activity was limited. In their study, they claim to identify the biosynthetic product of the gene cluster as 7-hydroxytropolone, which they were able to extract from cultures using EtOAc and verify its antifungal activity. These results stood somewhat in contrast with our own, as Ps652 was not noted as particularly antifungal, and the compound of interest appeared to be water-soluble rather than soluble in organic solvents. This cluster is investigated further in section 3.1.

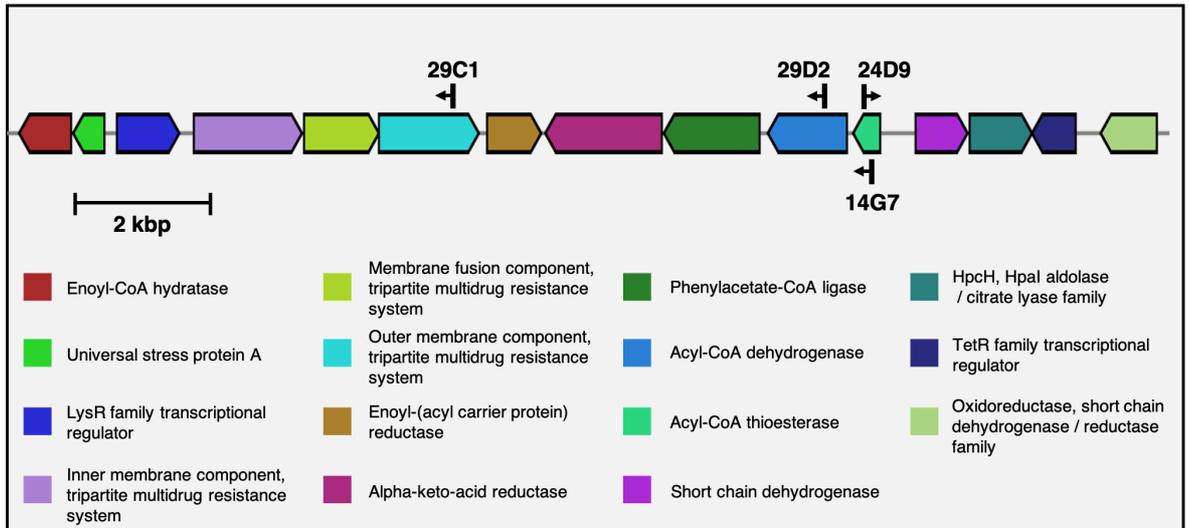


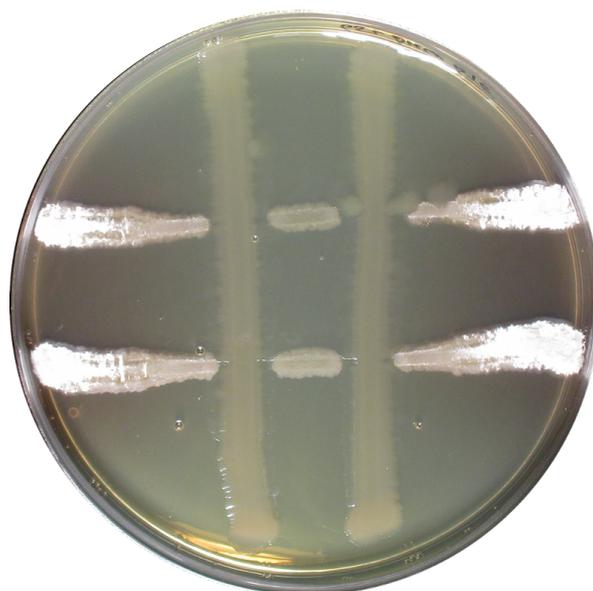
Figure 2.23. Putative biosynthetic gene cluster identified in Ps652 through amplicon sequencing of transposon mutants. Shown are the sites and directions of transposon insertions in the four selected transposon mutant strains. Genes are colour coded with the key showing their annotations. Reproduced with modifications with permission from Moffat *et al.* (337) (see appendix 1).

2.6 Validation of the Phenotype Using Clean Knockouts and Complementation

2.6.1 Deletion of the Tropolone Cluster Thioesterase

It was necessary to confirm that the phenotypes observed by transposon mutagenesis were really due to the genes where transposon had inserted, rather than through polar effects mediated by interruption of readthrough transcription or stimulation of genes downstream of the transposon, which has been known to occur with the pALMAR3 system (J Malone, personal communication). The most commonly hit gene in the transposon mutagenesis screen had been the predicted acyl-CoA thioesterase in the predicted tropolone cluster, where a transposon had inserted in each direction (Figure 2.23). This protein is has 72% identity to TdaD from *Phaeobacter inhibens* (accession AFO93377). Therefore this seemed the most appropriate candidate for targeted knockout creation, as it was the most likely to yield a phenotype. A clean deletion mutant of this strain was obtained by markerless allelic exchange and confirmed by Sanger sequencing. Given the lack of consistency in nomenclature for this cluster to date, and the lack of confirmed function for this gene, the resulting strain was named Ps652 Δ HCN Δ TE.

This strain was tested for activity against *S. scabies* 87-22 using the cross-streak assay. No inhibition of *S. scabies* was observed, as the phytopathogen was able to grow between the two *Pseudomonas* streaks (Figure 2.24). This is indicative of loss of production of a diffusible, non-volatile natural product with antibacterial activity. This matched the results observed for transposon insertion mutants 14G7 and 24D9, confirming this cluster is directly involved in the biological activity displayed for Ps652 against potato pathogens.

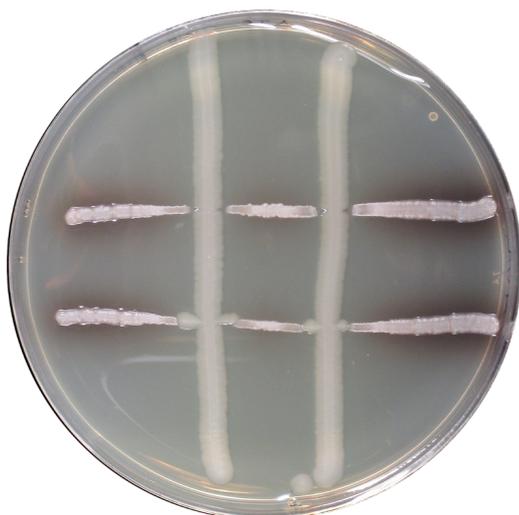


Ps652 Δ HCN Δ TE

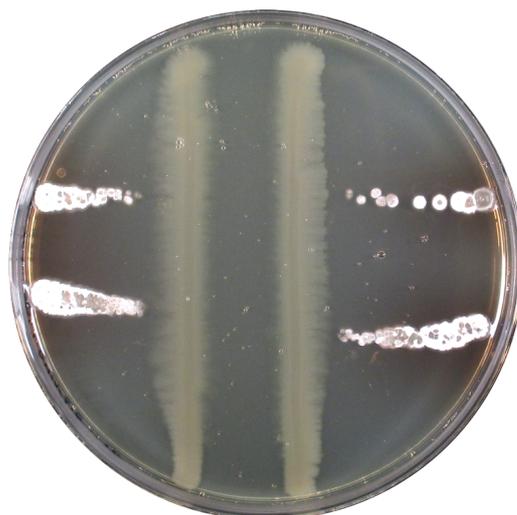
Figure 2.24. Cross-streak assay for Ps652 Δ HCN Δ TE against *S. scabies* 87-22. The markerless knockout of the acyl-CoA thioesterase from the Ps652 tropolone BGC was unable to inhibit the growth of *S. scabies*. Image is representative of three biological replicates.

2.6.2 Complementation of Acyl-Coa Thioesterase in Ps652 Δ HCN Δ TE

To confirm that the observed phenotype for the thioesterase mutant was due only to the absence of the thioesterase itself, the acyl-CoA thioesterase was cloned initially into pUC18-mini-Tn7-Gm. This vector is an integrative Tn7-based expression vector, which integrates into the Tn7 site in the *Pseudomonas* genome and allows for single-copy complementation of deletion strains *in trans* (342). This construct was successfully obtained, but all attempts at generating transformants failed, with no integration of the construct into the Ps652 genome. The acyl-CoA thioesterase was therefore cloned into a replicative plasmid, pME6032, which are indicated for use in Gram-negative plant-associated bacteria, and is a relatively low copy-number vector with approximately 15 copies per genome (343). The construct pME6032::TE was created, and transformed into Ps652 Δ HCN Δ TE, with an empty plasmid-containing strain made as a control. These were then tested for biological activity against *S. scabies* 87-22, to confirm whether the thioesterase mutant could be successfully complemented (Figure 2.25). Ps652 Δ HCN Δ TE pME6032::TE was able to inhibit the growth of the phytopathogen to the same extent as Ps652 Δ HCN, indicating the deletion could be successfully complemented, and cementing the role of the predicted acyl-CoA thioesterase in the biosynthesis of a natural product capable of inhibiting *S. scabies* 87-22.



**Ps652 Δ HCN Δ TE
pME6032**



**Ps652 Δ HCN Δ TE
pME6032::TE**

Figure 2.25. Complementation of the Ps652 tropolone cluster thioesterase. Cross-streak assays confirmed the thioesterase in the Ps652 tropolone BGC was able to be complemented *in-trans*. Images are representative of three biological replicates.

2.7 Discussion & Conclusions

2.7.1 Discussion

It is lamentable that in a large percentage of studies into soil bacteria capable of improving crop health either through inhibition of pathogens or stimulation of crop growth, the genetic factors governing these phenotypes are not elucidated. This is obviously complicated by multifactorial inhibition, where non-additive effects are observed for various molecules, making it difficult to make direct molecule-phenotype links (7). Again, a great deal of this might stem from our misunderstanding of how natural products are used in their native environment by the producing strains (22). However, the work presented here shows there is potential for deconstruction of the genetic factors determining biocontrol activity, at least in some strains.

Initial biological assays conducted as part of screening of a larger collection of environmental pseudomonads highlighted the promise of strain Ps652 for biocontrol applications. From an initial analysis of a short-read genome for this strain using prevalent genome mining tools, it was evident that there existed potential for discovery of previously unknown natural products involved in the strong inhibitory phenotypes exhibited by this strain against potato pathogens from different domains of life. Manual BLAST searches found the potential to synthesise hydrogen cyanide by glycine oxidation, which was subsequently confirmed with a visual biological assay. It was an unexpected finding, considering the strong history of biocontrol phenotypes being linked to cyanide production (126, 103), to discover the activity of Ps652 was not explained by this molecule nor other volatiles. Subsequent indication that the biological activity resulted from a solvent-extractable molecule made this strain a high priority for further work, with the opportunity to discover novel natural products that may be relevant to agriculture.

A higher quality, single-scaffold genome was obtained that could be used for further genetic investigations into this strain. This was made possible by advances in long-read sequencing and falling prices. The results highlight the potential of selective isolation and high-throughput phenotypic screening that was discussed in section 1.4, where novel

molecules of interest are still being found from studies of this kind. Crucially it is the dereplication step, using genome mining tools on a subset of strains with the biological activity of interest, that helps mitigate the issue of molecule rediscovery that has plagued natural product discovery since the end of the golden age (167). By prioritising strains that show a desired phenotype, but seem not to possess previously characterised BGCs, it is likely that new molecules with industrially, agriculturally, or societally relevant properties will be uncovered. The ever-growing repositories of BGC information simply augment the power of this approach (344). Following dereplication and prioritisation of a strain, an approach is needed to uncover the genetic basis of inhibition, to link any natural products that may be found by traditional chemical techniques to a BGC. Random mutagenesis approaches with transposon-based systems in genetically tractable isolates are a useful tool to this end, when combined with appropriate high-throughput screens that maximise the opportunities to generate a meaningful result. In this case, at the time the work was carried out, no natural product had been linked to the activity and the only possible readout of a high-throughput screen was the bioactivity itself. While high-throughput screens are incredibly powerful, the logistical constraints can often prove limiting.

In the work presented here, a robotics-based automated screen helped circumvent some of those issues, and screen sufficient colonies in a suitable format to uncover even single-gene determinants of inhibition based on the estimated size of the accessory genome in the average environmental pseudomonad (335). This led to the discovery of a gene cluster that had previously been linked to production of tropolone natural products (345), also originally found through a transposon mutagenesis screen like the one presented here (346), and rediscovered by others using the same approach with different screen outputs (341, 340). The multiple independent discoveries of this cluster validate the approach, but also highlights the importance of well-curated databases of natural product BGCs, to prevent multiple rediscoveries of a BGC. At the time this work was performed, the tropolone BGC was not found by antiSMASH (328), and the existing literature did not report the BGC and product in the same publication until mid-2018, when the transposon mutagenesis work presented here had already been performed (347). The cluster was not linked to biocontrol activity until the work of Muzio *et al.* (340), which was published while the manuscript for the present work was in preparation (337), meaning the BGC was not

searched for directly in the Ps652 genome. While manual BLAST searches were carried out to look for natural product gene clusters that are known to be involved in biocontrol but may not have been picked up by automated genome mining tools, a systematic approach was not taken to search for all genes reported to have such phenotypes. While this is certainly feasible, it is arguable that a random mutagenesis approach allows for both known and unknown genetic determinants to be found, while a systematic approach to manual BLAST searches might only serve as further initial dereplication, and targeted mutants would need to be made in any observed clusters to rule them out. It is likely that transposon mutagenesis was more time-efficient and likely to generate a meaningful result in this case. Confirmation of the observed phenotypes by ruling out growth defects, as well as clean knockouts and complementation experiments, added further validation to the workflow and made clear that the identified cluster was genuine.

This work has now linked a gene cluster responsible for the production of class of 7-membered ring natural products to the inhibition of commercially-relevant potato pathogens, and validated that link with targeted mutagenesis and complementation. The outstanding questions were whether Ps652 indeed produces tropolone natural products as is suggested in the literature, and what the exact identities of those molecules are. Additionally, there remained unsolved questions about which genes are essential for the production of those molecules, and what the genetic boundaries of the cluster are. Thus far, not all of the genes within the cluster have proposed functions in the biosynthesis of these molecules, making it unclear what the minimum BGC is.

2.7.2 Conclusions

This study indicates there is life yet in the 'new-strain, new-molecules' approach to natural product discovery, and that this approach has value to the probing of microbial interactions in an agricultural or biocontrol setting. By combining modern approaches like genome mining with stalwarts of the field like selective isolation and phenotypic screening, it is possible to discover new determinants of antimicrobial activity. This work investigated the biosynthetic potential of a promising field isolate from the *Pseudomonas* genus, and has linked the production of tropolonoid natural products to inhibition of phytopathogens

through a random mutagenesis-based screen. However, there remain unsolved questions around the biosynthetic pathways and the true application of these findings to agricultural settings.

Chapter 3: Production of Tropolones by Ps652

3.1 Biosynthesis of Tropolone

3.1.1 Chapter Aims

The objectives of this chapter were to investigate the biosynthesis of tropolone natural products by the environmental *Pseudomonas* isolate Ps652. Firstly, the intention was to ascertain whether Ps652 produces 7-hydroxytropolone, like other pseudomonads with the cluster found by transposon mutagenesis work (337), and confirm the activity of 7-hydroxytropolone towards *Streptomyces scabies* 87-22. Secondly, given the lack of clarity around tropolone biosynthesis in *Pseudomonas* at this time, gene knockouts combined with mass spectrometry approaches were to be used to establish which genes are essential for the production of these molecules. Thirdly, as tropolone-containing molecules are structurally diverse, and it is not known whether 7-hydroxytropolone is the end product of the *Pseudomonas* pathway, any additional metabolites from this cluster were to be purified and their contributions toward biological activity investigated.

3.1.2 Overview of Tropolone Natural Products in Bacteria

Tropolones are a class of hydroxylated forms of tropone (cyclohepta-2,4,6-trienone) with a seven membered, heteroatomic, conjugated ring structure (Figure 3.1A). They are found widely in nature being made by plants, fungi, and bacteria (348). As diverse as their producers, the structures of tropolone-containing molecules are highly variable, with some being simple hydroxylated forms of the basic tropolone skeleton (shown in Figure 3.1B), and others containing sulphur groups, or larger scaffolds with a tropolone motif (Figure 3.1C). A large range of bioactivities is observed for these molecules, including antibacterial, antifungal, insecticidal, antiviral, and anticancer (349). As such, bioactivity guided screening approaches have yielded a bounty of tropolone motif-containing natural products (348). The various activities of these molecules are often linked to their metal-binding and redox properties. The structures, ecological roles, and biosynthesis of many of these compounds are elaborated upon in previous comprehensive reviews (349, 350, 348). The following

sections will cover examples of tropolonoid natural products in bacteria, as well as what is known about their biosynthesis.

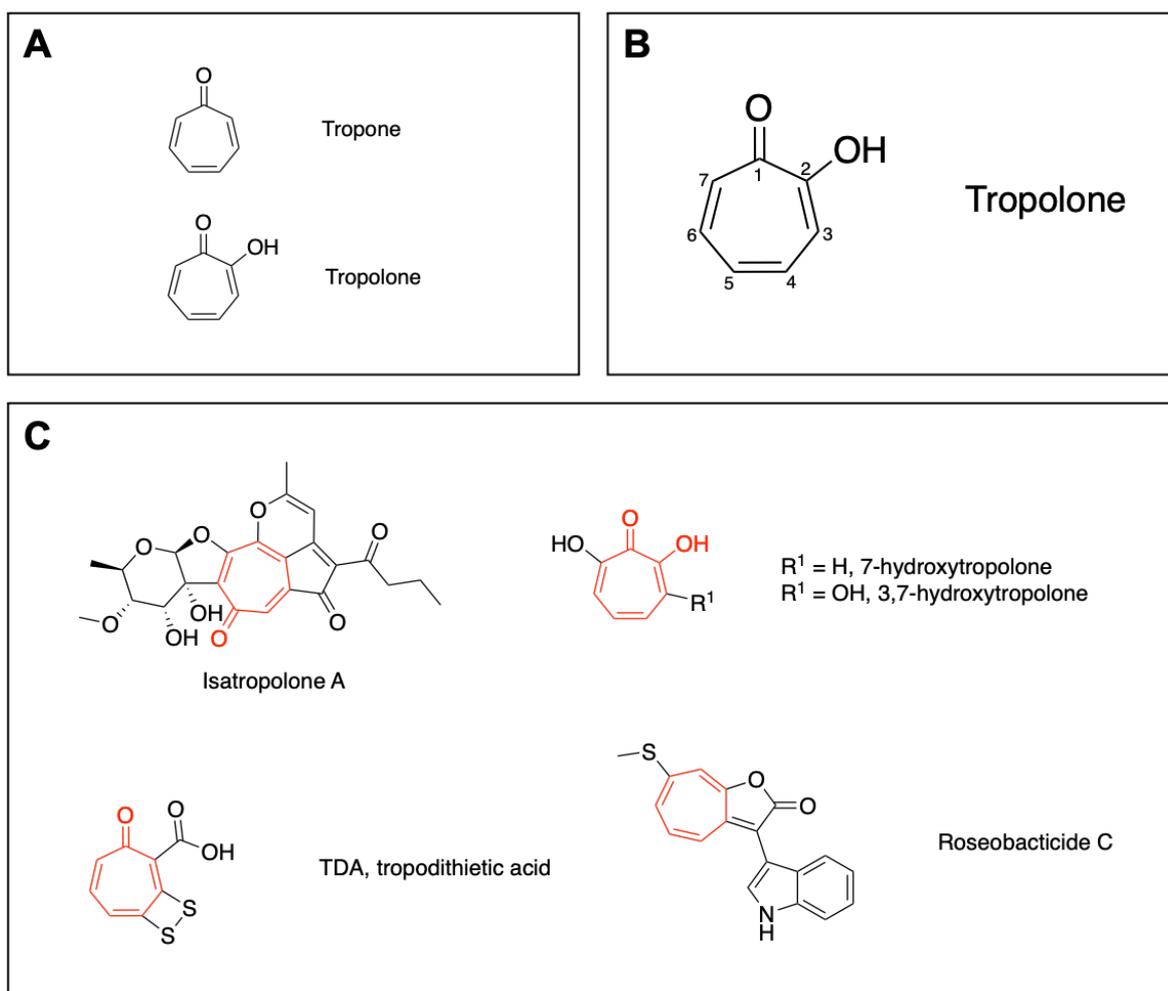


Figure 3.1. Tropolone natural products. (A) Tropone is the basic cyclohepta-2,4,6-trienone, and additional of an alcohol group creates tropolone. (B) Numbering of carbon atoms in the tropolone skeleton with the ketone in the C1 position. In tropolone, the alcohol is in the C2 position. (C) Diverse tropolone natural products, with the tropone / tropolone motif highlighted in red.

3.1.3 Tropolone Examples in Bacteria

3.1.3.1 Isatropolones

The isatropolones are a group of tropolonoid natural products from *Streptomyces*, first described in *Streptomyces* Gö66, which have a glycosylated tropolone core (Figure 3.1C) (351). They were initially noted for their ability to react with cellular nucleophiles, being amine-scavenging, and isatropolones A and C were demonstrated to have significant activity against *Leishmania donovani*, the causative agent of visceral leishmaniasis (351). This activity was linked to the 1,5-diketone moiety, which the authors noted can readily form pyridine rings with amine groups from protein amino acid residues like lysine. Of particular interest was the specificity towards *L. donovani*, making it appear to be a suitable drug candidate, but it was ineffective in *Leishmania*-infected macrophages. Very little antibacterial activity was noted originally, but isatropolone C was later observed to be active against the commercially-relevant plant pathogen *Streptomyces scabies* (321). The isatropolones are precursors of the rubrolones, which were originally discovered from *Streptomyces echinoruber* and proposed to be suitable candidates as food colouring agents given their red colour and lack of toxicity or antimicrobial activity (352).

3.1.3.2 Roseobacticide and Tropodithietic acid

Phaeobacter inhibens is a member of the *Roseobacter* clade that lives in a proposed symbiosis with an algal host in surface waters. Under normal conditions, the bacterium produces growth factors and a sulphur containing tropolonoid antibiotic called tropodithietic acid (TDA) as shown in Figure 3.1C. TDA displays a variety of activities, including bactericidal activity against Gram-positive and Gram-negative bacteria (353). This activity can be attributed to the proposed ability of TDA to disrupt the proton motive force (354), which has also led to its investigation as an anticancer agent. *Phaeobacter* exhibits resistance to this mode of action only during production of the compound itself, due to co-regulation of the production and immunity pathways (350). Additionally, it seems that TDA can act as a quorum sensing molecule in *Phaeobacter*, producing similar effects to acyl-homoserine lactone, affecting motility and biofilm formation (355). In a fascinating twist, it

was noted that upon sensing production of p-coumaric acid by their algal host, signalling senescence, the *Phaeobacter* switches to a parasitic life stage and produces the roseobacticides (356), potent toxins that kill the algal partner at nanomolar to micromolar concentrations. This duality was referred to as 'Jekyll and Hyde' chemistry. There are a number of different roseobacticide congeners, named roseobacticide A-K, with differing R groups at two variable sites, ranging from simple additions of benzyl groups, to larger modifications including an additional 7-membered ring in roseobacticides J and K (350). An example, roseobacticide C, is shown in Figure 3.1C.

3.1.3.3 7-hydroxytropolone

7-hydroxytropolone represents a minimally modified congener of the basic tropolone skeleton, differing only by the addition of a hydroxyl group in the C-7 position (Figure 3.1C). This modification renders the molecule symmetrical, like tropone. This molecule is well documented in a variety of bacteria, both in Gram-positives like *Streptomyces* and in Gram-negatives like *Pseudomonas*. It is documented to be a low-cost, non-fluorescent siderophore in *Pseudomonas*, with a relatively low affinity for iron compared to higher-cost traditional siderophores like the pyoverdines (345). Jiang *et al.* demonstrated that 7-hydroxytropolone forms a 2:1 complex with iron (Figure 3.2), and is produced at intermediate levels of iron but in larger quantities than other siderophores (345). This allows for consistent sequestration of iron under a variety of conditions, but switching to a higher affinity iron-sequestration tactic once iron levels drop below a given threshold. A dual benefit is that tropolone also offers competitive advantage through antagonism toward other bacteria, both in bacteriostatic and bactericidal manners (357). Additionally, there is documented activity against some insects and plant pathogenic fungi (358). While these activities were documented for tropolone rather than its 7-hydroxy derivative, it is likely this congener exhibits a similar spectrum of activity. In accordance with this hypothesis, Muzio *et al.* discovered 7-hydroxytropolone was responsible for the antifungal activity of *P. donghuensis* SVBP6 through a transposon mutagenesis screen (340). 7-hydroxytropolone is also documented to inhibit metalloproteases as well as aminoglycoside-2''-O-adenylyltransferase, leading to its investigation as an agent able to potentiate the use of aminoglycoside antibiotics against resistant bacterial strains (359).

Generally, it is assumed the activity of hydroxylated tropolones is due to inhibition of bimetallic enzymes, like the Mg^{2+} dependent HIV type I reverse transcriptase (350), or their ability to interfere with disrupt respiratory chains (348).

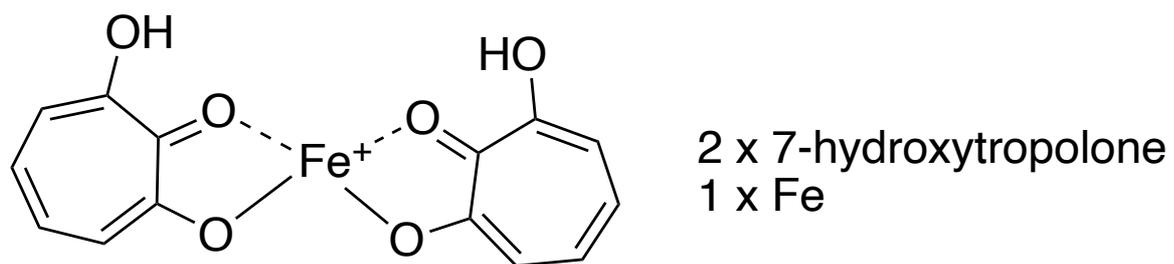


Figure 3.2. Formation of iron-chelates by 7-hydroxytropolone. 7-hydroxytropolone is known to form a 2:1 complex with iron, capable of being used as a siderophore in *Pseudomonas donghuensis*.

3.1.4 Biosynthesis of Tropolones in Bacteria

Their variety of structures is reflected in the differing modes of biosynthesis, as these molecules can be made by a variety of routes, but frequently rely on oxidative expansion of a 6-membered ring. The occurrence and biosynthesis of this class of molecules in bacteria is covered extensively in ref (350). To illustrate the variety with some examples, the isatropolones in *Streptomyces* species are biosynthesised by a type II polyketide synthase route. The same genus also produces simpler tropolones like 7-hydroxytropolone and 3,7-hydroxytropolone by a simpler mechanism involving a branching of the phenylacetic acid (paa) catabolic pathway, which was biochemically unravelled by Teufel and colleagues in 2011 (360). The paa pathway is observed in 16% of sequenced bacteria and involved processing of intermediates as CoA-thioesters with a seven-membered oxygen-containing ring intermediate. A mechanism thought to be a common pathway to producing the precursor for most tropolonoid natural products in bacteria, is the use of a PaaZ analogue, retaining the enoyl-CoA hydratase (ECH) activity but lacking the aldehyde dehydrogenase (ALDH) activity usually seen in PaaZ, is able to catalyse ring opening, generating this highly reactive intermediate that undergoes a spontaneous Knoevenagel condensation to the seven-membered ring (Figure 3.3). It has been demonstrated that a small amount of the highly reactive intermediate is formed even with the wild-type PaaZ containing both domains (361).

Accordingly, any bacteria that degrade phenylacetic acid are potential producers of tropolone natural products, and mutations in the ALDH domain of PaaZ can quickly shift the balance of production towards these molecules. A conserved protein in the *paa* gene cluster, PaaY, is a thioesterase with high substrate specificity to the CoA bound intermediate after Knoevenagel condensations, liberating CoA and preventing cellular CoA stores being depleted by a metabolic dead end, releasing 2-hydroxy-cyclohepta-1,3,5-trienecarboxylic acid (361). A subsequent decarboxylation and dehydrogenation could feasibly produce basic tropolone molecules (Figure 3.4). This is similar to how biosynthesis proceeds in *Streptomyces cyaneofuscatus* Soc7, where genes for increased precursor supply and downstream modification of the basic tropolone are encoded separately in the genome, but still rely on the paa pathway (Figure 3.5).

A flavoprotein dioxygenase (TdaE), conserved in *Phaeobacter* and *Burkholderia* tropolone BGCs, was recently characterised in *Burkholderia* (362). Independent of the PaaY thioesterase mentioned above, TdaE performs multiple steps including dehydrogenation of the CoA thioester intermediate, thioester cleavage, and epoxidation, to form an epoxide. This epoxide may either undergo spontaneous loss of CO₂ to form tropolone, or reaction with nucleophiles to form sulphur-containing congeners like tropodithietic acid (362) (Figure 3.4).

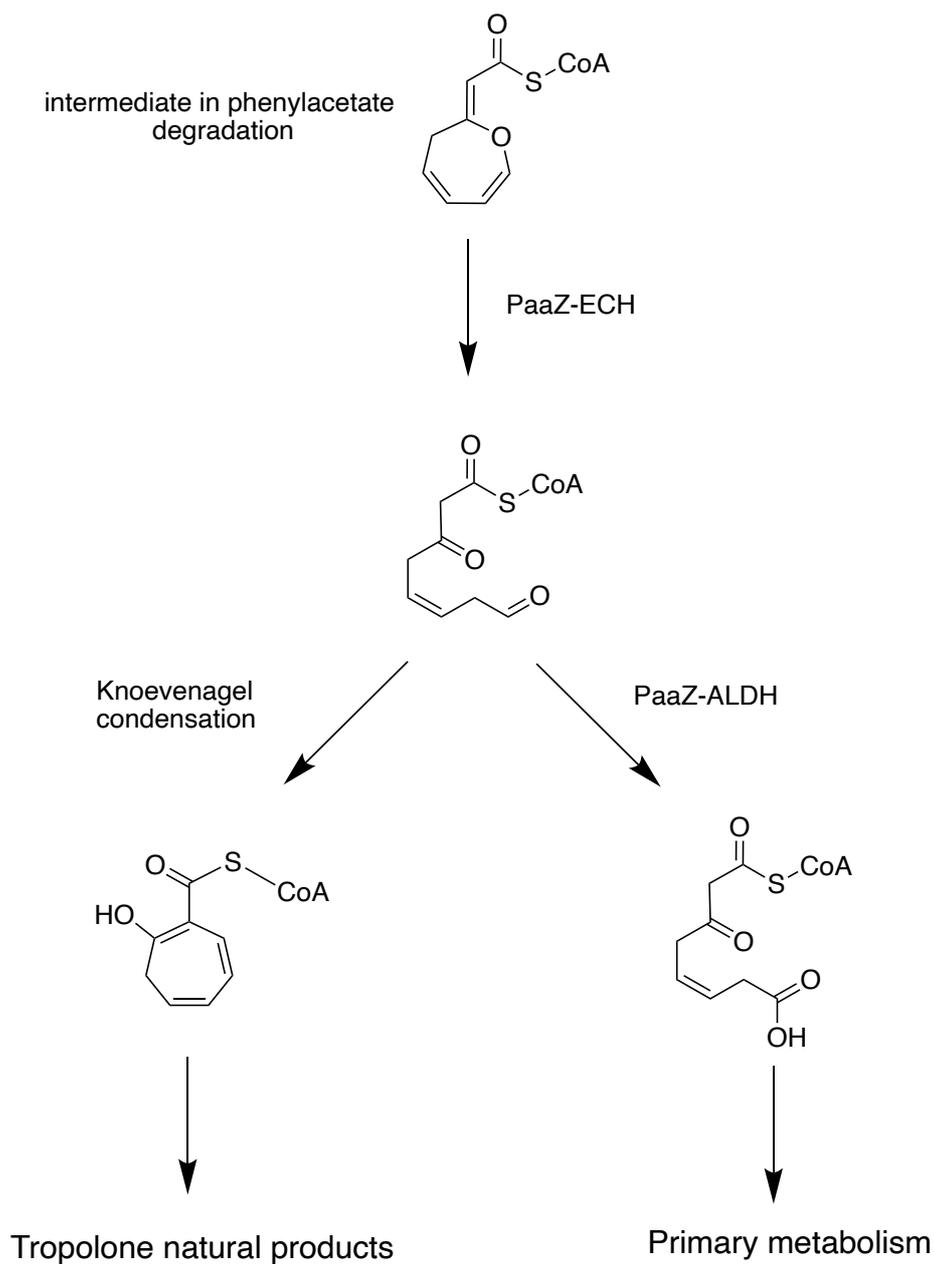


Figure 3.3. Branch point between phenylacetate degradation and tropolone biosynthesis. The enoyl-CoA domain of PaaZ can produce a highly reactive intermediate that is usually processed primarily by the aldehyde dehydrogenase domain of the same protein, but may proceed by spontaneous Knoevenagel condensation to tropolone natural products, particular in the case of standalone ECH domains.

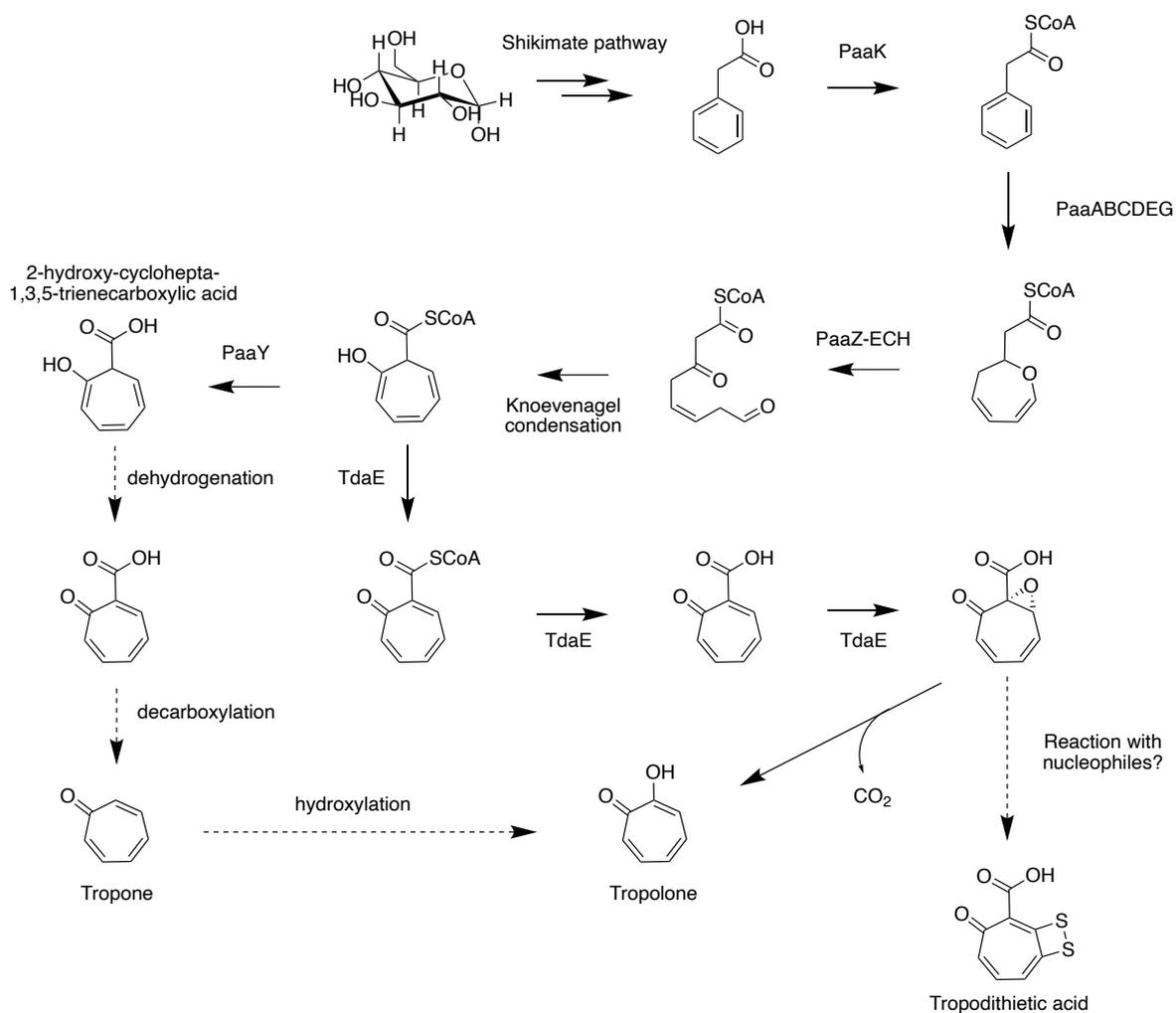


Figure 3.4. Mechanisms of producing tropolone natural products. Primary metabolism supplies precursors which are processed by the phenylacetate catabolon, until the PaaZ branch point. The endogenous PaaY from the *paa* cluster can then cleave the 2-hydroxy-cyclohepta-1,3,5-trienecarboxylic acid-CoA thioester. Subsequent dehydrogenation and decarboxylation may feasibly produce tropone, which could be modified by hydroxylation to form tropolone. Alternatively, TdaE cleaves the thioester in *Burkholderia* and generates a reactive epoxide that may undergo spontaneous decarboxylation to tropolone, or react with nucleophiles to produce sulphur-containing tropolone natural products like tropodithietic acid.

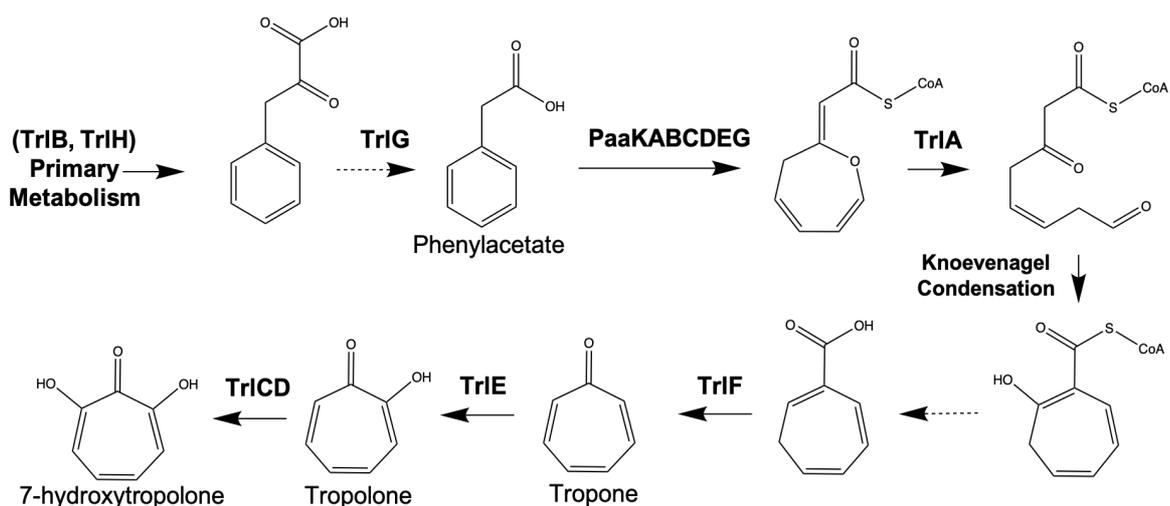


Figure 3.5. Tropolone and derivatives biosynthetic pathway in *Streptomyces*. Primary metabolism supplies the precursors for this pathway, converging through phenylacetate. TrIB and TrIH are involved in precursor supply, and TrIG shunts the products towards phenylacetate. The endogenous phenylacetate catabolic pathway is required for the next steps, and represents a branching point from primary to secondary metabolism where a truncated single domain protein TrIA, homologous to the enoyl-CoA hydratase domain of PaaZ, catalyses ring opening. A spontaneous Knoevenagel condensation then occurs to form a 7-membered ring, directing the pathway towards tropolonoid production. It is unknown whether cleavage from CoA occurs spontaneously or requires an unknown enzyme. TrIF then catalyses formation of tropone, which is hydroxylated sequentially by TrIE and TrICD to form tropolone and 7-hydroxytropolone respectively.

3.1.5 Tropolones in *Pseudomonas*

In 1985, tropolone was found in a culture extract from '*Pseudomonas plantarii*', now *Burkholderia plantarii*, as a virulence factor that acted to repress root growth in rice plants (363). Since that original publication and the reclassification of *P. plantarii* to *B. plantarii*, there has been little published on tropolonoid natural products in the *Pseudomonas* genus. That was until it was discovered in a screen for iron chelators (siderophores) in a new species of this genus, *P. donghuensis*, of which there are now four sequenced strains: P482 (341), SVBP6 (364), HYS (365), and 22G5 (366). Using a transposon mutagenesis approach, the investigators identified a cluster of genes in *P. donghuensis* HYS that was responsible for the production of a nonfluorescent siderophore (346). This was then later identified as being 7-hydroxytropolone (345), which is produced at intermediate iron levels as a low-cost iron scavenging mechanism, instead of the more metabolically expensive pyoverdine as described above. This same cluster was then independently identified in the present work in Ps652, also using a transposon mutagenesis screen, in an effort to find determinants of biocontrol activity (see chapter 2).

The research on 7-hydroxytropolone in *Pseudomonas* subsequently turned to its regulation in these bacteria, where it is understood to be under the control of the Gac/Rsm system and positively regulated by GacA/GacS (346). Other inputs on production of 7-hydroxytropolone are, as expected, the iron levels in the culture, which act through repression of LysR and TetR/AcrR family transcriptional regulators that flank the biosynthetic gene cluster (347). Maximal production of 7-hydroxytropolone was determined to occur between 1 μ M – 8 μ M ferrous iron in culture medium. The regulatory input to this system was later expanded upon by Matuszewska *et al.* who linked differing carbon sources to varying levels of expression of 7-hydroxytropolone biosynthetic genes using RT-PCR in *P. donghuensis* P482 (367). They observed that switching from glucose as a carbon source to glycerol lowered the expression of 7-hydroxytropolone related genes by almost 40 times, and this was evident as a decrease in the biological activity of the strain against the tested bacterial pathogens. On the whole, this class of molecules has demonstrated potential against a wide range of phytopathogens like *Dickeya solani* and

Pseudomonas syringae (367), *Macrophomina phaseolina* (340), *Verticillium dahliae* (366), and *Streptomyces scabies* (337).

The original discovery of 7-hydroxytropolone in *Pseudomonas* as a nonfluorescent siderophore also identified four genes responsible for its production in this genus (346), but this set was subsequently expanded upon to include some regulators mentioned above that flank these genes and form a two operon structure (347). This cluster was also independently identified through a transposon mutagenesis approach by Muzio *et al.* in 2020 and linked to 7-hydroxytropolone production in a search for loss of antifungal activity (340), and an additional gene encoding the outer member component of a tripartite multidrug resistance system was linked to 7-hydroxytropolone production / export. They were also able to predict three target sequences of the Gac-Rsm cascade both within and flanking the cluster. Despite the fairly detailed understanding we now have of the regulation of this cluster and 7-hydroxytropolone in *Pseudomonas*, the biosynthesis is not well understood at this stage. The gene cluster bears no similarity to that found in Gram-positive producers like *Streptomyces*, and many genes have only predicted functions, with little demonstration of pathway intermediates or a predicted biosynthetic scheme, with the exception of the recently characterised flavoprotein dioxygenase, TdaE, in *Burkholderia* discussed above (362). While the molecule itself is small and could easily be synthesised by processing a shunt metabolite of the paa pathway, the diversity of biosynthetic pathways leading to the tropolone motif in bacteria means we cannot take anything for granted, and a detailed investigation into the precise biochemical transformations is warranted. Finally, little is known about the activities and roles of additional congeners with additional modifications, like the extra hydroxylations observed in the *Streptomyces cyaneofuscatus* pathway (368), which should also be investigated.

3.2 Predicted Tropolone Biosynthesis in Ps652

3.2.1 Pathway Comparison by MultiGeneBlast

In order to investigate whether it is likely Ps652 produces 7-hydroxytropolone, it was first necessary to compare the gene cluster in Ps652 to the other reported clusters that encode

for the production of tropolones. Firstly, the Ps652 cluster was compared to the one reported in *P. donghuensis* SVBP6 (340). For this purpose, MultiGeneBlast was used (369). MultiGeneBlast is a bioinformatics tool that allows sequence comparisons using a gene cluster as input, rather than individual genes. Accordingly, it is possible to determine which genes are highly conserved, which are potentially dispensable for biosynthesis, and where the boundaries of the gene cluster lie in a putative BGC. This may give information on the evolutionary development of a gene cluster, depending on the search space. The Ps652 gene cluster was compared to the *P. donghuensis* SVBP6 gene cluster using default parameters, which revealed an extremely high level of similarity (Figure 3.6). Both clusters displayed the same genes in the same architecture, with remarkably high identity and coverage scores (Table 3.1). The lowest observed score was 92% amino acid identity, across 100% of the protein for the predicted short chain dehydrogenase.

Given these results, it was concluded that the Ps652 and *P. donghuensis* SVBP6 clusters likely produced the same product. The next step was to investigate the plausibility of the claims that this cluster encodes for the biosynthesis of 7-hydroxytropolone. Firstly, the Ps652 cluster was used as the query searching against the genome of *Streptomyces cyaneofuscatus* Soc7, which produces hydroxylated tropolones including 7-hydroxytropolone (368). Surprisingly, there was no similarity of any of the genes from the Ps652 cluster with any in the *S. cyaneofuscatus* Soc7 genome, even when lowering the lower identity threshold in MultiGeneBlast to 15% (data not shown). This led to the hypothesis that tropolone biosynthesis may have evolved separately in Gram-positive and Gram-negative bacteria. Multiple Proteobacteria are known to produce similar compounds, including the aforementioned *P. inhibens*, as well as the plant pathogen *Burkholderia plantarii* (370, 371, 363). The Ps652 cluster was then used as a search against the genomes of *P. donghuensis* SVBP6, *Phaeobacter inhibens* DSM 17395, and *Burkholderia plantarii* ZJ171 in MultiGeneBlast (Figure 3.6).

Table 3.1. Similarity between 7-Hydroxytropolone cluster from *P. donghuensis* SVBP6 and the cluster identified in Ps652.

<i>Pseudomonas donghuensis</i> SVBP6 Gene	Genbank Accession no.	<i>Pseudomonas</i> Ps652 Amino Acid Identity / Coverage
Enoyl-CoA hydratase	COO64_RS11460	99/100
Universal stress protein A	COO64_RS11465	96/100
LysR family transcriptional regulator	COO64_RS11470	100/100
Inner membrane component, tripartite multidrug resistance system	COO64_RS11475	99/100
Membrane fusion component, tripartite multidrug resistance system	COO64_RS11480	96/100
Outer membrane component, tripartite multidrug resistance system	COO64_RS11485	94/100
Enoyl-(acyl carrier protein) reductase	COO64_RS11490	98/100
Alpha-keto-acid reductase / decarboxylase	COO64_RS11495	98/100
Phenylacetate-CoA ligase	COO64_RS11500	98/96
Acyl-CoA dehydrogenase	COO64_RS11505	99/100
4-hydroxybenzyl-CoA thioesterase	COO64_RS11510	99/100
Short chain dehydrogenase	COO64_RS11515	92/100
HpcH, Hpal aldolase / citrate lyase family	COO64_RS11520	96/100
TetR family transcriptional regulator	COO64_RS11525	97/100
Oxidoreductase, short chain dehydrogenase / reductase family	COO64_RS11530	93/100

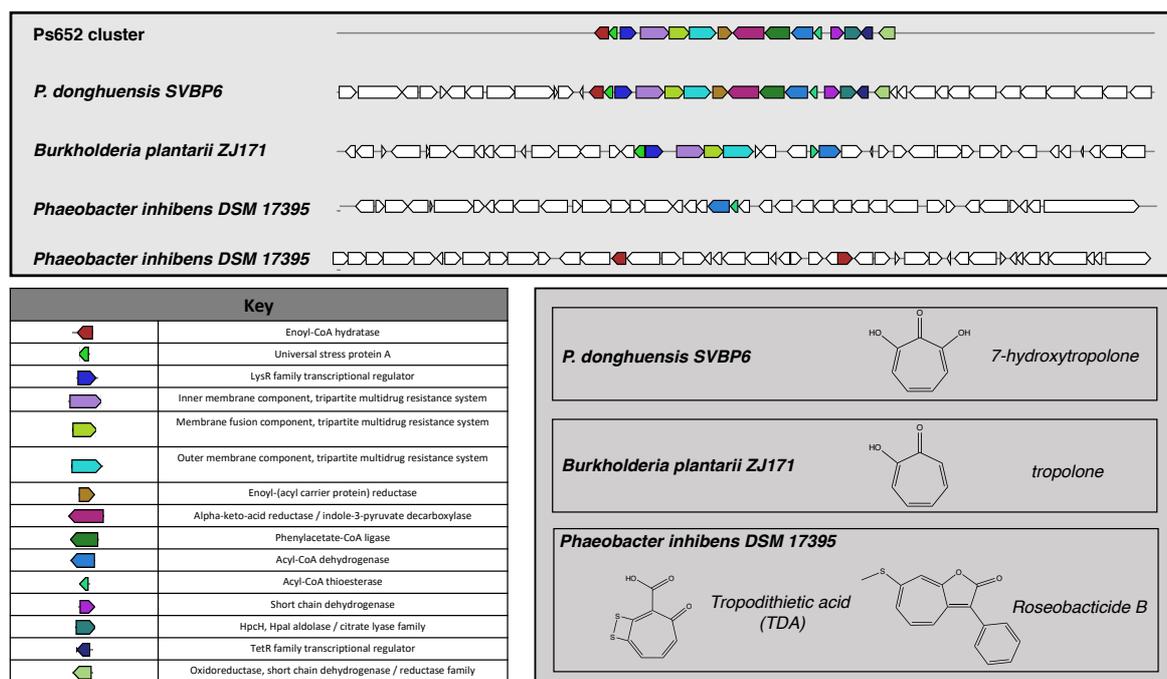


Figure 3.6. Comparison of Biosynthetic Gene Clusters in Tropolone-producing Bacteria. The gene cluster identified through transposon mutagenesis was used as a MultiGeneBlast search against the genomes of various tropolone-producing bacteria. No similarity was found with the best characterised cluster found in *Streptomyces cyaneofuscatus* Soc7. The *P. donghuensis* SVBP6 cluster is near-identical to the 652 gene cluster. *Burkholderia plantarii* is known to produce tropolone as a virulence factor, and shares a number of genes with the recently identified 7-HT cluster, including the characterised flavoprotein dioxygenase (TdaE). *Phaeobacter inhibens* produces multiple tropolonoid natural products, and shares the acyl-CoA thioesterase and a TdaE homolog, as well as having two highly similar enoyl-CoA hydratases (EchA) at a different locus. The natural products produced by each species are also shown (bottom right).

The output revealed some surprises. The only features shared between all tropolone producing Proteobacteria were the acyl-CoA thioesterase and acyl-CoA dehydrogenase / flavoprotein dioxygenase (TdaE) that were the sites of transposon insertion in mutants 14G7, 24D9, and 29D2 (see section 2.5.5) (337). These are TdaD and TdaE, respectively, in both *P. inhibens* and *B. plantarii*. Additionally, the genomes of both *P. donghuensis* SVBP6 and *P. inhibens* DSM 17395 encode homologs of the enoyl-CoA hydratase found in the Ps652 cluster. Curiously, it was noted that despite the lack of sequence similarity, the pathway in *Streptomyces* also features an enoyl-CoA hydratase protein, TrlA, which initiates tropolone biosynthesis by branching from primary metabolism.

3.2.2 Proposed Biosynthetic Pathway for 7-Hydroxytropolone in *Pseudomonas*

Based on these results, it seems plausible that the gene cluster identified in Ps652 and *P. donghuensis* SVBP6 may encode the production of 7-hydroxytropolone. In addition, the genome of Ps652 also contains a complete set of genes (*paaKABCDEFGZ*) encoding the phenylacetic acid catabolic pathway, which is required in *S. cyaneofuscatus* Soc7 as well as other Proteobacteria. Further lending weight to the hypothesis that tropolone biosynthesis might occur similarly in Proteobacteria as it does in streptomycetes, the Ps652 cluster contains a duplicated *paaK*, encoding a phenylacetate-CoA ligase, potentially involved in precursor supply. A similar theme is observed in the Soc7 pathway, where TrlH, TrlB, and potentially TrlG are involved in directing primary metabolism towards precursors for the phenylacetate catabolon and ultimately tropolone biosynthesis (Figure 3.5). Combining all this information, a preliminary proposed biosynthetic pathway for how tropolone biosynthesis might occur in Ps652 is presented in Figure 3.7. Similar to the already described pathways, it is proposed here that primary metabolism supplies the precursors for the pathway, leading to phenylacetic acid. This is then coupled to CoA by the phenyl-CoA ligase (PaaK) encoded in the cluster, allowing the normal phenylacetate catabolon to proceed until generation of 2-oxepin-2(3H)-ylideneacetyl-CoA. At this stage, the enoyl-CoA hydratase, which is highly homologous to the C terminus of the PaaZ protein in the endogenous *paa* cluster, then catalyses ring opening to 3-oxo-5,6-dehydrosuberil-CoA semialdehyde. A spontaneous Knoevenagel condensation is then proposed to occur as is described in *S. cyaneofuscatus* Soc7 to form the 7-membered tropolonoid ring. Two

possible routes then plausibly exist to produce tropolone natural products. The acyl-CoA thioesterase and decarboxylase (IpdC) may then effect cleavage of CoA followed by conversion to tropolone, possibly with a dehydrogenation by the annotated dehydrogenase with 59% identity to the characterised TdaE in *Burkholderia*. Alternatively, it is possible that the TdaE homologue itself performs a number of these steps, including thioester cleavage, as it does in *Burkholderia*, followed by spontaneous loss of CO₂ to form tropolone. Not all genes in the cluster are accounted for in this scheme, with particular emphasis on the lack of monooxygenases that might catalyse hydroxylation of tropolone to 7-hydroxytropolone. While TrLCD produce 7-hydroxytropolone and 3,7-dihydroxytropolone from tropolone in the *Streptomyces* pathway, some hydroxylated tropolones were also observed for which it was not possible to identify the responsible monooxygenase (368). It is possible a similar phenomenon occurs in Ps652, with additional enzymes encoded elsewhere in the genome that are capable of modifying the basic tropolone skeleton.

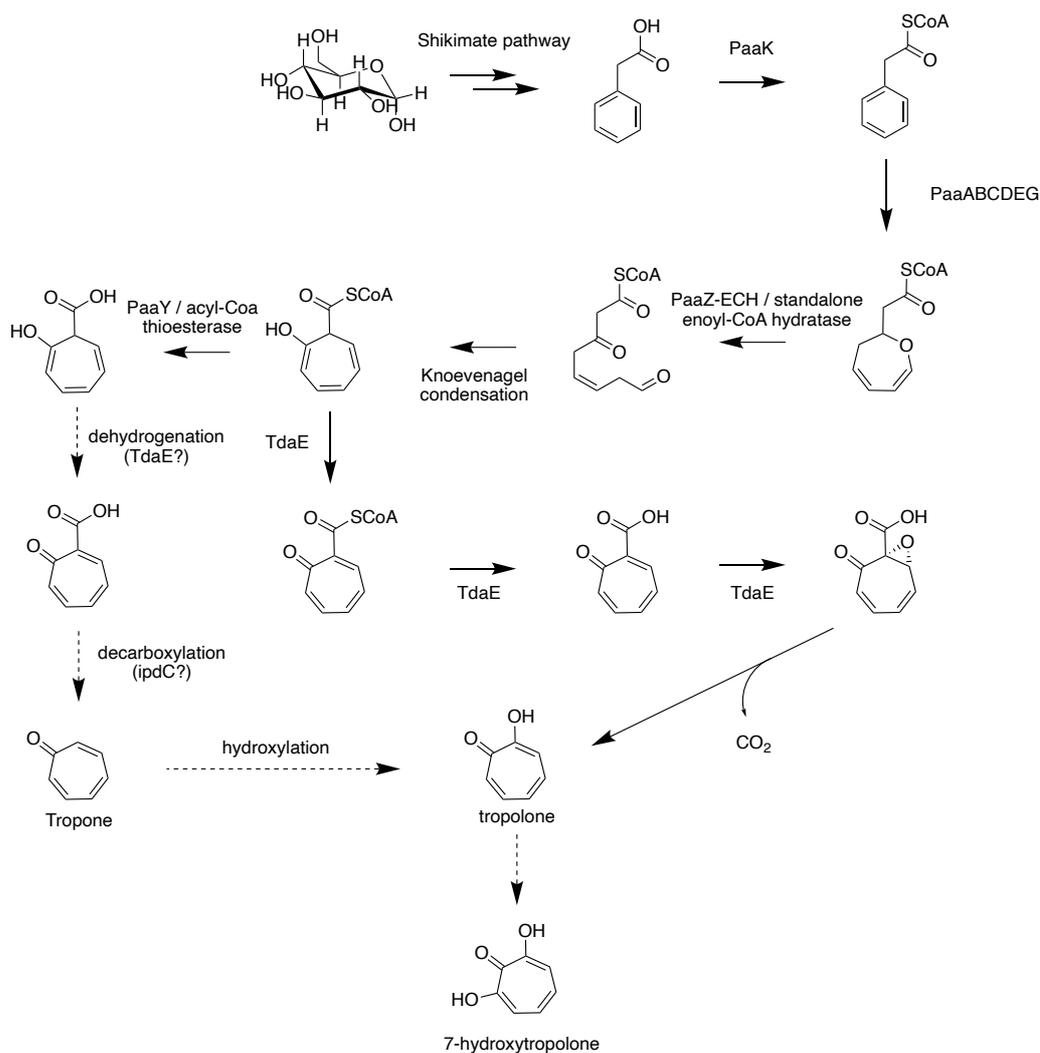


Figure 3.7. Proposed biosynthetic pathway for tropolone production in *Ps652*. From the documented bacterial tropolone BGCs in bacteria, and the bioinformatic analysis of the *Ps652* gene cluster, presented is a prediction of the possible routes to tropolone natural products in *Ps652*. Primary metabolism supplies the precursors, aided by the additional copy of PaaK. Phenylacetate catabolism proceeds until the enoyl-CoA hydratase branch point, where the standalone enoyl-CoA hydratase may increase flux towards tropolones. Two possible routes to tropolones are then hypothesised, one following the documented activity of TdaE, a homologue of which is present in the *Ps652* BGC (*AcdA*), with spontaneous loss of CO₂ from the epoxide to form tropolone. The other route may involve PaaY or the additional TdaD homologue acyl-CoA thioesterase (TE), followed by dehydrogenation (possibly by *AcdA*), decarboxylation (possibly by the encoded putative alpha-keto-acid reductase / decarboxylase), followed by hydroxylations to tropolone and then 7-hydroxytropolone by unknown enzymes.

3.3 Isolation of a Dihydroxylated Tropolone from Ps652

3.3.1 Detection of Tropolones in Ps652 Culture Extracts

As a result of the previous genetic work, mass spectrometry analysis was performed on Ps652 and selected knockout strains in order to investigate whether this strain produces tropolones. Using a low mass scan range, m/z 50-500, it was possible to detect the presence of 7-hydroxytropolone in culture extracts from modified King's Broth (MKB) and MKB Glucose media, but not in tropolone cluster mutants, by comparing to a synthetic standard of 7-hydroxytropolone provided by the Murelli group (City University of New York) (Figure 3.8A). Additionally, using a higher mass scan range, m/z 200-2000, ions were identified that were predicted to be tropolone-iron chelates (Figure 3.8B). These form a 2:1 ratio of hydroxytropolone molecules to iron, as demonstrated by Jiang *et al.* (345). Additionally, ions were observed that corresponded to the masses of iron-tropolone complexes containing dihydroxytropolones, suggesting Ps652 also produced tropolones with additional hydroxylations (Figure 3.8C). While the identity of the dihydroxylated tropolone was initially a mystery, it was later identified as 3,7-dihydroxytropolone (see section 3.3.5), and so the complexes represent either 7-hydroxytropolone—Fe—3,7-dihydroxytropolone or 3,7-dihydroxytropolone—Fe—3,7-dihydroxytropolone. These tropolone-iron complexes were also absent from tropolone cluster mutants (not shown).

The 3,7-dihydroxytropolone monomer was subsequently observed in mass spectral data, with a shift in retention time consistent with additional hydroxylations conferring a more polar character (Figure 3.9), and targeted for purification following the method outlined in ref (345) with some modifications.

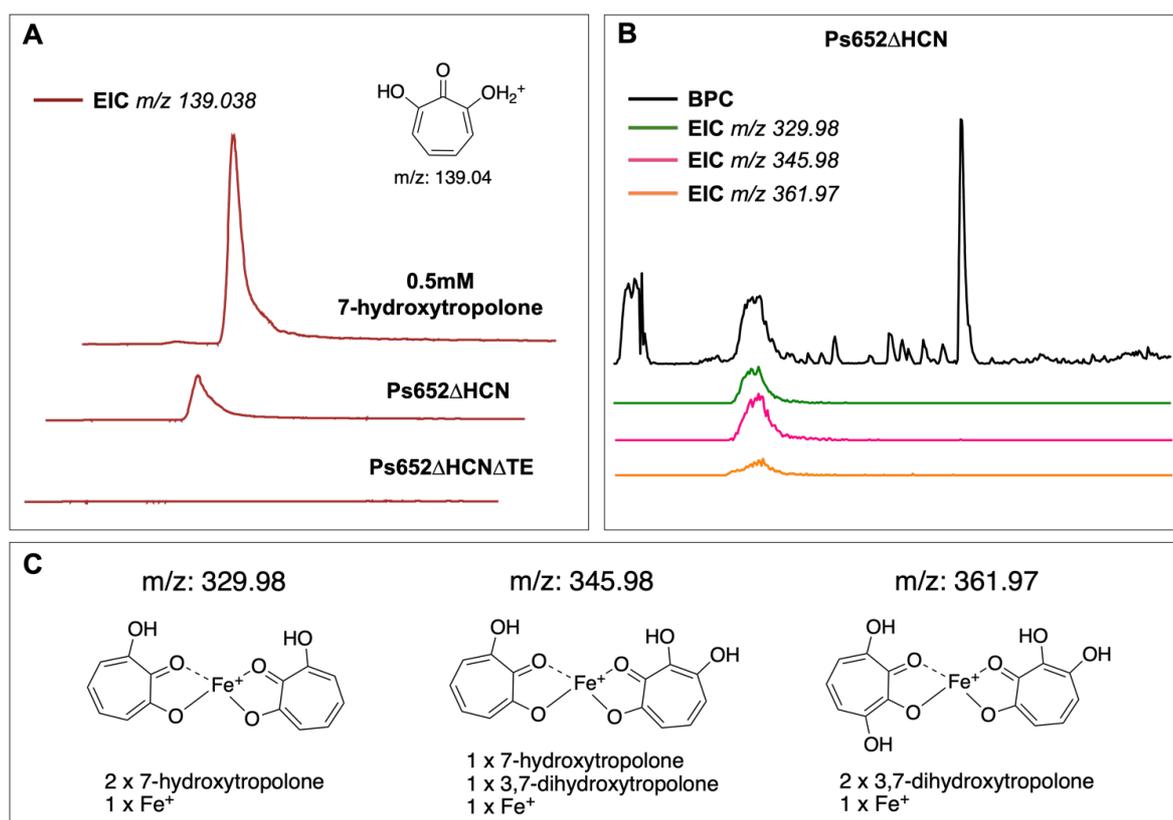


Figure 3.8. Detection of 7-hydroxytropolone and related iron chelates in mass spectral data from Ps652 fermentations. (A) Detection of 7-hydroxytropolone in Ps652 Δ HCN cultures in modified King's Broth. Comparison to a synthetic standard of 7-hydroxytropolone confirmed the presence of this metabolite in Ps652 Δ HCN cultures but was not observed in a tropolone cluster mutant. (B) Evidence for tropolone-iron chelates in mass spectral data. Three ions were observed in mass spectral data for Ps652 Δ HCN but not tropolone cluster mutants that corresponded to the predicted masses of iron chelates of 7-hydroxytropolone and a dihydroxylated tropolone. (C) Observed tropolone-iron chelates and their predicted m/z values in positive mode are shown. Chelates were observed that indicated the presence of a dihydroxylated tropolone congener. EIC, extracted ion chromatogram; BPC, base peak chromatogram.

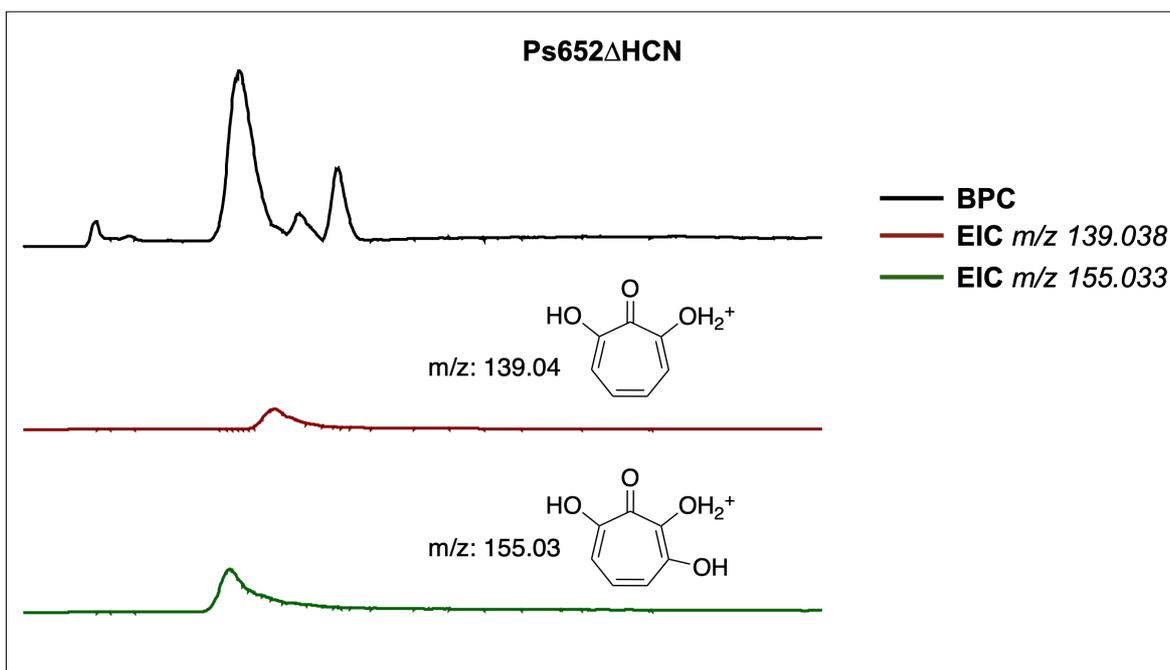


Figure 3.9. Detection of a dihydroxylated tropolone in mass spectral data from Ps652ΔHCN fermentations. Ions with m/z corresponding to the predicted masses of both 7-hydroxytropolone and what was predicted to be 3,7-hydroxytropolone were observed. Consistent with the addition of an extra hydroxyl group, the predicted dihydroxylated tropolone showed increased polarity and less interaction with a reverse phase HPLC stationary phase. BPC, base peak chromatogram; EIC, extracted ion chromatogram.

3.3.2 Initial Extraction

Ps652 Δ HCN Δ Pyo was chosen for the production of hydroxytropolones, as the pyoverdine from this strain had already been purified (see section 2.4). By using a pyoverdine deficient strain, the intention was to simplify the purification process, and limit siderophore production to tropolone molecules. Based on initial assays with extracts from both MKB and a version with glucose as a carbon source rather than glycerol, MKB Glucose was chosen for production cultures. It was observed that the biological activity was not ethyl acetate extractable (Figure 3.10A), despite previous reports of extraction of 7-hydroxytropolone with this solvent (345, 340). Switching carbon source from glycerol to glucose solved this issue, with tropolone cluster-related activity being extractable from MKB Glucose cultures with organic solvents (Figure 3.10B). Attempts to scale up cultures from 10 mL to 500 mL were unsuccessful and resulted in a loss of the above biological activity. Ps652 Δ HCN Δ Pyo was therefore grown in 200 x 10 mL aliquots in 50 mL falcon tubes, a combined volume of 2 L. All cultures were combined before removal of cells by centrifugation and extraction of supernatant with ethyl acetate as per the methods. The combined organic fractions were dried by rotary evaporation and redissolved in methanol.

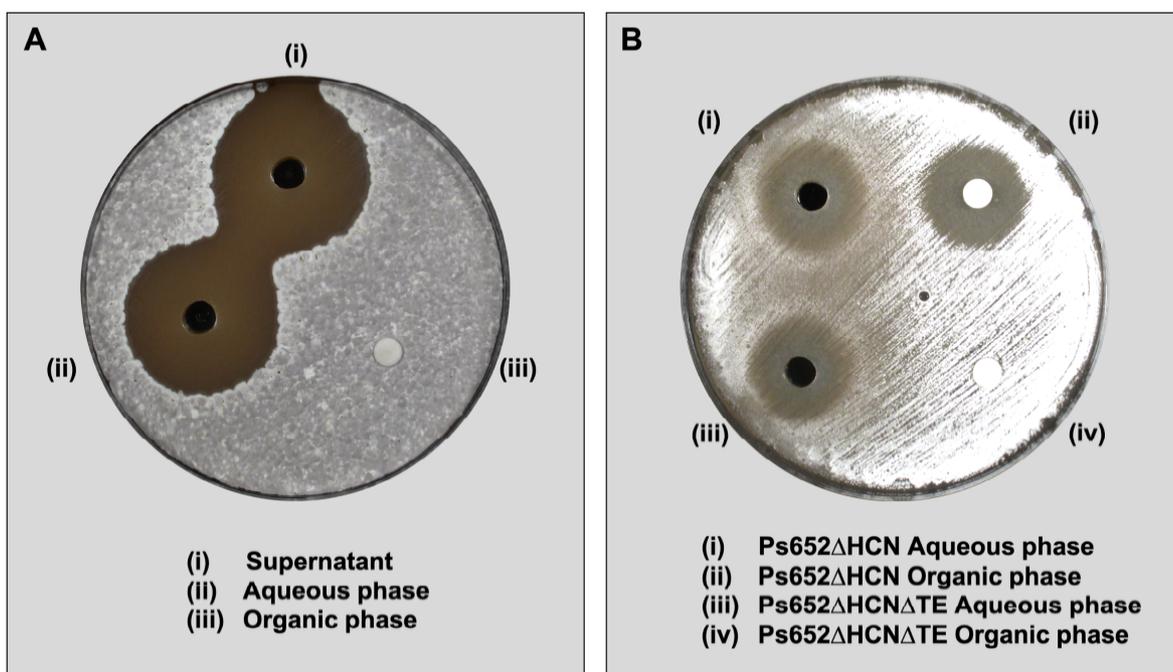


Figure 3.10. Extraction of bioactive molecules from fermentations of Ps652ΔHCN. (A) Extracts from Ps652ΔHCN grown in modified King's broth with glycerol as a carbon source. Biological activity in this media was not extractable with organic solvents. (B) Extracts from Ps652ΔHCN and Ps652ΔHCNΔTE grown in modified King's broth with glucose as a carbon source. Molecules predicted to be tropolone natural products are extractable with organic solvents, and not present in the tropolone cluster mutant. Presence of a zone of inhibition against *S. scabies* 87-22 for residual aqueous phase indicates the presence of a second biologically active molecule.

3.3.3 Size-exclusion Chromatography

Size-exclusion chromatography using Sephadex LH-20 has been shown to be effective for purification of tropolones, as the extraction of supernatant with ethyl acetate and then size-separation of molecules constitute orthogonal separation methods. Small polar molecules like sugars and small hydrophilic peptides are not extracted by the organic solvent, which then allows highly efficient separation of small hydrophobic molecules from larger hydrophobic molecules in the size-exclusion step. The combined organic extracts from section 3.3.2 were processed in two injections as per the methods section, with monitoring of tropolones by UV absorbance at 330 nm, and refractive index used to monitor non-absorbing contaminants. Fractions showing absorbance at 330 nm were collected and analysed by LC-MS using a low mass scan range (m/z 50-500), as any molecules larger than this range would be found in earlier fractions. Fractions were observed that contained mostly 7-hydroxytropolone, mostly 3,7-dihydroxytropolone, or a mixture of the two (Figure 3.11).

3.3.4 Semi-preparative HPLC

Fractions were further purified by a step of semi-preparative scale HPLC. Fractions containing 3,7-dihydroxytropolone were pooled, resulting in 5.72 mg of material. This was redissolved in 1 mL methanol and processed by HPLC in five injections. A polar C18 stationary phase with water/methanol with added 0.5% formic acid mobile phase allowed for efficient separation of 7-hydroxytropolone and 3,7-dihydroxytropolone (Figure 3.12). The second peak showed an identical mass and retention time to a 7-hydroxytropolone standard and so was identified as such. The first peak corresponded to the more polar dihydroxytropolone, and this peak from all five injections were pooled and dried on a Genevac EZ-2 Plus system using the HPLC lyophilisation setting. The resulting material was 2.83 mg of a white powder with a slight pink hue.

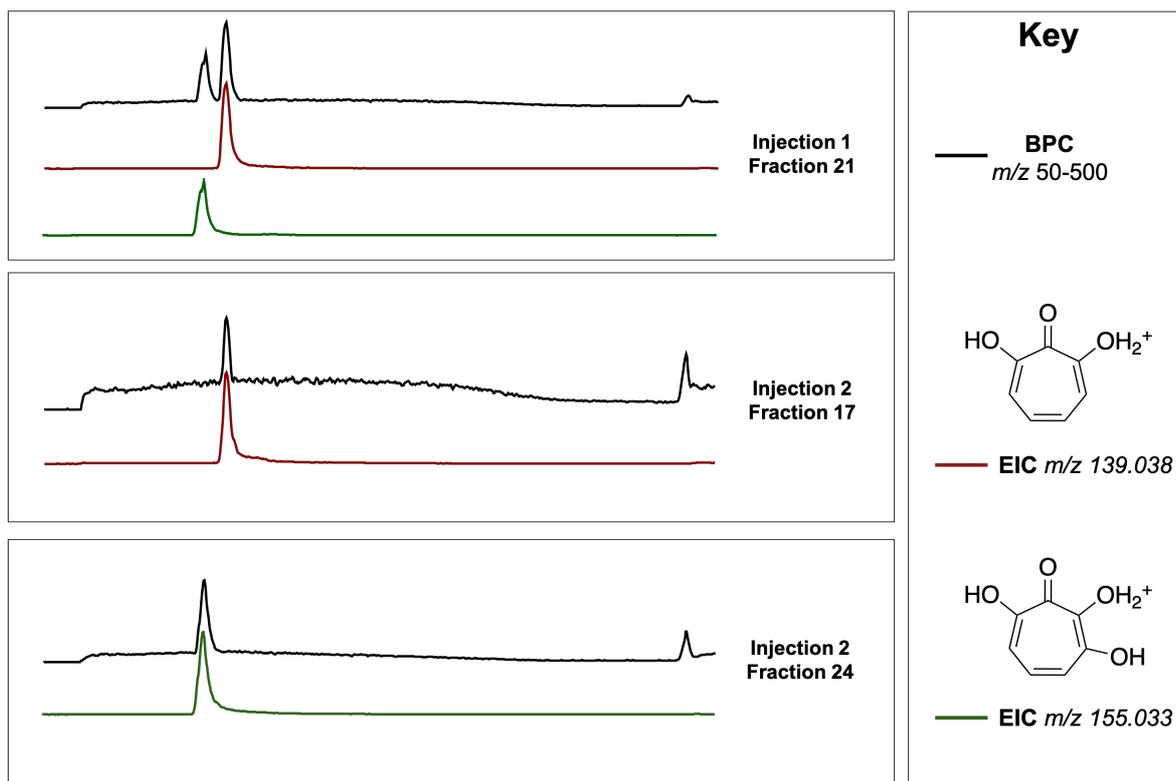


Figure 3.11. LC-MS chromatograms for Sephadex LH-20 fractions in tropolone purification. Fractions were observed that contained a mixture of 7-hydroxytropolone and 3,7-dihydroxytropolone (top panel), mainly 7-hydroxytropolone (middle panel), or mainly 3,7-dihydroxytropolone (bottom panel). BPC, base peak chromatogram; EIC, extracted ion chromatogram.

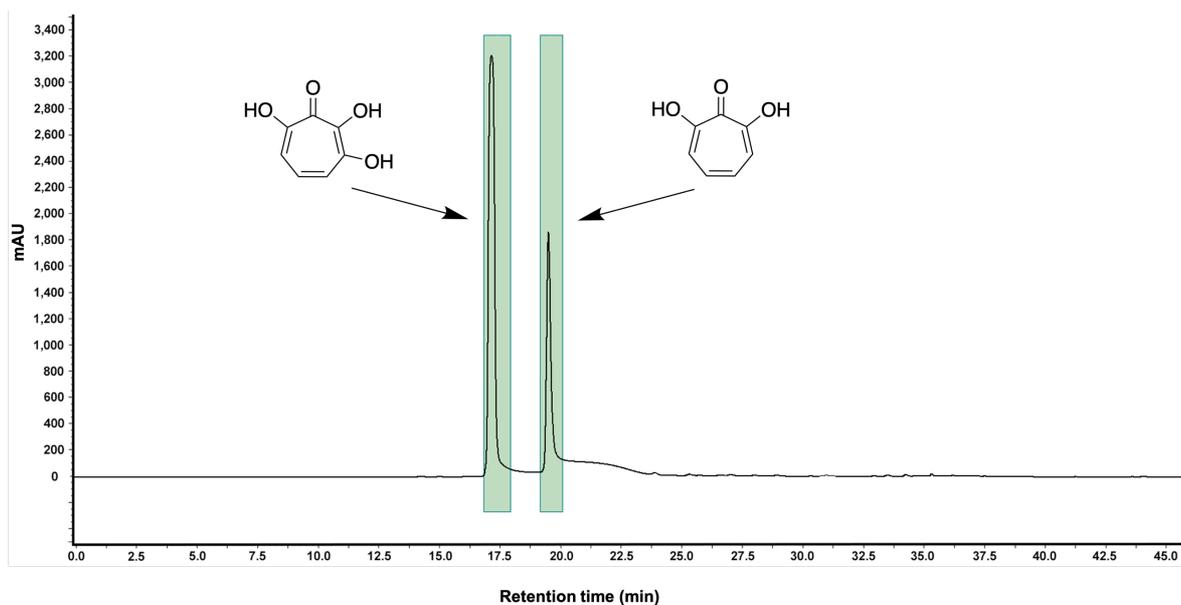


Figure 3.12. Separation of 7-hydroxytropolone and 3,7-dihydroxytropolone by reversed phase HPLC using a polar C18 stationary phase. Trace indicates UV absorbance at 327nm. Consistent with predictions and previous data, the dihydroxylated tropolone eluted earlier due to less interaction with the mainly non-polar stationary phase. The retention time of 7-hydroxytropolone matched that of a previously injected synthetic standard. Shaded green boxes represent the collected fractions.

3.3.5 Characterisation of 3,7-dihydroxytropolone by ^1H and ^{13}C NMR

Ahead of NMR analysis, the dihydroxylated tropolone was redissolved in 500 μL deuterated methanol (CD_3OD), dried on a Genevac EZ-2 Plus using the low BP setting, and redissolved again in 500 μL CD_3OD to ensure full exchange of hydroxyl protons with deuterium. 1D and 2D spectra were obtained on a 600 MHz instrument, but only the 1D data was necessary to confirm the structure as 3,7-hydroxytropolone by matching to previously published spectra (372). Two distinct signals were observed in ^1H NMR, at 4.78 and 6.98 ppm. The 4.78 ppm signal was assigned as an MeOH contaminant peak by comparison to reference tables, and the multiplet at 6.98 ppm matched the signal observed by Takeshita *et al.*, which they assigned as the three non-hydroxyl protons (Figure 3.13) (372). They reported a singlet at 7.04 ppm, whereas a multiplet was observed in the data presented here, a difference attributable to the spectrum in ref (372) being acquired on a 100 MHz instrument, compared to the 600 MHz instrument used here. Four distinct signals were observed in ^{13}C NMR at 119.24, 129.44, 158.11, and 158.94 ppm (Figure 3.14) while the published spectrum from Takeshita *et al.* reported values of 119.3, 129.4, 158.4, and 159.1 ppm. Two smaller peaks with different chemical shift values were also observed in ^{13}C NMR that showed coupling in HSQC data with the three non-hydroxyl protons; these were attributed to a difference resonance form of 3,7-hydroxytropolone (Figure 3.15). Crystal data from ref (368) indicates 3,7-dihydroxytropolone is the predominant form, whereas the 3,4-dihydroxytropolone form is less prevalent, hence the less intense peaks for this resonance form in ^{13}C data.

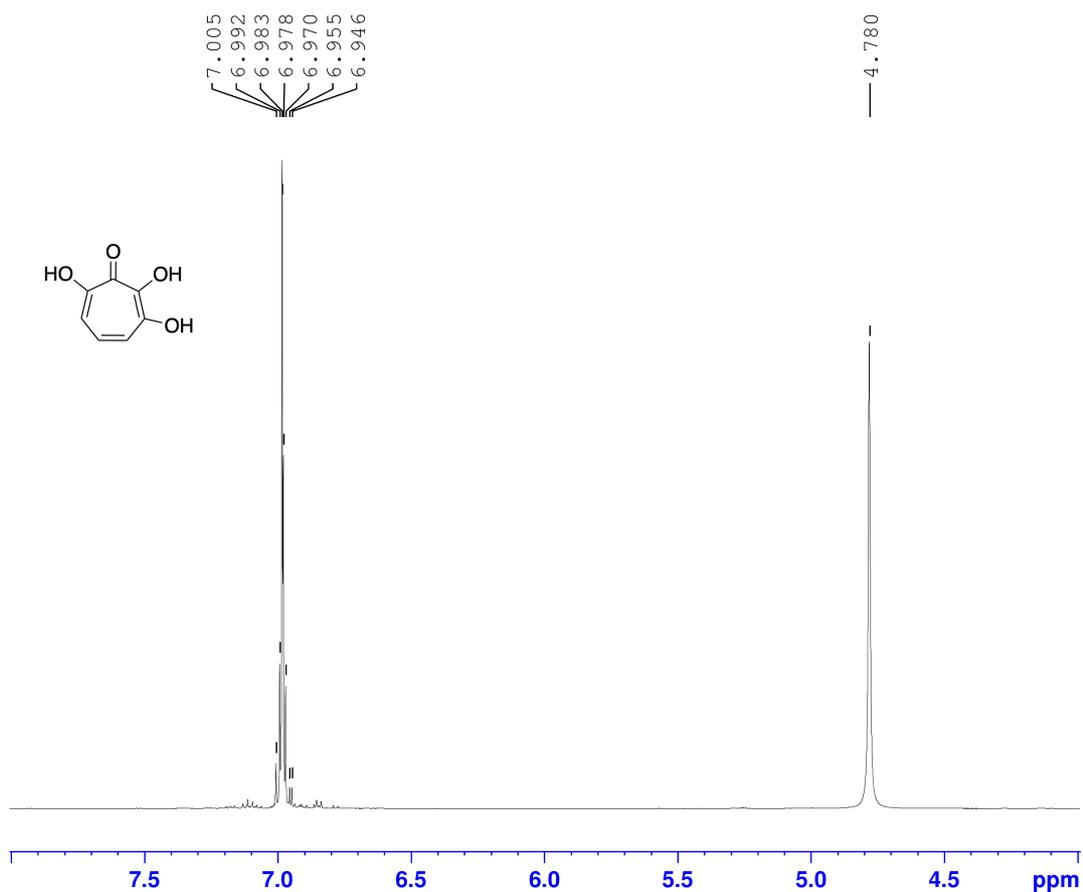


Figure 3.13. ¹H NMR in CD₃OD for the purified 3,7-dihydroxytropolone. The multiplet at 6.978ppm was identified as the 3 CH₂ protons, matched to a previously published spectrum for this compound, where a singlet was observed due to the lower resolution instrument used for that study. The broad singlet peak at 4.78ppm was identified as a methanol contaminant in the CD₃OD solvent.

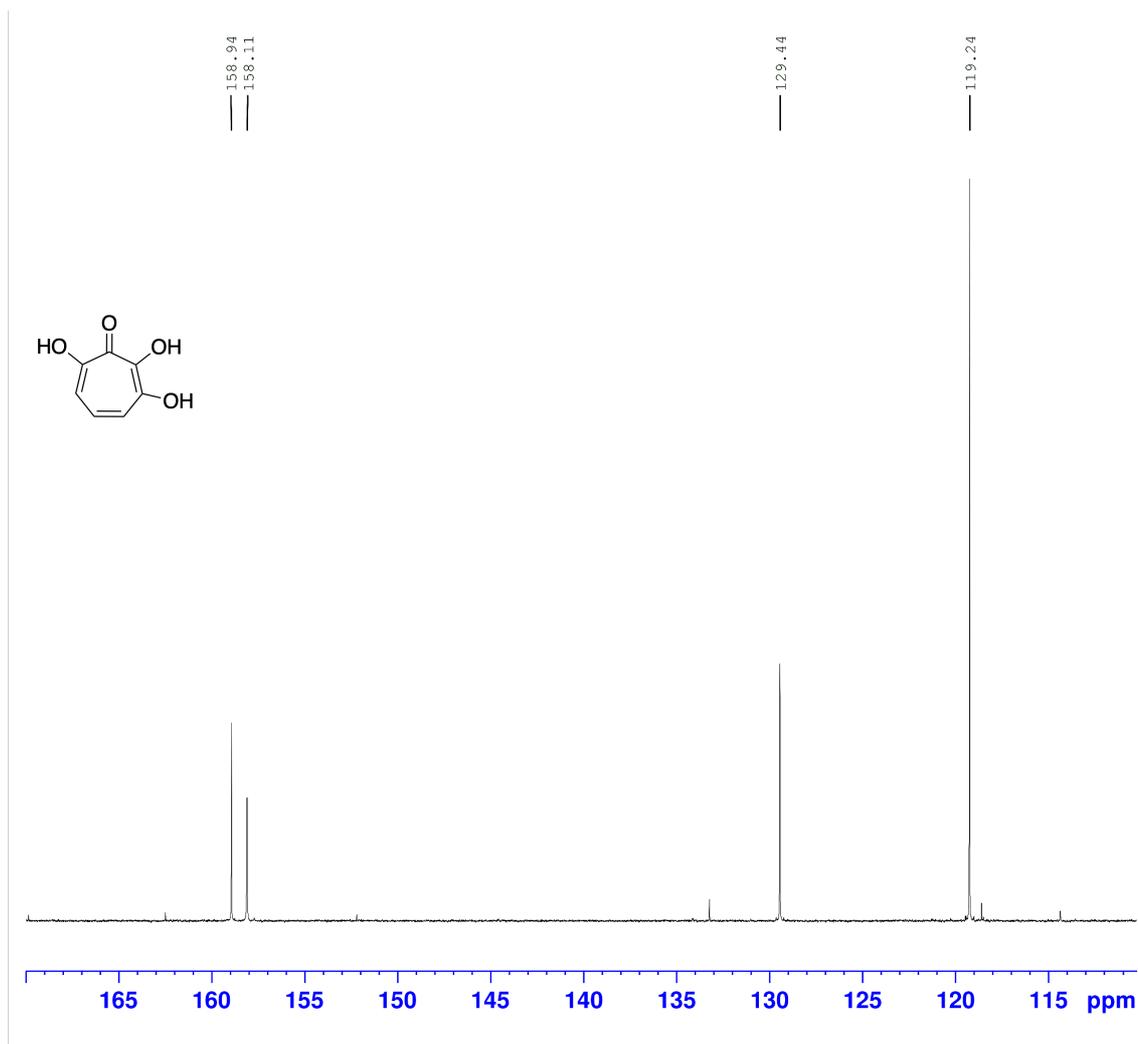
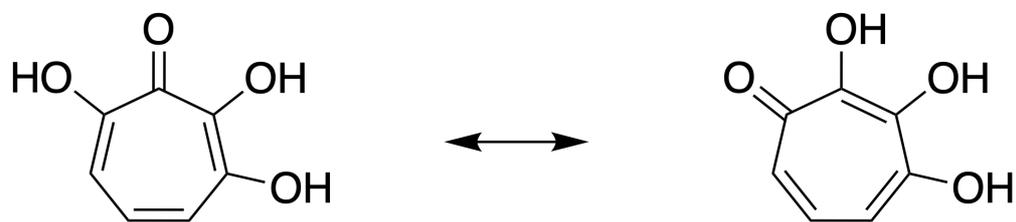


Figure 3.14. ^{13}C NMR in CD_3OD for the purified 3,7-dihydroxytropolone. Four high intensity peaks were observed at 119.24ppm, 129.44ppm, 158.11ppm, and 158.94ppm, matching a previously published spectrum for this molecule. Lower intensity peaks are also observed which may correspond to a different resonance form of the molecule, 3,4-dihydroxytropolone, as they also showed coupling to CH_2 protons in 2D data.

Resonance forms of 3,7-dihydroxytropolone



3,7-dihydroxytropolone

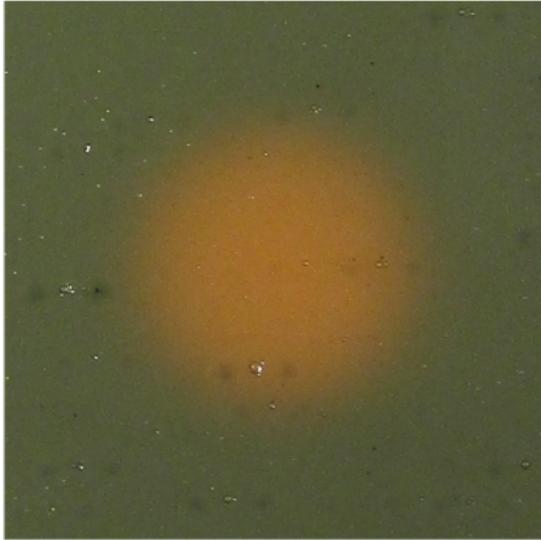
3,4-dihydroxytropolone

Figure 3.15. Resonance forms of the purified dihydroxytropolone. As shown, the purified dihydroxylated tropolone is predicted to exist in two resonance forms, 3,7-dihydroxytropolone and 3,4-dihydroxytropolone. Crystal data from Chen *et al.* (368) indicates 3,7-dihydroxytropolone is the dominant form, and therefore this name is used for the molecule.

3.3.6 Iron Binding of Hydroxytropolones

Given that 7-hydroxytropolone is known to act as a siderophore in *P. donghuensis* (345), as well as the observation of tropolone-iron chelates in mass spectral data, an obvious next step was to assess the ability of the purified dihydroxylated tropolone to bind iron. A qualitative rather than quantitative approach was chosen due to the limited supply of 3,7-dihydroxytropolone, using the CAS agar plate system for siderophore assessment (373). This medium contains a dye complex of chrome azurol S and hexadecyltrimethylammonium bromide (HDTMA), which bind ferric iron to produce a blue colour. Siderophores are able to liberate the iron from the CAS/HDTMA complex, resulting in a colour change from blue to orange that is visible to the naked eye. The ability of 7-hydroxytropolone to act as a siderophore and produce this colour change has previously been demonstrated in producing strains of *P. donghuensis*, but not with the compound directly (346). It was anticipated that addition of hydroxytropolones to CAS agar plates would result in a blue to orange colour change. 5 µg of pure 7-hydroxytropolone and 5 µg of pure 3,7-dihydroxytropolone were added to a CAS agar plate in 5 µL methanol each; 5 µL methanol was used as a control. Both 7-hydroxytropolone and 3,7-dihydroxytropolone were able to bind iron and produce a blue to orange colour change, indicating that 3,7-dihydroxytropolone is likely also able to act as a siderophore (Figure 3.16). This data was in concordance with the observed metal chelates in the mass spectral data (Figure 3.8), where both 7-hydroxytropolone—Fe—3,7-dihydroxytropolone and 3,7-dihydroxytropolone—Fe—3,7-dihydroxytropolone chelates were observed.

7-hydroxytropolone



3,7-dihydroxytropolone



Figure 3.16. Binding of 7-hydroxytropolone and 3,7-dihydroxytropolone to iron on CAS agar plates. 5 μg each of 7-hydroxytropolone and 3,7-dihydroxytropolone were added to the surface of a CAS agar plate in 5 μL methanol. Methanol alone shows no iron binding as anticipated and is not shown here. Both molecules displayed the ability to bind iron as indicated by the blue to orange colour change.

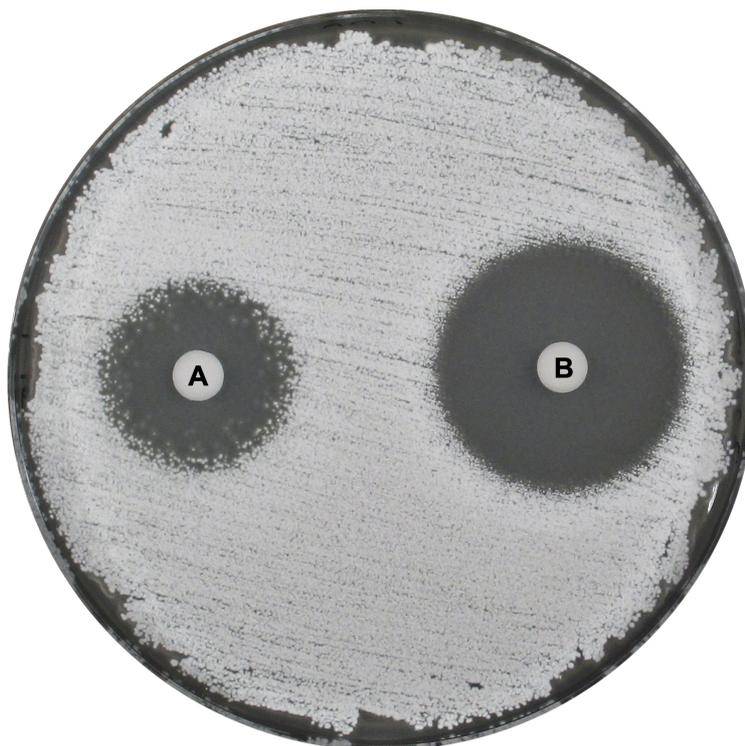
3.3.7 Biological Activity against *Streptomyces scabies* 87-22

The inability of tropolone cluster mutants (see section 2.5.3) to inhibit the growth of *S. scabies* 87-22 suggested that 7-hydroxytropolone was responsible for the biological activity of this strain. However, 3,7-dihydroxytropolone has not been previously reported from a *Pseudomonas* isolate, and it is not known whether it contributes towards the inhibition of phytopathogens. Disk diffusion assays were employed to this end, given that the molecules were soluble in methanol and unlikely to remain bound to the disk, which has been known for polar molecules (T. Scott, personal communication). Both compounds were dissolved to a concentration of 1 mg/mL in methanol, and 10 µg of 7-hydroxytropolone and 10 µg of 3,7-dihydroxytropolone were added to paper disks and allowed to dry before placing on the surface of a lawn of *S. scabies* spores on IMA. After three days at 30 °C, clear zones of inhibition were visible around each compound, indicating inhibition of *S. scabies* growth (Figure 3.17). These results confirm that the purified 3,7-dihydroxytropolone contributes towards the biological activity of Ps652 against *Streptomyces scabies*, and is the first such report. Interestingly, the activity of 3,7-dihydroxytropolone towards *S. scabies* appears significantly greater than 7-hydroxytropolone, by the larger zone of inhibition for this compound.

3.3.8 Biological Activity against *Phytophthora infestans*

7-hydroxytropolone and 3,7-dihydroxytropolone were tested against *Phytophthora infestans* 88069 as for *Streptomyces scabies*. 10 µg of each were added to paper disks in 10 µL methanol each and allowed to dry before placing on the surface of a PDA plate with a plug taken from the growing edge of a *P. infestans* mycelium. After 10 days at ~18 °C, no visible zones of inhibition were evident for either compound, with the *P. infestans* mycelium growing onto the paper disks containing the compound (Figure 3.18). These results suggested, rather unexpectedly, that other natural products were responsible for the anti-*Phytophthora* effects of Ps652.

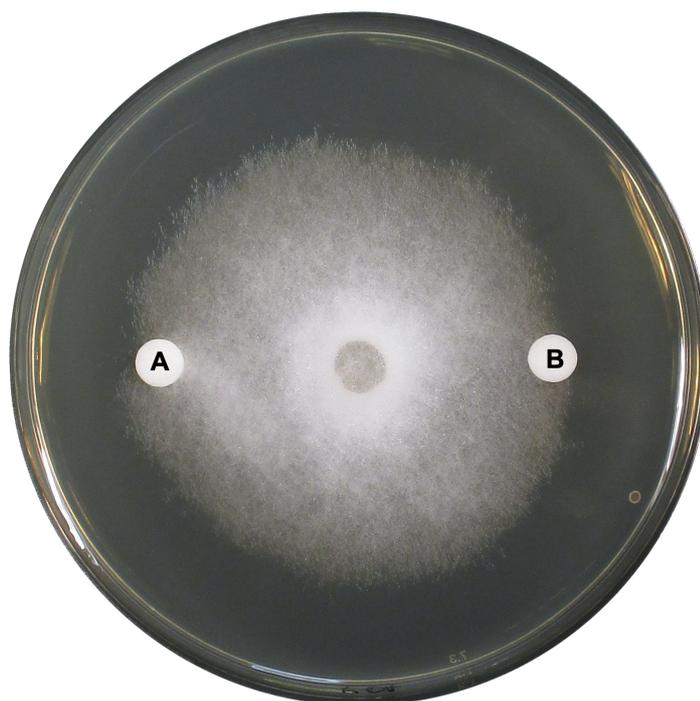
Streptomyces scabies 87-22



(A) 7-hydroxytropolone (B) 3,7-dihydroxytropolone

Figure 3.17. Biological activity of 7-hydroxytropolone and 3,7-dihydroxytropolone against *Streptomyces scabies* 87-22 *in vitro*. 10 μg of each compound was added to a 3 mm Whatman paper disk in 10 μL methanol and allowed to dry before being placed onto the agar surface with a lawn of *Streptomyces* spores as per the methods. Both compounds show the ability to inhibit the phytopathogen, but 3,7-dihydroxytropolone shows increased activity compared to the monohydroxylated form. Plates were incubated for three days at 30 $^{\circ}\text{C}$ before imaging.

Phytophthora infestans 88069



(A) 7-hydroxytropolone (B) 3,7-dihydroxytropolone

Figure 3.18. Biological activity of 7-hydroxytropolone and 3,7-dihydroxytropolone against *Phytophthora infestans* 88069 *in vitro*. 10 μg of each compound was added to a 3 mm Whatman paper disk in 10 μL methanol and allowed to dry before being placed onto the agar surface. Neither compound was able to inhibit the growth of the phytopathogen after 10 days. Plates were incubated 10 days at 18 $^{\circ}\text{C}$ before imaging.

3.4 Effect of Gene Deletions in the 3,7-dihydroxytropolone Biosynthetic Gene Cluster

By comparing the Ps652 gene cluster with the *P. donghuensis* SVBP6 cluster, the conservation of genes appeared significant. However, it was not evident which of these were essential for the observed biocontrol phenotype. Various genes in the pathway have now been investigated by others, including two regulatory elements that are LysR and TetR type regulators (347) and the flavoprotein dioxygenase (TdaE) that was recently characterised in *Burkholderia* (362), a homolog of which exists in this BGC with 59% identity. To further evaluate the genes important for tropolone biosynthesis and the associated phenotype in Ps652, in-frame deletion strains were made for numerous genes in the pathway as illustrated in (Figure 3.19). In addition, BLAST analysis identified an additional copy of the enoyl-CoA hydratase (*echA8*) elsewhere in the Ps652 genome, which was also deleted. Various assays were then used to investigate whether each had a role in the production of tropolone natural products, iron binding activity, and inhibition of *S. scabies*.

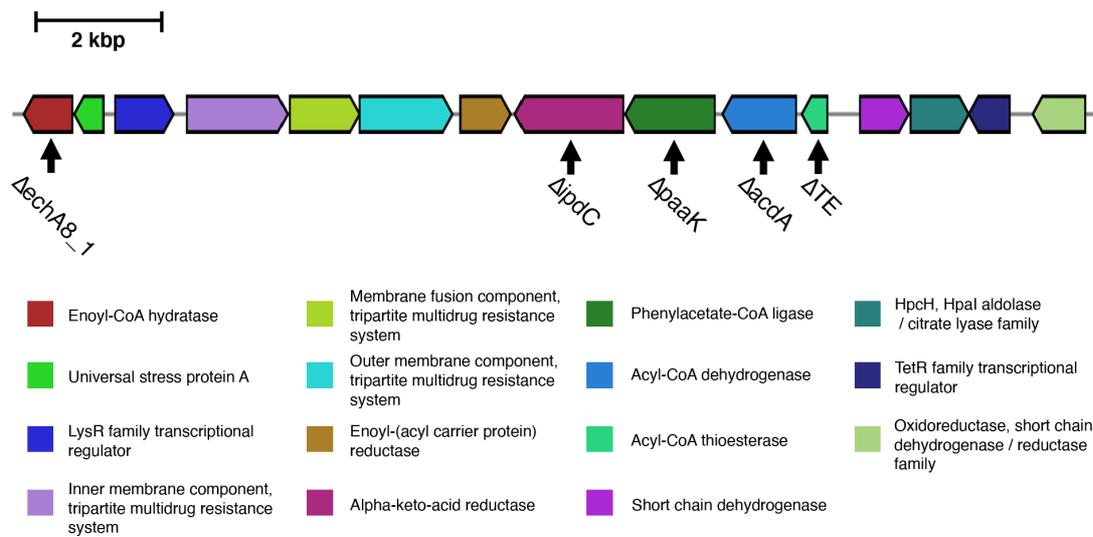


Figure 3.19. In-frame gene deletions made in the Ps652 tropolone biosynthetic gene cluster. Arrows indicate the genes deleted, and the name of the resulting mutant.

3.4.1.1 Cross Streak Assays

Given that the original discovery of Ps652 as a potential biocontrol strain was uncovered through the cross-streak assay against *S. scabies*, it seemed fitting to revisit this assay for tropolone cluster mutants, particularly considering tropolone molecules were biologically active against this pathogen. This data could also be combined with downstream analyses such as probing whether mutants made tropolones or only a limited set of pathway congeners. By establishing the biological activity of the mutants at this stage, it would be possible to link molecules to phenotypes later. As already described in section 2.6, the predicted acyl-CoA thioesterase in the cluster had already been deleted and verified to affect the phenotype in cross-streak assays, which could be complemented *in trans* (see section 2.6). Cross-streak assays were performed for all mutants in triplicate against *S. scabies* 87-22, and representative plates are shown in Figure 3.20. Contrary to initial expectations, deletion of the gene encoding the phenylacetate-CoA ligase (*paak*), which is anticipated to be involved in precursor supply, had no effect on inhibition of *S. scabies*; this is likely due to there being an additional copy in the *paa* cluster, as it relates to primary metabolism. Likewise, despite the conserved nature of the predicted alpha-keto-acid reductase / decarboxylase (*lpdC*), its deletion had no measurable impact on the phenotype. While no biosynthetic role has yet been proposed for this gene, it may also relate to precursor supply. Deletion of the enoyl-CoA hydratase *EchA8_1*, which was anticipated to be involved in a key step in the pathway, where primary metabolism branches away from phenylacetate degradation to lead to production of tropolone natural products (Figure 3.3), had no effect on inhibition of *S. scabies*. It was subsequently discovered using BLAST, that a paralog of this gene existed outside of this cluster in the Ps652 genome, and might be capable of complementing the deletion. A double deletion strain Ps652 Δ HCN Δ EchA8_1 Δ EchA8_2 was generated to test this hypothesis and, despite appearing to have incurred a fitness penalty evidenced by poorer on-plate growth, inhibition of *S. scabies* growth between the two *Pseudomonas* streaks was not abolished. However, it has been documented previously that PaaZ itself is not 100% efficient, and a small amount of its substrate undergoes Knoevenagel condensation to the 7-membered CoA-bound ring structure, potentially limiting the role of standalone enoyl-CoA hydratase domains to nothing more than increasing flux through the pathway (360).

3.4.1.2 Iron-binding Assays on CAS Agar

Given that hydroxytropolones had been shown to chelate ferric iron in CAS agar plates (Figure 3.16), overnight cultures of Ps652 Δ HCN and tropolone cluster mutants in 10 mL Lennox broth were spotted onto CAS agar plates and incubated overnight at 28 °C. Large halos were observed for Ps652 Δ HCN, but only small halos were observed for tropolone cluster mutants, suggesting loss of production of these siderophores (Figure 3.21). Complementation of the thioesterase mutant restored the Ps652 Δ HCN phenotype, indicating that differences in production of hydroxytropolones were responsible for the differences in phenotype. However, this contrasts with cross-streak data against *S. scabies*, where only a subset of mutants lost activity that had been attributed to hydroxytropolones. The residual small halos for tropolone cluster mutants are anticipated to be due to production of pyoverdine or further uncharacterised siderophores or additional excreted molecules that bind ferric iron with a higher affinity than CAS/HDTMA. Further analysis using more specific techniques was necessary to establish whether production of hydroxytropolones had been genuinely lost in all mutants.

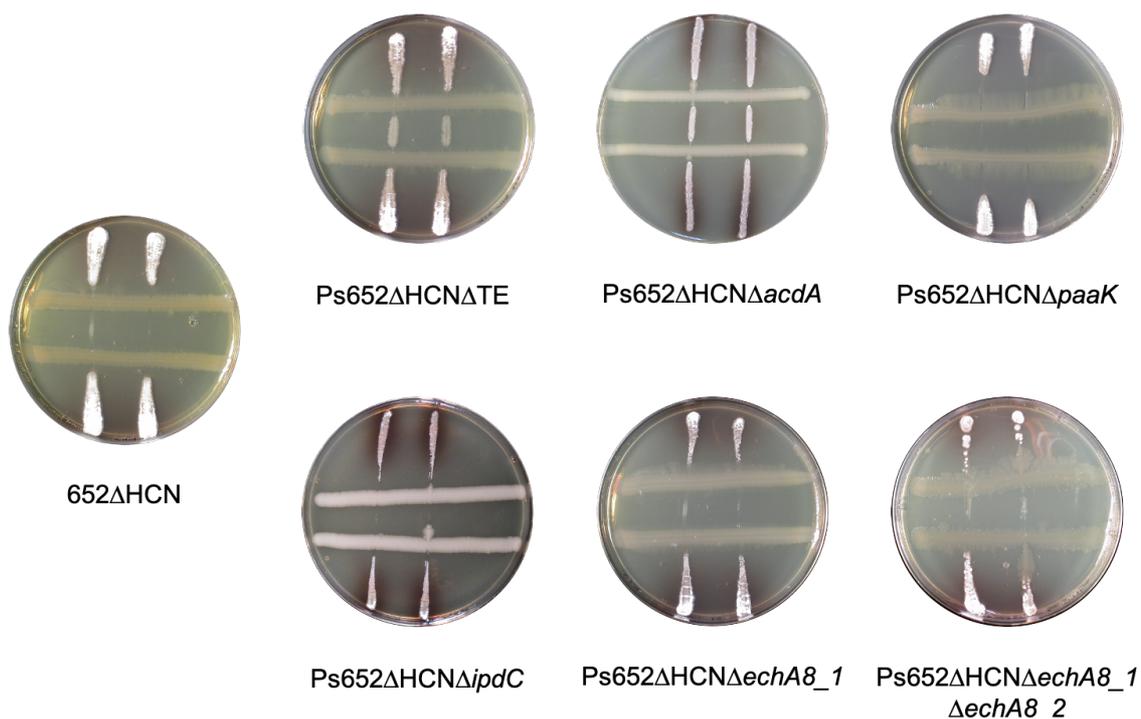


Figure 3.20. Cross-streak assays for Ps652 tropolone cluster mutants against *S. scabies* 87-22. Ps652ΔHCN was used as the parent strain for these mutants, as it was previously shown that hydrogen cyanide contributes in part to the biological activity of this strain. Thioesterase and dehydrogenase mutants, Ps652ΔHCNΔTE and Ps652ΔHCNΔacdA respectively, show loss of biological activity consistent with abolition of tropolone production. The other cluster mutants show no loss of activity compared to Ps652ΔHCN. Enoyl-CoA hydratase double mutant Ps652ΔHCNΔechA8_1ΔechA8_2 incurred a fitness penalty with less dense growth on plates. All images are representative of three biological replicates.

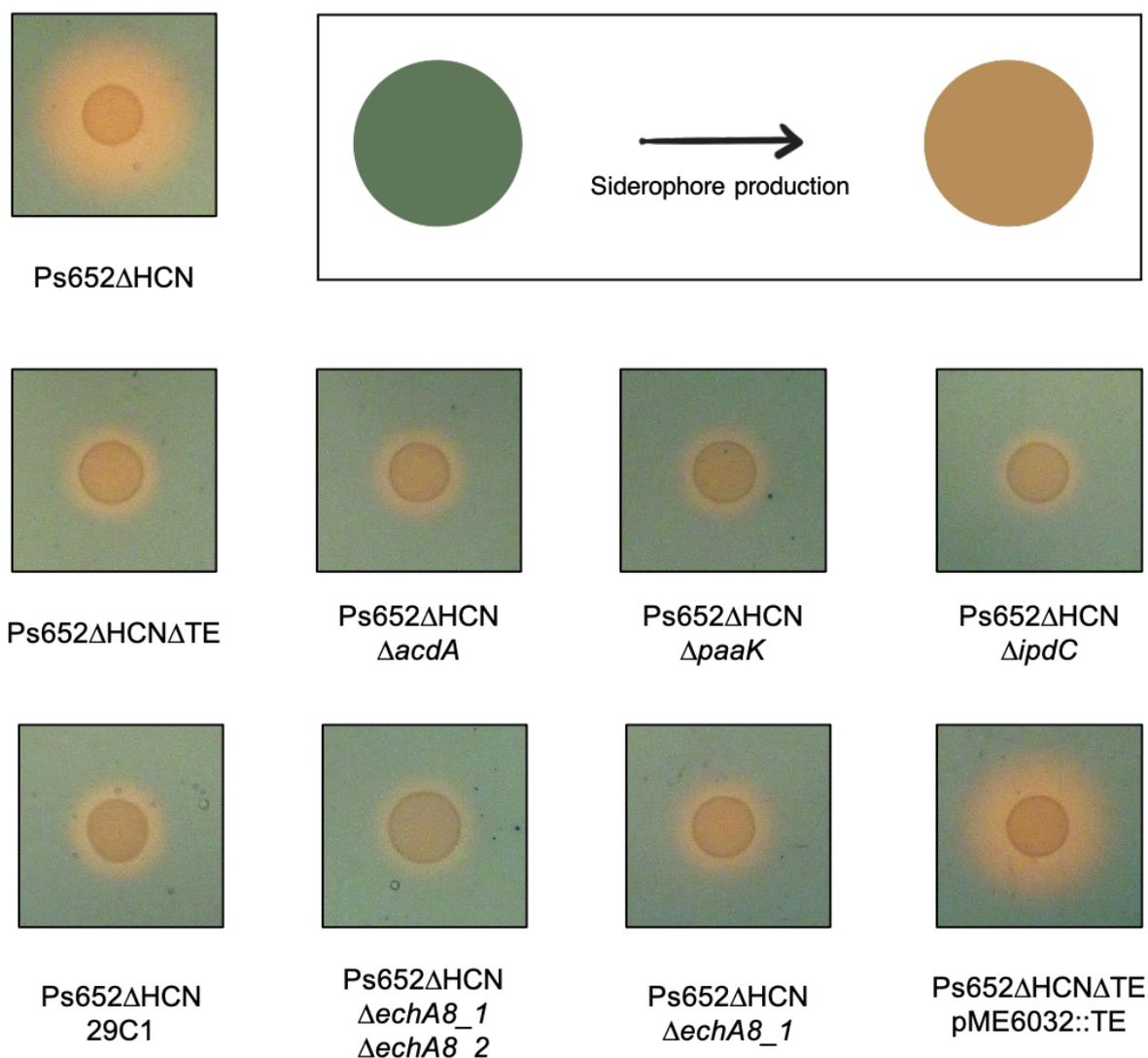


Figure 3.21. Iron-binding assay for Ps652ΔHCN and mutant derivatives using CAS agar plates. Shown are phenotypes on CAS agar plates for PsΔHCN and mutants in tropolone BGC genes. Production of siderophores liberates iron from the chrome azurol S and HDTMA complex, producing a colour change from blue/green to orange. Ps652ΔHCN shows a large zone, consistent with production of hydroxytropolones, whereas all tropolone BGC knockouts show loss of this large halo. The halo is restored in the complemented thioesterase mutant, indicating hydroxytropolones are responsible for the large halo. Tropolone BGC knockouts show a residual, smaller halo, likely due to production of pyoverdine. Images are representative of biological duplicates.

3.4.1.3 Production of Tropolones

To link the phenotypes observed in cross-streak assays and CAS agar iron-binding plates for the tropolone cluster mutants, LC-MS analysis was performed on EtOAc extracts of mutants grown in MKB Glucose, using a mass scan range of m/z 50-500. Included in the panel was the transposon mutant 29C1, which has a transposon insertion in the membrane fusion component of a tripartite multidrug resistance system encoded in the cluster. Production of both 7-hydroxytropolone and 3,7-dihydroxytropolone was abolished in all tested mutants, and restored in the complemented thioesterase mutant (Figure 3.22). This data matches the observed phenotypes on CAS agar plates (Figure 3.21), but is not in perfect accordance with the phenotypes observed for these mutants against *S. scabies* 87-22 in the cross-streak assay (Figure 3.20). However, this trend was consistent across the biological duplicates.

3.4.1.4 Searching for Tropolone Biosynthetic Intermediates by LC-MS

To probe the biosynthesis of tropolones, LC-MS analysis was performed on the same EtOAc extracts above in positive and negative mode with a mass scan range of m/z 50-500. The aim was to detect biosynthetic intermediates in the tropolone biosynthetic pathway that might inform as to the roles of each of the genes. This was particularly relevant for the predicted acyl-CoA thioesterase that is highly conserved in tropolone clusters in Proteobacteria, essential for production of these molecules, but currently has no assigned or hypothesised role. However, no intermediates of tropolone biosynthesis were observed for this mutant by mass spectrometry. 2-hydroxy-1,4,6-cycloheptatriene-1-carboxylic acid, the product of PaaY, was found to be substantially enriched in enoyl-CoA hydratase mutant extracts (calculated m/z = 151.0395; observed m/z 151.0387), and absent in all other strains with the exception of Ps652 Δ Hcn and the complemented thioesterase mutant (Figure 3.23). This suggests this molecule is a shunt metabolite of the pathway.

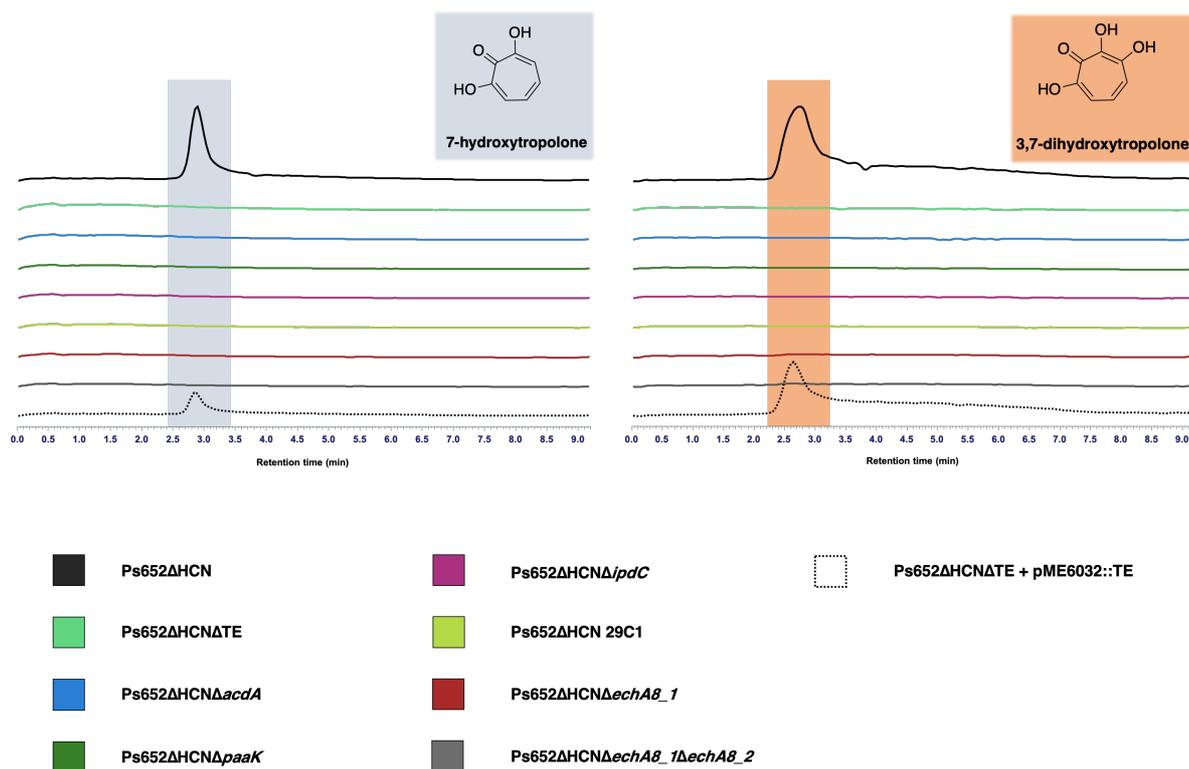


Figure 3.22. Loss of tropolone production in tropolone cluster mutants. Presented are LC-MS chromatograms for 7-hydroxytropolone (m/z 139.04, $M + H^+$) and 3,7-dihydroxytropolone (m/z 155.03, $M + H^+$) in ethyl acetate extracts of Ps652ΔHCN and biosynthetic gene cluster mutant derivatives, acquired in positive mode with a mass scan range of m/z 50-500. Production of hydroxytropolones is lost in all mutants, and is partially restored by complementation in the thioesterase mutant completed in trans. Data are representative of biological duplicate cultures. Chromatogram intensities are standardised relative to Ps652ΔHCN.

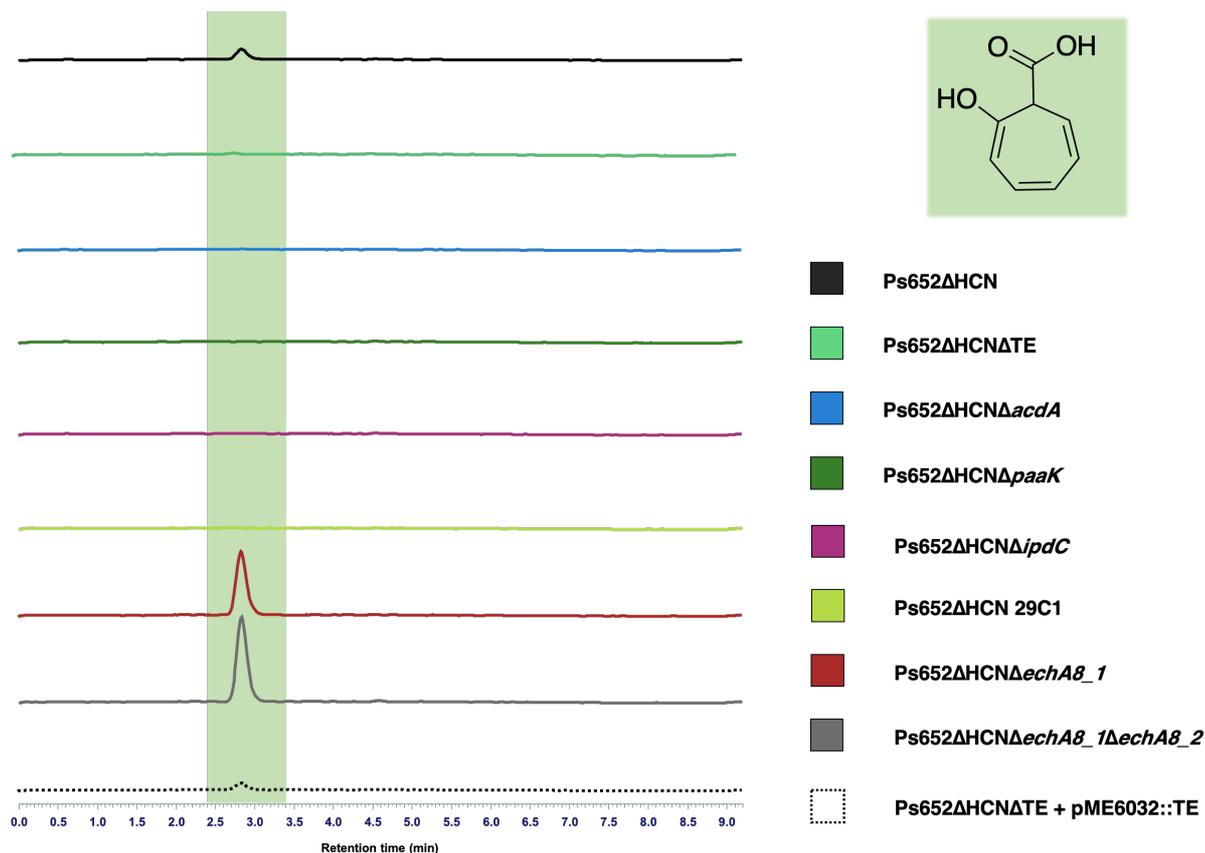


Figure 3.23. Accumulation of a shunt metabolite in tropolone cluster mutants. Presented are LC-MS chromatograms for m/z 151.04 (M-H) in ethyl acetate extracts of Ps652ΔHCN and biosynthetic gene cluster mutant derivatives, acquired in negative mode with a mass scan range of m/z 50-500. This matches the mass of the predicted PaaY product, 2-hydroxy-1,4,6-cycloheptatriene-1-carboxylic acid (calculated m/z 151.0395, observed m/z 151.0387; error = -5.29ppm), and accumulates in enoyl-CoA hydratase mutants Ps652ΔHCNΔechA8_1 and Ps652ΔHCNΔechA8_1ΔechA8_2. Low levels are observed in Ps652ΔHCN and complemented thioesterase mutant, but undetectable in other mutants. Data are representative of biological duplicate cultures. Chromatogram intensities are standardised relative to Ps652ΔHCN.

3.5 Field Trails for Ps652 against Potato Diseases

3.5.1 *Phytophthora infestans*

A common stumbling block in development of biocontrol strains for use in the field is that initial promise in the laboratory and in pot trials seldom materialises into genuine capacity to protect plants in the field. Given that pot trials are often not indicative of how well a strain will perform in the more varied and unpredictable setting of an agricultural field, the decision was taken to proceed directly to field trials in collaboration with VCS Potatoes. Strain inocula were prepared by lyophilisation, comprising 10^9 CFU/mL, verified by dilution plating of a sample of lyophilised material to confirm the number of viable cells. All subsequent field work was performed by VCS Potatoes. Strains were sprayed directly onto the crop once weekly from 19 days after planting, with the final application 59 days after planting. Plants were either left untreated or treated with Ps652 or Oxathiapiprolin in a Randomised Complete Block Design (RCBD) as per the methods section. A natural blight infection was recorded 47 days after planting as single lesions each on an untreated block and a block treated with Ps652. Observations were made 47, 54, and 66 days after planting, which was the end point of this trial. Severity of blight infection was measured as percentage of leaf area covered by *Phytophthora*.

At 54 days post planting, Ps652 offered a statistically significant protection from *P. infestans* when measuring % leaf coverage compared to untreated plants. Leaf coverage for untreated plants stood at an average of 15% across the 4 blocks, whereas the value for Ps652-treated blocks was 6.75%. However, the industry standard chemical fungicide, oxathiapiprolin, offered much greater protection as there were no visible lesions on plants treated with oxathiapiprolin after 54 days (Figure 3.24A). At 66 days post planting, the protection offered by Ps652 had been broken, with no statistically significant difference between the untreated plants and Ps652-treated plants at this time point, at averages of 85% and 78.75% respectively. However, oxathiapiprolin still offered substantial protection at this stage with only 1.5% leaf coverage with *Phytophthora* on average (Figure 3.24B).

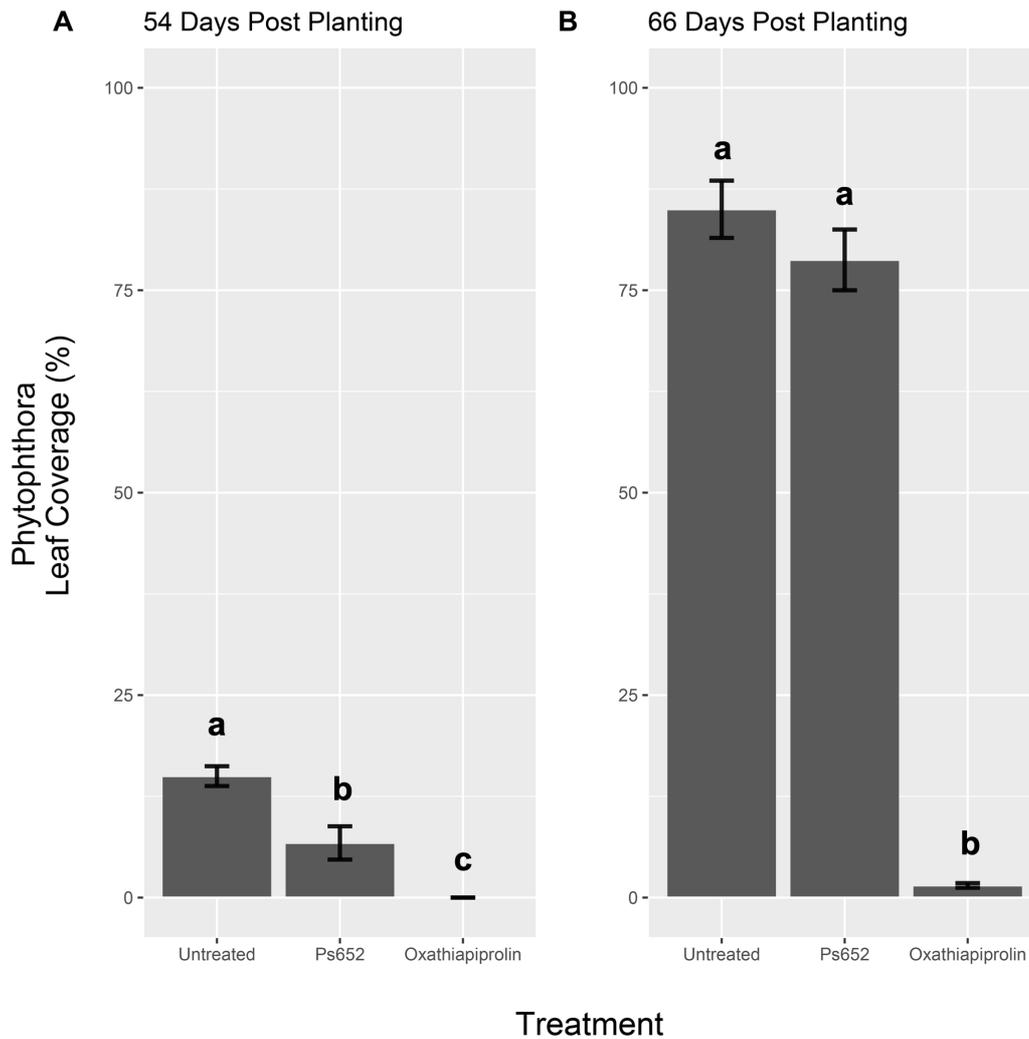


Figure 3.24. Performance of Ps652 as a biocontrol agent in field trials for protection of potato crops against *Phytophthora*. Shown is the percentage leaf coverage with *Phytophthora* in different treatment groups at 54 days (A) and 66 days (B) after planting. Ps652 offers early protection against late blight symptoms caused by *Phytophthora*, that is later overwhelmed in a natural infection scenario. Oxathiapiprolin offers significantly better protection at all time points. Treatments that share a letter are not statistically different at $p = 0.05$. $n = 4$ for each treatment.

3.5.2 *Streptomyces scabies*

Field trials were attempted for prevention of common scab disease caused by *Streptomyces* sp. In the 2021 spring-summer growing season using Ps652 or Ps925 as biological treatments as described in methods section 5.10.2, along with a no-treatment control. In short, seed potatoes were planted in parallel beds and irrigated using drip irrigation as per standard practice. 20 tubers per biological treatment / irrigation regime combination (Ps652 and Ps925) were dipped by hand in an $OD_{600} = 1.0$ suspension of either strain immediately before planting, and the rest were planted as normal with no dipping. Drip lines were used to apply 50 mL of $OD_{600} = 0.2$ culture per plant for each treatment on 17/05/2021 and again on 28/05/2021. Additionally, plants were split into three irrigation categories in order to account for fluctuations in environmental conditions: normal irrigation, half irrigation, and no irrigation. This aimed to ensure at least some common scab on no-treatment control plants in the event of a wet growing season limiting the development of common scab symptoms. Unfortunately, an excessively wet season with rainfall throughout the scab control period meant that upon sampling of individual plants on 20/09/2021, insufficient scab lesions were observed to draw meaningful conclusions. Lesions were only observed on tubers from plants that received no biological or chemical treatment, and no irrigation, and still had only a maximum of one lesion per tuber (data not shown).

3.6 Discussion & Conclusions

3.6.1 Summary / Discussion of Results

While data indicated that Ps652 harboured a gene cluster that was proposed to encode for the production of tropolone natural products, it had not been established whether Ps652 was capable of producing these molecules. This chapter has established firmly that the gene cluster originally identified by Yu *et al.* (346) as encoding a non-fluorescent siderophore, later identified as 7-hydroxytropolone (345), and linked to phytopathogen inhibition by others (341, 340), is responsible for the production of tropolones in Ps652. 7-hydroxytropolone was identified in mass spectral data from Ps652, as were iron-chelates that support its previously annotated role as a siderophore. Additionally, a dihydroxylated

tropolone, was purified and the structure elucidated as 3,7-dihydroxytropolone. This molecule has not been previously observed in *Pseudomonas* or linked to this BGC. While production of this molecule has been observed in *Streptomyces* (368), it has not been described in Gram-negatives to date. Given the levels of production, comparable with that of 7-hydroxytropolone, shows this is a significant product of the pathway and, given its additional hydroxylations compared to its monohydroxylated congener, may represent the true final product of the pathway.

Further experiments confirmed that both 7-hydroxytropolone and 3,7-dihydroxytropolone contribute towards the biological activity against *S. scabies* 87-22 *in vitro*. Interestingly, the dihydroxylated congener exhibited greater activity towards the pathogen than the monohydroxylated tropolone, despite both being able to bind iron in equal ratios (2:1 complexes). Similar trends are observed for relative activities of both molecules against various bacteria and fungi (374). This hints at an additional role of hydroxyl groups in tropolone biological activity, or possible improved solubility and uptake by *Streptomyces* of the dihydroxylated tropolone. It was not possible to assess the role of 3,7-dihydroxytropolone as a siderophore by itself using chemical complementation, as attempts to generate the double siderophore mutant strain Ps652 Δ HCN Δ Pyo Δ TE were repeatedly unsuccessful.

However, it was surprising to note that neither molecule showed activity towards *Phytophthora infestans* in the assay used, suggesting possible additional natural products contributing towards the antagonism of Ps652. Matuszewska *et al.* noted in their analysis of tropolone BGC expression under different carbon source availabilities, that *P. donghuensis* P482 harboured an additional putative biosynthetic gene cluster that contributed towards the antibacterial activity of that strain (367). They named the cluster “cluster 17”, which was only identified in antiSMASH 2.0 (375) and not in later versions of the tool, but it appeared in their estimations to be loosely associated with polyketide biosynthesis. A BLAST searched subsequently revealed this cluster is conserved in Ps652, where it may also contribute towards antimicrobial activity and explain the antagonism towards *P. infestans*.

The tropolone BGC in *Pseudomonas* has now been the subject of a number of publications as described above, and a number of the genes in the pathway have been characterised, including regulatory elements (347). A conserved annotated dehydrogenase, TdaE, found in *Pseudomonas*, *Burkholderia*, and *Phaeobacter*, was recently characterised as a flavoprotein-dependent dioxygenase, that can catalyse a number of steps in tropolone biosynthesis (362). A homologue of this gene is found in Ps652 with 59% amino acid identity, AcdA, and was identified here as being essential to production of these molecules. However, analysis of the literature suggests there may be multiple alternative routes to produce tropolones in bacteria, and it is unclear whether AcdA functions the same way in *Pseudomonas* as TdaE in *Burkholderia*. Furthermore, other genes are identified as part of the cluster both here and by others, that are required for the biosynthesis of tropolones but are given no proposed role in biosynthetic schemes. One such gene is that encoding a predicted acyl-CoA thioesterase that was observed in the present work to be critical for production of both 7-hydroxytropolone and 3,7-dihydroxytropolone, and also found to be important by Muzio *et al.* (340). Oddly, given the recent evidence that TdaE is capable of cleaving the thioester, as is PaaY, it is proving difficult to assign a role to this gene in the biosynthetic scheme.

A host of other genes not yet given roles in the pathway were also observed to abolish hydroxytropolone production when deleted, in both mass spectral data and iron-binding assays. These include the phenylacetate CoA ligase, PaaK, which is predicted only to improve precursor supply, but appears to eliminate production of 7-hydroxytropolone and 3,7-dihydroxytropolone, as well as the siderophore phenotype that those molecules produce when it is absent. Oddly this did not appear to adversely affect biological activity against *S. scabies*, which it has not been possible to fully explain at this stage. Another is the predicted alpha-keto-acid decarboxylase, whose involvement is curious given the absence of predicted alpha-keto-acid-containing intermediates. It was hypothesised this might function again by increasing precursor supply, converting phenylpyruvate to phenylacetate, before the ligation of the latter to CoA and downstream processing to tropolones. In this way, it would function similar to Lor1 in *P. inhibens* (348). The involvement of both IpdC and PaaK in the production of tropolones might be explained by the hypothesis that cellular phenylacetyl-CoA levels need to be high before phenylacetate

catabolism, and therefore tropolone production, can take place. Phenylacetyl-CoA was previously shown to inactivate PaaX in *E. coli*, allowing derepression of the phenylacetate catabolon (376). As this is required for tropolone biosynthesis, lower cellular phenylacetyl-CoA levels would reduce expression of PaaABCDEG, and prevent production of tropolones. By extension, if IpdC is responsible for increasing levels of phenylacetate by conversion from phenylpyruvate, its absence could indirectly lower phenylacetyl-CoA levels, and again repress PaaABCDEG expression and tropolone biosynthesis. In this way, mutants affected in these genes may be selectively unable to produce tropolones in media where the paa pathway is naturally inactive, in the absence of a mechanism to derepress the catabolon and promote tropolone biosynthesis. Thus, tropolone production may be compromised in these mutants in MKB or CAS agar, but not in MYM on which the *S. scabies* assays were performed.

It was hoped that searching for biosynthetic intermediates by mass spectrometry in relevant deletion strains might help uncover the roles of these mystery proteins in the biosynthetic pathway. An updated predicted scheme for biosynthesis of tropolones in *Pseudomonas* is presented in Figure 3.25. Only a single intermediate was observed in the dataset, which was identified as 2-hydroxy-1,4,6-cycloheptatriene-1-carboxylic acid. This indicated that tropolone biosynthesis may not proceed by the same route as that proposed for *Burkholderia*, where TdaE catalyses the majority of the steps. This intermediate only accumulated in enoyl-CoA hydratase (EchA8) mutants, which were proposed to act at an earlier stage to the formation of this product. In the absence of EchA8, which possibly somehow directs metabolites through a TE / AcdA-dependent pathway, the product of the Knoevenagel condensation may be preferentially acted upon by PaaY instead, producing the observed biosynthetic intermediate. If this intermediate is itself biologically active, that may explain the lack of phenotype observed for $\Delta echA8$ strains in the cross-streak assays. However, CAS agar plate data suggests this molecule is either retained intracellularly or does not bind iron, and either way hints that tropolone activity against *Streptomyces* may be independent of metal-binding properties.

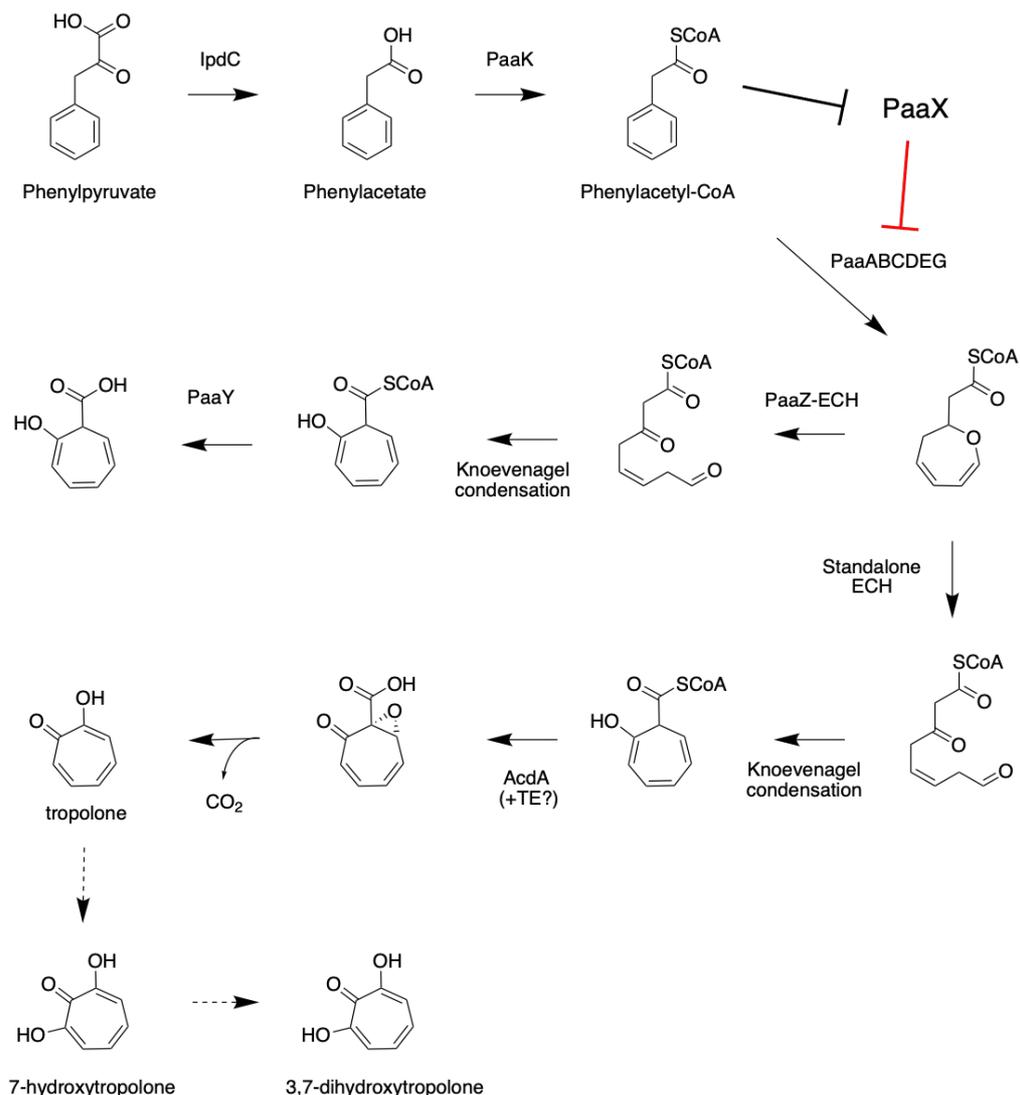


Figure 3.25. Updated proposed biosynthetic pathway for biosynthesis of tropolones in *Pseudomonas*. A putative decarboxylase (IpdC) is proposed to convert phenylpyruvate to phenylacetate, which is then attached to CoA by the phenylacetate-CoA ligase (PaaK). Under conditions where phenylacetate supply is naturally low, this could increase cellular phenylacetyl-CoA levels sufficiently to inhibit the paa pathway repressor PaaX, relieving repression of the paa pathway, and allow for the initial stages of phenylacetate degradation by PaaABCDEG. After the formation of 2-oxepin-2(3H)-ylideneacetyl-CoA, two possible routes exist for the pathway to follow. Either ring opening by the ECH domain of PaaZ followed by a Knoevenagel condensation to form 2-hydroxy-cyclohepta-1,3,5-trienecarboxylic acid-CoA, then cleavage from CoA by PaaY to form 2-hydroxy-1,4,6-cycloheptatriene-1-carboxylic acid. Alternatively, where a standalone ECH exists, this could ensure that the product of the Knoevenagel condensation is acted upon directly by AcdA (TdaE homolog), with subsequent spontaneous loss CO₂ to form tropolone. The conserved acyl-CoA thioesterase (TE) contains a conserved hotdog fold, and is predicted to act on CoA-bound substrates catalysing thioester cleavage, so may be involved in this step. The enzymes that perform the hydroxylations are unknown at this stage.

Finally, field trials designed to assess the suitability of Ps652 as a biological control agent in real agricultural settings were modestly successful. Pot trials were not attempted as they are not always a reliable predictor of success in the field. The trials presented here constitute a first attempt at using this strain for control of potato pathogens, and it is likely significant improvements are to be found in optimising both strain preparation and delivery. For example, the number of cells reaching the plant leaf then colonising and surviving in that environment was not assessed for late blight trials, and improving the formulations and delivery may enhance the crop protection offered. Regardless, Ps652 did show some success in controlling late blight in the short term when compared to an untreated control, indicating an ability to control *Phytophthora in planta*. Unfortunately, field trials assessing Ps652 in biological control of common scab caused by *Streptomyces* sp. were unsuccessful due to an atypically wet growing season and need to be repeated. Pot trials with mutant strains may be used in future to determine, where possible, the relative contributions of each natural product to this phenotype. But, as discussed previously, biocontrol phenotypes can often be more than the sum of their parts, and tricky to disentangle (7).

3.6.2 Future Work

Despite the ground gained in understanding bacterial biosynthesis of tropolone natural products to date, this work highlights the fact that there remain several unknowns about production of these molecules in *Pseudomonas* species. Various conserved genes are observed in the *Pseudomonas* tropolone BGC that have no predicted roles in biosynthetic schemes, yet appear important for production of these hydroxylated 7-membered ring natural products. In order to properly elucidate the biosynthetic pathway, purification of the encoded proteins followed by *in vitro* characterisation of their activity is necessary. This approach has proven successful for other steps in the pathway (362), and may be the only definitive way of solving tropolone biosynthesis in Proteobacteria. Particular emphasis should be given to the acyl-CoA thioesterase that is conserved across all tropolone-producing Proteobacteria, as recent characterisation of TdaE has left little need for a thioesterase (362), and its activity remains a mystery. Furthermore, the involvement of the putative decarboxylase and phenylacetate-CoA ligase in potentially derepressing the paa

pathway to allow tropolone biosynthesis should be investigated by analysing the expression of the paa pathway in different media / strain combinations with the use of a reporter system, and chemical complementation where appropriate.

Work with Ps652 has also highlighted the multifactorial nature of biocontrol phenotypes in bacteria, with both HCN and hydroxytropolones involved. However, there is enigmatic biological activity remaining, including an unidentified volatile produced under certain conditions, and an apparently diffusible natural product responsible for anti-oomycete activity. Investigations into this phenotype should include probing the mysterious “cluster 17” identified by Matuszewska *et al.*, which may encode for the production of a unique group of natural products (367).

What is clear is that we are only just beginning to unravel the intricacies of biological control of phytopathogens and the molecular weapons involved in those interactions. The approaches discussed herein appear to constitute a useful strategy for investigating, prioritising, and developing biocontrol bacterial strains which have the potential to transform our crop protection strategies.

Chapter 4: Bioinformatically Identified RiPP-like Clusters

4.1 Introduction & Aims

The original basis for this work is described in Stefanato et al (324). The aim of that work was to selectively isolate *Pseudomonas* strains from agricultural soils (RG Abrey Farms, Norfolk), and perform high-throughput screening for inhibitory activity against the potato pathogens *Phytophthora infestans* and *Streptomyces scabies*. Additionally, many high-throughput assays were conducted to investigate phenotypes relevant to the rhizosphere environment and potentially in biocontrol, which included swarming motility, secreted protease activity, and hydrogen cyanide production. Comparative genomics, including searching for known and novel natural product BGCs, was carried out using both existing bioinformatic tools and manual approaches by Dr A Truman. These were plotted against strain phylogeny and correlations analysis performed to link gene clusters to inhibition of phytopathogens (Figure 4.1). Gene clusters were identified by antiSMASH 3.0 (377) analysis that were listed as ‘bacteriocins’ in the output, and some were hypothesised on the basis of genetic architecture to be potentially novel RiPP gene clusters; they were subsequently named Pep1, Pep2, and Pep4 (Figure 4.2). These clusters showed correlation with inhibition of *P. infestans* (Pep1, $r = 0.47$; Pep2, $r = 0.37$; Pep4, $r = 0.05$) and *S. scabies* (Pep1, $r = 0.57$; Pep2, $r = 0.50$; Pep4, $r = 0.20$). However, despite being annotated as ‘bacteriocins’ by antiSMASH, no representatives of these putative BGCs have been characterised to date, and their function is unknown. Superficially, these clusters resemble RiPP BGCs, containing short peptides and putative tailoring enzymes. A more detailed analysis of these clusters is presented in section 4.2. The aim of the work presented here was to (i) investigate the hypothesis that these clusters represent RiPP biosynthetic gene clusters, (ii) evaluate their role in suppression of potato pathogens, and (iii) characterise any natural products encoded by these clusters. To this end, these BGCs were investigated in two closely related environmental *Pseudomonas* isolates, Ps706 and Ps708, originally isolated by Stefanato et al. (324).

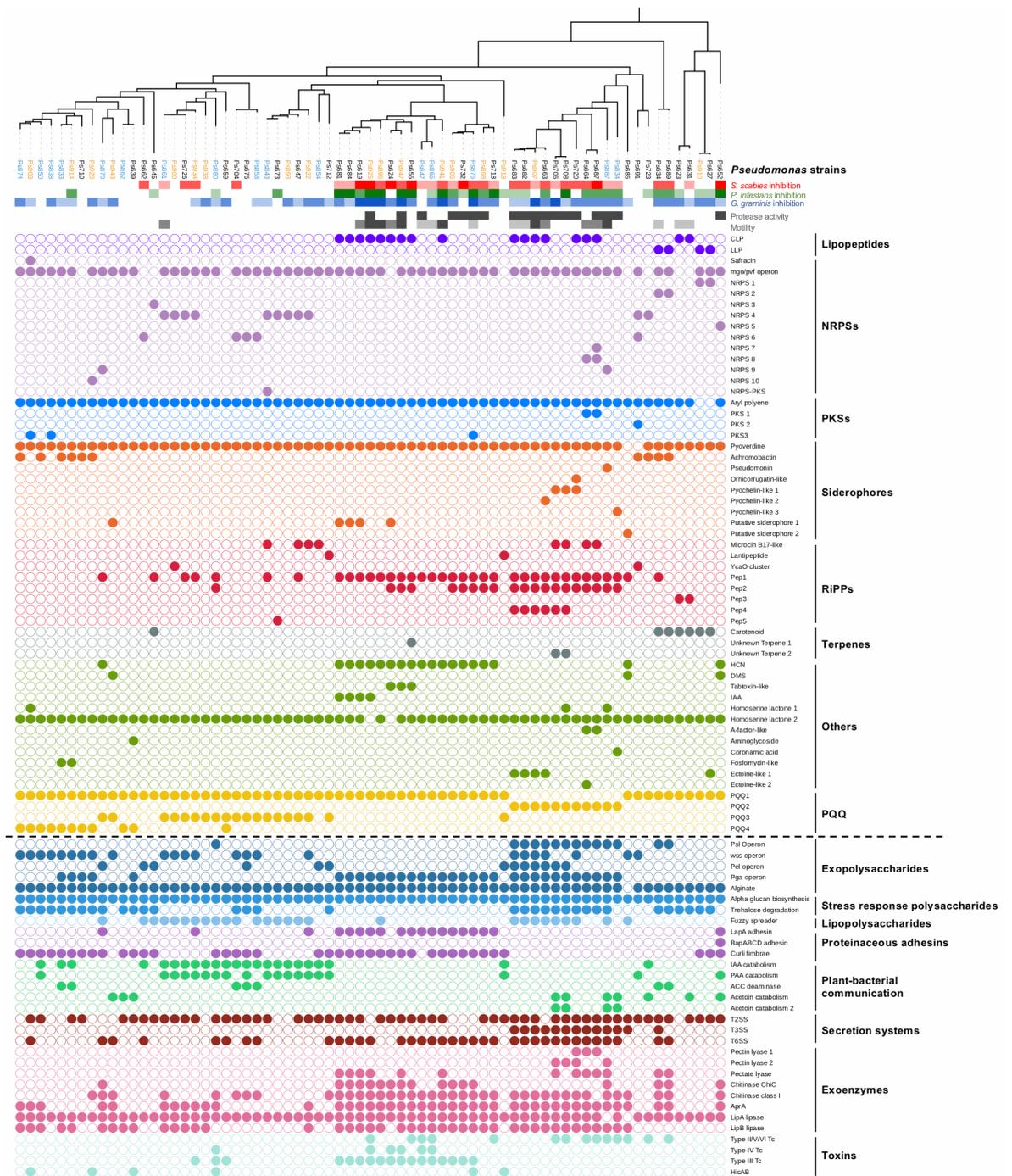


Figure 4.1. Comparative genomics for field isolated *Pseudomonads* with biocontrol potential.

Comparison of phylogeny, *S. scabies* suppression (red colour scales), *P. infestans* suppression (green colour scale), *Gaeumannomyces graminis pv. tritici* (take-all) suppression (blue colour scale), phenotypes (grey colour scales), natural product biosynthetic gene clusters (filled circles = presence of a gene or gene cluster) and the accessory genome (separated from BGCs by a dotted line). In the phylogenetic tree of *Pseudomonas* strains, blue strains were collected from irrigated plots while eight orange strains were collected from unirrigated plots. All other strains were collected from the pre-irrigation plots. Reproduced from Stefanato *et al.* (324) with permission.

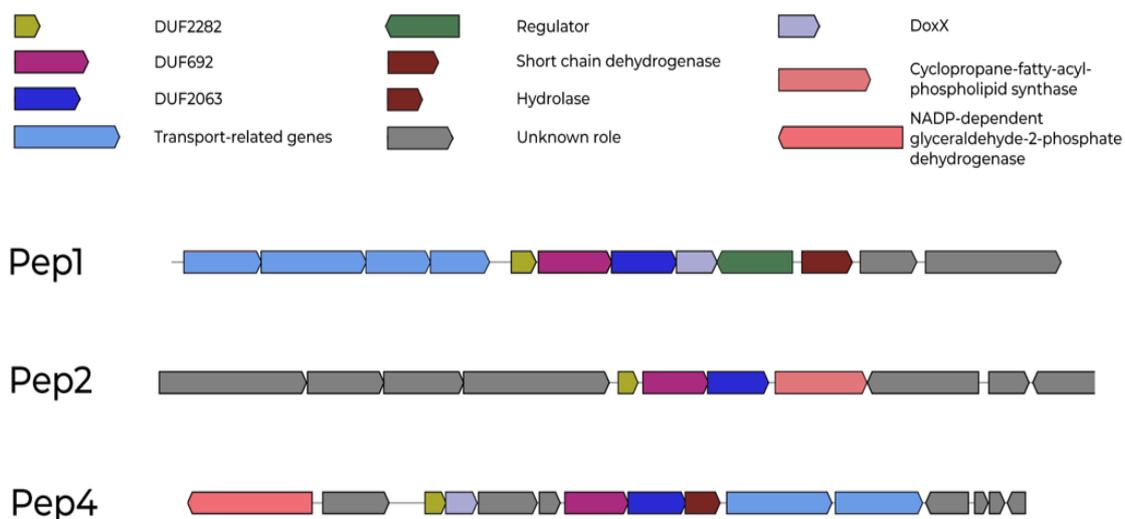


Figure 4.2. Pep1, Pep2, and Pep4 gene clusters as identified by antiSMASH. Genes and their annotations as per antiSMASH are shown. The DUF2282, DUF692, DUF2063 triplet is found in the Pep1 and Pep2 gene clusters, whereas DUF692 and DUF2063 are found adjacent to a different short peptide in the Pep4 cluster. Superficially, these clusters resemble those of Ribosomally Synthesised Post-translationally modified Peptides (RiPPs), with a short putative precursor peptide followed by a number of predicted tailoring enzymes. None of the clusters display architecture similar to known *Pseudomonas* bacteriocins.

4.2 Bioinformatic Analysis of Pep Gene Clusters

4.2.1 Overview of Pep Gene Clusters

These clusters resemble ribosomally synthesised post-translationally modified peptides (RiPPs) in that a putative precursor peptide is encoded alongside downstream putative peptide-modifying enzymes (233). Neither the precursor peptide nor the downstream hypothetical proteins had been characterised at this stage, indicating the clusters may be responsible for the production of as-yet unknown natural products. In each cluster is found a combination of hypothetical proteins with annotated domains of unknown function (DUF), including the putative precursor peptide. DUF2282 is found in the putative precursor peptide, and DUF692 and DUF2063 found in hypothetical proteins that we postulate may constitute peptide-modifying enzymes. At the outset of this project, little was known about these DUFs.

DUF2282 is listed by PFAM as a predicted integral membrane protein on the basis of sequence, but bioinformatics analysis presented here suggests this is not the case. To determine whether DUF2282 peptides are likely to be processed and secreted, SignalP 5.0 (378) was used to look for predicted signal peptide cleavage sites. A conserved signal peptidase cleavage motif (likelihood value = 0.9965) is observed in 88 of 91 analysed sequences as the amino acid sequence A-H-A, with a histidine to glutamine substitution in 3 of 88. This supports its role as a secreted peptide rather than an integral membrane protein. DUF2063 has a two-domain structure observed from crystallised members of this family (379), and is thought to constitute a transcriptional regulator, containing the putative DNA-binding domain PF09836. However, there is little verification of this hypothesis.

DUF692, until relatively recently, had not been associated with or had any function proposed for members of this family. Kenney *et al.* recently demonstrated a member of this family, MbnB, to be responsible for installation of an oxazolone and a thioamide motif, through a four-electron oxidation, in the precursor peptide MbnA of the RiPP methanobactin (380). An alignment of DUF692 from the Ps706 Pep1 cluster with MbnB

from *Methylosinus trichosporium* OB3b reveals that the Pep1 DUF692 contains 5 of the 6 residues identified as important for heterodimerisation, with only glutamate 249 absent (Figure 4.3). However, mutants lacking this residue were demonstrated to be inactive by Kenney *et al.* (380). Further work has since shown another DUF692 to be involved in a rather unique biosynthetic modification of a RiPP precursor peptide (381). In that example, TglB catalyses addition of a cysteine residue to the C-terminal alanine of the precursor peptide TglA, and the DUF692-containing TglH catalyses excision of the cysteine beta carbon. After further processing, 3-thiaglutamate is produced by cleavage from the precursor peptide, which is itself recycled. Again, in this case, TglH acts through a four-electron oxidation. As an operon, DUF2282, DUF2063, and DUF692 had previously been linked to oxidative and metal stress through correlation analyses in multiple species (382-384). However, considering recent evidence for DUF692 involvement in RiPP biosynthesis, the function of these clusters was still considered undetermined.

In order to ascertain whether these cluster likely related to the biosynthesis of RiPPs, a bioinformatic analysis was conducted looking for sequence conservation of short peptides associated with DUF692 proteins that may be RiPP precursor peptides, co-evolution between the two, and conservation of the surrounding genetic architecture. The workflow is described in ref (238) after being originally being developed and validated by Santos-Aberturas *et al.* (171). An overview of the biological logic of the workflow is presented in Figure 4.4, and a description of the process is presented in Figure 4.5.

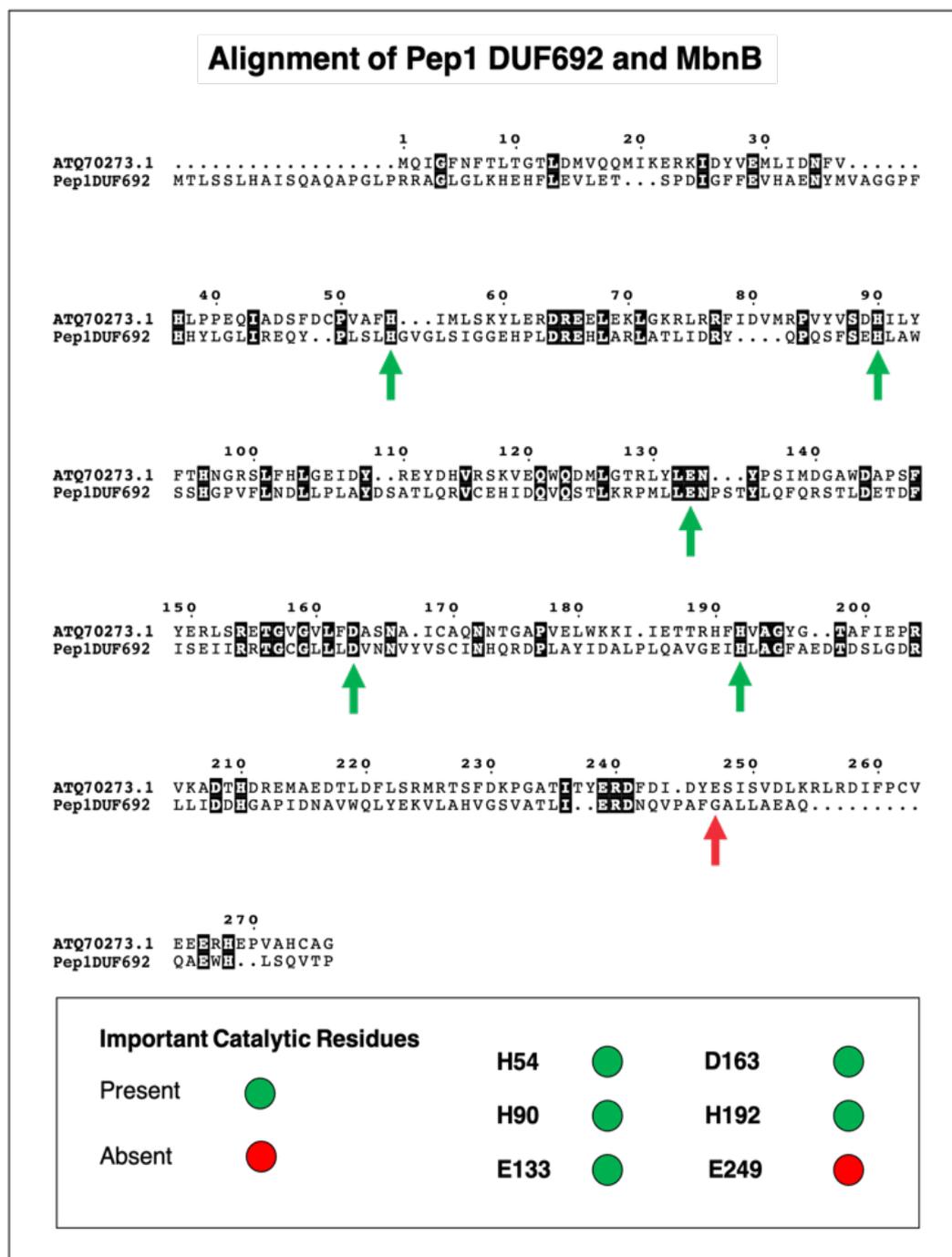


Figure 4.3. Alignment of MbnB from *Methylosinus trichosporium* OB3b with DUF692 protein from the Ps706 Pep1 cluster. Arrows highlight the location of amino acid residues important for the diiron catalytic centre, with green indicating the residue is conserved in Ps706 Pep1 DUF692, and red indicating absence. Based on the absence of E249, it is anticipated Ps706 Pep1 DUF692 would still form heterodimers, but may not be catalytically active or function in the same way as MbnB.

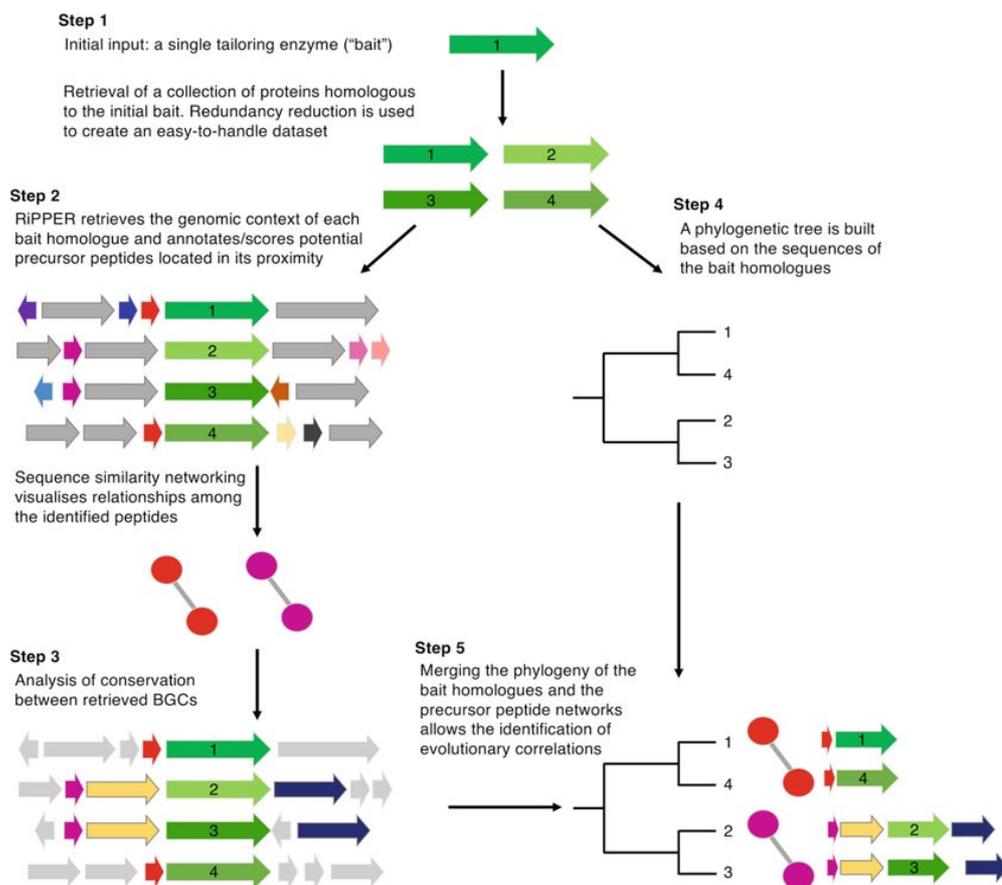


Figure 4.4. Overview of the biological logic of the RiPPER genome mining tool. The workflow involves searching databases for similar proteins to one the user has predicted to be involved in RiPP biosynthesis, and searching for co-evolution of nearby short peptides, as well as conservation of the surrounding genes. This process can rapidly distinguish good RiPP cluster candidates from coincidental co-occurrence of a predicted RiPP tailoring enzyme and a possible precursor peptide. Reproduced from Moffat *et al.* with permission (238) (see appendix 2).

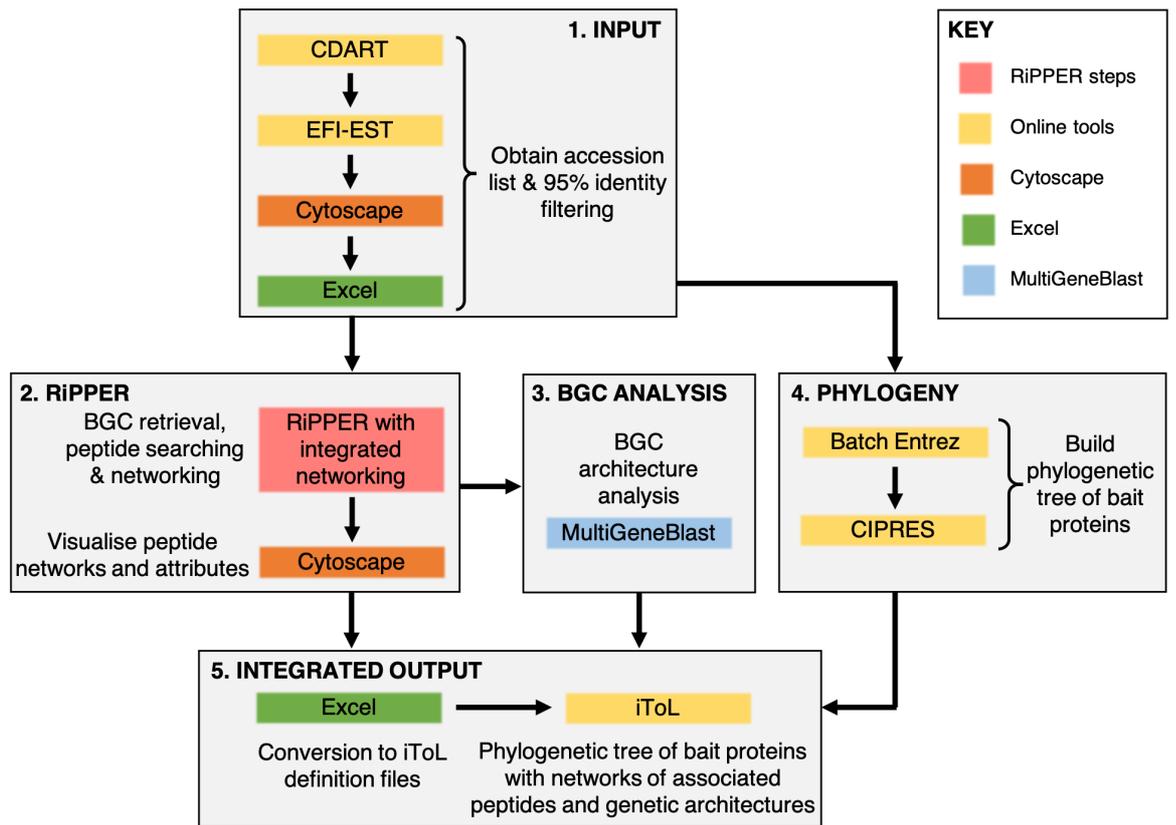


Figure 4.5. The steps and tools in the RiPPER genome mining workflow. Conceptual steps in the process are numbered, and the tools or programs used are represented by coloured boxes. The nature of the tools or programs is shown in the key. Some data manipulation or reformatting may be necessary between numbered steps. For full details of the process see the accompanying publication in appendix 2. Reproduced with permission from Moffat *et al.* (238) (see appendix 2).

4.2.2 Searching for Conserved DUF692-associated Short Peptides using RiPPER & EGN

The RiPPER tool was used with the standard parameters as described in the publication of the workflow (see appendix 2) to evaluate whether short peptides co-occur with DUF692 proteins, testing the hypothesis that the Pep clusters are RiPP or RiPP-like. Genbank files were retrieved for each protein accession from a 95% reduced list containing 907 sequences from the *Pseudomonas* genus, including the six DUF692 proteins found in *Pseudomonas* Ps706 and Ps708, and searched for short peptides occurring within 8 kbp of DUF692 proteins. The output list of short peptides was extensive, so sequence similarity networking was performed using the EGN tool (385), generating 279 networks of closely related peptides. Short peptides from each network were then aligned to others from the same network, and a consensus sequence generated for each of the largest five networks (Figure 4.6). It was observed that short peptides with strong sequence conservation are frequently found in association with DUF692 proteins. The largest network of short peptides contained the DUF2282-containing sequences (Figure 4.2) that were proposed to be putative precursor peptides for RiPP biosynthesis in the Pep gene clusters. It was noted that these short peptides contained a number of highly conserved cysteine, threonine, and serine residues, which may be sites of post-translational modifications. Conserved cysteines of this kind are found in methanobactins, where oxazolone and thioamide modifications are installed (386). The second largest network of short peptides contained peptides found in the Pep1 gene clusters from 706 and 708, but not found in the Pep2 or Pep4 gene clusters from these strains. These short peptides were divergently transcribed from the rest of the cluster, an arrangement not commonly seen in RiPP systems, and contained a DUF2288 motif with very strong sequence conservation (Figure 4.7). Members of the DUF2288 family of bacterial proteins have no known function at present.

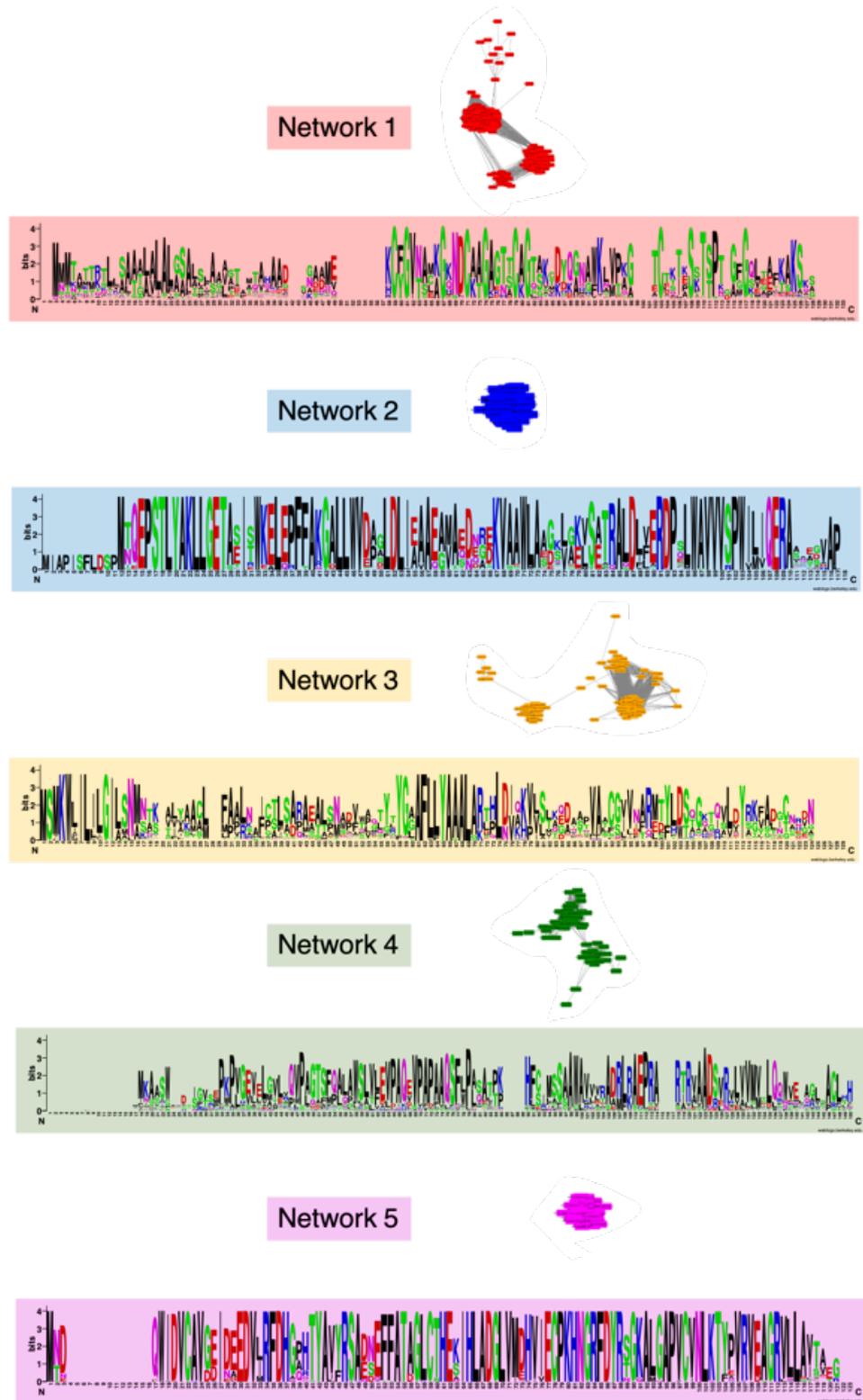


Figure 4.6. Networks of Short Peptides Retrieved from RiPPER-EGN and Aligned Sequences. The five largest networks are shown from the RiPPER-EGN output, along with their corresponding sequence conservation. Peptides belonging to networks 1 and 2 occur in the Pep gene clusters from strains Ps706 and Ps708.

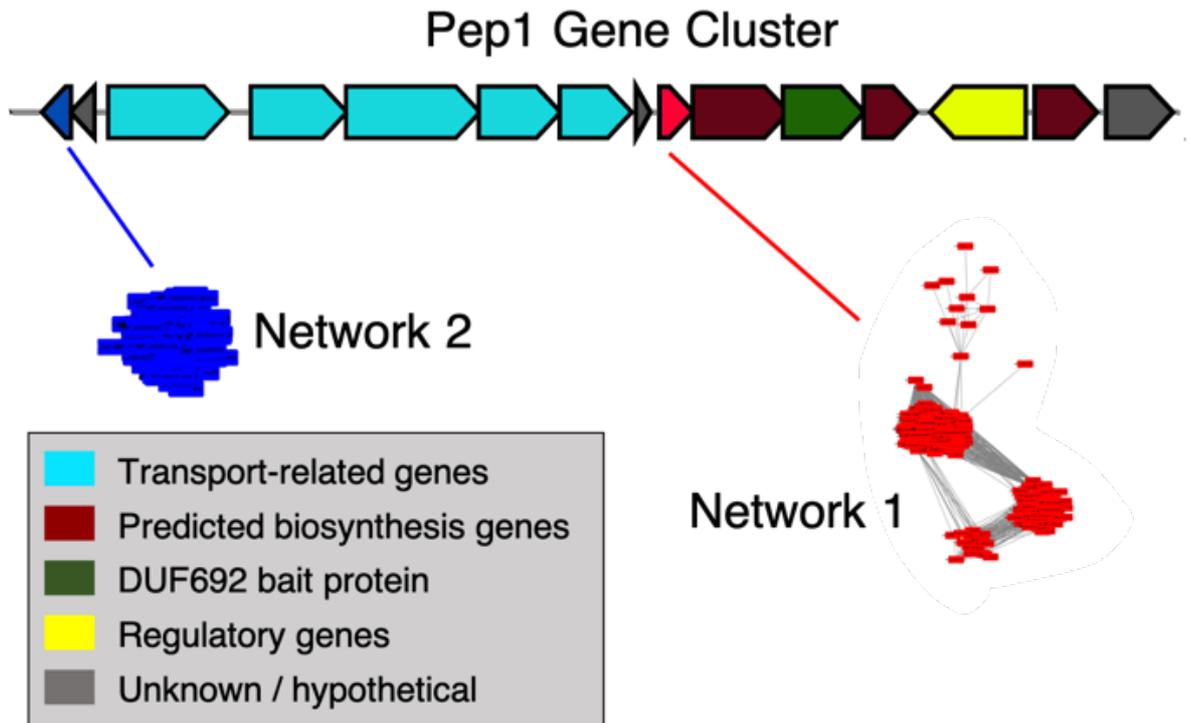


Figure 4.7. Pep1 Gene Cluster with Short Peptides Belonging to Networks 1 and 2. Short peptides belonging to network 2 were confined to a monophyletic group within the large clade that also harbour network 1 peptides. Network 1 peptides contain a DUF2282 domain and network 2 peptides contain a DUF2288 domain with high sequence conservation across the length of the peptide. DUF2288 represents a family of bacterial hypothetical proteins with no known function.

4.2.3 Phylogenetics of DUF692 Proteins

A list of nearly 3000 proteins containing DUF692 domains was retrieved using CDART (387) and filtered to various identity thresholds using EFI-EST (388). Firstly, to evaluate the evolutionary relationship between the recently reported DUF692 protein found in *Pseudomonas syringae maculicola* es4326 (TglH) described above and the DUF692 proteins found in Pep1 BGCs, a phylogenetic tree was constructed using a dataset containing DUF692 proteins with less than 85% identity. This was deemed sufficient to evaluate evolutionary relationships on a broad scale, while helping keep the number of sequences sufficiently low for computational analysis. This analysis revealed that, while DUF692 proteins from Ps706 and Ps708 form part of the same large monophyletic group (133 sequences), there is significant evolutionary distance between them and TglH, given the long branch length for the *P. syringae* protein. This protein appears to be very distantly related to other DUF692 proteins found using this strategy, suggesting that it may be fairly unique within sequenced *Pseudomonas* isolates, and its function may be quite distinct (Figure 4.8). This is consistent with the unusual reaction catalysed by TglH, which was the first record of such activity for a DUF692 protein, and the first record of such a biosynthetic mechanism in RiPP natural products (237).

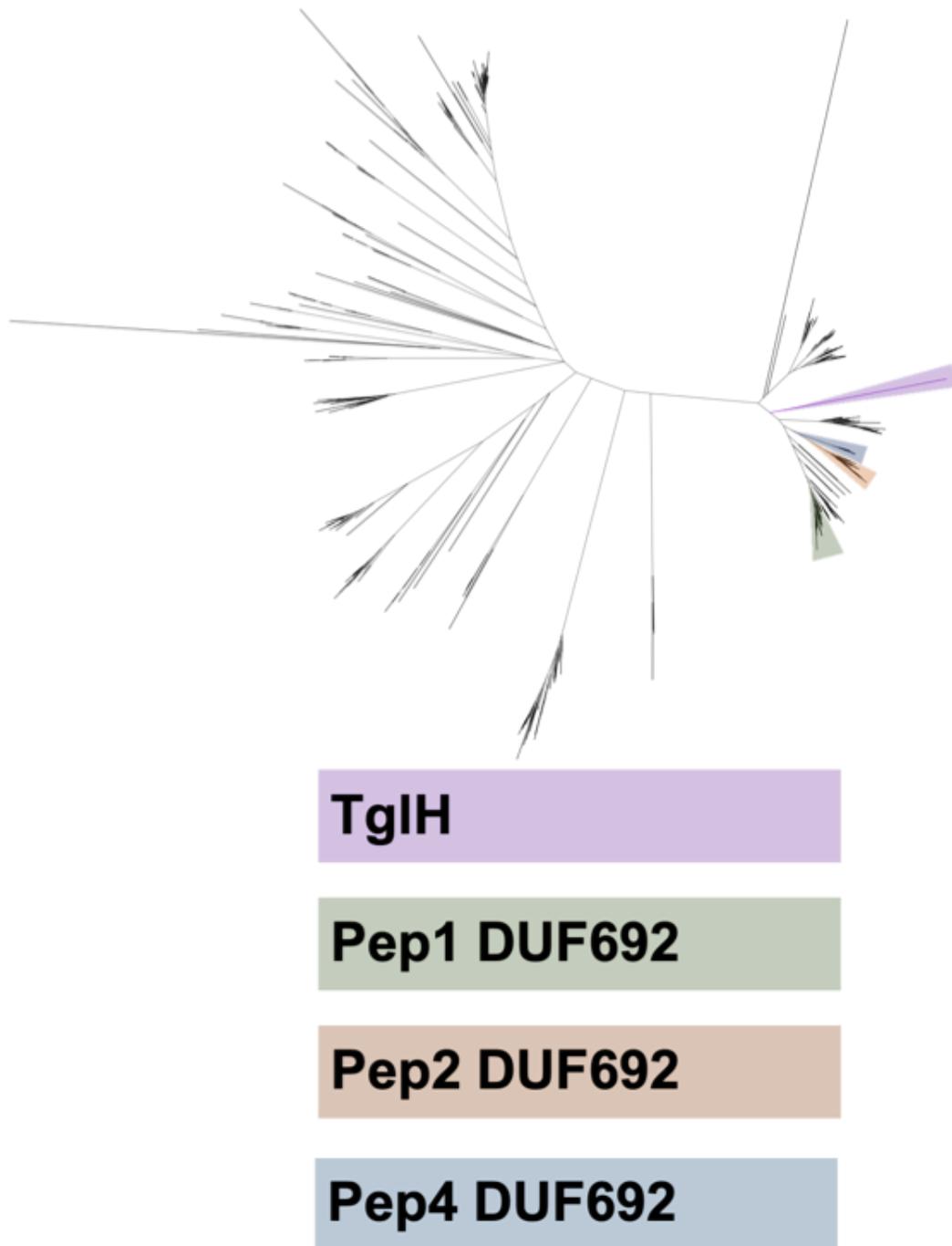


Figure 4.8. 85% Identity Unrooted Maximum Likelihood Tree of DUF692 Proteins in *Pseudomonas*. The positions of the DUF692 proteins from the environmental *Pseudomonas fluorescens* isolates Ps706 and Ps708 are shown. The DUF692 protein from the 3-thiaglutarate cluster of *P. syringae maculicola* es4326, TgIH, occurs as a long single branch with no closely related proteins in sequenced pseudomonads. 482 sequences were used from an 85% identity reduced dataset.

Another phylogenetic tree was built using a 95% identity filtered dataset to visualise evolutionary trends on a finer scale. The representative proteins from Pep2 clusters formed a clade with WP_044288082.1 from *Pseudomonas simiae*. Networks of RIPPER-identified short peptides were mapped to this phylogenetic tree for the largest five networks, which mapped to well-defined clades (Figure 4.9). Network 1 mapped to a large clade containing DUF692 proteins from all three Pep cluster types, and consisted of a large paraphyletic group. Network 2 mapped to a monophyletic group, representing a subset of those sequences contained within the network 1 clade. This arrangement suggests co-evolution between DUF692 and their associated short peptides, supporting the hypothesis that DUF692 proteins may constitute RiPP tailoring enzymes that act on genetically co-located precursor short peptides.

MultiGeneBlast (369) analysis was conducted to visualise the level of gene conservation across the different clades of DUF692 proteins (Figure 4.9). Here, the Ps706 Pep1 gene cluster Genbank file from antiSMASH (328) was used as the input to search for homology. Representatives from network 1 and 2 clades show strong conservation of the entire gene cluster. Representatives from the other network clusters show very little conservation of other genes from the Pep1 gene cluster. Curiously, with default MultiGeneBlast settings, the DUF692 protein itself was not annotated as homologous for the distantly related cluster 4 and cluster 5 clades. Overall, these results indicate DUF692 proteins co-occur with co-evolving short peptides, which is strongly indicative of a RiPP-like cluster, further supported by strong conservation of the surrounding genes in a subset of DUF692 proteins.

Further MultiGeneBlast analysis suggested the Pep1 BGC may be larger than previously anticipated.

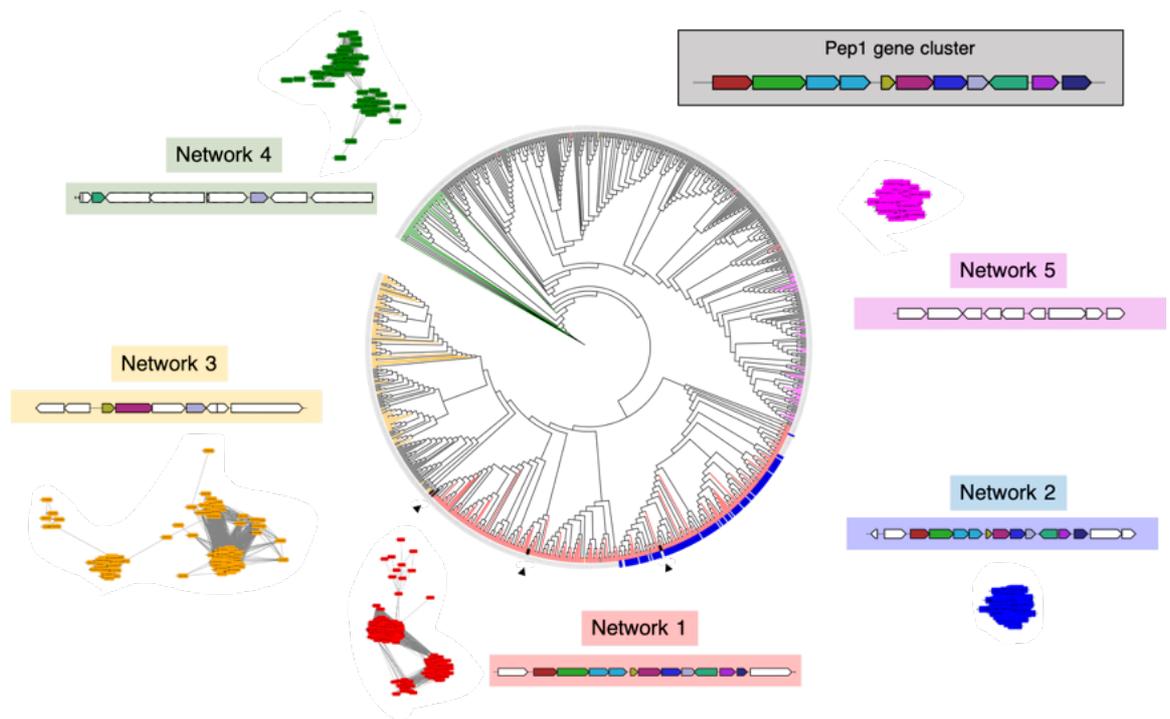


Figure 4.9. Evolutionary Relationship Between DUF692 Proteins and Networks of RiPPER-identified Short Peptides. Short peptide networks 1-5 from the EGN output were mapped back to their respective DUF692 proteins in the phylogenetic tree, identifying clades that share similar short peptides. Network 1 (red) mapped to a large paraphyletic group, which also contained a large monophyletic group which also contain network 2 (blue) short peptides. MultiGeneBlast analysis is shown for each, highlighting strong gene conservation amongst the network 1 and 2 clades. Black wedges indicate the positions in the tree for Pep2, Pep4, and Pep1 cluster DUF692 proteins (left to right).

4.2.4 Establishing BGC Boundaries using MultiGeneBlast

In order to gain an understanding of gene conservation to inform further work, MultiGeneBlast output from the RiPPER workflow was investigated for the gene clusters that contained network 2 short peptides. These represented a monophyletic group of DUF692 proteins that included those from the Ps706 and Ps708 Pep1 gene clusters, and had a highly conserved gene architecture, which differed from BGCs containing more distantly-related DUF692 proteins. The Ps706 Pep1 gene cluster file as downloaded from the antiSMASH analysis detailed above was used as the input, and all network 2 peptide-containing BGCs were searched for homology. The resulting output showed a high degree of conservation of the gene cluster (Figure 4.10A). The genetic context for the cluster in each strain is highly similar, with some small variations as would be expected. It was noted that a set of three genes adjacent to the cluster appeared to be highly conserved but had not been annotated as part of the cluster by antiSMASH analysis (Figure 4.10B). These genes encode an NAD(P)/FAD-dependent oxidoreductase (WP_010212374.1), a PLP-dependent cysteine synthase family protein (WP_154847011.1), and a Na⁺ / H⁺ antiporter NhaA (VVN91338.1). It was not known whether these were likely to be involved in natural product biosynthesis. The NAD(P)/FAD-dependent oxidoreductase was conserved in 100% of the clusters, but not predicted by antiSMASH (377) as a natural product biosynthetic protein.

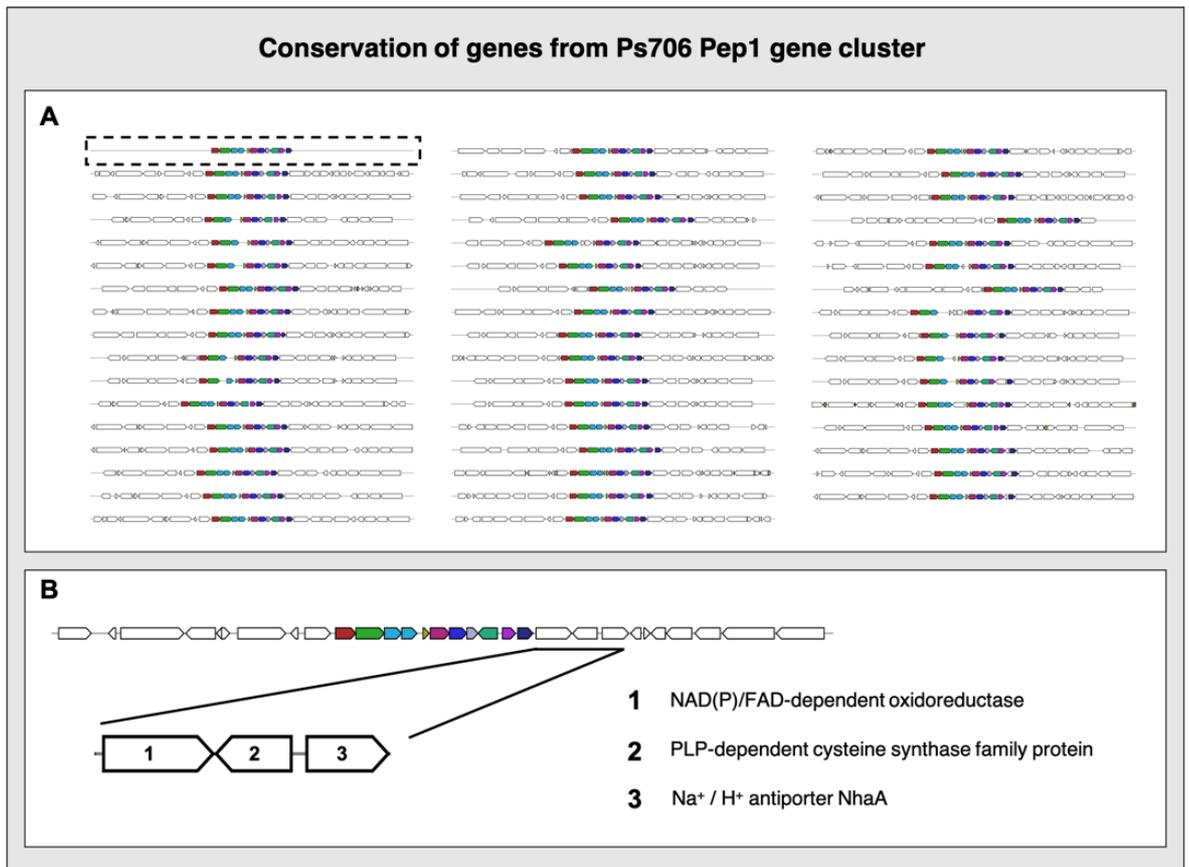


Figure 4.10. Conservation of genes from the Ps706 Pep1 gene cluster by Multigeneblast analysis. **(A)** MultiGeneBlast output for network cluster 2 peptide-containing gene clusters identified by RiPPER analysis showing strong conservation of the cluster as identified by antiSMASH. The hatched box indicates the Ps706 cluster that was used as the query for MultiGeneBlast. **(B)** Additional conserved genes not annotated by antiSMASH that are conserved in all but one of the gene clusters. The NAD(P)/FAD-dependent oxidoreductase is present in all identified gene clusters.

4.3 Gene Deletion Experiments

4.3.1 Deletion of Pep Gene Clusters

Given the link between genotype and phenotype from initial investigations, deletions of the Pep gene clusters was a natural first step in investigating the role of the bioinformatically identified gene clusters in the inhibition of various phytopathogens. The Pep1, Pep2, and Pep4 gene clusters were selected for deletion in the environmental isolates Ps706 and Ps708. Deletions were performed by markerless allelic exchange as described in the methods and ref. (181). Deletion of Pep1 and Pep2 gene clusters were performed both individually and as double mutants in Ps706 and Ps708 by Dr. F Stefanato. Deletion of the Pep4 gene cluster was later performed both singly and in a Δ Pep1 Δ Pep2 background for both strains using the same methodology by myself. A list of strains created is shown in Table 4.1.

Table 4.1. Pep Cluster Deletion Strains

Strain Name	Genes Deleted From Cluster	Parent Strain
706ΔPep1	DUF2282, DUF692, DUF2063	706
706ΔPep2	DUF2282, DUF692, DUF2063	706
706ΔPep4	DUF2282 → short chain dehydrogenase	706
706ΔPep1ΔPep2	DUF2282, DUF692, DUF2063	706 Δ Pep1
706ΔPep1ΔPep2ΔPep4	DUF2282 → short chain dehydrogenase	706 Δ Pep1 Δ Pep2
708ΔPep1	DUF2282, DUF692, DUF2063	708
708ΔPep2	DUF2282, DUF692, DUF2063	708
708ΔPep4	DUF2282 → short chain dehydrogenase	708
708ΔPep1ΔPep2	DUF2282, DUF692, DUF2063	708 Δ Pep1
708ΔPep1ΔPep2ΔPep4	DUF2282 → short chain dehydrogenase	708 Δ Pep1 Δ Pep2

4.3.2 Investigating the Link between DUF2282-DUF692-DUF2063 Operon & Oxidative Stress

Some sources have linked the DUF2282-DUF692-DUF2063 set of genes to oxidative stress tolerance, or heavy metal tolerance (389, 382-384), possibly indirectly through oxidative stress tolerance. To test this, serial dilution spot assays were performed on LB agar plates containing sodium chlorite added as an oxidant. Preliminary experiments had suggested a concentration of 0.25 mM sodium chlorite precluded growth of all strains, whereas a concentration of 0.01 mM had no effect on growth (not shown). Accordingly, a concentration of 0.05 mM sodium chlorite was used. Ps706, and all mutant derivatives, were grown overnight, standardised to an $OD_{600} = 1.0$, followed by serial dilution from 10^{-1} to 10^{-6} . No visually significant differences could be observed under the tested conditions between the wild type and any mutant that would validate these clusters being truly involved in oxidative stress tolerance (Figure 4.11).

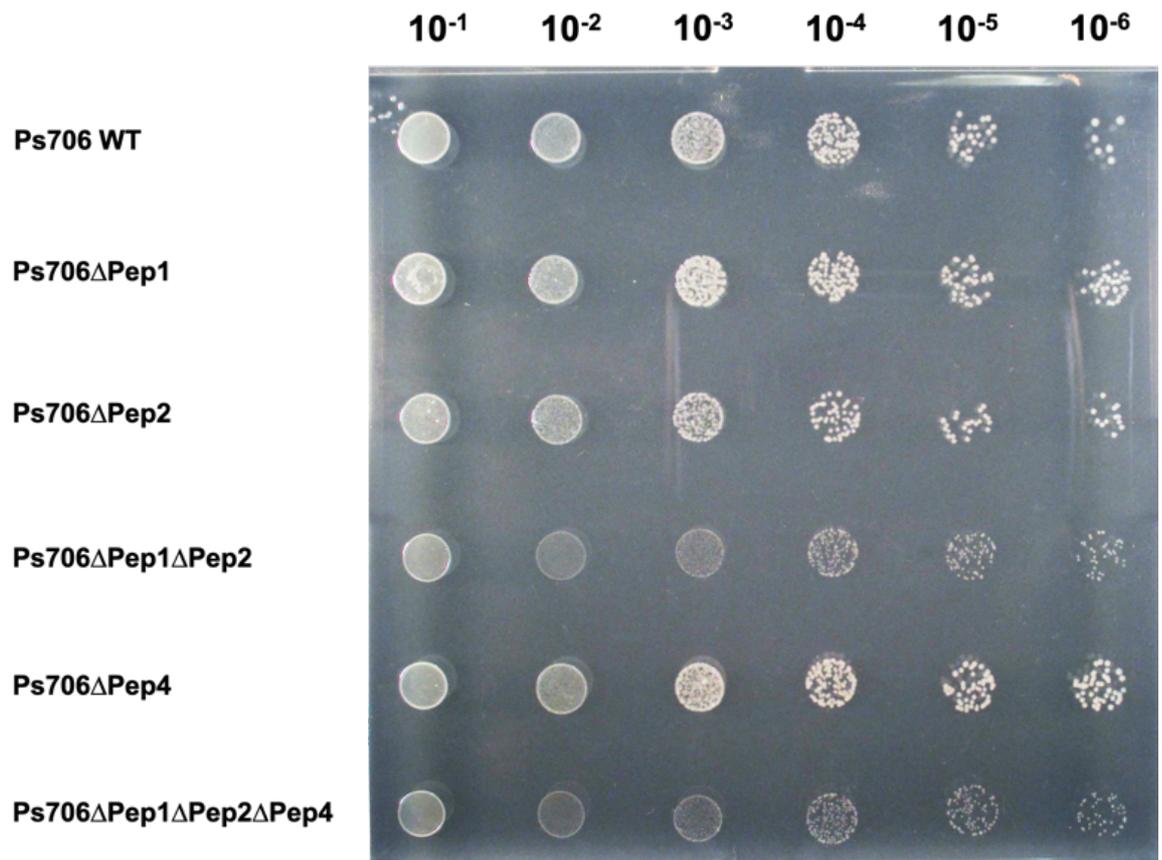


Figure 4.11. Serial dilution spot assay on LB plates with 0.05 mM sodium Chlorite for Ps706 and Pep cluster mutants. No visually significant differences could be observed for any of the strains at any of the tested dilution factors.

4.3.3 Bioassays with Deletion Strains

Once the deletion strains for environmental isolates Ps706 and Ps708 had been obtained (Table 4.1), an obvious next step was to test the effect of these gene deletions on the strains' inhibitory activity against phytopathogens. Again, the same pathogens were used as in Stefanato *et al.* (324). *Streptomyces scabies* 87-22 was used as the bacterial phytopathogen, *Gaeumannomyces graminis* var. *tritici* as the fungal phytopathogen, and *Phytophthora infestans* 6_A1 as the oomycete phytopathogen.

4.3.3.1 *Streptomyces scabies*

The cross-streak assay was used to test whether there was any effect of Pep cluster deletion on inhibitory activity against *Streptomyces scabies* 87-22. This assay allows visualisation of various modes of inhibition simultaneously. Contact-based inhibition may be observed where the two bacterial strains meet, inhibition caused by diffusible compounds may be observed by zones of inhibition, and inhibition due to volatile compounds produced by one strain may cause plate-wide effects. As such, the assay has significant potential to establish whether the investigated gene clusters are involved in a biocontrol phenotype. Ps706 and Ps708 had shown anti-*scabies* activity in initial assays, but contained no BGCs expected to produce this phenotype, like those encoding hydrogen cyanide or lipopeptide antibiotics (Figure 4.1). In this case, no effect of any of the BGC deletions was visible for strain Ps706 on MYM or SFM agar (Figure 4.12) or for strain Ps708 on MYM or SFM agar (Figure 4.13). In light of these results, it does not appear the bioinformatically identified Bac/pep gene clusters contribute to biocontrol activity *in vitro* against *S. scabies* 87-22.

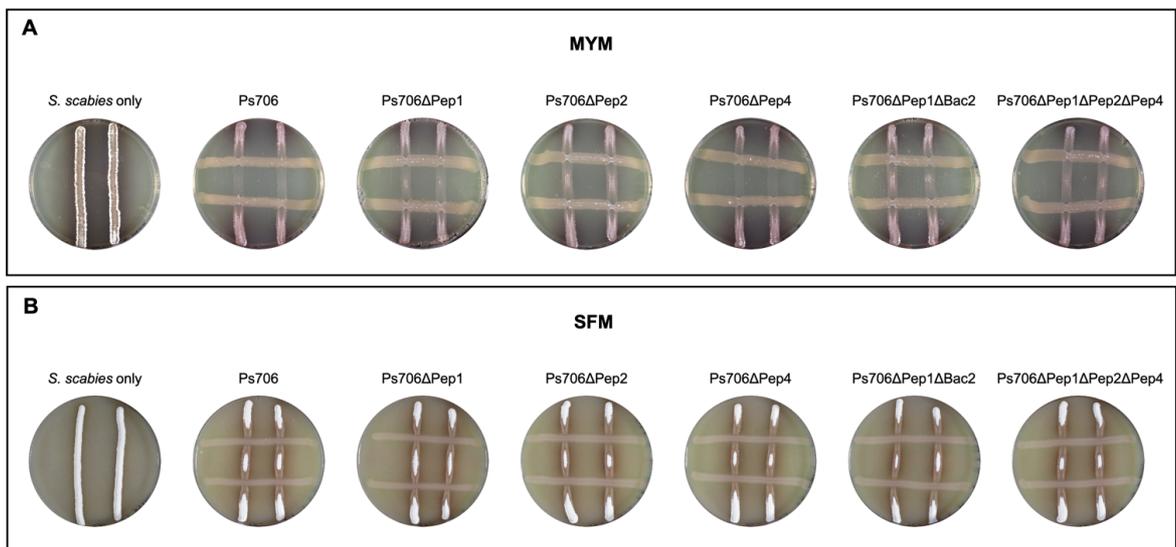


Figure 4.12. Cross streaks of environmental isolate Ps706 and Pep cluster mutants against *Streptomyces scabies* 87-22. When compared to the wild type, no significant difference in inhibitory phenotype towards the phytopathogen was observed on either MYM agar (top panel) or SFM agar (bottom panel). All plates are representative of three biological replicates. *Ss*, *Streptomyces scabies* 87-22.

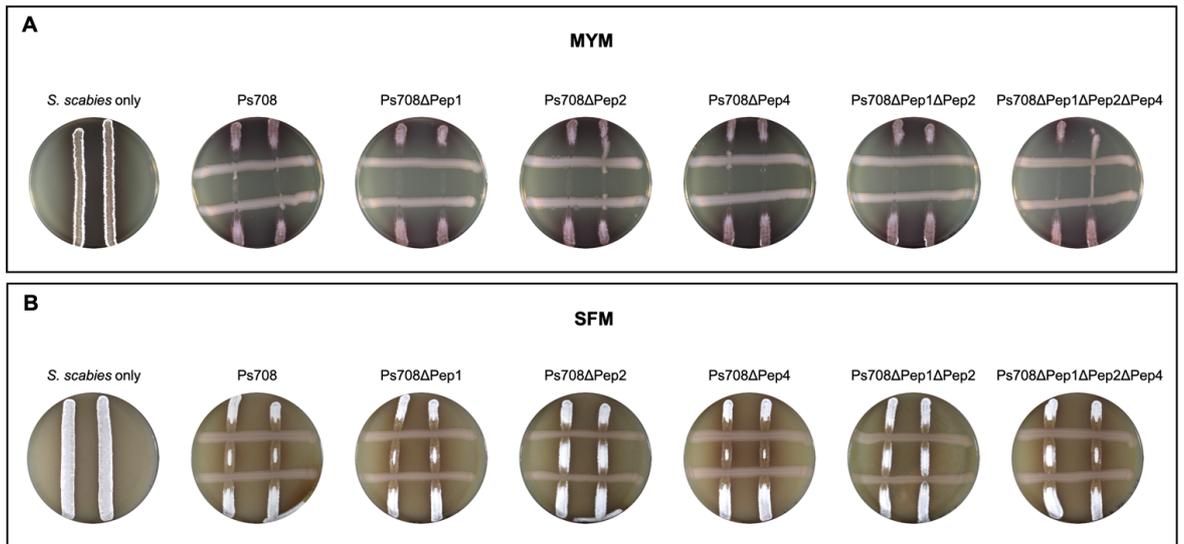


Figure 4.13. Cross streaks of environmental isolate Ps708 and Pep cluster mutants against *Streptomyces scabies* 87-22. When compared to the wild type, no significant difference in inhibitory phenotype towards the phytopathogen was observed on either MYM agar (top panel) or SFM agar (bottom panel). All plates are representative of three biological replicates. *Ss*, *Streptomyces scabies* 87-22.

4.3.3.2 *Phytophthora infestans*

Assays against *Phytophthora infestans* were performed on rye sucrose agar plates (RSA). Results for Ps706 and mutants thereof are described, with no data presented for Ps708 and derivatives. Neither Ps706 nor Ps708 were observed to be strong inhibitors of this phytopathogen in preliminary assays. While Ps706 was able to grow alongside the pathogen without being overgrown, there was no inhibition of *P. infestans* at a distance, evidenced by the lack of zone of inhibition. Consistent with data for inhibition of *S. scabies* 87-22, no difference in phenotype is observed between the wild type Ps706 and any Pep cluster mutant strains, with all lacking zones of inhibition but resisting being overgrown by the pathogen (Figure 4.14).

4.3.3.3 *Gaeumannomyces graminis*

Despite not being a potato pathogen, *G. graminis* var. *tritici* was selected as a model fungal phytopathogen against which to screen for antifungal activity of any potential products of the Pep gene clusters. Again, only strain Ps706 was used in this assay, with Ps706 and Ps708 showing similar levels of inhibition against this pathogen in earlier work. While not the strongest inhibitors of fungal pathogens in the collection of 240 *Pseudomonas* isolates, the strains were able to limit the growth of *G. graminis* more so than they were *P. infestans*. Even after 10 days growth, the fungal mycelium did not reach the *Pseudomonas* colonies, indicative of a zone of inhibition. However, it was observed that deletion of the Pep gene clusters had no effect on this phenotype, with mutants showing the same limitation of fungal growth as the wild type Ps706 (Figure 4.15).

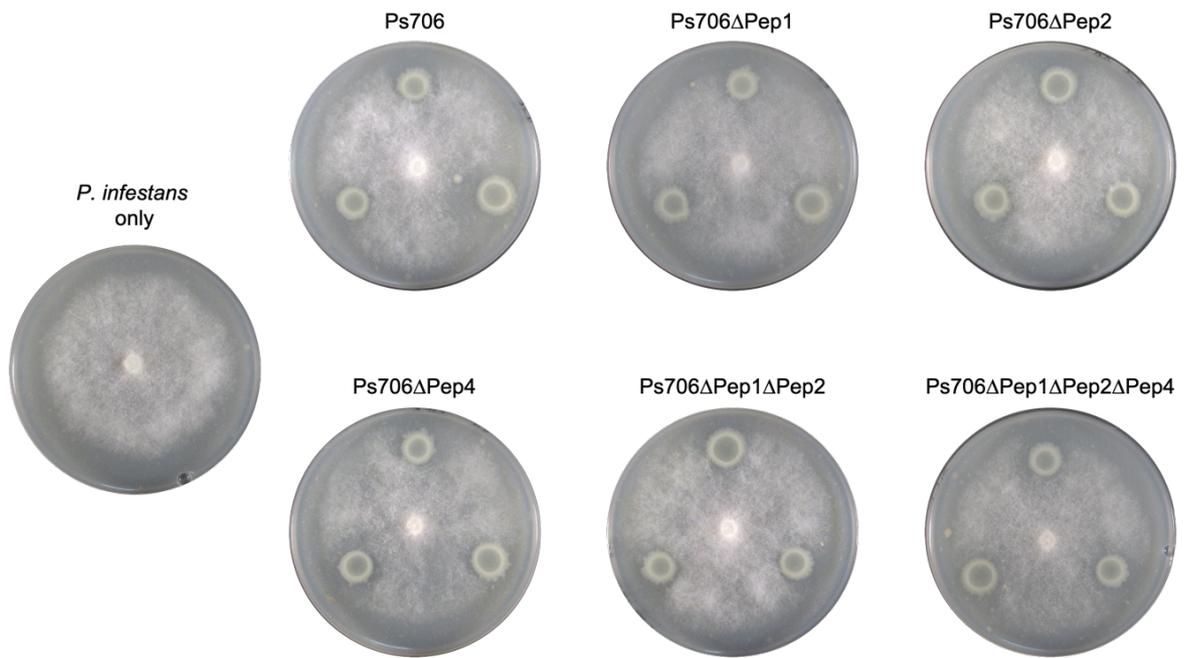


Figure 4.14. Bioassays of *Ps706* and *Pep* gene cluster mutants against *Phytophthora infestans*. No significant inhibitory activity was observed for the wild type *Ps706*, and there was no observable difference in this phenotype when comparing any of the mutant strains to the wild type. All images are taken at 10 days and representative of three biological replicates.

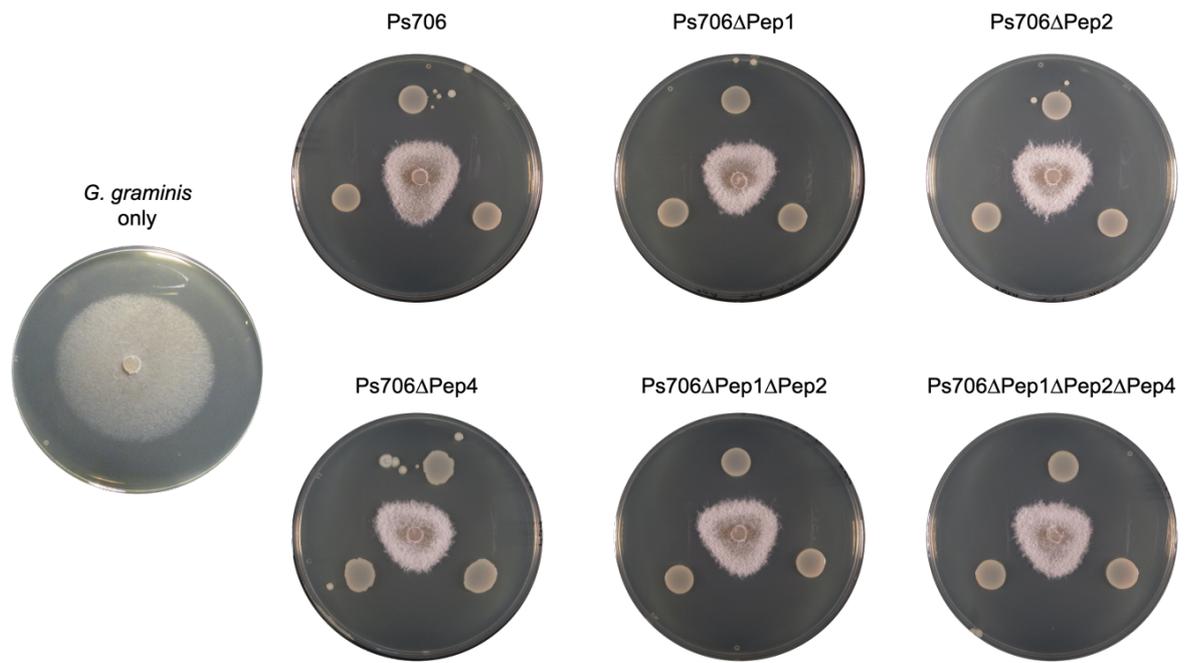


Figure 4.15. Bioassays of Ps706 and Pep gene cluster mutants against *Gaeumannomyces graminis* var. *tritici*. A visible zone of inhibition was observed for the wild type Ps706 when compared with the control (pathogen alone). Deletion of Pep gene clusters had no effect on this phenotype, with all mutants still able to produce the observed zone of inhibition. All images were taken at 10 days and are representative of three biological replicates.

4.4 Heterologous Expression of Pep1 Gene Cluster

4.4.1 Cloning the Pep1 Gene Cluster

Given the lack of phenotype observed for any of the Pep cluster deletion mutants, and the publication of two different biosynthetic gene clusters with DUF692 proteins involved in their pathways (380, 237), a heterologous expression approach, common to natural products discovery studies, was attempted. A minimal set of genes for the Pep1 biosynthetic gene cluster from Ps706 (Figure 4.16), were selected and cloned in to the pJH10TS expression vector, which contains a promoter for the constitutive expression of genes in *Pseudomonas*. This represented the conserved region of the gene cluster identified by MultiGeneBlast analysis in section 4.2.4, with the exclusion of the adjacent transporters and the NAD(P)/FAD-dependent oxidoreductase that was not predicted by antiSMASH analysis to be involved in biosynthesis. The transporters were excluded for technical reasons, as PCR amplification of such a large segment of DNA proved intractable. However, one of the chosen host strains expresses many transporters that allow export of xenobiotics, theoretically bypassing this issue (240). The resulting Pep1 expression vector, pJH10TS::Pep1, was expressed constitutively in *Pseudomonas* strains Ps723, KT2440, and EM383 for further characterisation.

4.4.2 Expression in Strain Ps723 (Bioassays)

Another environmental isolate from Stefanato *et al.* (324), Ps723, was chosen as a suitable host in which to attempt heterologous expression, as it is closely related to Ps706 and Ps708 (Figure 4.1) but does not show strong inhibition of the tested pathogens, and lacks any Pep BGCs. The Ps723 pJH10TS::Pep1 expression strain was assayed for activity against *Streptomyces scabies* and *G. graminis*, and compared to the wild type and empty vector controls. No increase in inhibitory activity against *S. scabies* was observed on SFM medium (Figure 4.17A), and a marginal increase in inhibition was seen on MYM for the empty vector control and the Pep1 expressing strain (Figure 4.17B); this activity was attributed to increased pyoverdine production based on later mass spectrometric analysis, and it can be clearly seen that the empty vector control exhibits a different growth morphology to the

wild type. Similarly, no increased inhibition against the fungus *G. graminis* was observed (Figure 4.17C). Activity against *P. infestans* was not tested at this time, as the organism was not available.

Pep1 gene cluster

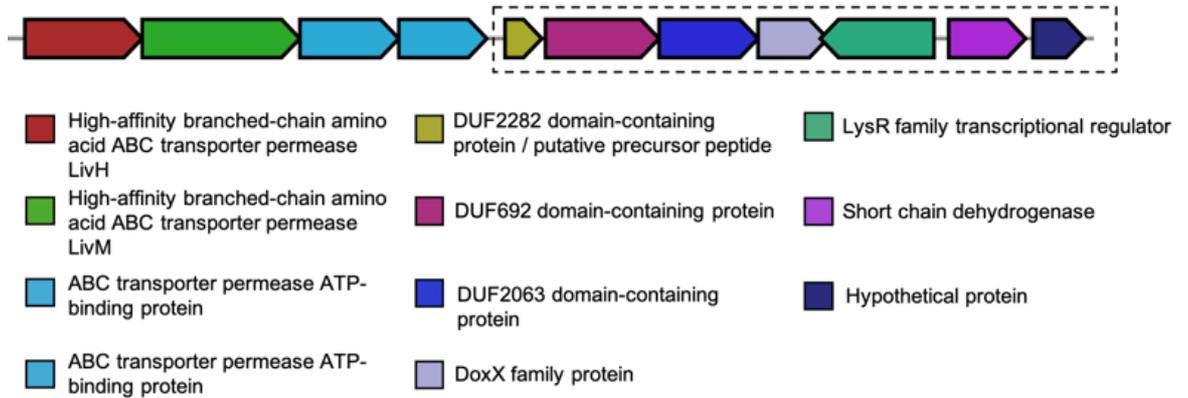


Figure 4.16. Schematic of the Pep1 gene cluster. Shown is the conserved set of genes for the Pep1 gene cluster with network 1 peptides (yellow) identified from the phylogenetics and RIPPER analysis. Coloured boxes indicate the annotations for each gene. Genes cloned in to the pJH10TS expression vector are shown in the hatched box.

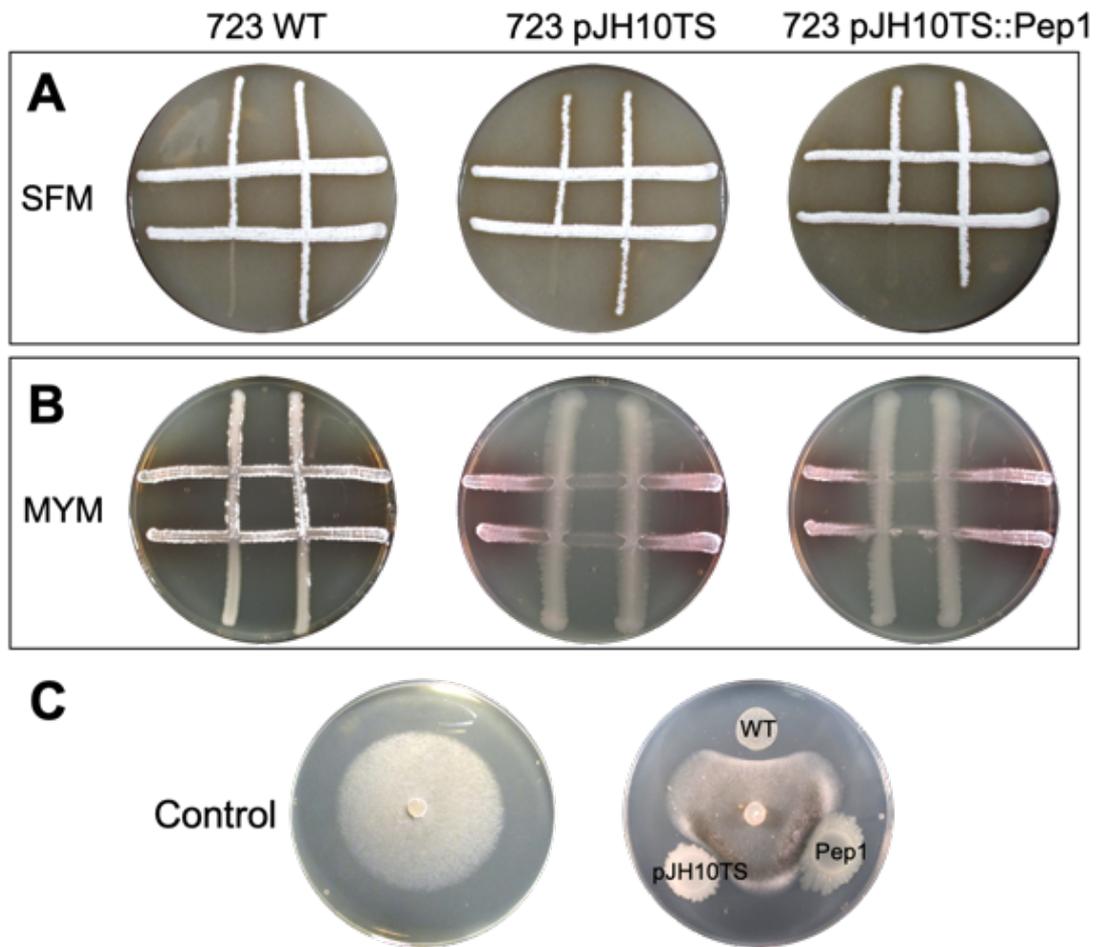


Figure 4.17. Effect of heterologous expression of Pep1 gene cluster in strain Ps723 on the inhibition of *Streptomyces scabies* 87-22 and *Gaeumannomyces graminis* var. *tritici*. (A) Wild type Ps723, Ps723 containing empty vector pJH10TS, and Ps723 containing pJH10TS::Pep1 cross-streaked against *S. scabies* 87-22 on SFM agar. (B) Wild type Ps723, Ps723 containing empty vector pJH10TS, and Ps723 containing pJH10TS::Pep1 cross-streaked against *S. scabies* 87-22 on MYM agar. (C) *G. graminis* var. *tritici* alone (left) and the three strains assayed against *G. graminis* var. *tritici* (right). All plates are representative of three biological replicates.

4.4.3 Expression in Strain Ps723 (Metabolomics)

Heterologous expression studies often employ metabolomics-based approaches to probe the comparative range of molecules produced by strains that are either expressing or not expressing the gene(s) of interest. By these means, it is possible to link molecules to genes, uncovering the biosynthetic product of a natural product gene cluster. Metabolomic analysis was therefore performed by liquid chromatography-mass spectrometry (LC-MS) to investigate any new metabolites. Ps723 pJH10TS and Ps723 pJH10TS::Pep1 were grown overnight in a range of media (Table 4.2), centrifuged to remove cells and the supernatant collected, and this supernatant analysed by LC-MS. Given the lack of information available on heterologous expression of natural product gene clusters in pseudomonads, media were chosen based on carbon/nitrogen sources, complexity, availability of reagents, and those that others had previous success with (N. Miguel-Vior, personal communication). Differences in metabolic profile were only visible in two media, King's B and YPD medium, with the change clearest in YPD medium. Chromatograms for all replicates expressing the Pep1 gene cluster showed a prominent peak at a retention time of ~3.5 minutes compared to the control (empty vector pJH10TS). This peak is significantly enriched in the presence of the gene cluster, and corresponds to a doubly charged ion hereby referred to as ' m/z 583.25(2)' (Figure 4.18). In order to confirm this phenotype was linked to the Pep1 gene cluster, the experiment was repeated in a different host strain. *Pseudomonas putida* KT2440 had been proposed as a suitable heterologous expression platform (390-392) due to numerous favourable properties that make it suitable for industrial and biotechnological approaches (240). This confirmed the same ion, m/z 583.25(2), was present in the KT2440 cultures expressing the Pep1 cluster, showing its presence is linked to the expression of this cluster. However, LC-MS chromatograms of crude supernatant were quite complex, as KT2440, and likely the environmental strain Ps723, produce many more molecules besides those potentially encoded by the Pep1 gene cluster (240). Further work would use a KT2440 derivative strain optimised for heterologous expression, providing increased production and a cleaner background.

Table 4.2. Media used for screening of heterologous expression of Pep1 gene cluster.

Medium	New metabolite(s) observed (Y/N)
King's B	Y
LB	N
Nutrient Broth	N
Nutrient Broth + Glucose	N
M9 + Cas Amino Acids	N
Obafluorin Production Medium	N
YPD	Y

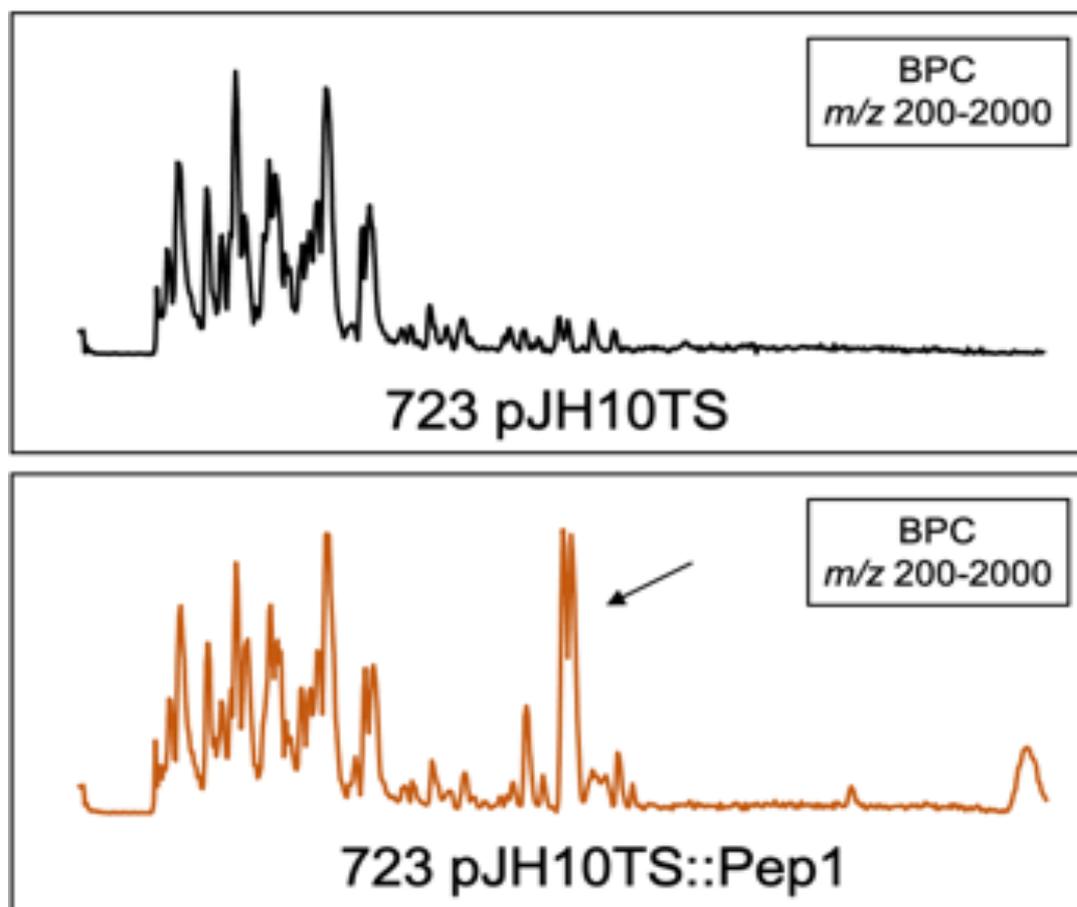


Figure 4.18. Chromatograms from LC-MS of Ps723 heterologous expression strains in YPD medium. Compared to the plasmid only negative control (top), the Ps723 strain expressing the Pep1 biosynthetic gene cluster shows a very prominent double peak, corresponding to a m/z of 583.25(2) (arrow). Chromatograms are representative of three biological replicates. BPC, base peak chromatogram.

4.4.4 Heterologous Expression in Strain EM383

4.4.4.1 Overview of Strain EM383 as a Heterologous Host

Pseudomonas putida strain EM383 is a highly genome edited derivative of the original KT2440 isolate, produced by deleting 300 genes in 11 non-adjacent regions of the KT2440 genome, amounting to 4.3% of the entire genome (240). These modifications, which include elimination of transposons, prophages, DNA restriction-modification systems, the flagellar system, and metabolic pathways perceived as a drain on cell energy and reducing-agent supply were demonstrated by Martinez-Garcia *et al.* to increase a number of characteristics relevant to heterologous expression platforms. EM383 was demonstrated to have a shorter lag time, and increased biomass production under specific circumstances, as well as improved tolerance of various stressors. Combined, all these parameters made it likely that the streamlined EM383 would enhance production of Pep1 cluster-related molecules, as well as a potentially cleaner background for LC-MS analyses.

4.4.4.2 Expression of Pep1 in YPD Medium (Metabolomics)

The metabolomic analysis described in section 4.4.3 was repeated with strain EM383 in YPD media alone. Given the cleaner background of this streamlined strain, it was possible to observe the m/z 583.25(2) ion in negative control samples, albeit at a much lower level than in samples expressing the Pep1 biosynthetic gene cluster (Figure 4.19A). This data indicated that the observed metabolite represented by the ion m/z 583.25(2) is not a true product of the Pep1 biosynthetic gene cluster. The most abundant ionic species in the data, this molecule could represent either a stress response molecule produced as a consequence of presence of the cluster's true product, or an endogenous product that may be overproduced as a result of promiscuous activity of unidentified enzymes in the Pep1 cluster.

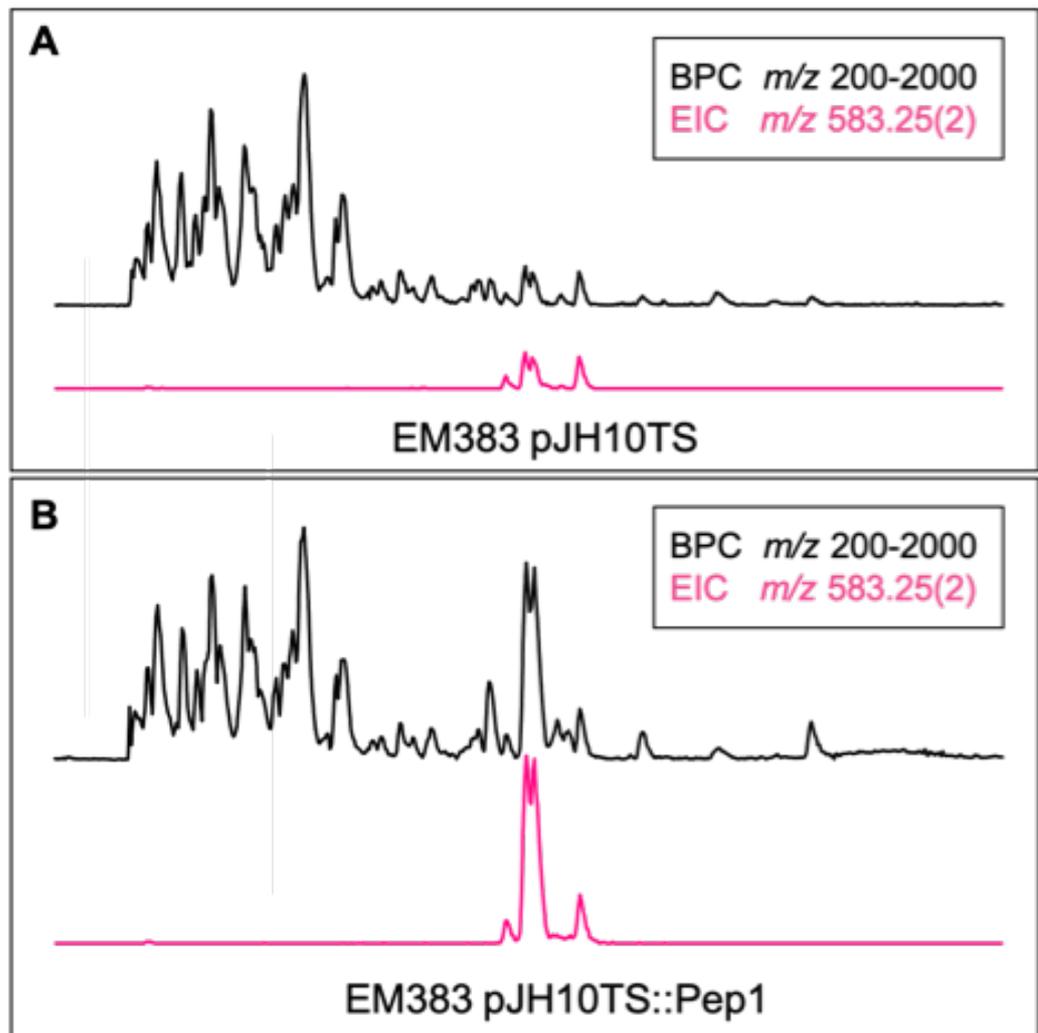


Figure 4.19. Expression of Pep1 gene cluster in *Pseudomonas putida* EM383. (A) EM383 containing an empty vector control, with the previously identified ion m/z 583.25(2) peak visible in pink. (B) EM383 containing pJH10TS::Pep1 vector, with the m/z 583.25(2) ion peak visible in pink. BPC, base peak chromatogram; EIC, extracted ion chromatogram.

4.4.4.3 Cloning of Pep1 Δ DUF2282

In order to investigate whether *m/z* 583.25(2) was a stress molecule arising from toxicity of the cluster's true product, or a result of promiscuous activity of putative tailoring enzymes in the Pep1 biosynthetic gene cluster, the gene cluster was cloned without the putative precursor peptide (DUF2282) by PCR amplification starting downstream of the DUF2282 gene. The resulting cluster was named Pep1 Δ 2282 and cloned in to the pJH10TS expression vector as previously for the entire Pep1 cluster (Figure 4.20) to create the expression construct pJH10TS::Pep1 Δ 2282. The DUF2282 protein was also cloned separately into a gentamycin-selectable derivative of the pME6032 plasmid (393), pGME6032, to create the expression vector pGME6032::DUF2282. This plasmid is intended to be used as an inducible expression vector in Gram-negative plant-associated bacteria (343), in this case allowing for control of the expression of the putative precursor peptide.

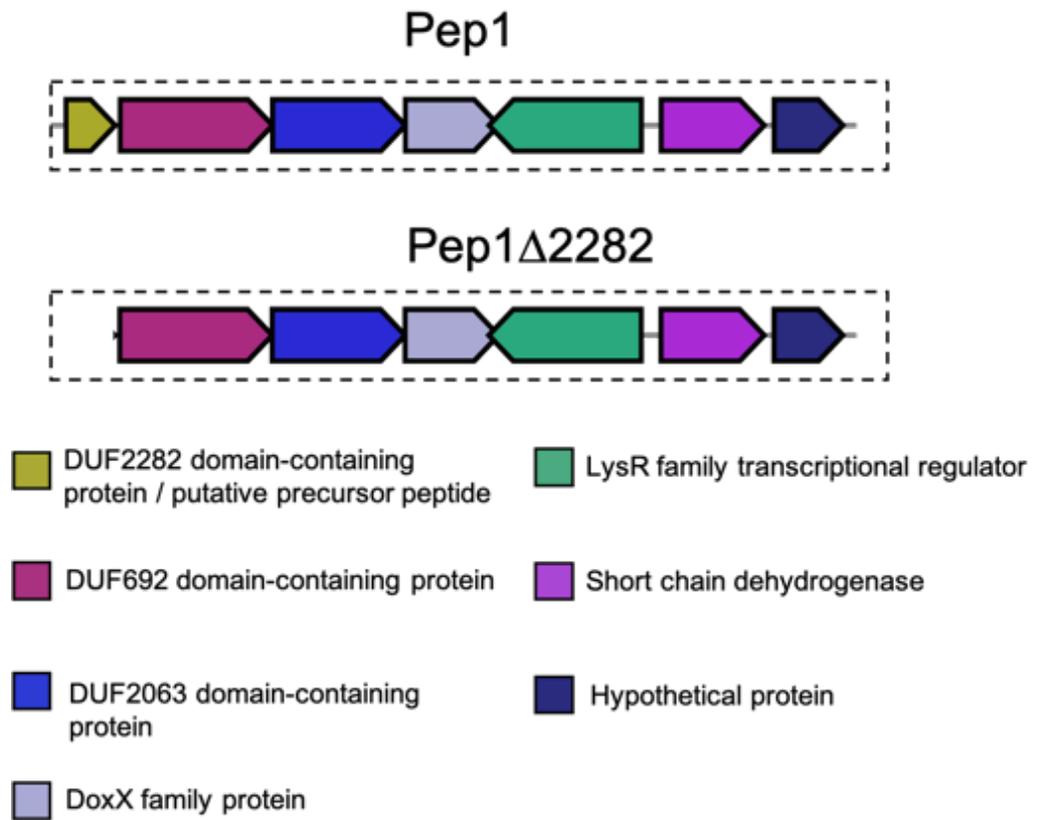


Figure 4.20. Schematic of Pep1 and Pep1 Δ 2282 gene clusters heterologously expressed in *Pseudomonas putida* EM383. Shown are the genes included in each cluster, with a key indicating the hypothesised role of each gene from the antiSMASH output.

4.4.4.4 Expression of Pep1 Δ DUF2282 in *Pseudomonas putida* EM383

Expression of Pep1 Δ 2282 and DUF2282 were performed in *P. putida* EM383 as for the whole Pep1 cluster described above. Firstly, Expression of the DUF2282 protein alone in pGME6032 did not increase the peak area for m/z 583.25(2), indicating the unmodified precursor peptide alone was not responsible for this peak (Figure 4.21A). This result is not surprising, as an unmodified precursor peptide from a RiPP cluster would likely be degraded back to its constituent amino acids in the absence of any modifications to protect it from proteolysis. However, this also suggests this short peptide is unlikely to play a regulatory role in the cell, as no other significant metabolic changes were observed (Figure 4.21B). This result remained unchanged with or without induction of DUF2282 in pGME6032. Next, expression of the Pep1 Δ 2282 cluster displayed a peak height and area for m/z 583.25(2) somewhere intermediate between the empty vector control and the complete Pep1 gene cluster (Figure 4.21). Complementation with pGME6032::DUF2282 failed to restore levels of m/z 583.25(2) to those observed in Pep1 expression cultures. Curiously, without induction of the DUF2282 protein, the peak area of m/z 583.25(2) returned to that of the empty vector control. These results failed to conform to either hypothesis postulated for the increase in m/z 583.25(2) with the expression of the Pep1 gene cluster.

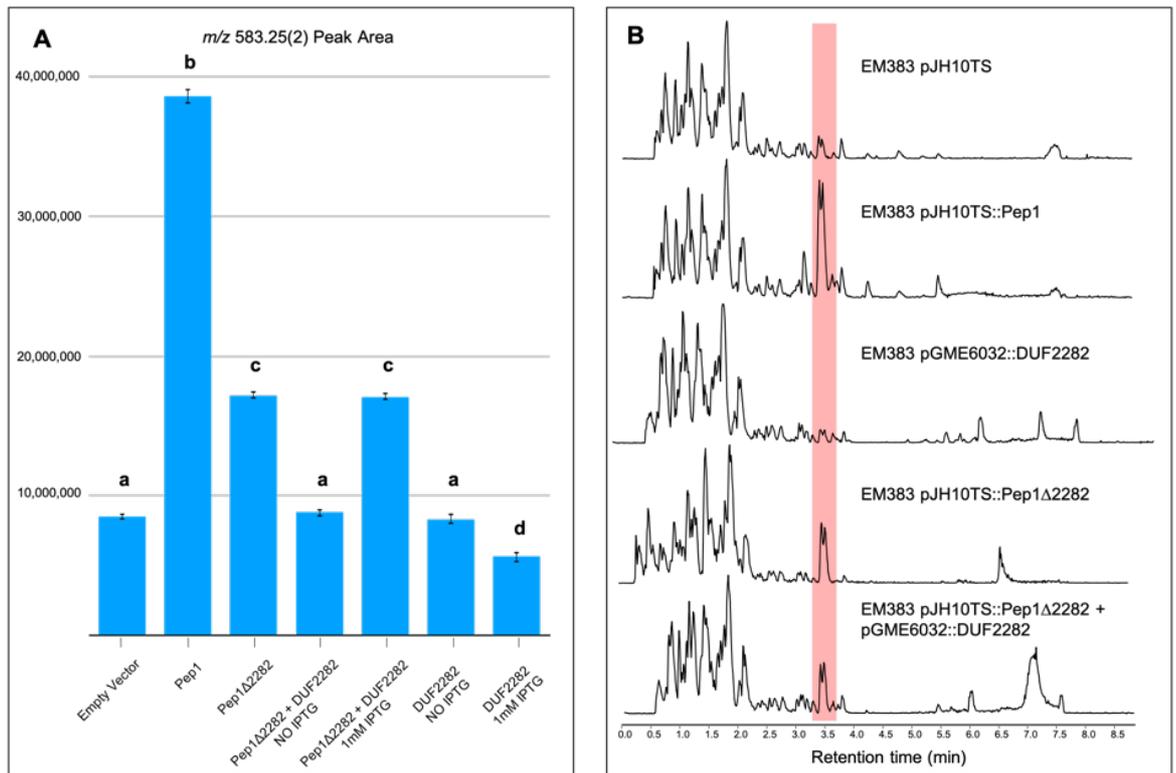


Figure 4.21. Profiles of *m/z* 583.25 in *Pseudomonas putida* EM383 heterologous expression cultures. All experiments were performed in YPD medium. (A) Peak area calculated in Shimadzu LabSolutions software Browser, for all EM383 heterologous expression conditions. Letters denote statistically significant differences at $p = 0.05$, where bars that do not share a letter are statistically significantly different. (B) LC-MS chromatograms for selected EM383 heterologous expression conditions. The profile of *m/z* 583.25(2) (red box) peak height did not correspond to that expected by either of the postulated hypotheses. All data are representative of a minimum three biological replicates.

4.4.4.5 XCMS Analysis

As the expression profile of m/z 583.25(2) proved difficult to interpret under the tested conditions, a more unbiased approach to attempt to uncover the true product of the Pep1 cluster was considered. For a higher throughput unbiased analysis of the metabolomic data, spectra were processed using the XCMS online portal (394) (see methods). Pairwise analyses were performed for empty vector (pJH10TS) versus Pep1 expression strains, and for complemented strains expressing both Pep1 Δ 2282 and DUF2282 without IPTG induction versus with IPTG induction for DUF2282 expression. 13 metabolites met these criteria in the former case (Figure 4.22A), and 34 in the latter (Figure 4.22B). Common to both was a family of molecules between m/z 1165-1169. Only 1165.49 was observed in induced vs uninduced controls, upregulated 31-fold ($p < 0.01$). In Pep1 expression strains vs empty vector controls, a family of metabolites was observed, ranging from m/z 1165.55 (25-fold) to 1167.49 (116-fold). However, it must be noted that maximum intensity was still low at \sim 340,000 ion counts. Given the mass and retention times, these likely represent the singly charged counterparts of m/z 583.25(2). Their identity is still unknown, but their masses match closely with the pyoverdines reported for many pseudomonads (331), and UV absorbance at 405nm was observed consistent with pyoverdine (Figure 4.23). This finding lends weight to the suggestion that expression of these clusters is eliciting a stress response, matching results from expression of the Pep1 cluster in Ps723 in Section 4.4.2. Under the conditions tested it has not been possible to identify a novel natural product encoded by the Pep1 BGC using heterologous expression and metabolomic analyses.

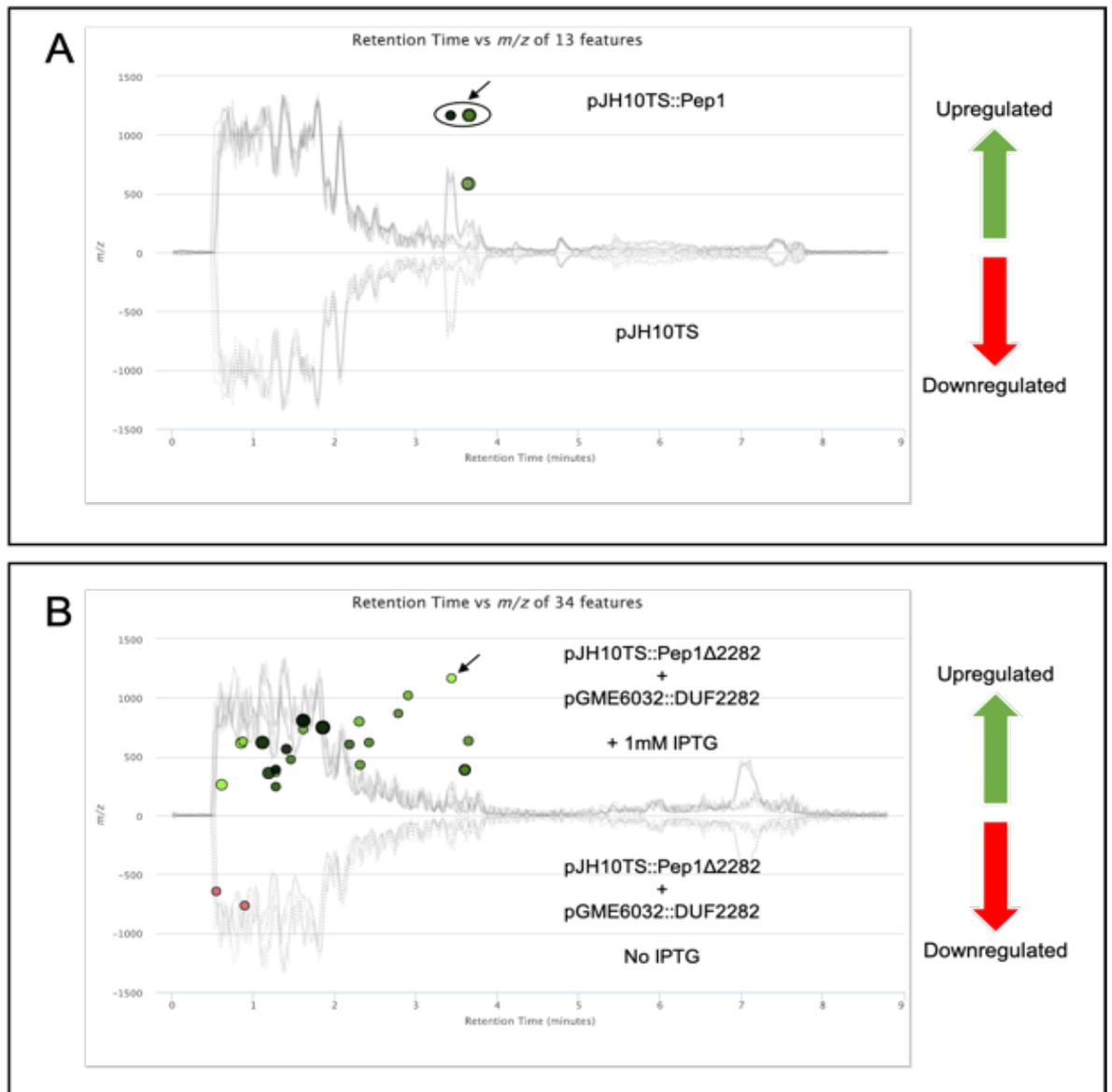


Figure 4.22. XCMS Analysis of Pep1 Expression Cultures. Presented are annotated Metabolomic Cloud Plots from XCMS Online with total ion chromatograms overlaid (grey). **(A)** Comparison of cultures expressing Pep1 constitutively, versus an empty vector control. **(B)** Comparison of uninduced cultures of a strain expressing the Pep1 cluster with the predicted precursor peptide on an inducible expression vector versus equivalent cultures induced with 1mM IPTG. A family of molecules around 1165 Daltons was common to both analyses (arrows). Ions upregulated in expression cultures versus their negative controls are graphed above the x-axis, while those downregulated versus negative controls are graphed below the x-axis. Darkness of spots inversely correlates with p-value.

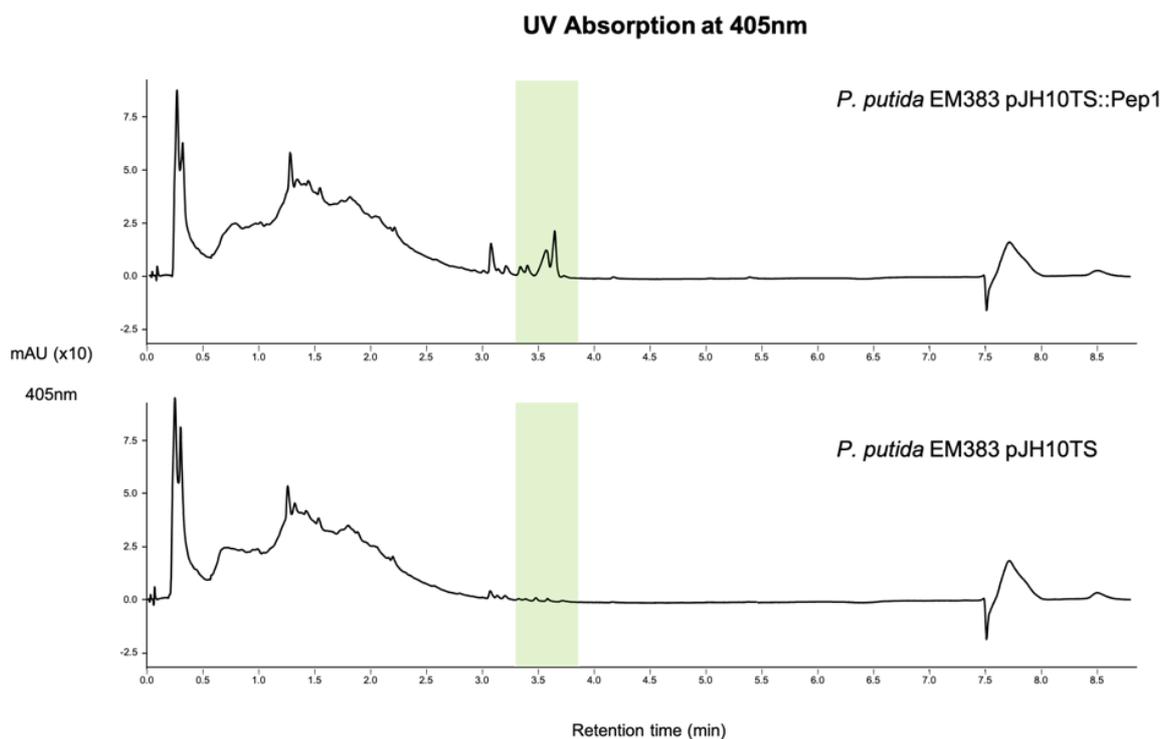


Figure 4.23. UV Absorption spectrum at 405nm for Pep1 expression cultures. UV absorbance is observed for *P. putida* EM383 expressing the Pep1 cluster but not the empty vector control. The retention time of these peaks is consistent with the observed ions in mass spectrometry. Chromatograms are representative of three biological replicates.

4.5 Discussion & Conclusions

4.5.1 Summary / Discussion of Results

Based on initial assessment of the investigated clusters, it appeared likely that they represented biosynthetic gene clusters encoding for the production of RiPP natural products that were involved in the inhibition of the potato pathogen *Streptomyces scabies*. While previous work has linked a subset of the genes found in the clusters to tolerance of various environmental stresses, biochemical evidence for how they might do this is lacking at present (383, 384, 389, 382, 395). The oxidative stress assays presented here suggested, at least under the conditions tested, that the clusters do not play a role in oxidative stress tolerance. Heavy metal stress tolerance was not tested directly, which may have produced a different result. However, two well-characterised examples of DUF692 proteins being involved in RiPP biosynthetic pathways exist (380, 237). Combined with the presented bioinformatic assessment of precursor peptide sequence conservation, their co-evolution with DUF692 proteins, and conservation of surrounding genetic architecture, it seemed plausible that these clusters encoded the necessary machinery to produce as-yet-unknown RiPP natural products. The biological logic and computational steps involved in the RiPPER genome mining workflow had previously been validated experimentally and resulted in the discovery of a novel family of RiPP natural products, the thiovarsolins (171). In that example, the authors searched for proteins homologous to enzymes involved in installing thioamide bonds in RiPP precursor peptide backbones, and investigated a cluster from *Streptomyces varsoviensis* that displayed a novel gene architecture and precursor peptide, which was conserved across a range of strains. Unfortunately, attempts here to replicate that success with the DUF692 protein in the set of environmental *Pseudomonas* isolates have proven unsuccessful, as it was not possible to detect novel RiPP natural products despite observing marked metabolomic changes.

Initial analysis conducted by Dr. A Truman on data acquired by Dr. F Stefanato had suggested that these clusters might be linked to the ability of environmental isolates from a commercial potato field to inhibit potato pathogens. However, it is possible that a number of confounding factors made this appear more promising than it might otherwise

have been. Firstly, linking genes or gene clusters to a given phenotype, like genome-wide-association-studies (GWAS) frequently used in human biology (396), is notoriously tricky in bacteria (397). The clonality of bacterial lineages mean that adjustments have to be made in statistical calculations, and modifications have to be made to account for the presence or absence entirely of certain genes due to recombination, which is not a considerable issue in human GWAS studies that typically rely on single nucleotide polymorphisms (398). These modifications to the methodology are possible, and increase the odds of finding true associations (399). It is possible that the lineage effects that confound bacterial GWAS studies, along with the large accessory genomes of *Pseudomonas* strains (335), led to false positive associations of Pep clusters and inhibition of the chosen pathogens. Visually, inhibition of *S. scabies* does appear to be restricted mainly to particular clades in the phylogenetic tree (Figure 4.1), suggesting possessing Pep BGCs and suppressing phytopathogens may simply be co-occurring traits. This perspective is also supported by the absence of observable effect when deleting these gene clusters individually or in combination, despite conducting the assays on the same growth medium as for the original phenotype assessment.

Comparative transcriptomics in *Salinispora* has shown us that it can be difficult or impossible to predict whether a given biosynthetic gene cluster will be active and produce the final molecule just by looking at the gene clusters, with inactivation of regulatory elements being a key issue in otherwise identical clusters (400). In the Pep BGCs discussed here, given that no similar cluster is currently characterised, there is theoretically no reliable means of predicting whether a particular strain will produce a hypothesised natural product under a given set of circumstances, or whether the cluster is functional. It is plausible that a only subset of strains will contain functional clusters, confounding the analysis. Following that logic, it is possible that the Pep clusters do represent genuine RiPP BGCs, but only a single example was used in heterologous expression work, and this example may not have been capable of producing the encoded natural product. However, the observation of a phenotype in heterologous expression work, while not what was initially expected, suggests the cluster was expressed and functional. In hindsight, including the NAD(P)/FAD-dependent oxidoreductase in the expressed set of genes may have resulted in a detectable final product, if somehow involved in installing post-translational

modifications that protect against proteolysis of the precursor peptide. A number of other approaches could have been attempted to fully explore whether the clusters encode a novel RiPP or RiPP-like natural product.

Firstly, a reporter fusion, RT-PCR, or RNA-Seq experiment may have proven useful in confirming that the cluster was being actively expressed in the heterologous hosts, or under the conditions in which the original phenotype data was acquired. By ensuring that the gene clusters were expressed in all cases where phytopathogen inhibition was observed, the correlation analysis would gain robustness. However, as mentioned above, the observation of a phenotype in heterologous expression work suggests the clusters were expressed successfully.

Secondly, a larger number of growth media might have been trialled for production of the compound in heterologous hosts, as well as attempting production in a pyoverdine-null background. A pyoverdine deletion of KT2440 derivative EM42 was generated, and it was hoped this might prove a cleaner background in which to have investigated the metabolome of the expression strains. It is possible that by preventing production of the siderophore pyoverdine, other metabolic changes would be more evident. However time constraints precluded heterologous expression work in this strain.

Thirdly, a wider range of representative Pep1 gene clusters could have been cloned and expressed, to account for possible differences affecting production of the compound, as documented in the *Salinispora* example above. However, this is less likely to offer a significant advantage compared to other options, as deletion of the cluster in a second strain, Ps708, exhibited the same phenotype as Ps706 from which the cluster was cloned for heterologous expression work. Alternatively, a genetic refactoring approach might have helped solve issues of relative expression levels, ensuring optimal expression of each gene in the cluster, which proved confusing in the attempted complementation of DUF2282 in Pep1 Δ 2282 expression strains.

Finally, it is worth considering alternative roles for the Pep gene clusters, beyond antimicrobial RiPP biosynthesis. While the bioinformatic analysis using RiPPER and other

tools were strongly suggestive of the Pep gene clusters representing RiPP BGCs, the observed stress response in heterologous expression work may indicate a regulatory role of the clusters. While expression of the short peptide alone did not appear to elicit a significant response, additional genes in the cluster may be required. However, the identified signal peptidase cleavage site does suggest an extracellular role. A possible explanation linking a RiPP natural product, heavy metal stress tolerance, and the extracellular role, is that the product is a secreted modified peptide capable of binding heavy metals. This may allow biocontrol strains possessing the cluster to grow in agricultural soils under significant pathogen pressure that are treated with copper-based strategies for prevention of crop diseases (401). This could explain the co-occurrence of these clusters in strains that are also capable of inhibiting various plant-pathogens.

4.5.2 Conclusions and Future Perspectives

The BGCs identified by antiSMASH (377) do not appear, contrary to the initial hypothesis, to be responsible for inhibition of the tested plant pathogens. While in some bacteria the presence of a subset of these genes appears linked to heavy metal or oxidative stress tolerance, our data did not support this conclusion, at least for oxidative stress. However, bioinformatic analysis was highly indicative of a conserved BGC, with co-evolving short peptides that may constitute RiPP precursor peptides. This supports the hypothesis that these clusters encode RiPPs. Unfortunately, attempts to characterise the products of these clusters by heterologous expression were unsuccessful. A number of the strategies discussed above might be employed to uncover the natural products encoded by these clusters, with heterologous expression in a pyoverdine-null background being one of the more promising options, with a reporter fusion to confirm that the clusters are being expressed. A wider range of media should also be tested for fermentations. Finally, the factors governing phytopathogen inhibition by Ps706 and Ps708 are unknown, and warrant further investigation given the lack of characterised BGCs identified in these strains.

Chapter 5: Materials and Methods

5.1 Materials

5.1.1 Strains

All bacterial, fungal, and oomycete strains were grown in the media indicated in the relevant points in the text and section 5.2, with selection where necessary as indicated in section 5.1.3.2. A full list of strains is provided in Table 5.1.

Table 5.1. Strains used in this study.

Strain	Genotype / description	Application	Reference
Ps706	Environmental isolate	Environmental biocontrol isolate	(324)
Ps706 Δ Pep1	Δ Pep1	RiPP-like gene cluster investigation	F Stefanato
Ps706 Δ Pep2	Δ Pep2	RiPP-like gene cluster investigation	F Stefanato
Ps706 Δ Pep1 Δ Pep2	Δ Pep1 Δ Pep2	RiPP-like gene cluster investigation	F Stefanato
Ps706 Δ Pep4	Δ Pep4	RiPP-like gene cluster investigation	This work
Ps706 Δ Pep1 Δ Pep2 Δ Pep4	Δ Pep1 Δ Pep2 Δ Pep4	RiPP-like gene cluster investigation	This work
Ps708	Environmental isolate	Environmental biocontrol isolate	(324)
Ps708 Δ Pep1	Δ Pep1	RiPP-like gene cluster investigation	F Stefanato
Ps708 Δ Pep2	Δ Pep2	RiPP-like gene cluster investigation	F Stefanato
Ps708 Δ Pep1 Δ Pep2	Δ Pep1 Δ Pep2	RiPP-like gene cluster investigation	F Stefanato

Strain	Genotype / description	Application	Reference
Ps708 Δ Pep4	Δ Pep4	RiPP-like gene cluster investigation	This work
Ps708 Δ Pep1 Δ Pep2 Δ Pep4	Δ Pep1 Δ Pep2 Δ Pep4	RiPP-like gene cluster investigation	This work
Ps723	Environmental isolate	Host for Pep1 heterologous expression	(324)
Ps652	Environmental isolate	Environmental biocontrol isolate	(324)
Ps652 Δ HCN	Δ <i>hcnABC</i>	Investigation of Ps652 bioactivity	This work
Ps652 Δ Pyo	Pyoverdine null mutant	Investigation of Ps652 bioactivity	This work
Ps652 Δ NRPS5	Δ NRPS5	Investigation of Ps652 bioactivity	This work
4A5	Ps652 Δ <i>panB</i> transposon insertion mutant	Ps652 transposon mutant	This work
4B3	Ps652 Δ <i>ilvH</i> transposon insertion mutant	Ps652 transposon mutant	This work
10B3	Ps652 Δ <i>leuB</i> transposon insertion mutant	Ps652 transposon mutant	This work
14B12	Ps652 Δ <i>argB</i> transposon insertion mutant	Ps652 transposon mutant	This work
14G7	Ps652 Δ <i>acdA1</i> transposon insertion mutant	Ps652 transposon mutant	This work
15C7	Ps652 Δ <i>trpA</i> transposon insertion mutant	Ps652 transposon mutant	This work
17A1	Ps652 Δ <i>leuB</i> transposon insertion mutant	Ps652 transposon mutant	This work
17B4	Ps652 Δ <i>ubiE</i> transposon insertion mutant	Ps652 transposon mutant	This work
19C1	Ps652 Δ <i>recA</i> transposon insertion mutant	Ps652 transposon mutant	This work

Strain	Genotype / description	Application	Reference
20A6	Ps652 Δ <i>bioC</i> transposon insertion mutant	Ps652 transposon mutant	This work
20C5	Ps652 Δ <i>hisF</i> transposon insertion mutant	Ps652 transposon mutant	This work
21B6	Ps652 Δ <i>mgtC/sapB</i> family protein transposon insertion mutant	Ps652 transposon mutant	This work
21E5	Ps652 Δ <i>accB</i> transposon insertion mutant	Ps652 transposon mutant	This work
21F1	Ps652 Δ <i>lcfB</i> transposon insertion mutant	Ps652 transposon mutant	This work
21G8	Ps652 Δ <i>trpA</i> transposon insertion mutant	Ps652 transposon mutant	This work
23D8	Ps652 Δ <i>argB</i> transposon insertion mutant	Ps652 transposon mutant	This work
24D3	Ps652 Δ <i>metF</i> transposon insertion mutant	Ps652 transposon mutant	This work
24D9	Ps652 Δ Acyl-Coa thioesterase transposon insertion mutant	Ps652 transposon mutant	This work
25D12	Ps652 Δ <i>trpB</i> transposon insertion mutant	Ps652 transposon mutant	This work
26H7	Ps652 Δ <i>dnaJ</i> transposon insertion mutant	Ps652 transposon mutant	This work
28E3	Ps652 Δ <i>rffh</i> transposon insertion mutant	Ps652 transposon mutant	This work
29C1	Ps652 Δ <i>oprM1</i> transposon insertion mutant	Ps652 transposon mutant	This work
29D2	Ps652 Δ <i>acdA1</i> transposon insertion mutant	Ps652 transposon mutant	This work
Ps652 Δ Paak	Δ <i>paak</i> phenylacetate-CoA ligase mutant	Investigation of Ps652 gene cluster	This work
Ps652 Δ HCN Δ Paak	Δ <i>hcnABC</i> Δ <i>paak</i>	Investigation of Ps652 gene cluster	This work

Strain	Genotype / description	Application	Reference
Ps652ΔHCNΔTE	<i>ΔhcnABC Δacyl-CoA thioesterase</i>	Investigation of Ps652 gene cluster	This work
Ps652ΔHCNΔEchA8_1	<i>ΔhcnABC ΔechA8_1</i> enoyl-CoA hydratase mutant	Investigation of Ps652 gene cluster	This work
Ps652ΔHCNΔEchA8_1 ΔEchA8_2	<i>ΔhcnABC ΔechA8_1 ΔechA8_2</i> enoyl-CoA hydratase double mutant	Investigation of Ps652 gene cluster	This work
Ps652ΔHCNΔipdC	<i>ΔhcnABC ΔipdC</i> indole pyruvate decarboxylase mutant	Investigation of Ps652 gene cluster	This work
Ps652ΔHCNΔTE::pME6032	Empty vector control for Acyl-Coa thioesterase complementation	Investigation of Ps652 gene cluster	This work
Ps652ΔHCNΔTE::TE	<i>ΔhcnABCΔacyl-CoA thioesterase complemented with Acyl-CoA thioesterase in multicopy expression vector pME6032</i>	Investigation of Ps652 gene cluster	This work
Ps652ΔHCNΔacdA	<i>ΔhcnABC ΔacdA</i> acyl-CoA dehydrogenase mutant	Investigation of Ps652 gene cluster	This work
<i>Escherichia coli</i> DH5α	<i>fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	Plasmid host for molecular cloning	New England Biolabs
<i>Escherichia coli</i> DH5α pTS1	Carrying pTS1 plasmid	Maintenance of pTS1 plasmid	This work
<i>Escherichia coli</i> DH5α pTS1::ΔPep4	DH5α carrying pTS1::ΔPep4 vector	Maintenance of pTS1::ΔPep4 plasmid	This work
<i>Escherichia coli</i> DH5α pTS1::ΔHCN	DH5α carrying pTS1::ΔHCN vector	Maintenance of pTS1::HCN plasmid	This work
<i>Escherichia coli</i> DH5α pTS1::Δ652Pyo	DH5α carrying pTS1::Δ652Pyo vector	Maintenance of pTS1::Δ652Pyo plasmid	This work

Strain	Genotype / description	Application	Reference
<i>Escherichia coli</i> DH5 α pTS1:: Δ EM42Pyo	DH5 α carrying pTS1:: Δ EM42Pyo vector	Maintenance of pTS1:: Δ EM42Pyo plasmid	This work
<i>Escherichia coli</i> DH5 α pTS1:: Δ NRPS5	DH5 α carrying pTS1:: Δ NRPS5 vector	Maintenance of pTS1::NRPS5 plasmid	This work
<i>Escherichia coli</i> DH5 α pTS1:: Δ paaK	DH5 α carrying pTS1:: Δ paaK vector	Maintenance of pTS1:: Δ paaK plasmid	This work
<i>Escherichia coli</i> DH5 α pTS1:: Δ TE	DH5 α carrying pTS1:: Δ TE vector	Maintenance of pTS1::TE plasmid	This work
<i>Escherichia coli</i> DH5 α pTS1:: Δ EchA8_1	DH5 α carrying pTS1:: Δ EchA8_1 vector	Maintenance of pTS1:: Δ EchA8_1 plasmid	This work
<i>Escherichia coli</i> DH5 α pTS1:: Δ EchA8_2	DH5 α carrying pTS1:: Δ EchA8_2 vector	Maintenance of pTS1:: Δ EchA8_2 plasmid	This work
<i>Escherichia coli</i> DH5 α pTS1:: Δ ipdC	DH5 α carrying pTS1:: Δ ipdC vector	Maintenance of pTS1:: Δ ipdC plasmid	This work
<i>Escherichia coli</i> DH5 α pTS1:: Δ acdA	DH5 α carrying pTS1:: Δ acdA vector	Maintenance of pTS1:: Δ acdA plasmid	This work
<i>Escherichia coli</i> DH5 α pJH10TS	DH5 α carrying pJH10TS plasmid	Maintenance of pJH10TS plasmid	This work
<i>Escherichia coli</i> DH5 α pJH10TS::Pep1	DH5 α carrying pJH10TS::Pep1 vector	Maintenance of pJH10TS::Pep1 plasmid	This work
<i>Escherichia coli</i> DH5 α pJH10TS::Pep1 Δ 2282	DH5 α carrying pJH10TS::Pep1 Δ 2282 vector	Maintenance of pJH10TS::Pep1 Δ 2282 plasmid	This work
<i>Escherichia coli</i> DH5 α pGME6032	DH5 α carrying pGME6032 plasmid	Maintenance of pGME6032 plasmid	This work
<i>Escherichia coli</i> DH5 α pGME6032::DUF2282	DH5 α carrying pGME6032::DUF2282 vector	Maintenance of pGME6032::DUF2282 plasmid	This work

Strain	Genotype / description	Application	Reference
<i>Escherichia coli</i> DH5 α pME6032	DH5 α carrying pME6032 plasmid	Maintenance of pME6032 plasmid	This work
<i>Escherichia coli</i> DH5 α pUC18::TE	DH5 α carrying pUC18::TE vector	Maintenance of pUC18::TE vector for subcloning of thioesterase	This work
<i>Escherichia coli</i> DH5 α pME6032::TE	DH5 α carrying pME6032::TE vector	Maintenance of pME6032::TE vector	This work
<i>Pseudomonas putida</i> EM42	Genome reduced <i>P. putida</i> KT2440 derivative for heterologous expression	Heterologous expression of <i>Pseudomonas</i> gene clusters	(240)
<i>Pseudomonas putida</i> EM42 Δ Pyo	Pyoverdine null mutant of EM42	Heterologous expression in Δ Pyo background	This work
<i>Pseudomonas putida</i> EM383	<i>recA</i> mutant of EM42	Heterologous expression of <i>Pseudomonas</i> gene clusters	(240)
<i>Pseudomonas putida</i> EM383 pJH10TS	EM383 with empty pJH10TS vector	Empty vector control for heterologous expression	This work
<i>Pseudomonas putida</i> EM383 pJH10TS::Pep1	EM383 with pJH10TS::Pep1 vector	Heterologous expression of Pep1 gene cluster	This work
<i>Pseudomonas putida</i> EM383 pJH10TS::Pep1 Δ 2282	EM383 with pJH10TS::Pep1 Δ 2282 vector	Heterologous expression of Pep1 gene cluster without DUF2282 protein	This work
<i>Pseudomonas putida</i> EM383 pGME6032	EM383 with empty pGME6032 vector	Empty vector control for heterologous expression	This work

Strain	Genotype / description	Application	Reference
<i>Pseudomonas putida</i> EM383 pGME6032::DUF2282	EM383 with pGME6032::DUF2282 vector	Heterologous expression of DUF2282 protein (inducible)	This work
<i>Pseudomonas putida</i> EM383 pJH10TS::Pep1Δ2282 pGME6032::DUF2282	EM383 with pJH10TS::Pep1Δ2282and pGME6032::DUF2282vectors	Heterologous expression of complemented Pep1 gene cluster (inducible)	This work
<i>Streptomyces scabies</i> 87-22	Wild type	Test bacterial potato pathogen for biological control assays	(293)
<i>Streptomyces coelicolor</i> M145	<i>Streptomyces coelicolor</i> A3(2) wild type without SCP1 and SCP2 plasmids	Automated screening of transposon mutants	(402)
<i>Gaeumannomyces graminis</i> var. <i>tritici</i>	Wild type	Test fungus for biological control assays	J. Malone (JIC)
<i>Phytophthora infestans</i> strain 6_A1	Wild type	Test oomycete potato pathogen for original biological control assays for Ps652, Ps706, Ps708	J. Jones (TSL)
<i>Phytophthora infestans</i> strain 88069	Wild type	Test oomycete potato pathogen for Ps652 cluster mutants and tropolone assays	(403)

5.1.2 Plasmids

All plasmid used in this work are described in Table 5.2 for the purposes described therein.

Table 5.2. Plasmids and vectors used in this study

Plasmid	Application	Reference
pTS1	Suicide vector for in-frame deletions in <i>Pseudomonas</i>	T. Scott (JIC) (404)
pTS1::ΔPep4	In-frame deletion of Pep4 gene cluster	This work
pTS1::ΔHCN	In-frame deletion of <i>hcnABC</i> gene cluster	This work
pTS1::Δ652Pyo	In-frame deletion of pyoverdine encoding NRPS in Ps652	This work
pTS1::ΔEM42Pyo	In-frame deletion of pyoverdine encoding NRPS in <i>Pseudomonas putida</i> EM42	This work
pTS1::ΔNRPS5	In-frame deletion of L-threonine encoding NRPS in Ps652	This work
pTS1::ΔpaaK	In-frame deletion of <i>paaK</i> found in Ps652 cluster	This work
pTS1::ΔTE	In-frame deletion of acyl-CoA thioesterase in Ps652 cluster	This work
pTS1::ΔEchA8_1	In-frame deletion of <i>echA8_1</i> enoyl-CoA hydratase in Ps652 cluster	This work
pTS1::ΔEchA8_2	In-frame deletion of <i>echA8_2</i> duplicated enoyl-CoA hydratase in Ps652 genome	This work
pTS1::ΔipdC	In-frame deletion of <i>ipdC</i> indole pyruvate decarboxylase in Ps652 cluster	This work
pTS1::ΔacdA	In-frame deletion of <i>acdA</i> acyl-CoA dehydrogenase in Ps652 cluster	This work
pJH10TS	High copy-number replicative plasmid for heterologous expression in <i>Pseudomonas</i>	T. Scott (JIC) (404)
pJH10TS::Pep1	Heterologous expression of Pep1 RiPP-like gene cluster	This work
pJH10TS::Pep1Δ2282	Heterologous expression of Pep1 RiPP-like gene cluster without DUF2282 protein	This work
pME6032	Kanamycin-selective replicative plasmid for inducible heterologous expression in <i>Pseudomonas</i>	(343)

Plasmid	Application	Reference
pME6032::TE	Complementation of acyl-CoA thioesterase in Ps652ΔHCNΔTE	This work
pGME6032	Gentamycin-selective derivative of pME6032	N. Miguel-Vior (JIC)
pGME6032::DUF2282	Inducible heterologous expression of DUF2282 and complementation of Pep1Δ2282	This work
pALMAR3	Mariner-transposon carrying plasmid with tetracycline marker	(333)
pUC18-mini-Tn7	Mini-Tn7 vector for subcloning	(405)
pUC18_LacZ	Subcloning host for creation of pUC18_LacZ_IV	This work
pUC18_LacZ_IV	Integrative Tn7 vector for lacZ labelling of <i>Pseudomonas</i> strains	This work
pUC18::TE	Integrative Tn7 vector for complementation of Ps652 acyl-CoA thioesterase mutants	This work

5.1.3 Chemicals, Reagents, and Media

5.1.3.1 Chemicals & Reagents

Unless otherwise specified, all chemicals were obtained from Sigma-Aldrich (Merck). All solvents for extractions and chromatographic applications were supplied by Fisher Scientific. All enzymes were obtained from New England Biolabs (NEB) unless otherwise specified. DNA purification kits were obtained from Qiagen. Ultrapure H₂O was obtained using a Milli-Q purification system from Merck.

5.1.3.2 Growth Media

All media used are defined below; all formulae are for 1 L of de-ionised water unless specified otherwise. All media were standardised to pH 7.2 unless otherwise stated. All media ingredients were sourced from Sigma-Aldrich (Merck) unless otherwise stated.

Antibiotics were added where necessary at the following concentrations: Tetracycline (15µg/mL for strain Ps723 and DH5α, 25 µg/mL for KT2440 and 652ΔHCN, 100 µg/mL for EM383 and EM42); Gentamycin, 25 µg/mL.

Table 5.3. Growth media used in this work

Medium	Ingredient	Per Litre	Application
Lennox Broth and Lennox Agar (L)	Tryptone	10 g	Maintenance and growth of <i>Pseudomonas</i> strains
	Yeast extract	5 g	
	NaCl	5 g	
	D-glucose	1 g	
	Agar (Formedium)	10 g	
Lysogeny Broth (LB) or Lysogeny agar (LB agar)	Bacto-tryptone	10 g	Maintenance and growth of <i>E. coli</i>
	Yeast extract	5 g	
	NaCl	10 g	
	Agar (Formedium)	10 g	
Soy-Flour Mannitol (SFM)	Soya flour	20 g	Cross-streak assays, split-plate assays
	Mannitol	20 g	
	Lab M #1 agar	20 g	
Soy-Flour Mannitol + Yeast Extract (SFM+YE)	Soya flour	20 g	Cross-streak assays, split-plate assays
	Mannitol	20g	
	Yeast extract	4 g	
	Lab M #1 agar	20 g	
Malt Extract-Yeast Extract Maltose Medium (MYM)	Malt extract	10 g	Cross-streak assays, split-plate assays
	Yeast extract	4 g	
	Maltose	4 g	
	Bacteriological agar	18 g	
	Tap water	1 L	

Medium	Ingredient	Per Litre	Application
Potato Dextrose Agar (PDA)	Potato Dextrose Agar (Formedium PDA0102S)	41 g	Biological assays of Ps652 against <i>G. graminis</i>
Rye Sucrose Agar (RSA)	Rye grains; pre-germinated, ground using stick blender, heated for 3 hours at 50 °C	60 g	Biological assays of Ps652 against <i>P. infestans</i> , automated screening of Ps652 transposon mutants
	Sucrose	20 g	
Instant Mash Agar (IMA)	Instant mashed potato (Smash, Batchelors)	20 g	Biological assays of Ps652 fractions against <i>S. scabiei</i> 87-22
	Lab M #1 agar	20 g	
King's Broth (KB)	Proteose peptone	20 g	Hydrogen cyanide production testing in Ps652
	Glycerol	15 g	
	K ₂ HPO ₄ (anhydrous)	1.6 g	
	1M MgSO ₄ (after autoclaving)	5 mL	
Nutrient Broth	Difco Nutrient Broth powder	4 g	Pep1 Expression Cultures
Nutrient Broth Glucose	Difco Nutrient Broth powder	4 g	Pep1 Expression Cultures
	D-Glucose	20 g	
M9 + Casamino acids	Na ₂ HPO ₄	6 g	Pep1 Expression Cultures
	KH ₂ PO ₄	3 g	
	NaCl	0.5 g	
	NH ₄ Cl	1 g	
	Casein hydrolysate	20 g	
Obafluorin production medium (OPM)	Yeast extract	5 g	Pep1 Expression Cultures
	D-glucose	5 g	
	MgSO ₄ ·7H ₂ O	0.1 g	
	FeSO ₄	0.1 g	

Medium	Ingredient	Per Litre	Application
Yeast-Peptone-Dextrose (YPD)	Yeast extract	10 g	Pep1 Expression Cultures
	Proteose peptone	20 g	
	D-glucose	20 g	
Modified King's Broth (MKB)	Glycerol	15 mL	5L growth cultures for Ps652 pyoverdine purification, Transposon mutant metabolomics
	K ₂ HPO ₄ (anhydrous)	2.5 g	
	Casein hydrolysate	5 g	
	MgSO ₄ ·7H ₂ O	2.5 g	
	Adjust to pH 6.7 with HCl		
Modified King's Broth + glucose (MKB Glucose)	Glucose	15 g	Comparative metabolomics and production of tropolones in Ps652 and mutants
	K ₂ HPO ₄ (anhydrous)	2.5 g	
	Casein hydrolysate	5 g	
	MgSO ₄ ·7H ₂ O	2.5 g	
	Adjust to pH 6.7 with HCl		
CAS	Prepared in accordance with ref (406)		Iron binding assays

5.1.4 Oligonucleotides

Oligonucleotides were obtained (desalted) from Sigma-Aldrich and redissolved in Milli-Q H₂O. All oligonucleotides used in this work are described in Table 5.4.

Table 5.4. Oligonucleotides used in this study.

Primer	Sequence 5'-3'	Application
Arb1b	GGCCAGCGAGCTAACGAGACNNNGATAT	Locating transposon insertion site
Arb-PCR	CGCAAACCAACCCTTGGCAG	Locating transposon insertion site
Arb1	GGCCAGCGAGCTAACGAGAC	Locating transposon insertion site

Primer	Sequence 5'-3'	Application
Almar3-seq	ACATATCCATCGCGTCCGCC	Locating transposon insertion site
Pep4_UP_F	GCGAGCTCCCAGGAAAGCTTCACCCC	Amplification of Upstream Pep4 fragment (forward)
Pep4_UP_R	GCTCTAGAGCCGAAGCACTTTTCTGGG	Amplification of Upstream Pep4 fragment (reverse)
Pep4_DOWN_F	GCTCTAGAGGCTGGTCAGCTTGAACAC	Amplification of downstream Pep4 fragment (forward)
Pep4_DOWN_R	GCCATATGCCCAAAGTCAACGCACCCTGG	Amplification of downstream Pep4 fragment (reverse)
HCN_UP_F	CAGAGAGGAGCACATTCACG	Amplification of Upstream HCN fragment (forward)
HCN_UP_R	CAGTTGATCATGGTCGTTGC	Amplification of upstream HCN fragment (reverse)
HCN_DOWN_F	CCTTCCACTGCGATACCCTG	Amplification of downstream HCN fragment (forward)
HCN_DOWN_R	CCTGCTTGTTTTCCACAGCC	Amplification of downstream HCN fragment (reverse)
Pep1_F_XhoI	GCGGCGCTCGAGGCTGGAGATTCAACATGACTACC	Amplification of Pep1 gene cluster + restriction site (forward)
Pep1_R_NdeI	GCGGCGCATATGCCTCAATGGCGGTTTCC	Amplification of Pep1 gene cluster +

Primer	Sequence 5'-3'	Application
		restriction site (reverse)
Pep1_Seq_1F	GCTGGAGATTCAACATGACTAC	Pep1 construct sequencing
Pep1_Seq_1R	AATCGTTCTTGCCAGCTTC	Pep1 construct sequencing
Pep1_Seq_2F	GGTACAGAGCACCTTGAAGC	Pep1 construct sequencing
Pep1_Seq_3F	ACAACTGGTGGGCGAG	Pep1 construct sequencing
Pep1_Seq_4F	CCACAACGCCATCACTC	Pep1 construct sequencing
Pep1_Seq_5F	GATGCACCAGGTACATC	Pep1 construct sequencing
Pep1_Seq_6F	GTGCAGCAATACCACG	Pep1 construct sequencing
Pep1_Seq_7F	AATGGTGCCTGGATG	Pep1 construct sequencing
Pep1_Seq_8F	GACCTGCAACAACCTCTAC	Pep1 construct sequencing
ShortDPep4F	CCATGAAGCGTCTTACG	PCR screening of Pep4 mutants (forward)
Pep4_Ext_R	CGGTAGCGTTAATCAGCACT	PCR screening of Pep4 mutants (reverse)
Pep1_d2282_GF	TTAACTGCGCTAGCACCTCTCGAGAGGAGGGAGTAC TGATGACTCTTTCATCCCTGCA	Pep1Δ2282 Cloning (forward)
Pep1_d2282_GR	CTAGAATCGATGGTACCTTAATGCCATATGCCTCAAT GGCGGTTTCC	Pep1Δ2282 Cloning (reverse)
DUF2282_F	CAATTTCACACAGGAAACAGAATTCGCTGGAGATTC AACATGACTACC	DUF2282 Cloning
DUF2282_R	TGATCCGCTAGTCCGAGGCCTCGAGTCAGGACTTGG CAGTGAAGG	DUF2282 Cloning

Primer	Sequence 5'-3'	Application
UP_Pyo_fwd	CCATGGTTGGCTAGCACCTCTCGAGGCACCGGCAAG GTATGG	Deletion of Pyoverdine NRPS in EM42
UP_Pyo_rev	CAAGTCGATCATTGGCTGG	Deletion of Pyoverdine NRPS in EM42
DOWN_Pyo_fwd	AATCGTCCAGCCAATGATCGACTTGGGAAAGGTCCA ACATCTCG	Deletion of Pyoverdine NRPS in EM42
DOWN_Pyo_rev	GGTGAGAATGGCAAAGCTCATATGGCTGTGCTGGA CGATTGC	Deletion of Pyoverdine NRPS in EM42
Pyo_Seq_1F	GGTCACCATGGTTGGCTAGC	Sequencing of pTS1::ΔEM42Pyo
Pyo_Seq_Rev	CTGAATCCGGTGAGAATGG	Sequencing of pTS1::ΔEM42Pyo
Pyo_del_fwd	CAGGATACTGCGGTTGTCTG	PCR screening of EM42ΔPyo
Pyo_del_rev	AATGGTTCGACTATGAAGCC	PCR screening of EM42ΔPyo
TE_del_fwd	ACCGAGAGCAGCAGTTCC	PCR screening of Ps652 thioesterase mutant
TE_del_rev	ATCAACAGCACTCCTTGAGG	PCR screening of Ps652 thioesterase mutant
TE_del_seq	GACCACTTCCTTGCTGTCTG	Sequencing of Ps652 thioesterase mutants
Paak_del_fwd	CGTAGGTGGTCAAGTACC	PCR screening of Ps652 phenylacetate- CoA ligase mutant

Primer	Sequence 5'-3'	Application
Paak_del_rev	AGGATGCATTGGGCTTCG	PCR screening of Ps652 phenylacetate-CoA ligase mutant
NRPS5_del_fwd	GAACTCGCGATCACTAACC	PCR screening of Ps652 NRPS5 mutants
NRPS5_del_rev	CTGGAATCGGATCGAAGTCG	PCR screening of Ps652 NRPS5 mutants
EchA_del_fwd	CTTACCAGTTTCTGCAGG	PCR screening of Ps652 enoyl-CoA hydratase ($\Delta echA8_1$) mutants
EchA_del_rev	GGTGCGATTCATATCG	PCR screening of Ps652 enoyl-CoA hydratase ($\Delta echA8_1$) mutants
$\Delta EchA8_2$ _fwd	AGGACGGTGCCGATGC	PCR screening of Ps652 enoyl-CoA hydratase ($\Delta echA8_2$) mutants
$\Delta EchA8_2$ _rev	ATGTTTCGCGTTCCGGTAGC	PCR screening of Ps652 enoyl-CoA hydratase ($\Delta echA8_2$) mutants
TE_comp_F	TCATGCATGAGCTCACTAGTGGATCATCAACAGCACT CCTTGAGG	Cloning acyl-CoA thioesterase into pUC18-mini-Tn7-Gm
TE_comp_R	AGGAATTCCTGCAGCCCGGGGATCTCAGCCTTCCT GGAACG	Cloning acyl-CoA thioesterase into pUC18-mini-Tn7-Gm
TE_comp_PME6 032_F	AAGCCGCCACGGCGATATCGGATCCATCAACAGCAC TCCTTGAGG	Cloning acyl-CoA thioesterase into pME6032

Primer	Sequence 5'-3'	Application
TE_comp_PME6 032_R	CCATGGTACCCGGGAGCTCGAATTCTCAGCCTTCCTG GAACG	Cloning acyl-CoA thioesterase into pME6032
ipdC_UP_Fwd	GCTCGGTACCCGGGGATCCTCTAGAATGGCCGAGCG TGAATGG	Deletion of ipdC gene in Ps652ΔHCN
ipdC_UP_Rev	GGAATTCGGCATCCGCCATCTGTTTGCCGAGTTCGAC TACG	Deletion of ipdC gene in Ps652ΔHCN
ipdC_DOWN_Fw d	ATGGCGGATGCCGAATTCC	Deletion of ipdC gene in Ps652ΔHCN
ipdC_DOWN_Rev v	GGTGAGAATGGCAAAGCTCATATGGTCCTGACCAC CCTGAGC	Deletion of ipdC gene in Ps652ΔHCN
ΔipdC_Fwd_2	GACATCAGCGACGATGACTGG	PCR screening of Ps652 ipdC mutants
ΔipdC_Rev	TGAGCAAGGAAGCACTGC	PCR screening of Ps652 ipdC mutants
ΔacdA_Fwd	G TTCACGATGAGCAACAAGG	PCR screening of Ps652 acdA mutants
ΔacdA_Rev	GACCACGGTCGGTTTGC	PCR screening of Ps652 acdA mutants

5.2 General methods

5.2.1 E. coli Methods

Escherichia coli DH5α was grown in liquid LB medium at 30 °C or 37 °C for 16-20 hours as required. *E. coli* DH5α was grown on LB agar at 37 °C until colonies were visible to the naked eye and large enough to facilitate manual colony picking. Plates were stored at 4-8 °C. Stocks of DH5α strains were stored at -70 °C in 25% glycerol.

5.2.2 *Pseudomonas* Strain Methods

5.2.2.1 Growth Conditions

Pseudomonas strains were grown in Lennox medium (environmental isolates and derivatives) or LB medium (KT2440, EM42, EM383 and derivatives) at 30 °C for 16-20 hours unless otherwise specified in the text. *Pseudomonas* strains were grown on solid media at 30 °C until visible colonies formed, large enough to be visible to the naked eye and facilitate manual colony picking. Plates were stored at 4-8 °C, and stocks of *Pseudomonas* strains were stored at -70 °C as 925 µL overnight culture + 75 µL DMSO, or 500 µL overnight culture + 500 µL 50% glycerol.

5.2.2.2 DNA Extraction

DNA was extracted from *Pseudomonas* strains either by addition of a single colony to 50 µL Milli-Q H₂O and incubation for 10 minutes at 98 °C in a thermocycler (for routine PCR reactions), or where necessary genomic DNA was purified using Fast DNA™ SPIN Kit for Soil (MP Biomedicals, Loughborough, UK).

5.2.3 Transformations

5.2.3.1 *Pseudomonas* Transformations

Pseudomonas strains were transformed by electroporation as follows. The desired strain was streaked from glycerol stocks onto Lennox agar and grown at 30 °C until visible colonies appeared. Single colonies were used to inoculate 10 mL Lennox broth per transformation, with a no-DNA sample for each strain-antibiotic condition. The following instructions are per transformation sample. Liquid cultures were grown overnight before being centrifuged at 4000 x g for 8 minutes and resuspended in 2 mL sterile 300 mM sucrose. 1 mL was then transferred each to a 2 mL microcentrifuge tube and centrifuged at 11,000 x g for 1 minute. Supernatant was removed by decanting, 1 mL 300 mM sucrose added, and resuspended by gentle vortex mixing. This was centrifuged at 11,000 x g for 1 minute and the supernatant removed as before. This wash was repeated twice more, and the cells from both tubes

resuspended in a total of 100 μ L 300 mM sucrose. 2 μ L of DNA was added to each 100 μ L aliquot (2 μ L Milli-Q H₂O for controls), and mixed carefully with a pipette tip. Samples were then added to electroporation cuvettes (2 mm), ensuring no bubbles formed. Electroporation was performed using an Eppendorf Eporator at a setting of 2500 V. 900 μ L of Lennox broth was added immediately after electroporation and the samples mixed by gentle pipetting. Electroporated cells were incubated at 30 °C for 1 hour with shaking at 250 rpm. 100 μ L of electroporated cells was then plated on Lennox agar or LB agar with antibiotics as required and incubated at 30 °C until colonies were visible.

5.2.3.2 *E. coli* DH5 α

NEB 5-alpha Competent *E. coli* (High Efficiency) C2987H cells were used for all transformations of *E. coli* and performed as per manufacturer instructions.

5.2.4 Cloning and Sequencing

5.2.4.1 Polymerase Chain Reaction (PCR)

PCR reactions were performed as outlined in the tables below, following manufacturer recommendations. For larger high-fidelity PCR reactions, reaction volume was increased to 100 μ L and up to 3 μ L dNTPs used to prevent this limiting the amount of product for longer template sequences. Annealing temperatures were calculated using the online NEB TM calculator tool (<http://tmcalculator.neb.com/>) with the appropriate enzyme selected. High-fidelity PCR reactions were performed with Q5 DNA Polymerase (NEB) (Table 5.5), and colony PCR reactions using GoTaq 2x Master Mix (Promega, WI, U.S.A) (Table 5.6).

Table 5.5. High-fidelity PCR with Q5 DNA Polymerase conditions

Reaction components		Thermocycler conditions		
Reagent	Volume (μ L)	Temperature (°C)	Time (s)	
5X Q5 Reaction Buffer	10	98	120	
10mM dNTPs	1	98	15	x 35
Each primer	2.5	Annealing Temperature	30	

Template	2	72	20 / kbp	
Q5 High-Fidelity DNA Polymerase	0.5	72	300	
Milli-Q H ₂ O	31.5			

Table 5.6. Colony PCR with 2x GoTaq Master Mix conditions

Reaction components		Thermocycler conditions		
Reagent	Volume (μL)	Temperature (°C)	Time (s)	
GoTaq 2x Master Mix	7.5	95	480	
Each primer	0.5	95	30	x 35
Template	Colony	Annealing Temperature	30	
Milli-Q H ₂ O	6.5	72	60 / kbp	
		72	300	

5.2.4.2 Agarose Gel Electrophoresis

0.8% agarose gels with 3 μL ethidium bromide per 100 mL were used for electrophoresis of DNA. All gels were run for 45 minutes at 110 V in 1 x TBE (Tris/Borate/EDTA) buffer. Samples were loaded with 1 μL 6x loading dye (NEB) per 5 μL sample. Gels were visualised under UV light, and DNA fragments compared to NEB 2-log or 1KB plus ladders (NEB).

5.2.4.3 Purification of DNA from Agarose Gels and Reaction Mixtures

10 μL loading dye was added to Q5 PCR mixtures, linear fragment DNA digests, and plasmid digests (per 50 μL) before loading onto agarose gels. DNA fragments and plasmids were excised from agarose gels using a razor blade and purified using QIAquick Gel Extraction Kit (Qiagen) according to manufacturer instructions and eluted in 30 μL Milli-Q H₂O. DNA fragments were purified from PCR and DNA ligation reaction mixtures using a QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer instructions and eluted in 20-50 μL Milli-Q H₂O as necessary.

5.2.4.4 DNA Digestions

All restriction enzymes were sourced from New England Biolabs (NEB), used with the accompanying CutSmart® Buffer, and incubated at 37 °C for 1-3 hours as necessary. Reaction mixtures were prepared as below, and immediately gel purified or heat inactivated at 65 °C for 20 minutes.

Table 5.7. Reaction mixture for diagnostic restriction digests of vectors.

Component	Volume (µL)
Each enzyme	1
10X CutSmart® Buffer	2
Milli-Q H ₂ O	XµL to make up to 20µL
Plasmid Miniprep	10
Total	20

Table 5.8. Reaction mixture for restriction digest of plasmids.

Component	Volume (µL)
Each enzyme	2
10X CutSmart® Buffer	5
Milli-Q H ₂ O	XµL to make up to 50µL
Plasmid	10
Total	50

Table 5.9. Reaction mixture for restriction digest of linear DNA fragments.

Component	Volume (µL)
Each enzyme	2
10X CutSmart® Buffer	5-7
Milli-Q H ₂ O	21
DNA fragment	20
Total	50

5.2.4.5 DNA Ligations

Ligations were performed in 20 μL volumes, with 2 μL 10X Invitrogen T4 Buffer, 1 μL Invitrogen T4 DNA ligase, X μL DNA fragments with a 3-fold molar excess of insert to vector, and $17 - X$ μL Milli-Q H_2O .

5.2.4.6 Gibson Assembly

Gibson assembly was performed using Gibson Assembly[®] Master Mix or NEBuilder[®] HiFi DNA Assembly mix according to manufacturer instructions for 2-3 fragment assemblies, with minor modifications in reaction time. 0.02-0.5 pmols of DNA, comprising 50-100 ng vector and a 3-fold molar excess of inserts, with 10 μL of 2x Gibson Assembly Master Mix and make up to 20 μL with Milli-Q H_2O . Reactions were incubated in a thermocycler at 50 $^{\circ}\text{C}$ for 60 minutes.

5.2.4.7 Sequencing

All sequencing of vector and amplicon sequencing was performed as Sanger Sequencing by Eurofins Genomics using a Mix2Seq kit as per manufacturer instruction with the appropriate sequencing primers (see Table 5.4). Sequencing results were analysed using a text editor and a local installation of BLAST to account for possible PCR / sequencing errors.

5.2.5 Making In-frame Deletion Strains

In-frame deletion mutants of *Pseudomonas* strains were constructed by markerless 2-step allelic exchange (181). Mutant alleles were designed by selecting 500 bp sequences upstream and downstream of the region to be deleted, in-frame with the gene(s) to be deleted. These were amplified by PCR and assembled by Gibson assembly, or concatenated in a word processor and synthesised by Twist Biosciences Ltd and cloned into the pTS1 suicide vector. *Pseudomonas* strains were transformed as per section 5.2.3.1 with the pTS1 plasmid containing the shortened, mutant allele and 100 μL of cells plated on Lennox agar supplemented with the appropriate concentration of tetracycline. Plates were incubated at 28 $^{\circ}\text{C}$ until visible colonies formed. Individual colonies were inoculated into 10 mL Lennox

broth and grown for 16-18 hours at 28 °C with 250 rpm shaking. These cultures were then diluted to 10^{-6} and plated on Lennox agar supplemented with 10% sucrose, and incubated at 28 °C until visible colonies formed. These colonies were checked by colony PCR with GoTaq 2X Master Mix for the mutant allele, with the parent strain as the control. Positive clones were restreaked to single colonies, tested again, and then grown 16-18 hours at 28 °C with 250 rpm shaking in Lennox broth supplemented with tetracycline. 75 μ L DMSO was added to 925 μ L of culture and this was stored at -70 °C.

5.3 Metabolomics and LC-MS

5.3.1 Sample Preparation (General)

Except where otherwise stated, LC-MS samples were prepared from bacterial cultures by 2 rounds of centrifugation at 17,900 x g for 2.5 minutes to remove cells; 500 μ L cell-free supernatant was then transferred to 1.5 mL glass vials and stored at -30 °C until required. LC-MS samples of purification fractions were analysed directly by transferring to 1.5 mL glass vials and storing at -30 °C until required.

5.3.1.1 Preparation of Tropolone BGC Mutant Extracts

All strains were streaked from stocks to single colonies on Lennox agar plates with selection if necessary, and grown overnight in biological duplicate in 5 mL MKB Glucose in 50 mL centrifuge tubes with sterilised foam bungs (no selection). 5 mL EtOAc was added, and tubes agitated for 1 hour at 250 rpm. These were then centrifuged at 4,000 x g for 5 minutes to pellet any cell debris and ensure proper separation of aqueous and organic phases. 1.5 mL samples of the organic phase were taken for each, and dried fully by centrifugation under reduced pressure at 30 °C on a Genevac EZ-2 Plus system using the 'Low BP' setting, before being resuspended in 1.5 mL 100% MeOH and stored at -30 °C until use.

5.3.2 LC-MS Conditions

For studies of bioinformatically identified RiPP-like gene clusters (chapter 4), LC-MS/MS analysis was performed on a Shimadzu Nexera X2 UHPLC connected to a Shimadzu ion-trap time-of-flight (IT-TOF) mass spectrometer and analysed using LabSolutions software (Shimadzu). For investigations of the metabolites produced by Ps652, including purification of tropolone natural products, LC-MS/MS analysis was performed on a Q-Exactive mass spectrometry system (Thermo Fisher Scientific, U.S.A). A Luna Omega 1.6 μ m Polar C18 100Å (50 x 2.1mm) column (Phenomenex) was used for compounds produced by strain 652 Δ H₂CN using 0.1% formic acid in H₂O and MeOH. A Kinetex 2.6 μ m XB-C18 100Å (50 x 2.1mm) column (Phenomenex) was used for assessment of metabolites produced in Pep1 gene cluster work using 0.1% formic acid in H₂O and acetonitrile (MeCN). For Kinetex C-18, gradients were linear 5-95% organic solvent from 1-6 minutes, followed by 100% organic solvent until 8.8 minutes to wash off any material still associated with the column. For Polar C18, gradients were linear 0-95% organic solvent from 1-6 minutes, followed by 100% organic solvent until 8.8 minutes to wash off any material still associated with the column. Spectra were acquired in negative mode with MS/MS, or positive mode with MS/MS as stated in the text. Unless otherwise stated, the mass scan range used was *m/z* 200-2000.

5.3.3 Accurate Mass

High resolution masses were acquired on a Q-Exactive mass spectrometer (Thermo Fisher Scientific, U.S.A).

5.3.4 Analysis of Results

LC-MS(/MS) spectra were analysed in a number of ways. Where relevant, spectra were analysed by visual inspection of the chromatograms and displaying the extracted ion chromatograms (EIC) of individual ions, comparing samples directly. Any associated UV absorbance or charged aerosol detection (CAD) data was analysed in the same way. For comparison of Pep1 expression cultures, the XCMS online platform was used.

5.3.4.1 XCMS Online Pairwise Analysis

Metabolites of interest were identified from MS data using the pairwise comparison tool at XCMS online (394). The metabolomic cloud plot tool was used with the following values to identify metabolites showing different profiles in mutants versus the parent strain: p-value = 0-0.5, fold change = 10+, m/z range = 200-1115, retention time = 0.6 minutes, intensity = 100000+. Candidate masses were then further investigated in LCMS Solutions software (Shimadzu).

5.4 Biological Assays

5.4.1 Cross Streaks

Cross-streak assays against *S. scabies* 87-22 were performed on MYM or SFM agar as stated in the text. *Streptomyces* spores were streaked in two parallel lines on the plate using a sterile toothpick, which were then 'cross streaked' at a 90° angle with overnight cultures of the *Pseudomonas* strain of interest. This method allows visualisation of contact-dependent inhibition, soluble compound mediated inhibition, and volatile mediated inhibition. Plates were incubated at 28 °C for 5 days, or until interspecies interactions were visible. Plates were then imaged using a Canon Ixus 175 digital camera (Canon, Japan). All assays were performed in biological triplicate.

5.4.2 Spot Assays Against *Phytophthora infestans* and *Gaeumannomyces graminis* var. *tritici*

A 5mm diameter plug was taken from the outer edge of fungal or oomycete mycelium using a heat and ethanol sterilised number 3 metal corer, and placed in the centre of a standard petri dish on the appropriate medium. Rye Sucrose Agar (RSA) was used for *P. infestans* and Potato Dextrose Agar used for *G. graminis*. 15µL of overnight liquid culture of the *Pseudomonas* strain to be tested was then spotted in triplicate as an equilateral triangle of vertex 5cm around the central plug. Spotted liquid culture was then allowed to dry in a Biosafety level 2 cabinet, before the plate was transferred to either 18 °C or 20 °C and

incubated for 10 days. Plates were then imaged using a Canon Ixus 175 digital camera (Canon, Japan).

5.4.3 Split-plate Assays

Split-plate assays were employed to test for volatile compounds inhibiting *Streptomyces scabies* and *Phytophthora infestans*. *P. infestans* was placed centrally as a 5mm diameter plug (as above) on one side of a barrier plate on RSA, while the other side was inoculated with the *Pseudomonas* isolate either on Lennox agar or RSA. Plates were incubated at 18 °C until the control plate (*P. infestans* only) had grown to the outer edge of the plate. For assays against *S. scabies*, assay was performed as the cross-streak method defined above, with each bacterium streaked perpendicular to the other on opposite sides of the barrier. The intention being to directly replicate the cross-streak assay in all ways except limiting contact and the diffusion of soluble compounds. These plates were incubated for 5 days. In both cases, Plates were imaged using a Canon Ixus 175 digital camera (Canon, Japan) at the relevant stage.

5.4.4 Testing Extracts of *Pseudomonas* Cultures & Purification Fractions

Testing *Pseudomonas* culture extracts for biological activity against *Streptomyces scabies* 87-22 was performed by creating a lawn of *S. scabies* on instant mash agar (IMA). 10 µL of spore stock was resuspended in 1 mL Milli-Q H₂O, before 35 µL of this mixture was spread thoroughly onto the IMA using a sterile cotton bud. A 5 mm diameter metal corer was sterilised using heat and ethanol, then used to remove plugs of agar to create wells. Extracts or fractions to be tested were concentrated 10x, unless otherwise stated, and redissolved in 50-100 µL Milli-Q H₂O. This was then added to the agar wells. Less polar fractions / extracts were redissolved in 50-100 µL MeOH and applied to filter paper disks, which were allowed to dry completely before being added to the IMA plate. Plates were incubated at 28 °C for 2 days or until biological activity of extracts / fractions could be observed. Plates were imaged using a Canon Ixus 175 digital camera (Canon, Japan).

5.4.5 Iron Binding Assays

Iron binding was tested using CAS agar plates as originally defined in ref (373). The binding of Chrome Azurol S (CAS) / hexadecyltrimethylammonium (HDTMA) in the medium complexes with ferric iron to produce a blue colour. Siderophores are able to compete for the binding of iron, liberating it from CAS/HDTMA, producing a colour change from blue to orange. 5 μ L of overnight culture or 5 μ g of purified compounds dissolved in MeOH were added as spots to the plate and allowed to dry. If overnight culture was used, plates were incubated for 1 day at 30 °C before being imaged. For pure compounds, the solvent was allowed to dry, plates were then stored at room temperature overnight to allow for diffusion of applied molecules before being imaged. Assays were imaged using a Canon Ixus 175 digital camera (Canon, Japan).

5.5 Purification and Partial Characterisation of Pyoverdine from Ps652

5.5.1 Scaled-up Production Cultures

Ps652 Δ HCN was grown in 5 x 1 L batches of MKB liquid medium, with glycerol as a carbon source, for 24 hours at 30 °C. All batches were checked for biological activity by concentrating 1 mL and testing against *S. scabiei* 87-22 on plates – no activity was observed for an MKB only negative control. All 1 L cultures were combined and processed by rotary evaporation to a volume of 1 L.

5.5.2 Initial Ethyl Acetate Extraction

The concentrated 1 L sample was extracted with 3 x 500 mL EtOAc using a separating funnel. A 10 mL samples of the defatted aqueous fraction was subsequently adjusted to pH 2 using 5 M HCl, before being extracted a further 3 times with 3 x 5 mL EtOAc. All phases were tested for biological activity as above, with the pre-acidification aqueous fraction retaining the most activity. This was then concentrated to ~700 mL by rotary evaporation and freeze-drying was attempted, but was not successful.

5.5.3 Sample Debulking by Flash Chromatography

5.5.3.1 500g XAD-4

To reduce the amount of material sufficiently to allow for separation of the target compound by HPLC, flash chromatography was performed using a Biotage Isolera system (Biotage, Sweden) with large cartridges to permit direct loading of the entire sample. Initially, a 500 g XAD-4 cartridge was used. The entire ~700 mL extract was loaded directly onto the cartridge through a silica frit to remove undissolved debris, and the immediate flowthrough collected. A solid-phase extraction approach was then followed, with 500 mL fractions collected at the following concentrations of acetone/water: 0%/100%, 5%/95%, 10%/90%, 15%/85%, 20%/80%, 30%/70%, 50%/50%, 100%/0%. 1 mL samples of each fraction were dried on a Genevac EZ-2 Plus and redissolved in 100 μ L Milli-Q H₂O, followed by testing of biological activity. Assays revealed the active compound was mainly in the flowthrough and the first fraction, suggesting minimal interaction with the XAD-4 stationary phase.

5.5.3.2 120g C18

The post XAD-4 flowthrough fraction was lyophilised, 1/5th redissolved in 50 mL Milli-Q H₂O, and loaded onto a 120 g Sfär C18 Cartridge (Biotage, Sweden). Equilibration was performed pre-sample loading with 100% de-ionised H₂O without additives. After sample loading, 2 CV of H₂O were passed and 0.1 CV fractions collected. This was followed by a 0%-20% MeCN gradient over 1 CV, collecting 0.1 CV fractions. 2 CV of 20% MeCN were then passed, collecting into a single vessel. Finally the cartridge was washed with 2 CV of 100% MeCN, again collecting into a single vessel. All fractions were concentrated to 16 mL, and 1 mL samples taken for assessment of biological activity as above, the majority of which was found in the 20% MeCN fraction. This fraction was analysed by mass spectrometry using a 0%-95% MeOH gradient on a Polar C18 column (Phenomenex) with a scan range of *m/z* 200-2000 in positive and negative mode. The base peak chromatogram revealed the sample to still be highly complex. The remaining 4/5th of the post-XAD-4 flowthrough were processed as above but only 4 fractions were collected: 0% MeCN (2CV), 1%-19% MeCN (1

CV), 20% MeCN (2 CV), 21%-100% MeCN (1.5 CV). Again, biological activity was observed in the 20% MeCN fraction. All 20% MeCN fractions were combined and lyophilised, amounting to 2.23 g of dried material.

5.5.3.3 Further C18 Flash Chromatography

To reduce the 2.23 g of dried material to an amount suitable for preparative scale HPLC, further separation on a Sfär 120 g C18 cartridge as performed, with finer-level separation within the region of interest. The sample was dissolved in 60 mL Milli-Q H₂O and loaded onto the cartridge after an initial equilibration of 2 CV de-ionised H₂O. Mobile phase consisted of MeCN/H₂O without additives and was passed as follows: 0% MeCN (3 CV), 0%-18% MeCN (2 CV), 18%-20% MeCN (2 CV), 20%-100% MeCN (1.5 CV), 100% MeCN (0.5 CV). 0.1 CV fractions were collected for the 18%-20% MeCN step; all other steps were collected into single vessels per step. All fractions were concentrated to 16 mL and 2% of this used for assessment of biological activity. Fractions 5-10 of the 18%-20% MeCN step contained the majority of the activity, so were combined and dried on a Genevac EZ-2 Plus using the 'HPLC' setting at 30 °C. The dried combined fractions amounted to 531 mg of material.

5.5.4 Preparative C18 HPLC

All 531 mg was redissolved in 1 mL Milli-Q H₂O and processed in 5 x 200 µL injections on a Dionex Ultimate 3000 preparative scale HPLC (Dionex, U.S.A) with a Gemini-NX 5µm C18 110Å 150 x 21.1 mm column. Solvents were 100% de-ionised H₂O and 100% MeCN with no additives. A multistep gradient was used as shown in Table 5.10.

Table 5.10. Multistep gradient for preparative scale C18 HPLC.

Retention Time (min)	% H ₂ O	% MeCN
0	100	0
3	100	0
18	60	40
19	0	100

Retention Time (min)	% H ₂ O	% MeCN
22	0	100
23	100	0
30	100	0

6 x 5 minute fractions were collected, allowing broad scale separation of material. 4 mL samples were taken for activity assessment as above. The majority of the activity was observed in fraction 3 (10-15 min), which was dried by rotary evaporation and Genevac EZ-2 Plus on 'HPLC' setting at 30 °C to 142 mg of material. This step was repeated using the 142 mg of material dissolved in 900 µL Milli-Q H₂O in 2 injections, collecting 10 x 20 mL fractions between 5-15 minutes. Activity was found predominantly in fraction 6, comprising 60 mg of dried material.

5.5.5 Semi-preparative HPLC

Semi-preparative scale HPLC separation was performed on a Dionex Ultimate 3000 (Dionex, U.S.A) with a flow rate of 5 mL/min. UV absorbance was monitored at 254 nm

5.5.5.1 C18

A Luna 5 µm C18(2) 100Å 250 x 10 mm column (Phenomenex) was used with Milli-Q H₂O and MeCN solvents, with 0.5% formic acid added after the biological activity of Ps652 in MKB was shown to be acid stable. A multistep gradient was used as shown in Table 5.11. The 60 mg of material was redissolved in Milli-Q H₂O and processed in 4 injections.

Table 5.11. Multistep gradient for semi-preparative scale C18 HPLC.

Retention Time (min)	% H ₂ O + 0.5% Formic Acid	% MeCN + 0.5% Formic Acid
0	100	0
4	100	0

Retention Time (min)	% H ₂ O + 0.5% Formic Acid	% MeCN + 0.5% Formic Acid
5	88	12
25	75	25
26	0	100
30	0	100
31	100	0
35	100	0

Fractions were collected by time (30s) between 5 and 25 minutes. Fractions from all 4 injections were pooled and dried on a Genevac EZ-2 Plus using the 'HPLC' setting at 30 °C. The UV absorbance chromatogram at 254 nm revealed a dominant peak (fraction 8), which amounted to 5 mg of dried material. The scarcity of sample at this stage precluded activity assessment, so this fraction was analysed by mass spectrometry as previously. The data indicated the sample was not pure, but the majority of ions observed related to a single large molecule over 1300 Daltons.

5.5.5.2 Pentafluorophenyl (PFP)

To improve separation of molecules in the sample, an orthogonal pentafluorophenyl (PFP) stationary phase was used in the form of a Luna 5 µm PFP(S) 100Å 250 x 10 mm column (Phenomenex), with Milli-Q H₂O and 100% MeOH as solvents with 0.5% formic acid. A multistep gradient was used as shown in Table 5.12. The 5 mg of material from semi-preparative scale C18 HPLC was redissolved in Milli-Q H₂O and processed in a single injection. Fractions were collected by peak.

Table 5.12. Multistep gradient for semi-preparative scale pentafluorophenyl HPLC.

Retention Time (min)	% H ₂ O + 0.5% Formic Acid	% MeOH + 0.5% Formic Acid
0	100	0
5	100	0
6	80	20

Retention Time (min)	% H ₂ O + 0.5% Formic Acid	% MeOH + 0.5% Formic Acid
30	0	100
35	0	100
36	100	0
47	100	0

6 fractions were collected in total, with a single dominant peak and 5 minor contaminant peaks. Fraction 4 contained the peak of interest, which was analysed by LC-MS as above. Data indicated further purification was necessary. Fractions were dried on a Genevac EZ-2 Plus system using the 'HPLC' setting at 30 °C.

5.5.5.3 Polar C18

A Luna Omega 5 µm Polar C18 100Å 250 x 10 mm column (Phenomenex) was used with Milli-Q H₂O and 100% MeOH solvents with 0.5% formic acid added. Flow rate was lowered to 4 mL/min to reduce pressure-related issues. A multistep gradient was used as shown in Table 5.13.

Table 5.13. Multistep gradient for semi-preparative scale polar C18 HPLC.

Retention Time (min)	% H ₂ O + 0.5% Formic Acid	% MeOH + 0.5% Formic Acid
0	100	0
5	100	0
6	85	15
30	50	50
35	0	100
40	0	100
41	100	0
46	100	0

Fractions were collected by peak at 327 nm, resulting in 2 fractions. The dominant peak was dried down on a Genevac EZ-2 Plus using the 'HPLC' setting at 30 °C to 1.82 mg of a yellow-brown water soluble resin. The sample appeared sufficiently pure for NMR analysis and high-resolution mass spectrometry.

5.5.6 High-resolution mass spectrometry (HRMS)

High-resolution mass spectrometry data was acquired on a Q-Exactive mass spectrometer (Thermo Fisher Scientific, USA) using a Luna Omega 1.6 μm Polar C18 100 \AA (50 x 2.1 mm) column (Phenomenex), with Milli-Q H₂O and MeOH as solvents with 0.1% formic acid. A 0%-95% gradient was used as specified in section 5.3.

5.6 Purification of Tropolones from Ps652

5.6.1 Production Cultures

In order to avoid reisolating pyoverdine when using UV absorption to track tropolone molecules throughout purification, a pyoverdine-null mutant (Ps652 Δ HCN Δ Pyo) was used for production and purification of tropolones. Preliminary attempts at scaling up production cultures were unsuccessful. Accordingly, Ps652 Δ HCN Δ Pyo was grown for 16 hours in 2 x 10 mL MKB Glucose in plastic universal 30 mL tubes. This was then used to inoculate 2L of MKB Glucose in 200 x 10 mL aliquots in 50 mL centrifuge tubes with foam bungs, using 100 μL seed culture per 10 mL, and then grown for 24 hours at 30 °C. Culture from all 200 tubes were then combined into a single 2 L volume.

5.6.2 Extraction of Tropolones from 2L Culture

Cells were removed by centrifugation at 8,000 x g for 8 minutes at 4 °C in a Sorvall Lynx 6000 centrifuge (Thermo Fisher Scientific). The supernatant was collected and cells discarded. The supernatant was concentrated by rotary evaporation to a volume of 1 L. Following the method of Jiang *et al.* (345), the supernatant was then extracted 3 times with 0.5 L of EtOAc before being acidified to pH 2 using HCl. The aqueous phase was then extracted 3 more times with 0.5 L to yield a total of 3 litres of organic extract. All organic extracts were combined and processed by rotary evaporation to dryness. The aqueous phase was discarded.

5.6.3 Size Exclusion Chromatography using Sephadex LH-20

The dried organic phase was dissolved in 3 mL MeOH and processed by size-exclusion chromatography using a Sephadex[®] LH-20 column on an ÄKTA Pure system (Cytiva) in 100% MeOH attached to an Optilab[®] differential refractive index detector (Wyatt Technology). Flow rate was 1 mL / min, and 10 mL fractions were collected from 350 min to 700 min. UV absorbance was monitored at 330 nm to detect tropolones. The 3 mL of organic extracted was processed in 2 injections. All fractions were dried on a Genevac EZ-2 Plus system using the 'Low BP' setting at 30 °C, and all fractions with strong absorbance at 330 nm were redissolved in MeOH, diluted 1/5, and 5 µL analysed by LC-MS on a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher) as per section 5.3.2. Several fractions showed high levels of 7-hydroxytropolone, a dihydroxylated tropolone, or a mixture of both. These fractions were selected for further purification.

5.6.4 Semi-preparative HPLC

Fractions containing hydroxylated tropolones were redissolved in 100% MeOH and combined with fractions with similar profiles by LC-MS. These fractions were processed by semi-preparative HPLC on a Dionex Ultimate 3000 system (Thermo Fisher Scientific) with a Luna Omega 5 µm Polar C18 100Å 250x10 mm column (Phenomenex) in 5 injections. Solvents were Milli-Q H₂O with 0.5% formic acid and 100% MeOH with 0.5% formic acid. Flow rate was 4 mL/min with a multistep gradient as per section 5.5.5.3, and UV absorbance was monitored at 254 nm and 327 nm. The chromatogram showed two clear peaks at 327 nm, with the earlier peak representing the dihydroxylated tropolone and the later peak representing the monohydroxylated tropolone. All fractions from each injection for each peak were combined and dried overnight on a Genevac EZ-2 Plus system using the 'HPLC Lyo' setting at 30 °C. 2.83 mg of 3,7-dihydroxytropolone was obtained and was a white powder with a slight pink hue.

5.6.5 NMR Analysis of Purified Tropolone

5.6.5.1 3,7-dihydroxytropolone

Peak 1 from section 5.6.4 was redissolved in 500 μL CD_3OD and shaken for 10 minutes to exchange OH protons for deuterium, and dried on a Genevac EZ-2 Plus system on the 'Low BP' setting at 30 °C. The sample was then again redissolved in 500 μL CD_3OD and analysed on a Bruker AVANCE 600 MHz Spectrometer. ^1H NMR δ = 6.98 (3H, m). ^{13}C NMR δ = 119.24 (2C), 129.44, 158.11 (2C), 158.94 (2C).

5.7 Bioinformatic Analysis of Pep gene Clusters

5.7.1 AntiSMASH

Genomes were analysed using the antiSMASH versions described in the text. Genbank (.gbk) files were used as input, and default parameters were used for the analysis. Results were analysed manually.

5.7.2 Alignment of MbnB and DUF692

The methanobactin DUF692-containing protein, MbnB, from *Methylosinus trichosporium* OB3b was retrieved from Genbank (ATQ70273.1). MbnB and the DUF692 protein from the Pep1 gene cluster in Ps706 were aligned using MUSCLE (407), visualised using the ESPRIPT server (408), and analysed manually for the presence of essential residues.

5.7.3 RiPPER Workflow

RiPPER analysis was performed as described in detail in the associated publication (238) (see appendix 2), which constitutes a user guide for the workflow. An overview is given here. The DUF692 protein, linked to RiPP biosynthesis in other organisms, was chosen as a 'bait' protein. A list of accessions of DUF692 proteins in *Pseudomonas* was obtained using CDART to retrieve proteins similar to the DUF692 protein in the Pep1 cluster in *Pseudomonas* strain 706. 2,904 sequences were obtained. This list was then filtered using

EFI-EST (388) to create a reduced dataset consisting of proteins that were less than 95% identical. The cut-off values used for network finalisation were similarity score: 40, length: 250-315. The 95% reduced dataset comprised 904 sequences. RiPPER (171) was then used with default settings to identify short peptides associated with these DUF692 proteins. Sequence similarity networks (SSNs) were then generated for the retrieved short peptides and Pep cluster peptides using EGN (385) with the following parameters: E-value threshold = 10, identity threshold = 40%, coverage = 40%, minimum hit length = 15 AA. The resulting network contained 271 clusters. A phylogenetic tree of DUF692 proteins was constructed using the CIPRES server (409), using MUSCLE for alignment with default parameters, and RAxML-HPC on XSEDE for tree building with default parameters. Trees were visualised in iTOL (410). Networks of short peptides were mapped to the DUF692 phylogenetic tree using iTOL definition files. An 85% identity tree was also generated using the same procedure with the exception of using the 85% similarity network from EFI-EST and the manual addition of DUF692 from the recently described 3-thiaglutamate cluster (237). Gene architecture conservation for each network of short peptides was assessed using MultiGeneBlast (369) with default values, and alignment of short peptides was performed using MUSCLE at the EMBL-EBI server (407). Final output figures were generated using Adobe Photoshop.

5.8 Bioinformatic Analysis of Ps652

5.8.1 Genome Assembly

A combined Illumina-Nanopore genome for Ps652 was obtained from MicrobesNG (Birmingham, U.K.). The supplied sequence was received as 5972118 bp in 5 contigs with 44.406099310161x coverage. The Genome was reordered using the Align & Reorder contigs tool in Mauve (411) using *Pseudomonas* sp. SNU WT1 as a reference. The exported reordered file was submitted to the MEDUSA server (412) for scaffolding. The output scaffold produced was 5972218 bp in 3 contigs (5966768 bp, 5323 bp, and 127 bp). Contig 2 (127 bp) and contig 3 (5325 bp) were both found by BLAST (329) to be part of contig 1 (5966768 bp). These were thus removed and 5966768 bp used at the single scaffold genome. This single scaffold genome was annotated using the Prokka tool (413). The final

genome has a size of 5.96 Mbp with 62.19% GC and 5371 genes. This genome was used for all bioinformatic analyses in this work.

5.8.2 AntiSMASH

The Ps652 genome was analysed using antiSMASH 5.0 (328) with the default parameters. The single scaffold genome of Ps652 was submitted as a .gbk file. Gene clusters identified as being similar to published or known gene clusters in the BigSCAPE database (414) were investigated individually for presence of the required genes for biosynthesis of the final product.

5.8.3 Manual BLAST

To search for homologues in Ps652 of genes involved in biosynthesis of *Pseudomonas* natural products identified from the literature, the most relevant genes for each BGC were identified from the relevant publications, and the sequence downloaded from Pseudomonas.com (415). The Ps652 genome described above was used to create a local nucleotide database using a local installation of BLAST (329), and searched using the nucleotide sequences of interest using blastn with default values.

5.9 Automated Transposon Mutagenesis Screen

This was performed as per the associated publication (337) and is reproduced here with modifications. Automation work was carried out in collaboration with A. Elliston at the Earlham Institute.

5.9.1 Creation of a Transposon Mutant Library of Ps652 Δ HCN

Ps652 Δ HCN was transformed by electroporation (as above) in 300 mM sucrose with the mariner transposon-containing plasmid pALMAR3 and 100 μ L was plated directly in single well NuncTM OmniTrayTM plates (242811, Thermo Scientific, UK) containing Lennox agar + 25 μ g/mL tetracycline with a plating density that enabled automated colony picking. Plates

were incubated overnight at 28 °C for 16 hours and resulting colonies were selected using the Hamilton Microlab STARplus (Hamilton, Bonaduz AG, Switzerland) equipped with one 96-channel pipetting head and eight individual single channels capable of liquid level detection, a light table and camera for colony picking, and grippers for plate movement steps. Methods were programmed using the Venus Three software (version 4.4.07740, Hamilton). The detection and selection of the maximum number of viable colonies was enabled by a modified picking method using the EasypickII software (Version 1.0.2, Hamilton). A maximum number of 500 colonies from each plate were selected to generate the library. Criteria for colony picking were established to avoid cross contamination and aberrant phenotype. Colonies were defined with a minimal circularity factor of 0.05 and an area of 0.4 - 15 mm². The liquid transfer and aspiration process were initially hampered by the presence of air bubbles formed by successive pipetting that hampered the autosensing of liquid levels, leading to inaccurate transfers. To overcome this, the pipette position was altered during aspiration and ejection of media. Suitable colonies were transferred using disposable 300 µL tips into liquid media (200 µL Lennox broth + 25 µg/mL tetracycline) in 96-well microplates (4ti-0117, Brooks Life Science, UK) Cultures were grown at 28 °C, 250 rpm for 16 hours. 15 µL DMSO was added to each well manually using a multichannel pipette, and plates were stored at -80 °C until required.

5.9.2 Automated Screening of Mutant Library Against *S. coelicolor* M145

Spores of *S. coelicolor* M145 were resuspended into Milli-Q water at a concentration of 35 µL spore stock per 80 mL H₂O. A Hamilton Microlab STARplus, as described above, was used to dispense 20 µL of spore suspension into the centre of each well of 120 24-well plates (Greiner Bio-One, 662102) containing 1 mL rye sucrose agar (RSA) per well. Spores were distributed using the on-deck shaker at 1,500 rpm for 5 seconds. The plates were dried for 1 hour in a level 2 biological safety cabinet before use in automated bioassays.

The library of *Pseudomonas* sp. Ps652ΔHCN mutants were spiked onto the *S. coelicolor* M145 plates using a Hamilton Microlab STARplus as described above. Pipette tips were used to transfer droplets by submersion without aspiration followed by a brief touch of the

agar surface using a liquid level setting of 0 mm. Plates were incubated at 28 °C for 5 days before photographic imaging (Canon Ixus 175, Canon, Tokyo, Japan).

5.9.3 Selection of Mutants for Further Verification

Mutants of interest were identified visually, by selecting wells of 24-well plates where *Streptomyces coelicolor* M145 was able to grow immediately adjacent to the *Pseudomonas* colony, and where the growth of the pseudomonad did not appear compromised. High and low confidence mutants are described in the text and shown in the associated figure.

5.9.4 Growth Curves

Mutants of interest were streaked by hand to single colonies and grown overnight in 10 mL LB. Growth curves were performed by manually inoculating 5 µL of overnight culture diluted to an optical density at 600 nm (OD_{600}) = 0.01 into 150 µL of LB medium in 96-well plates. Plates were incubated at 30 °C with shaking at 200 rpm for 48 hours with absorbance readings taken every 30 minutes. Data was acquired on a SPECTROstar Nano UV/Vis microplate reader (BMG Labtech, Ortenberg, Germany), and processed using Microsoft Excel and R (20). 3 replicates were used per mutant.

5.9.5 Amplicon Sequencing of Mutants

Sites of transposon insertion were determined by polymerase chain reaction (PCR) as described in (22). DNA flanking the insertion sites was enriched in two rounds of amplification using primers specific to the ends of the mariner transposon element (Arb1b 5'-GGCCAGCGAGCTAACGAGACNNNGATAT-3' and Arb-PCR 5'-CGCAAACCAACCCTTGGCAG-3') with an initial denaturation at 95 °C for 3 minutes followed by 30 cycles of: 95 °C for 15 seconds, 38 °C for 30 seconds, and 72 °C for 90 seconds, followed by a final extension of 72 °C for 3 minutes. Products were purified, eluted in 30 µL sterile water and used as template for a second round of PCR with degenerate primers that anneal to chromosomal sequences flanking the transposon (Arb1 5'-GGCCAGCGAGCTAACGAGAC-3' and Almar3-seq 5'-ACATATCCATCGCGTCCGCC-3') with an

initial denaturation of 95 °C for 3 minutes followed by 30 cycles of 95 °C for 30 seconds, 56 °C for 30 seconds, and 72 °C for 90 seconds, followed by a final extension of 72 °C for 3 minutes. Amplicons were purified as described above, and were sequenced by Sanger sequencing using the Almar3-seq primer (5'-ACATATCCATCGCGTCCGCC-3').

5.10 Field Trials

5.10.1 Field Trials for Late Blight caused by *Phytophthora infestans*

5.10.1.1 Late blight Field Trial Study Design

Field trials to assess effectiveness of Ps652 as a biocontrol strain for protection of potato crops against late blight caused by *Phytophthora infestans* were conducted by VCS Potatoes (Suffolk, UK). The trial was carried out at a field site roughly 1 mile from Orford (Suffolk, UK). The potato variety used was Bambino, and the crop was planted on 27/07/2020. Fertilisation was 120kg of nitrogen per hectare in the form of N 26S liquid (Omex). The crop was irrigated every 6-7 days from 90% emergence until rain started in late September. No inoculation of plants with *Phytophthora* took place, as a natural infection was first noted on 11/09/2020. Crops were sprayed at 7-day intervals from 14/08/2020 (early rosette stage, approximately 10% ground cover) to 23/09/2020 with inocula of the biocontrol strain, an industry standard chemical oxathiapiprolin, or left untreated. Spraying was carried out using a knapsack plot sprayer. Ps652 was inoculated on plants at a density of 10^9 CFU/m². Treatments were organised in a Randomised Complete Block Design as shown in Figure 5.1, with 4 replicates per treatment. Plants were assessed by estimation of total % leaf coverage with *Phytophthora*, using a leaf area grid and counting the squares with more than 50% leaf area infected.

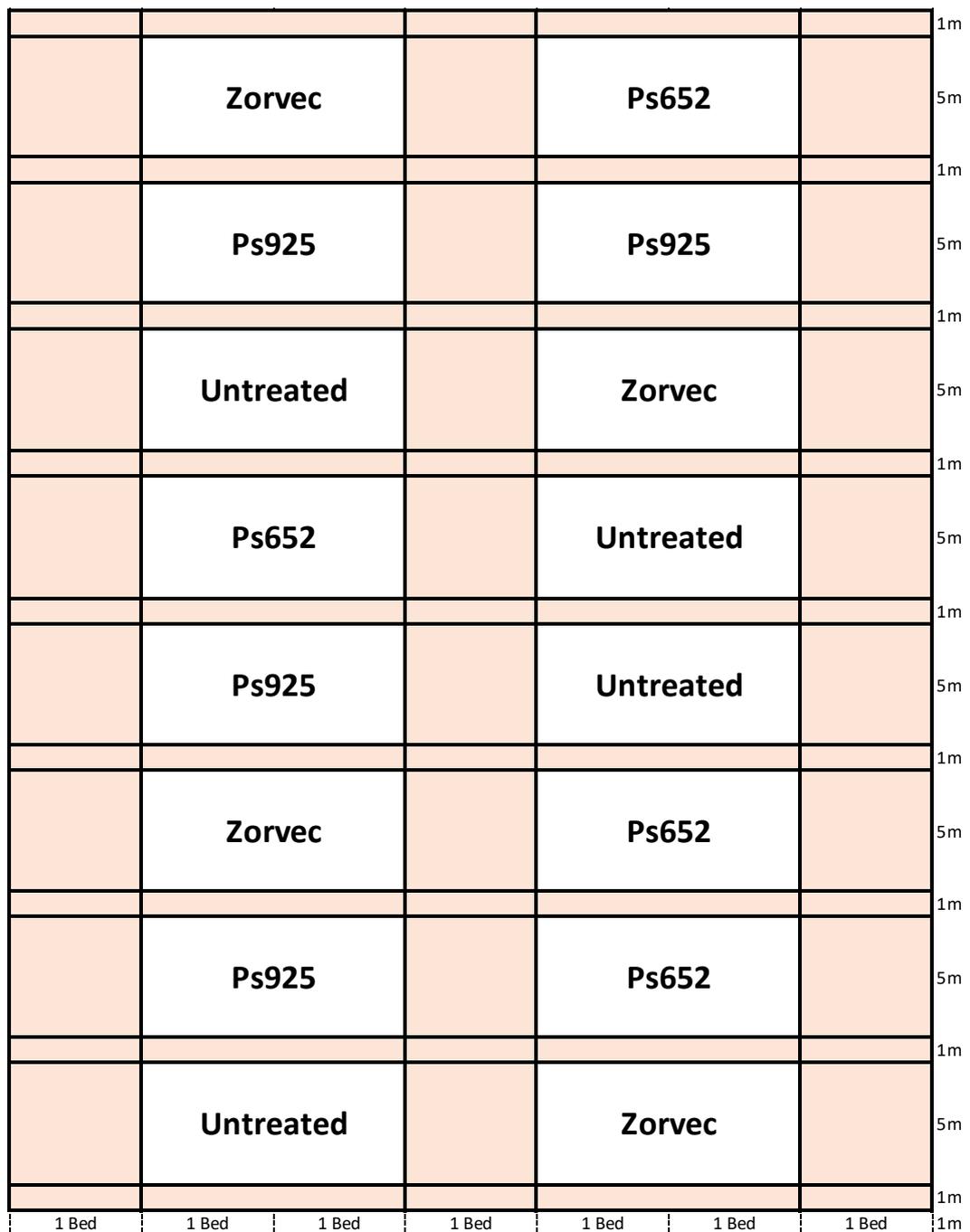


Figure 5.1. Randomised complete block design (RCBD) for late blight field trials. 4 replicates plots were used per strain, and distances are shown in metres. Zorvec is an oxathiapiproiin-based commercially available product for controlling late blight.

5.10.1.2 Preparation of Biocontrol Strain Inocula

Strains were prepared by lyophilisation of cells. Ps652 or Ps925 were grown overnight in King's B medium to $OD_{600} = 1.6$ and CFU/mL enumerated at this stage using serial dilution plates in triplicate. Cells collected by centrifugation at $8,000 \times g$ for 8 minutes, supernatant removed, and cells resuspended in an equivalent volume of PBS. This step was repeated, and then cells were again collected by centrifugation as above, and resuspended in 120 mL per 500 mL culture of 24% sucrose solution in a round bottom flask. The bacterial suspension was then frozen slowly at $-30\text{ }^{\circ}\text{C}$ with a rotation of 90° every 15 minutes until solid. This was then transferred to $-80\text{ }^{\circ}\text{C}$ for a minimum of 1 hour, before being freeze dried overnight. Dried material was then collected and weighed. Dry weight was divided by the original culture volume to get the weight per mL of original culture. CFU per mL of original culture were then enumerated as above, after redissolving and resuspending the relevant weight of dry material in 1 mL Milli-Q H_2O and leaving to stand for 10 minutes. Dried bacterial cells were divided into aliquots of 88×10^9 cells in parafilm-sealed falcon tubes, representing $10^9/\text{m}^2$ of area to be sprayed plus 10%. These were stored at $4\text{ }^{\circ}\text{C}$, and only removed to room temperature within 2 hours of application to the field site. Dried cells were added to the water reservoir of a knapsack plot sprayer with an equivalent of 200 L water per hectare and mixed thoroughly, before being sprayed onto the field site.

5.10.2 Field trials for Potato Common Scab Caused by *Streptomyces* sp.

5.10.2.1 Common Scab Field Trial Study Design

Field trials to assess effectiveness of Ps652 as a biocontrol strain for protection of potato crops against common scab caused by *Streptomyces* sp. were conducted by VCS Potatoes (Suffolk, UK) as shown in Figure 5.2. The trial was carried out at a field site at Tuddenham (Suffolk, UK). The potato variety used was Maris Piper, and the crop was planted on 06/04/2021. The crop was fertilised throughout the season with the following total amounts of each nutrient: Nitrogen, 250 kg / Ha; Potassium (K_2O), 170 kg / Ha; Phosphate (P_2O_5), 70 kg / Ha; Magnesium (MgO), 60 kg / Ha; Sulphur (SO_3), 120 kg / Ha. Organic manure was also used. The crop was irrigated every 6-7 days from 90% emergence for

normal irrigation treatment groups. Seed potatoes were planted in parallel beds with 30 cm spacing, and irrigated using drip irrigation as per standard practice. 20 tubers per biological treatment / irrigation regime combination (Ps652 and Ps925) were dipped by hand in an $OD_{600} = 1.0$ suspension of either strain immediately before planting, and the remainder were planted as normal with no dipping. Drip lines were used to apply approximately 50 mL of $OD_{600} = 0.2$ culture per plant for each treatment (Ps652 and Ps925) on 17/05/2021 and 28/05/2021 in addition to the standard irrigation schedule. Plants were split into 3 irrigation categories to account for fluctuations in environmental conditions: normal irrigation, half irrigation, and no irrigation. Severity of disease was established at the time of harvest by visually screening tubers for common scab lesions.

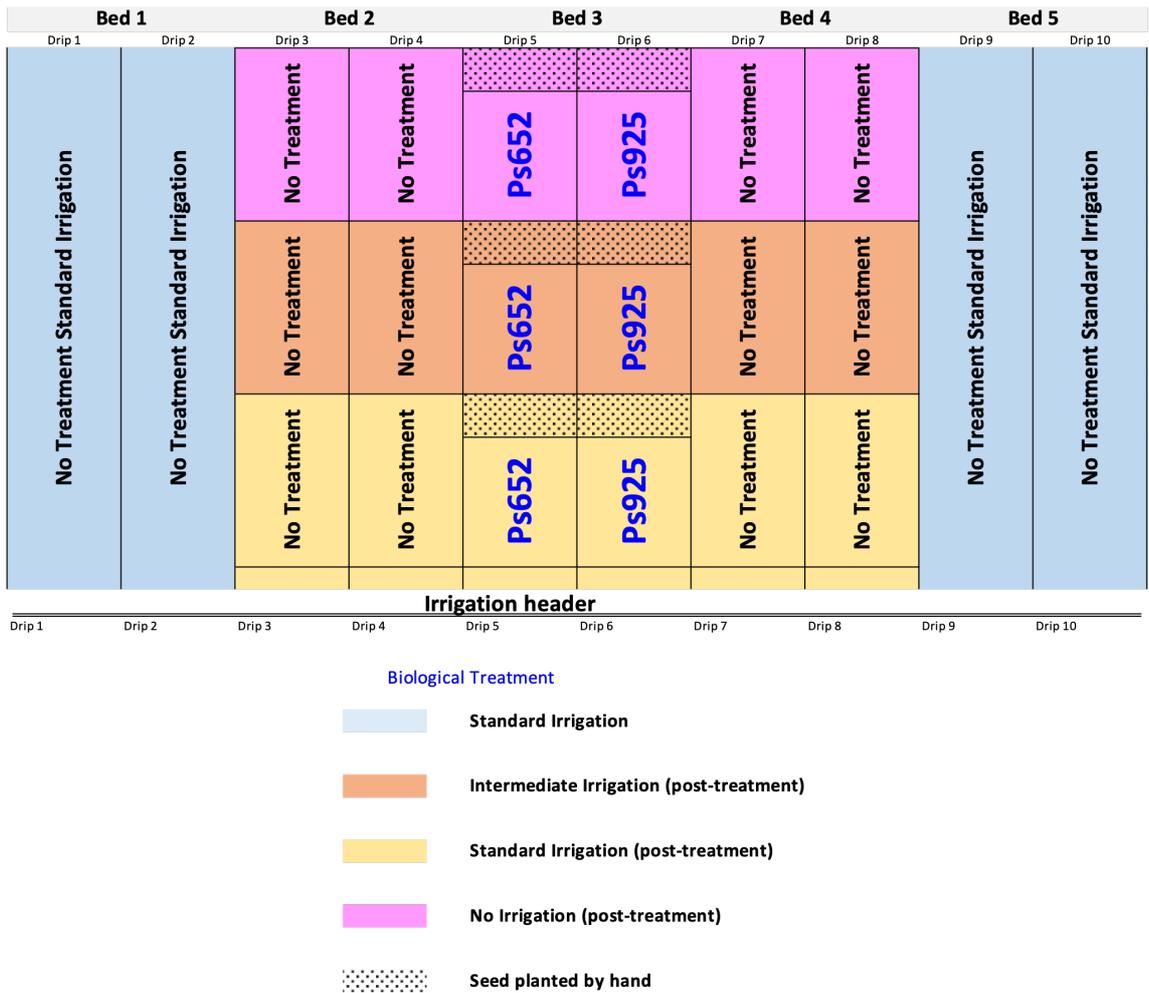


Figure 5.2. Field trial study design for control of potato common scab using Ps652 and Ps925 as biological treatments. Each bed is 210 m in length, totalling 70 m per irrigation treatment group. Seed tubers were planted with 30 cm spacing.

5.10.2.2 Preparation of Inocula for Dipping of Seed Tubers

Strains were prepared as liquid cultures to allow for the dipping of seed potatoes prior to planting. Both strains were streaked from stocks to single colonies on King's B plates, which were used to inoculate 5 x 200 mL King's broth for each strain, and grown for 16 hours at 28 °C with shaking at 250 rpm. Cells were harvested by centrifugation at 8,000 x g for 8 minutes at 4 °C, and the supernatant discarded. Cells were resuspended in 1 L PBS per strain, centrifuged as above, and the supernatant discarded. The PBS wash step was repeated once more, before cells were resuspended in PBS to an optical density at 600 nm of 1.0 ($OD_{600} = 1.0$). Cells were stored at 4-8 °C before being used the following day. For tuber dipping, each 1 L cell suspension was added to 4 L of tap water, and 200 tubers dipped before planting.

5.10.2.3 Preparation of Inocula for Drip Irrigation

Strains were prepared as concentrated liquid cultures for ease of application to drip irrigation lines. 50 mL of $OD_{600} = 0.2$ culture per plant had previously shown success for other strains in pot trials (A. Pacheco-Moreno, personal communication), and so was chosen for this work. Strains Ps652 and Ps925 were streaked from stocks to single colonies on King's B plates, which were used to inoculate 8 L of King's broth per strain as 1 L cultures in 2 L conical flasks. Strains were grown for 16 hours at 28 °C with shaking at 250 rpm. Cells were harvested by centrifugation at 8,000 x g for 8 minutes at 4 °C, and the supernatant discarded. As above, cells were washed twice in an equal volume of PBS, discarding the supernatant each time. Cells were subsequently resuspended in PBS a final $OD_{600} = 2.4$, and 3 L was used per strain. Cells were stored at 4-8 °C until use the following day.

References

1. Nichols, D., et al. (2010) Use of Ichip for High-Throughput In Situ Cultivation of "Uncultivable" Microbial Species. *Appl. Environ. Microbiol.*, 76, 2445-2450.
2. Sherpa, R. T., Reese, C. J. and Aliabadi, H. M. (2015) Application of iChip to Grow "Uncultivable" Microorganisms and its Impact on Antibiotic Discovery. *J Pharm Pharm Sci*, 18, 303-315.
3. Daniel, R. (2005) The metagenomics of soil. *Nature Reviews Microbiology*, 3, 470-478.
4. Stubbendieck, R. M. and Straight, P. D. (2016) Multifaceted Interfaces of Bacterial Competition. *J. Bacteriol.*, 198, 2145-2155.
5. Ryan, R. P., et al. (2008) Bacterial endophytes: recent developments and applications. *FEMS Microbiol. Lett.*, 278, 1-9.
6. Silby, M. W., et al. (2011) Pseudomonas genomes: diverse and adaptable. *FEMS Microbiol. Rev.*, 35, 652-680.
7. Köhl, J., Kolnaar, R., Ravensberg, W.J. (2019) Mode of Action of Microbial Biological Control Agents Against Plant Diseases: Relevance Beyond Efficacy. *Front. Plant. Sci.*, 10, 845.
8. Raaijmakers, J. M., et al. (2009) The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant Soil*, 321, 341-361.
9. Berendsen, R. L., Pieterse, C. M. and Bakker, P. A. (2012) The rhizosphere microbiome and plant health. *Trends Plant Sci.*, 17, 478-486.

10. Nicholson, A. J. (1954) An Outline of the Dynamics of Animal Populations. *Aust. J. Zool.*, 2, 9-65.
11. Hibbing, M. E., et al. (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, 8, 15-25.
12. Ghoul, M. and Mitri, S. (2016) The Ecology and Evolution of Microbial Competition. *Trends Microbiol.*, 24, 833-845.
13. Bais, H. P., et al. (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu. Rev. Plant Biol.*, 57, 233-266.
14. Leinweber, A., Weigert, M. and Kummerli, R. (2018) The bacterium *Pseudomonas aeruginosa* senses and gradually responds to interspecific competition for iron. *Evolution*, 72, 1515-1528.
15. Nielsen, T. H., et al. (2005) Genes involved in cyclic lipopeptide production are important for seed and straw colonization by *Pseudomonas* sp. strain DSS73. *Appl. Environ. Microbiol.*, 71, 4112-4116.
16. Kim, W., et al. (2014) Importance of positioning for microbial evolution. *Proc Natl Acad Sci U S A*, 111, E1639-1647.
17. Dorosky, R. J., et al. (2017) *Pseudomonas chlororaphis* Produces Two Distinct R-Tailocins That Contribute to Bacterial Competition in Biofilms and on Roots. *Appl. Environ. Microbiol.*, 83, e00706-00717.
18. Raaijmakers, J. M. and Mazzola, M. (2012) Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu. Rev. Phytopathol.*, 50, 403-424.

19. Burr, T. J., et al. (1996) Effectiveness of bacteria and yeasts from apple orchards as biological control agents of apple scab. *Biol. Control*, 6, 151-157.
20. Milus, E. A. and Rothrock, C. S. (1997) Efficacy of Bacterial Seed Treatments for Controlling Pythium Root Rot of Winter Wheat. *Plant Dis.*, 81, 180-184.
21. Arseneault, T. and Fillion, M. (2017) Biocontrol through antibiosis: exploring the role played by subinhibitory concentrations of antibiotics in soil and their impact on plant pathogens. *Canadian Journal of Plant Pathology*, 39, 267-274.
22. Price-Whelan, A., Dietrich, L. E. P. and Newman, D. K. (2006) Rethinking 'secondary' metabolism: physiological roles for phenazine antibiotics. *Nat. Chem. Biol.*, 2, 71-78.
23. Okegbe, C., et al. (2017) Electron-shuttling antibiotics structure bacterial communities by modulating cellular levels of c-di-GMP. *P. Natl. Acad. Sci. USA*, 114, E5236-E5245.
24. López, D., et al. (2009) Structurally diverse natural products that cause potassium leakage trigger multicellularity in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences*, 106, 280-285.
25. West, S. A., et al. (2012) Quorum sensing and the confusion about diffusion. *Trends Microbiol.*, 20, 586-594.
26. Chernin, L., et al. (2011) Quorum-sensing quenching by rhizobacterial volatiles. *Environ Microbiol Rep*, 3, 698-704.
27. Fetzner, S. (2015) Quorum quenching enzymes. *J. Biotechnol.*, 201, 2-14.
28. Wang, Y. J. and Leadbetter, J. R. (2005) Rapid acyl-homoserine lactone quorum signal biodegradation in diverse soils. *Appl. Environ. Microbiol.*, 71, 1291-1299.

29. Augustine, N., Kumar, P. and Thomas, S. (2010) Inhibition of *Vibrio cholerae* biofilm by AiiA enzyme produced from *Bacillus* spp. *Arch. Microbiol.*, 192, 1019-1022.
30. Glick, B. R. (2012) Plant growth-promoting bacteria: mechanisms and applications. *Scientifica (Cairo)*, 2012, 963401.
31. Ortiz-Castro, R., et al. (2011) Transkingdom signaling based on bacterial cyclodipeptides with auxin activity in plants. *Proc Natl Acad Sci U S A*, 108, 7253-7258.
32. Zamioudis, C., et al. (2013) Unraveling root developmental programs initiated by beneficial *Pseudomonas* spp. bacteria. *Plant Physiol.*, 162, 304-318.
33. Zhang, H., et al. (2007) Rhizobacterial volatile emissions regulate auxin homeostasis and cell expansion in *Arabidopsis*. *Planta*, 226, 839-851.
34. Pii, Y., et al. (2015) Microbial interactions in the rhizosphere: beneficial influences of plant growth-promoting rhizobacteria on nutrient acquisition process. A review. *Biol. Fertility Soils*, 51, 403-415.
35. Awwad, F., et al. (2019) Auxin protects *Arabidopsis thaliana* cell suspension cultures from programmed cell death induced by the cellulose biosynthesis inhibitors thaxtomin A and isoxaben. *BMC Plant Biol.*, 19, 512.
36. Pieterse, C. M., et al. (2014) Induced systemic resistance by beneficial microbes. *Annu. Rev. Phytopathol.*, 52, 347-375.
37. De Vleeschauwer, D. and Höfte, M. (2009) Chapter 6 Rhizobacteria-Induced Systemic Resistance. In: *Adv. Bot. Res.* Academic Press, pp. 223-281
38. Glazebrook, J. (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu. Rev. Phytopathol.*, 43, 205-227.

39. van de Mortel, J. E., et al. (2012) Metabolic and transcriptomic changes induced in Arabidopsis by the rhizobacterium *Pseudomonas fluorescens* SS101. *Plant Physiol.*, 160, 2173-2188.
40. van Wees, S. C., et al. (2000) Enhancement of induced disease resistance by simultaneous activation of salicylate- and jasmonate-dependent defense pathways in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*, 97, 8711-8716.
41. Millet, Y. A., et al. (2010) Innate immune responses activated in Arabidopsis roots by microbe-associated molecular patterns. *Plant Cell*, 22, 973-990.
42. Deng, Y., et al. (2019) Endophyte *Bacillus subtilis* evade plant defense by producing lantibiotic subtilomycin to mask self-produced flagellin. *Commun Biol*, 2, 368.
43. Stevens, R. B. (1960) Plant Pathology, an Advanced Treatise. Plant Pathology, an Advanced Treatise. Academic Press, NY
44. Velasquez, A. C., Castroverde, C. D. M. and He, S. Y. (2018) Plant-Pathogen Warfare under Changing Climate Conditions. *Curr. Biol.*, 28, R619-R634.
45. Cheng, Y. T., Zhang, L. and He, S. Y. (2019) Plant-Microbe Interactions Facing Environmental Challenge. *Cell Host Microbe*, 26, 183-192.
46. Bencze, S., et al. (2017) Rising atmospheric CO₂ concentration may imply higher risk of Fusarium mycotoxin contamination of wheat grains. *Mycotoxin Res*, 33, 229-236.
47. Dutta, B., et al. (2017) A Risk Assessment Model for Bacterial Leaf Spot of Pepper (*Capsicum annuum*), Caused by *Xanthomonas euvesicatoria*, Based on Concentrations of Macronutrients, Micronutrients, and Micronutrient Ratios. *Phytopathology*, 107, 1331-1338.

48. Hassan, J. A., et al. (2018) Soil mixture composition alters Arabidopsis susceptibility to *Pseudomonas syringae* infection. *Plant Direct*, 2, e00044.
49. Devleeschauwer, B., et al. (2017) High relative humidity pre-harvest reduces post-harvest proliferation of *Salmonella* in tomatoes. *Food Microbiol.*, 66, 55-63.
50. Wei, Z., et al. (2017) Seasonal variation in the biocontrol efficiency of bacterial wilt is driven by temperature-mediated changes in bacterial competitive interactions. *J. Appl. Ecol.*, 54, 1440-1448.
51. Bocsanczy, A. M., et al. (2014) Proteomic comparison of *Ralstonia solanacearum* strains reveals temperature dependent virulence factors. *BMC Genomics*, 15, 280.
52. Narisawa, K., et al. (2005) Effects of Pathogen Density, Soil Moisture, and Soil pH on Biological Control of Clubroot in Chinese Cabbage by *Heteroconium chaetospora*. *Plant Dis.*, 89, 285-290.
53. Cray, J. A., et al. (2016) Biocontrol agents promote growth of potato pathogens, depending on environmental conditions. *Microb Biotechnol*, 9, 330-354.
54. Mauchline, T. H. and Malone, J. G. (2017) Life in earth - the root microbiome to the rescue? *Curr. Opin. Microbiol.*, 37, 23-28.
55. Mauchline, T. H., et al. (2015) An analysis of *Pseudomonas* genomic diversity in take-all infected wheat fields reveals the lasting impact of wheat cultivars on the soil microbiota. *Environ. Microbiol.*, 17, 4764-4778.
56. Landa, B. B., et al. (2002) Differential ability of genotypes of 2,4-diacetylphloroglucinol-producing *Pseudomonas fluorescens* strains to colonize the roots of pea plants. *Appl. Environ. Microbiol.*, 68, 3226-3237.

57. Ramette, A., Moenne-Loccoz, Y. and Defago, G. (2003) Prevalence of fluorescent pseudomonads producing antifungal phloroglucinols and/or hydrogen cyanide in soils naturally suppressive or conducive to tobacco black root rot. *FEMS Microbiol. Ecol.*, 44, 35-43.
58. Haas, D. and Defago, G. (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nature Reviews Microbiology*, 3, 307-319.
59. Stockwell, V. O. and Stack, J. P. (2007) Using *Pseudomonas* spp. for integrated biological control. *Phytopathology*, 97, 244-249.
60. Mendes, R., et al. (2011) Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science*, 332, 1097-1100.
61. Gross, H. and Loper, J. E. (2009) Genomics of secondary metabolite production by *Pseudomonas* spp. *Nat Prod Rep*, 26, 1408-1446.
62. Cho, J. C. and Tiedje, J. M. (2000) Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Appl. Environ. Microbiol.*, 66, 5448-5456.
63. Chaube, H. S. and Pundhir, V. S. (2005) Crop Diseases and their Management. Crop Diseases and their Management. PHI Learning,
64. Wallace, R. L., Hirkala, D. L. and Nelson, L. M. (2018) Mechanisms of action of three isolates of *Pseudomonas fluorescens* active against postharvest grey mold decay of apple during commercial storage. *Biol. Control*, 117, 13-20.
65. Wallace, R. L., Hirkala, D. L. and Nelson, L. M. (2017) Postharvest biological control of blue mold of apple by *Pseudomonas fluorescens* during commercial storage and potential modes of action. *Postharvest Biol. Technol.*, 133, 1-11.

66. Wang, Z., et al. (2021) Potential of Volatile Organic Compounds Emitted by *Pseudomonas fluorescens* ZX as Biological Fumigants to Control Citrus Green Mold Decay at Postharvest. *J. Agric. Food Chem.*, 69, 2087-2098.
67. Iglesias, M. B., et al. (2018) Efficacy of *Pseudomonas graminis* CPA-7 against *Salmonella* spp. and *Listeria monocytogenes* on fresh-cut pear and setting up of the conditions for its commercial application. *Food Microbiol.*, 70, 103-112.
68. Alegre, I., et al. (2013) Antagonistic effect of *Pseudomonas graminis* CPA-7 against foodborne pathogens in fresh-cut apples under simulated commercial conditions. *Food Microbiol.*, 33, 139-148.
69. Kim, M. K., et al. (2008) Effect of *Pseudomonas* sp P7014 on the growth of edible mushroom *Pleurotus eryngii* in bottle culture for commercial production. *Bioresour. Technol.*, 99, 3306-3308.
70. Jablonowski, N. D., et al. (2012) Long-term persistence of various C-14-labeled pesticides in soils. *Environ. Pollut.*, 168, 29-36.
71. Bento, C. P. M., et al. (2016) Persistence of glyphosate and aminomethylphosphonic acid in loess soil under different combinations of temperature, soil moisture and light/darkness. *Sci. Total Environ.*, 572, 301-311.
72. Beketov, M. A., et al. (2013) Pesticides reduce regional biodiversity of stream invertebrates. *P. Natl. Acad. Sci. USA*, 110, 11039-11043.
73. Whitehorn, P. R., Wallace, C. and Vallejo-Marin, M. (2017) Neonicotinoid pesticide limits improvement in buzz pollination by bumblebees. *Sci Rep-Uk*, 7, 15562.
74. Desneux, N., Decourtye, A. and Delpuech, J. M. (2007) The sublethal effects of pesticides on beneficial arthropods. *Annu. Rev. Entomol.*, 52, 81-106.

75. Cheng, Y. H., et al. (2017) PBPK/PD assessment for Parkinson's disease risk posed by airborne pesticide paraquat exposure. *Environ Sci Pollut Res Int*, 25, 5359-5368.
76. Aktar, M. W., Sengupta, D. and Chowdhury, A. (2009) Impact of pesticides use in agriculture: their benefits and hazards. *Interdisciplinary Toxicology*, 2, 1-12.
77. Powell, J. F., et al. (2000) Management of dollar spot on creeping bentgrass with metabolites of *Pseudomonas aureofaciens* (TX-1). *Plant Dis.*, 84, 19-24.
78. Sigler, W. V., et al. (2001) Fate of the biological control agent *Pseudomonas aureofaciens* TX-1 after application to turfgrass. *Appl. Environ. Microbiol.*, 67, 3542-3548.
79. Hardebeck, G. A., et al. (2004) Application of *Pseudomonas aureofaciens* Tx-1 through irrigation for control of dollar spot and brown patch on fairway-height turf. *HortScience*, 39, 1750-1753.
80. Burges, H. D. (2012)
) Formulation of Microbial Biopesticides: Beneficial microorganisms, nematodes and seed treatments. 1 Formulation of Microbial Biopesticides: Beneficial microorganisms, nematodes and seed treatments. Springer Netherlands, Dordrecht
81. Miller, F. C. and Spear, M. (1995) Very large scale commercial trial of a biological control for mushroom blotch disease. *Science and Cultivation of Edible Fungi, Vols 1 and 2*, doi: 635-642.
82. Eberl, L. and Vandamme, P. (2016) Members of the genus *Burkholderia*: good and bad guys. *F1000Res*, 5, F1000 Faculty Rev-1007.
83. Cedomon®. <http://www.bioagri.se/?p=30922> (29 March)
84. Cerall®. <http://www.bioagri.se/?p=30838> (29 March)

85. Cedress®. <http://www.bioagri.se/?p=30923> (29 March)
86. Alberts, J. F., van Zyl, W. H. and Gelderblom, W. C. (2016) Biologically Based Methods for Control of Fumonisin-Producing *Fusarium* Species and Reduction of the Fumonisin. *Front Microbiol*, 7, 548.
87. U.S. Environmental Protection Agency. (2017) Notice of pesticide registration: Howler, EPA reg. number 91197-3. Office of Pesticide Programs Biopesticides and Pollution Prevention Division (7511P). U.S. Environmental Protection Agency. Washington, D.C.
88. U.S. Environmental Protection Agency. (2014) Notice of pesticide registration: D7, EPA reg. number 71975-4. Office of Pesticide Programs Biopesticides and Pollution Prevention Division (7511P). U.S. Environmental Protection Agency. Washington, D.C.
89. European Food Safety Authority. (2012) Conclusion on the peer review of the pesticide risk assessment of the active substance *Pseudomonas* sp. strain DSMZ 13134. *EFSA Journal*, 10, 2954.
90. Santos Kron, A., et al. (2020) *Pseudomonas orientalis* F9 pyoverdine, safracin and phenazine mutants remain effective antagonists against *Erwinia amylovora* in apple flowers. *Appl. Environ. Microbiol.*, 86, e02620-02619.
91. Rodriguez, M., et al. (2020) Plant growth-promoting activity and quorum quenching-mediated biocontrol of bacterial phytopathogens by *Pseudomonas segetis* strain P6. *Sci Rep*, 10, 4121.
92. Plyuta, V., et al. (2016) Influence of volatile organic compounds emitted by *Pseudomonas* and *Serratia* strains on *Agrobacterium tumefaciens* biofilms. *Apmis*, 124, 586-594.

93. Raza, W., et al. (2016) Volatile organic compounds produced by *Pseudomonas fluorescens* WR-1 restrict the growth and virulence traits of *Ralstonia solanacearum*. *Microbiol. Res.*, 192, 103-113.
94. Elsayed, T. R., et al. (2020) Biocontrol of Bacterial Wilt Disease Through Complex Interaction Between Tomato Plant, Antagonists, the Indigenous Rhizosphere Microbiota, and *Ralstonia solanacearum*. *Front. Microbiol.*, 10, 2835.
95. St-Onge, R., et al. (2011) The ability of *Pseudomonas* sp. LBUM 223 to produce phenazine-1-carboxylic acid affects the growth of *Streptomyces scabies*, the expression of thaxtomin biosynthesis genes and the biological control potential against common scab of potato. *FEMS Microbiol. Ecol.*, 75, 173-183.
96. Bauer, J. S., et al. (2015) Biosynthetic Origin of the Antibiotic Pseudopyronines A and B in *Pseudomonas putida* BW11M1. *ChemBioChem*, 16, 2491-2497.
97. Li, W., et al. (2013) The antimicrobial compound xantholysin defines a new group of *Pseudomonas* cyclic lipopeptides. *PLoS One*, 8, e62946.
98. Mansfield, J., et al. (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.*, 13, 614-629.
99. Cotter, P. D., Hill, C. and Ross, R. P. (2005) Bacteriocins: developing innate immunity for food. *Nat. Rev. Microbiol.*, 3, 777-788.
100. Coutinho, H. D. M., et al. (2008) Peptides and proteins with antimicrobial activity. *Indian J Pharmacol*, 40, 3-9.
101. Principe, A., et al. (2018) Effectiveness of tailocins produced by *Pseudomonas fluorescens* SF4c in controlling the bacterial-spot disease in tomatoes caused by *Xanthomonas vesicatoria*. *Microbiol. Res.*, 212-213, 94-102.

102. Baltrus, D., et al. (2021) Prophylactic Application of Tailocins Prevents Infection by *Pseudomonas syringae*. *Phytopathology*, doi: 10.1094/PHYTO-06-21-0269-R
103. Hunziker, L., et al. (2015) *Pseudomonas* Strains Naturally Associated with Potato Plants Produce Volatiles with High Potential for Inhibition of *Phytophthora infestans*. *Appl. Environ. Microbiol.*, 81, 821-830.
104. Weisskopf, L. (2013) The potential of bacterial volatiles for crop protection against phytopathogenic fungi. In: Méndez-Vilas, A., *Microbial pathogens and strategies for combating them: science, technology and education*. Formatex Research Center, pp. 1352-1363
105. Rojas-Solis, D., et al. (2018) *Pseudomonas stutzeri* E25 and *Stenotrophomonas maltophilia* CR71 endophytes produce antifungal volatile organic compounds and exhibit additive plant growth-promoting effects. *Biocatal Agric Biote*, 13, 46-52.
106. Zhang, Y., et al. (2019) Volatile Organic Compounds Produced by *Pseudomonas chlororaphis* subsp. *aureofaciens* SPS-41 as Biological Fumigants To Control *Ceratocystis fimbriata* in Postharvest Sweet Potatoes. *J. Agric. Food Chem.*, 67, 3702–3710.
107. Sindhu, S. S. and Dadarwal, K. R. (2001) Chitinolytic and cellulolytic *Pseudomonas* sp. antagonistic to fungal pathogens enhances nodulation by *Mesorhizobium* sp. Cicer in chickpea. *Microbiol. Res.*, 156, 353-358.
108. Oliver, C., et al. (2019) *Pseudomonas putida* strain B2017 produced as technical grade active ingredient controls fungal and bacterial crop diseases. *Biocontrol Sci. Technol.*, 29, 1053-1068.
109. Daura-Pich, O., et al. (2020) No antibiotic and toxic metabolites produced by the biocontrol agent *Pseudomonas putida* strain B2017. *FEMS Microbiol. Lett.*, 367,

110. Miller, C. M., et al. (1998) Ecomycins, unique antimycotics from *Pseudomonas viridiflava*. *J. Appl. Microbiol.*, 84, 937-944.
111. Wang, S. L., Yieh, T. C. and Shih, I. L. (1999) Purification and characterization of a new antifungal compound produced by *Pseudomonas aeruginosa* K-187 in a shrimp and crab shell powder medium. *Enzyme Microb. Technol.*, 25, 439-446.
112. Yen, Y. H., et al. (2006) An antifungal protease produced by *Pseudomonas aeruginosa* M-1001 with shrimp and crab shell powder as a carbon source. *Enzyme Microb. Technol.*, 39, 311-317.
113. Muller, T., et al. (2018) Antagonistic Potential of Fluorescent *Pseudomonads* Colonizing Wheat Heads Against Mycotoxin Producing *Alternaria* and *Fusaria*. *Front Microbiol*, 9, 2124.
114. Ren, Y., et al. (2021) Isolation and characterization of a *Pseudomonas poae* JSU-Y1 with patulin degradation ability and biocontrol potential against *Penicillium expansum*. *Toxicon*, 195, 1-6.
115. Zhang, J., et al. (2020) *Pseudomonas synxantha* 2-79 Transformed with Pyrrolnitrin Biosynthesis Genes Has Improved Biocontrol Activity Against Soilborne Pathogens of Wheat and Canola. *Phytopathology*®, 110, 1010-1017.
116. Grosch, R., et al. (2005) Effectiveness of 3 antagonistic bacterial isolates to control *Rhizoctonia solani* Kuhn on lettuce and potato. *Can J Microbiol*, 51, 345-353.
117. Schreiter, S., et al. (2018) Rhizosphere Competence and Biocontrol Effect of *Pseudomonas* sp RU47 Independent from Plant Species and Soil Type at the Field Scale. *Front. Microbiol.*, 9, 97.

118. He, Y. X., et al. (2004) Importance of 2,4-DAPG in the biological control of brown patch by *Pseudomonas fluorescens* HP72 and newly identified genes involved in 2,4-DAPG biosynthesis. *Soil Sci. Plant Nutr.*, 50, 1287-1293.
119. Yang, M. M., et al. (2011) Biological control of take-all by fluorescent *Pseudomonas* spp. from Chinese wheat fields. *Phytopathology*, 101, 1481-1491.
120. Yang, F. and Cao, Y. (2012) Biosynthesis of phloroglucinol compounds in microorganisms--review. *Appl. Microbiol. Biotechnol.*, 93, 487-495.
121. Raaijmakers, J. M. and Weller, D. M. (1998) Natural plant protection by 2,4-diacetylphloroglucinol - Producing *Pseudomonas* spp. in take-all decline soils. *Mol. Plant-Microbe Interact.*, 11, 144-152.
122. Neesemann, K., et al. (2018) Fluorescent pseudomonads pursue media-dependent strategies to inhibit growth of pathogenic *Verticillium* fungi. *Appl. Microbiol. Biotechnol.*, 102, 817-831.
123. Stockwell, V. O., et al. (2010) Control of Fire Blight by *Pseudomonas fluorescens* A506 and *Pantoea vagans* C9-1 Applied as Single Strains and Mixed Inocula. *Phytopathology*, 100, 1330-1339.
124. Kong, W. L., et al. (2020) Forest Tree Associated Bacterial Diffusible and Volatile Organic Compounds against Various Phytopathogenic Fungi. *Microorganisms*, 8, 590.
125. Sedlakova, V., et al. (2011) Effect of *Phytophthora infestans* on potato yield in dependence on variety characteristics and fungicide control. *Plant Soil Environ*, 57, 486-491.
126. Anand, A., et al. (2020) Contribution of Hydrogen Cyanide to the Antagonistic Activity of *Pseudomonas* Strains Against *Phytophthora infestans*. *Microorganisms*, 8, 1144.

127. Park, J. Y., et al. (2011) Production of the antifungal compounds phenazine and pyrrolnitrin from *Pseudomonas chlororaphis* O6 is differentially regulated by glucose. *Let. Appl. Microbiol.*, 52, 532-537.
128. Roquigny, R., et al. (2018) Transcriptome alteration in *Phytophthora infestans* in response to phenazine-1-carboxylic acid production by *Pseudomonas fluorescens* strain LBUM223. *BMC Genomics*, 19, 474.
129. El-Sayed, A. S. A., et al. (2018) A glucanolytic *Pseudomonas* sp. associated with *Smilax bona-nox* L. displays strong activity against *Phytophthora parasitica*. *Microbiol. Res.*, 207, 140-152.
130. Guerriero, G., et al. (2010) Chitin synthases from *Saprolegnia* are involved in tip growth and represent a potential target for anti-oomycete drugs. *PLoS Pathog*, 6, e1001070.
131. Sowanpreecha, R. and Rerngsamran, P. (2018) Biocontrol of Orchid-pathogenic Mold, *Phytophthora palmivora*, by Antifungal Proteins from *Pseudomonas aeruginosa* RS1. *Mycobiology*, 46, 129-137.
132. Kamoun, S., et al. (2015) The Top 10 oomycete pathogens in molecular plant pathology. *Mol. Plant Pathol.*, 16, 413-434.
133. Djonovic, S., Pozo, M. J. and Kenerley, C. M. (2006) Tvbg3, a beta-1,6-glucanase from the biocontrol fungus *Trichoderma virens*, is involved in mycoparasitism and control of *Pythium ultimum*. *Appl. Environ. Microbiol.*, 72, 7661-7670.
134. Teoh, M. C., Furusawa, G. and Veera Singham, G. (2021) Multifaceted interactions between the pseudomonads and insects: mechanisms and prospects. *Arch. Microbiol.*, 203, 1891-1915.

135. Flury, P., et al. (2017) Antimicrobial and Insecticidal: Cyclic Lipopeptides and Hydrogen Cyanide Produced by Plant-Beneficial *Pseudomonas* Strains CHA0, CMR12a, and PCL1391 Contribute to Insect Killing. *Front. Microbiol.*, 8, 100.
136. Pechy-Tarr, M., et al. (2008) Molecular analysis of a novel gene cluster encoding an insect toxin in plant-associated strains of *Pseudomonas fluorescens*. *Environ. Microbiol.*, 10, 2368-2386.
137. Ruffner, B., et al. (2015) Evolutionary patchwork of an insecticidal toxin shared between plant-associated pseudomonads and the insect pathogens *Photorhabdus* and *Xenorhabdus*. *BMC Genomics*, 16, 609.
138. Olcott, M. H., et al. (2010) Lethality and Developmental Delay in *Drosophila melanogaster* Larvae after Ingestion of Selected *Pseudomonas fluorescens* Strains. *Plos One*, 5, e12504.
139. Iwasaki, S., et al. (1984) Studies on Macrocyclic Lactone Antibiotics .7. Structure of a Phytotoxin Rhizoxin Produced by *Rhizopus-Chinensis*. *J. Antibiot.*, 37, 354-362.
140. Loper, J. E., et al. (2016) Rhizoxin analogs, orfamide A and chitinase production contribute to the toxicity of *Pseudomonas protegens* strain Pf-5 to *Drosophila melanogaster*. *Environ. Microbiol.*, 18, 3509-3521.
141. Brendel, N., et al. (2007) A cryptic PKS-NRPS gene locus in the plant commensal *Pseudomonas fluorescens* Pf-5 codes for the biosynthesis of an antimitotic rhizoxin complex. *Org Biomol Chem*, 5, 2211-2213.
142. Ruffner, B., et al. (2013) Oral insecticidal activity of plant-associated pseudomonads. *Environ. Microbiol.*, 15, 751-763.

143. Kang, B. R., Anderson, A. J. and Kim, Y. C. (2018) Hydrogen Cyanide Produced by *Pseudomonas chlororaphis* O6 Exhibits Nematicidal Activity against *Meloidogyne hapla*. *Plant Pathol J*, 34, 35-43.
144. Kang, B. R., Anderson, A. J. and Kim, Y. C. (2018) Hydrogen cyanide produced by *Pseudomonas chlororaphis* O6 is a key aphicidal metabolite. *Can J Microbiol*, 65, 185-190.
145. Vacheron, J., et al. (2019) T6SS contributes to gut microbiome invasion and killing of an herbivorous pest insect by plant-beneficial *Pseudomonas protegens*. *ISME J*, 13, 1318–1329.
146. Pineda, A., et al. (2010) Helping plants to deal with insects: the role of beneficial soil-borne microbes. *Trends Plant Sci.*, 15, 507-514.
147. Saravanakumar, D., et al. (2007) *Pseudomonas*-induced defence molecules in rice plants against leaffolder (*Cnaphalocrocis medinalis*) pest. *Pest Manage. Sci.*, 63, 714-721.
148. Gandhi, P. I., Gunasekaran, K. and Sa, T. (2006) Neem oil as a potential seed dresser for managing Homopterous sucking pests of Okra (*Abelmoschus esculentus* (L.) Moench). *J. Pest Sci.*, 79, 103-111.
149. Van Oosten, V. R., et al. (2008) Differential effectiveness of microbially induced resistance against herbivorous insects in *Arabidopsis*. *Mol. Plant-Microbe Interact.*, 21, 919-930.
150. Zehnder, G., et al. (1997) Induction of systemic resistance in cucumber against cucumber beetles (Coleoptera: Chrysomelidae) by plant growth-promoting rhizobacteria. *J. Econ. Entomol.*, 90, 391-396.
151. Flury, P., et al. (2016) Insect pathogenicity in plant-beneficial pseudomonads: phylogenetic distribution and comparative genomics. *Isme Journal*, 10, 2527-2542.

152. Keel, C. (2016) A look into the toolbox of multi-talents: insect pathogenicity determinants of plant-beneficial pseudomonads. *Environ. Microbiol.*, 18, 3207-3209.
153. Rastogi, G., Coaker, G. L. and Leveau, J. H. J. (2013) New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. *FEMS Microbiol. Lett.*, 348, 1-10.
154. Tao, C., et al. (2020) Bio-organic fertilizers stimulate indigenous soil *Pseudomonas* populations to enhance plant disease suppression. *Microbiome*, 8, 137.
155. Davies, J. and Ryan, K. S. (2012) Introducing the Parvome: Bioactive Compounds in the Microbial World. *Acs Chem Biol*, 7, 252-259.
156. Ling, L. L., et al. (2015) A new antibiotic kills pathogens without detectable resistance. *Nature*, 517, 455-459.
157. Som, N. F., et al. (2017) The Conserved Actinobacterial Two-Component System MtrAB Coordinates Chloramphenicol Production with Sporulation in *Streptomyces venezuelae* NRRL B-65442. *Front. Microbiol.*, 8, 1145.
158. Som, N. F., et al. (2017) The MtrAB two-component system controls antibiotic production in *Streptomyces coelicolor* A3(2). *Microbiol-Sgm*, 163, 1415-1419.
159. Choi, S. S., et al. (2015) Genome mining of rare actinomycetes and cryptic pathway awakening. *Process Biochem.*, 50, 1184-1193.
160. Ziemert, N., et al. (2014) Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc Natl Acad Sci U S A*, 111, E1130-1139.
161. Qin, Z., et al. (2017) Formicamycins, antibacterial polyketides produced by *Streptomyces formicae* isolated from African *Tetraponera* plant-ants. *Chem Sci*, 8, 3218-3227.

162. Heine, D., et al. (2018) Chemical warfare between leafcutter ant symbionts and a co-evolved pathogen. *Nat. Commun.*, 9,
163. Xu, L., et al. (2019) Two New Prenylated Indole Diterpenoids from *Tolypocladium* sp. and Their Antimicrobial Activities. *Chem. Biodivers.*, 16, e1900116.
164. Mori, T., et al. (2018) Single-bacterial genomics validates rich and varied specialized metabolism of uncultivated *Entotheonella* sponge symbionts. *Proc Natl Acad Sci U S A*, 115, 1718-1723.
165. Lincke, T., et al. (2010) Closthioamide: an unprecedented polythioamide antibiotic from the strictly anaerobic bacterium *Clostridium cellulolyticum*. *Angew Chem Int Ed Engl*, 49, 2011-2013.
166. Medema, M. H., et al. (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.*, 39, 339-346.
167. Hutchings, M. I., Truman, A. W. and Wilkinson, B. (2019) Antibiotics: past, present and future. *Curr. Opin. Microbiol.*, 51, 72-80.
168. Challis, G. L., Ravel, J. and Townsend, C. A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.*, 7, 211-224.
169. Stachelhaus, T., Mootz, H. D. and Marahiel, M. A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.*, 6, 493-505.
170. Haydock, S. F., et al. (1995) Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Lett.*, 374, 246-248.

171. Santos-Aberturas, J., et al. (2019) Uncovering the unexplored diversity of thioamidated ribosomal peptides in Actinobacteria using the RiPPER genome mining tool. *Nucleic Acids Res.*, 47, 4624-4637.
172. Tietz, J. I., et al. (2017) A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.*, 13, 470-478.
173. Merwin, N. J., et al. (2020) DeepRiPP integrates multiomics data to automate discovery of novel ribosomally synthesized natural products. *Proc Natl Acad Sci U S A*, 117, 371-380.
174. Rutledge, P. J. and Challis, G. L. (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.*, 13, 509-523.
175. Nai, C. and Meyer, V. (2018) From Axenic to Mixed Cultures: Technological Advances Accelerating a Paradigm Shift in Microbiology. *Trends Microbiol.*, 26, 538-554.
176. Tanaka, Y., Hosaka, T. and Ochi, K. (2010) Rare earth elements activate the secondary metabolite-biosynthetic gene clusters in *Streptomyces coelicolor* A3(2). *J Antibiot (Tokyo)*, 63, 477-481.
177. Mao, D., et al. (2020) Reporter-Guided Transposon Mutant Selection for Activation of Silent Gene Clusters in *Burkholderia thailandensis*. *ChemBioChem*, 21, 1826-1831.
178. Paterson, J., et al. (2017) The contribution of genome mining strategies to the understanding of active principles of PGPR strains. *FEMS Microbiol. Ecol.*, 93, fiw249.
179. Sagova-Mareckova, M., et al. (2015) Determination of Factors Associated with Natural Soil Suppressivity to Potato Common Scab. *Plos One*, 10, e0116291.
180. Garrido-Sanz, D., et al. (2016) Genomic and Genetic Diversity within the *Pseudomonas fluorescens* Complex. *Plos One*, 11, e0150183.

181. Hmelo, L. R., et al. (2015) Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nat Protoc*, 10, 1820-1841.
182. Nguyen, D. D., et al. (2016) Indexing the *Pseudomonas* specialized metabolome enabled the discovery of poeamide B and the bananamides. *Nat Microbiol*, 2, 16197.
183. Weller, D. M., et al. (2007) Role of 2,4-diacetylphloroglucinol-producing fluorescent *Pseudomonas* spp. in the defense of plant roots. *Plant Biol.*, 9, 4-20.
184. Ramette, A., et al. (2011) *Pseudomonas protegens* sp nov., widespread plant-protecting bacteria producing the biocontrol compounds 2,4-diacetylphloroglucinol and pyoluteorin. *Syst. Appl. Microbiol.*, 34, 180-188.
185. Masschelein, J., Jenner, M. and Challis, G. L. (2017) Antibiotics from Gram-negative bacteria: a comprehensive overview and selected biosynthetic highlights. *Nat Prod Rep*, 34, 712-783.
186. Brucker, R. M., et al. (2008) The identification of 2,4-diacetylphloroglucinol as an antifungal metabolite produced by cutaneous bacteria of the salamander *Plethodon cinereus*. *J. Chem. Ecol.*, 34, 39-43.
187. Banger, M. G. and Thomashow, L. S. (1999) Identification and characterization of a gene cluster for synthesis of the polyketide antibiotic 2,4-diacetylphloroglucinol from *Pseudomonas fluorescens* Q2-87. *J. Bacteriol.*, 181, 3155-3163.
188. Rezzonico, F., et al. (2007) Is the ability of biocontrol fluorescent pseudomonads to produce the antifungal metabolite 2,4-diacetylphloroglucinol really synonymous with higher plant protection? *New Phytol.*, 173, 861-872.

189. Yang, M., et al. (2020) Exploring the Pathogenicity of *Pseudomonas brassicacearum* Q8r1-96 and Other Strains of the *Pseudomonas fluorescens* Complex on Tomato. *Plant Dis.*, 104, 1026-1031.
190. Notz, R., et al. (2001) Biotic Factors Affecting Expression of the 2,4-Diacetylphloroglucinol Biosynthesis Gene *phIA* in *Pseudomonas fluorescens* Biocontrol Strain CHA0 in the Rhizosphere. *Phytopathology*, 91, 873-881.
191. Fischbach, M. A. and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem Rev*, 106, 3468-3496.
192. Yu, D., et al. (2012) Type III polyketide synthases in natural product biosynthesis. *IUBMB Life*, 64, 285-295.
193. Achkar, J., et al. (2005) Biosynthesis of phloroglucinol. *J. Am. Chem. Soc.*, 127, 5332-5333.
194. Zha, W. J., Rubin-Pitel, S. B. and Zhao, H. M. (2006) Characterization of the substrate specificity of PhID, a type III polyketide synthase from *Pseudomonas fluorescens*. *J. Biol. Chem.*, 281, 32036-32047.
195. Abbas, A., et al. (2002) Characterization of interactions between the transcriptional repressor PhIF and its binding site at the *phIA* promoter in *Pseudomonas fluorescens* F113. *J. Bacteriol.*, 184, 3008-3016.
196. Bottiglieri, M. and Keel, C. (2006) Characterization of PhIG, a hydrolase that specifically degrades the antifungal compound 2,4-diacetylphloroglucinol in the biocontrol agent *Pseudomonas fluorescens* CHA0. *Appl. Environ. Microbiol.*, 72, 418-427.

197. Schnider-Keel, U., et al. (2000) Autoinduction of 2,4-diacetylphloroglucinol biosynthesis in the biocontrol agent *Pseudomonas fluorescens* CHA0 and repression by the bacterial metabolites salicylate and pyoluteorin. *J. Bacteriol.*, 182, 1215-1225.
198. Laville, J., et al. (1992) Global control in *Pseudomonas fluorescens* mediating antibiotic synthesis and suppression of black root rot of tobacco. *Proc. Natl. Acad. Sci. USA*, 89, 1562-1566.
199. Schnider, U., et al. (1995) Amplification of the housekeeping sigma factor in *Pseudomonas fluorescens* CHA0 enhances antibiotic production and improves biocontrol abilities. *J. Bacteriol.*, 177, 5387-5392.
200. Sarniguet, A., et al. (1995) The sigma factor sigma s affects antibiotic production and biological control activity of *Pseudomonas fluorescens* Pf-5. *Proc. Natl. Acad. Sci. USA*, 92, 12255-12259.
201. Walsh, C. T. (2008) The chemical versatility of natural-product assembly lines. *Accounts Chem Res*, 41, 4-10.
202. Winn, M., et al. (2016) Recent advances in engineering nonribosomal peptide assembly lines. *Nat Prod Rep*, 33, 317-347.
203. Raaijmakers, J. M., et al. (2010) Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol. Rev.*, 34, 1037-1062.
204. Rokni-Zadeh, H., et al. (2012) Genetic and Functional Characterization of Cyclic Lipopeptide White-Line-Inducing Principle (WLIP) Production by Rice Rhizosphere Isolate *Pseudomonas putida* RW10S2. *Appl. Environ. Microbiol.*, 78, 4826-4834.
205. Alsohim, A. S., et al. (2014) The biosurfactant viscosin produced by *Pseudomonas fluorescens* SBW25 aids spreading motility and plant growth promotion. *Environ. Microbiol.*, 16, 2267-2281.

206. Bonnichsen, L., et al. (2015) Lipopeptide biosurfactant viscosin enhances dispersal of *Pseudomonas fluorescens* SBW25 biofilms. *Microbiol-Sgm*, 161, 2289-2297.
207. de Bruijn, I., et al. (2007) Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in *Pseudomonas* species. *Mol. Microbiol.*, 63, 417-428.
208. Boddy, C. N. (2014) Bioinformatics tools for genome mining of polyketide and non-ribosomal peptides. *J. Ind. Microbiol. Biotechnol.*, 41, 443-450.
209. Bozhüyük, K. A. J., et al. (2017) De novo design and engineering of non-ribosomal peptide synthetases. *Nature Chemistry*, 10, 275-281.
210. Muller, C. A., et al. (2015) Mining for Nonribosomal Peptide Synthetase and Polyketide Synthase Genes Revealed a High Level of Diversity in the Sphagnum Bog Metagenome. *Appl. Environ. Microbiol.*, 81, 5064-5072.
211. Machado, H., et al. (2015) Genome mining reveals unlocked bioactive potential of marine Gram-negative bacteria. *BMC Genomics*, 16, Article number: 158.
212. Takeda, R. (1959) *Pseudomonas* Pigments .3. Derivatives of Pyoluteorin. *B Agr Chem Soc Japan*, 23, 126-130.
213. Howell, C. R. and Stipanovic, R. D. (1980) Suppression of *Pythium-Ultimum*-Induced Damping-Off of Cotton Seedlings by *Pseudomonas-Fluorescens* and Its Antibiotic, Pyoluteorin. *Phytopathology*, 70, 712-715.
214. Maurhofer, M., et al. (1995) Influence of Plant-Species on Disease Suppression by *Pseudomonas-Fluorescens* Strain Chao with Enhanced Antibiotic Production. *Plant Pathol.*, 44, 40-50.

215. Blankenfeldt, W. and Parsons, J. F. (2014) The structural biology of phenazine biosynthesis. *Curr. Opin. Struct. Biol.*, 29, 26-33.
216. Dar, D., et al. (2020) Global landscape of phenazine biosynthesis and biodegradation reveals species-specific colonization patterns in agricultural soils and crop microbiomes. *Elife*, 9,
217. McDonald, M., et al. (2001) Phenazine biosynthesis in *Pseudomonas fluorescens*: Branchpoint from the primary shikimate biosynthetic pathway and role of phenazine-1,6-dicarboxylic acid. *J. Am. Chem. Soc.*, 123, 9459-9460.
218. Li, W., et al. (2016) Functional identification of phenazine biosynthesis genes in plant pathogenic bacteria *Pseudomonas syringae* pv. tomato and *Xanthomonas oryzae* pv. *oryzae*. *Journal of Integrative Agriculture*, 15, 812-821.
219. Guo, S., et al. (2020) Microbial Synthesis of Antibacterial Phenazine-1,6-dicarboxylic Acid and the Role of PhzG in *Pseudomonas chlororaphis* GP72AN. *J. Agric. Food Chem.*, 68, 2373-2380.
220. Schmidt, R., et al. (2015) Volatile affairs in microbial interactions. *Isme Journal*, 9, 2329-2335.
221. Vidal, D. M., et al. (2017) Long-Chain Alkyl Cyanides: Unprecedented Volatile Compounds Released by *Pseudomonas* and *Micromonospora* Bacteria. *Angew Chem Int Edit*, 56, 4342-4346.
222. Ryu, C. M., et al. (2004) Bacterial volatiles induce systemic resistance in *Arabidopsis*. *Plant Physiol.*, 134, 1017-1026.
223. Ryu, C. M., et al. (2003) Bacterial volatiles promote growth in *Arabidopsis*. *P. Natl. Acad. Sci. USA*, 100, 4927-4932.

224. Rudrappa, T., et al. (2010) The rhizobacterial elicitor acetoin induces systemic resistance in *Arabidopsis thaliana*. *Commun Integr Biol*, 3, 130-138.
225. Rui, Z., et al. (2014) Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *P. Natl. Acad. Sci. USA*, 111, 18237-18242.
226. Schulz, S. and Dickschat, J. S. (2007) Bacterial volatiles: the smell of small organisms. *Nat Prod Rep*, 24, 814-842.
227. Cheng, X., et al. (2016) Role of the GacS Sensor Kinase in the Regulation of Volatile Production by Plant Growth-Promoting *Pseudomonas fluorescens* SBW25. *Front. Plant. Sci.*, 7, 1706.
228. Fernando, W. G. D., et al. (2005) Identification and use of potential bacterial organic antifungal volatiles in biocontrol. *Soil Biology & Biochemistry*, 37, 955-964.
229. Audrain, B., et al. (2015) Role of bacterial volatile compounds in bacterial biology. *FEMS Microbiol. Rev.*, 39, 222-233.
230. Smith, D. and Spanel, P. (2005) Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. *Mass Spectrom Rev*, 24, 661-700.
231. Bean, H. D., Rees, C. A. and Hill, J. E. (2016) Comparative analysis of the volatile metabolomes of *Pseudomonas aeruginosa* clinical isolates. *J Breath Res*, 10, 047102.
232. Timm, C. M., et al. (2018) Direct Growth of Bacteria in Headspace Vials Allows for Screening of Volatiles by Gas Chromatography Mass Spectrometry. *Front Microbiol*, 9, 491.
233. Arnison, P. G., et al. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep*, 30, 108-160.

234. Ghequire, M. G. and De Mot, R. (2014) Ribosomally encoded antibacterial proteins and peptides from *Pseudomonas*. *FEMS Microbiol. Rev.*, 38, 523-568.
235. Elfarash, A., et al. (2014) Pore-forming pyocin S5 utilizes the FptA ferripyochelin receptor to kill *Pseudomonas aeruginosa*. *Microbiol-Sgm*, 160, 261-269.
236. Metelev, M., et al. (2013) Structure of Microcin B-Like Compounds Produced by *Pseudomonas syringae* and Species Specificity of Their Antibacterial Action. *J. Bacteriol.*, 195, 4129-4137.
237. Van der Donk, W. A., Funk, M., Ting, C. (2018) Catalytic Use of a Leader Peptide in the Biosynthesis of 3-Thiaglutamate. *BioRxiv*, 338681,
238. Moffat, A. D., et al. (2021) A User Guide for the Identification of New RiPP Biosynthetic Gene Clusters Using a RiPPER-Based Workflow. *Methods Mol Biol*, 2296, 227-247.
239. Hudson, G. A. and Mitchell, D. A. (2018) RiPP antibiotics: biosynthesis and engineering potential. *Curr. Opin. Microbiol.*, 45, 61-69.
240. Martinez-Garcia, E., et al. (2014) *Pseudomonas* 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microbial cell factories*, 13, 159.
241. Domrose, A., et al. (2017) Rapid generation of recombinant *Pseudomonas putida* secondary metabolite producers using γ TREX. *Synth Syst Biotechnol*, 2, 310-319.
242. FAO. (2019) World Food and Agriculture: Statistical Pocketbook 2019. World Food and Agriculture: Statistical Pocketbook 2019.
243. World Population Prospects 2019. <https://population.un.org/wpp/> (19/08/2021)

244. Erwin, D. C. and Ribeiro, O. K. (1996) *Phytophthora diseases worldwide*. Phytophthora diseases worldwide. APS Press, St. Paul, Minnesota, U.S.A.
245. Judelson, H. S. and Blanco, F. A. (2005) The spores of *Phytophthora*: weapons of the plant destroyer. *Nat. Rev. Microbiol.*, 3, 47-58.
246. Brasier, C. M. (1992) Evolutionary Biology of *Phytophthora* .1. Genetic System, Sexuality and the Generation of Variation. *Annu. Rev. Phytopathol.*, 30, 153-171.
247. Judelson, H. S. (1997) The genetics and biology of *Phytophthora infestans*: Modern approaches to a historical challenge. *Fungal Genet. Biol.*, 22, 65-76.
248. Turkensteen, L. J., et al. (2000) Production, survival and infectivity of oospores of *Phytophthora infestans*. *Plant Pathol.*, 49, 688-696.
249. Avrova, A. O., et al. (2003) Profiling and quantifying differential gene transcription in *Phytophthora infestans* prior to and during the early stages of potato infection. *Fungal Genet. Biol.*, 40, 4-14.
250. Haas, B. J., et al. (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*, 461, 393-398.
251. Raffaele, S., et al. (2010) Genome Evolution Following Host Jumps in the Irish Potato Famine Pathogen Lineage. *Science*, 330, 1540-1543.
252. Fry, W. (2008) *Phytophthora infestans*: the plant (and R gene) destroyer. *Mol. Plant Pathol.*, 9, 385-402.
253. Goss, E. M., et al. (2014) The Irish potato famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. *P. Natl. Acad. Sci. USA*, 111, 8791-8796.

254. Nowicki, M., et al. (2012) Potato and Tomato Late Blight Caused by *Phytophthora infestans*: An Overview of Pathology and Resistance Breeding. *Plant Dis.*, 96, 4-17.
255. Haverkort, A. J., et al. (2008) Societal Costs of Late Blight in Potato and Prospects of Durable Resistance Through Cisgenic Modification. *Potato Res*, 51, 47-57.
256. Haverkort, A. J., et al. (2016) Durable Late Blight Resistance in Potato Through Dynamic Varieties Obtained by Cisgenesis: Scientific and Societal Advances in the DuRPh Project. *Potato Res*, 59, 35-66.
257. Gisi, U. and Cohen, Y. (1996) RESISTANCE TO PHENYLAMIDE FUNGICIDES: a case study with *phytophthora infestans* involving mating type and race structure. *Annu. Rev. Phytopathol.*, 34, 549-572.
258. Pliakhnevich, M. and Ivaniuk, V. (2008) Aggressiveness and metalaxyl sensitivity of *Phytophthora infestans* strains in Belarus. *Zemdirbyste*, 95, 379-387.
259. Whisson, S. C., et al. (2016) The cell biology of late blight disease. *Curr. Opin. Microbiol.*, 34, 127-135.
260. Witek, K., et al. (2021) A complex resistance locus in *Solanum americanum* recognizes a conserved *Phytophthora* effector. *Nat Plants*, 7, 198-208.
261. Saville, A., et al. (2015) Fungicide Sensitivity of U.S. Genotypes of *Phytophthora infestans* to Six Oomycete-Targeted Compounds. *Plant Dis.*, 99, 659-666.
262. Lees, A. (2018) Comparison of sensitivity to a range of fungicides in contemporary genotypes of *Phytophthora infestans*. James Hutton Institute. Agriculture and Horticulture Development Board.
263. Schepers, H., et al. (2018) Reduced efficacy of fluazinam against *Phytophthora infestans* in the Netherlands. *Eur. J. Plant Pathol.*, 151, 947-960.

264. Witek, K., et al. (2016) Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nat. Biotechnol.*, 34, 656-660.
265. Jupe, F., et al. (2013) Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *Plant J.*, 76, 530-544.
266. Rodewald, J. and Trognitz, B. (2013) Solanum resistance genes against *Phytophthora infestans* and their corresponding avirulence genes. *Mol. Plant Pathol.*, 14, 740-757.
267. Gheysen, G. and Custers, R. (2017) Why Organic Farming Should Embrace Co-Existence with Cisgenic Late Blight-Resistant Potato. *Sustainability-Basel*, 9, 170.
268. Zitnak, A. and Johnston, G. R. (1970) Glycoalkaloid Content of B5141-6 Potatoes. *Am Potato J*, 47, 256.
269. Organisms, E. P. o. G. M. (2012) Scientific opinion addressing the safety assessment of plants developed through cisgenesis and intragenesis. *EFSA Journal*, 10, 2561.
270. Malcolmson, J. F. (1969) Races of *Phytophthora-Infestans* Occurring in Great Britain. *T Brit Mycol Soc*, 53, 417.
271. Beaumont, A. (1947) The dependence on the weather of the dates of outbreak of potato blight epidemics. *T Brit Mycol Soc*, 31, 45-53.
272. Smith, L. P. (1956) Potato Blight Forecasting by 90 Percent Humidity Criteria. *Plant Pathol.*, 5, 83-87.
273. Chapman, A. C. (2012) The changing *Phytophthora infestans* population: implications for late blight epidemics and control (Thesis). The University of Dundee. <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.585493>

274. Dancey, S. R., Skelsey, P., Cooke, D.E.L. Redefining the Smith Period for Potato Late Blight Management. The James Hutton Institute & the University of Dundee. Board, A. a. H. D.

275. Litschmann, T., Hausvater, E. and Dolezal, P. (2020) A new method of potato late blight forecasting in the Czech Republic. *Journal of Plant Protection Research*, 60, 134-140.

276. Fungicide Resistance Action Group UK. (2018). Fungicide Resistance Management in Potato Late Blight. FRAG-UK. [https://media.ahdb.org.uk/media/Default/Imported%20Publication%20Docs/AHDB%20Cereals%20&%20Oilseeds/Disease/FRAG%20Potato%20late%20blight%20guidelines%20\(May%202018\).pdf](https://media.ahdb.org.uk/media/Default/Imported%20Publication%20Docs/AHDB%20Cereals%20&%20Oilseeds/Disease/FRAG%20Potato%20late%20blight%20guidelines%20(May%202018).pdf)

277. Lambert, D. H. L., R. (1989) *Streptomyces scabies* sp. nov., nom. rev.†. *Int. J. Syst. Evol. Microbiol.*, 39, 387-392.

278. Lerat, S., Simao-Beaunoir, A. M. and Beaulieu, C. (2009) Genetic and physiological determinants of *Streptomyces scabies* pathogenicity. *Mol. Plant Pathol.*, 10, 579-585.

279. Loria, R., Kers, J. and Joshi, M. (2006) Evolution of plant pathogenicity in *Streptomyces*. *Annu. Rev. Phytopathol.*, 44, 469-487.

280. Kers, J. A., et al. (2005) A large, mobile pathogenicity island confers plant pathogenicity on *Streptomyces* species. *Mol. Microbiol.*, 55, 1025-1033.

281. Wanner, L. A. (2006) A survey of genetic variation in streptomyces isolates causing potato common scab in the United States. *Phytopathology*, 96, 1363-1371.

282. Kobayashi, A., et al. (2015) Community Analysis of Root- and Tuber-Associated Bacteria in Field-Grown Potato Plants Harboring Different Resistance Levels against Common Scab. *Microbes Environ.*, 30, 301-309.

283. King, R. R. and Calhoun, L. A. (2009) The thaxtomin phytotoxins: sources, synthesis, biosynthesis, biotransformation and biological activity. *Phytochemistry*, 70, 833-841.
284. Barry, S. M., et al. (2012) Cytochrome P450-catalyzed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis. *Nat. Chem. Biol.*, 8, 814-816.
285. Lerat, S., et al. (2010) Involvement of the plant polymer Suberin and the disaccharide cellobiose in triggering thaxtomin A biosynthesis, a phytotoxin produced by the pathogenic agent streptomyces scabies. *Phytopathology*, 100, 91-96.
286. Scheible, W. R., et al. (2003) An Arabidopsis mutant resistant to thaxtomin A, a cellulose synthesis inhibitor from Streptomyces species. *Plant Cell*, 15, 1781-1794.
287. Dauphin, A., et al. (2008) An early Ca²⁺ influx is a prerequisite to thaxtomin A-induced cell death in Arabidopsis thaliana cells. *J. Exp. Bot.*, 59, 4259-4270.
288. Fyans, J. K., et al. (2015) Characterization of the Coronatine-Like Phytotoxins Produced by the Common Scab Pathogen Streptomyces scabies. *Mol Plant Microbe Interact*, 28, 443-454.
289. Bignell, D. R. D., Cheng, Z. and Bown, L. (2018) The coronafacoyl phytotoxins: structure, biosynthesis, regulation and biological activities. *Antonie Van Leeuwenhoek*, 111, 649-666.
290. Bukhalid, R. A. and Loria, R. (1997) Cloning and expression of a gene from Streptomyces scabies encoding a putative pathogenicity factor. *J. Bacteriol.*, 179, 7776-7783.
291. Bukhalid, R. A., Chung, S.Y., Loria, R. (1998) nec1, a Gene Conferring a Necrogenic Phenotype, Is Conserved in Plant-Pathogenic Streptomyces spp. and Linked to a Transposase Pseudogene. *Mol. Plant-Microbe Interact.*, 11, 960-967.

292. Joshi, M., Rong, X., Moll, S., Kers, J., Franco, C., and Loria, R. (2007) *Streptomyces turgidiscabies* Secretes a Novel Virulence Protein, Nec1, Which Facilitates Infection. *Mol. Plant-Microbe Interact.*, 20, 599-608.
293. Seipke, R. F. and Loria, R. (2008) *Streptomyces scabies* 87-22 possesses a functional tomatinase. *J. Bacteriol.*, 190, 7684-7692.
294. Bouarab, K., et al. (2002) A saponin-detoxifying enzyme mediates suppression of plant defences. *Nature*, 418, 889-892.
295. Manulis, S., et al. (1994) Biosynthesis of indole-3-acetic acid via the indole-3-acetamide pathway in *Streptomyces* spp. *Microbiology*, 140 (Pt 5), 1045-1050.
296. Legault, G. S., et al. (2011) Tryptophan regulates thaxtomin A and indole-3-acetic acid production in *Streptomyces scabiei* and modifies its interactions with radish seedlings. *Phytopathology*, 101, 1045-1051.
297. Tegg, R. S., et al. (2008) Auxin-Induced Resistance to Common Scab Disease of Potato Linked to Inhibition of Thaxtomin A Toxicity. *Plant Dis.*, 92, 1321-1328.
298. Hill, J. and Lazarovits, G. (2005) A mail survey of growers to estimate potato common scab prevalence and economic loss in Canada. *Canadian Journal of Plant Pathology*, 27, 46-52.
299. Wilson, C. R. (2004) A summary of common scab disease of potato research from Australia, presented at: Proceedings of the International Potato Scab Symposium. Sapporo, Japan., 6-7 September 2004.
300. Stalham, M., Elphinstone, J., Wale, S., Steele, C. (2010) Managing the Risk of Common Scab. Potato Council, Agriculture & Horticulture Development Board.

301. Braun, S., et al. (2017) Potato Common Scab: a Review of the Causal Pathogens, Management Practices, Varietal Resistance Screening Methods, and Host Resistance. *Am J Potato Res*, 94, 283-296.
302. Wanner, L. A. and Kirk, W. W. (2015) Streptomyces - from Basic Microbiology to Role as a Plant Pathogen. *Am J Potato Res*, 92, 236-242.
303. Khatri, B. B., et al. (2011) Temporal association of potato tuber development with susceptibility to common scab and Streptomyces scabiei-induced responses in the potato periderm. *Plant Pathol.*, 60, 776-786.
304. Enciso-Rodriguez, F., et al. (2018) Genomic Selection for Late Blight and Common Scab Resistance in Tetraploid Potato (*Solanum tuberosum*). *G3 (Bethesda)*, 8, 2471-2481.
305. Braun, S. R., et al. (2017) Quantitative Trait Loci for Resistance to Common Scab and Cold-Induced Sweetening in Diploid Potato. *Plant Genome*, 10, doi: 10.3835/plantgenome2016.3810.0110.
306. Jansky, S., Douches, D. and Haynes, K. (2018) Germplasm Release: Three Tetraploid Potato Clones with Resistance to Common Scab. *Am J Potato Res*, 95, 178-182.
307. Ahn, Y. K. and Park, T. H. (2013) Resistance to common scab developed by somatic hybrids between *Solanum brevidens* and *Solanum tuberosum*. *Acta Agric. Scand. Sect. B-Soil Plant Sci.*, 63, 595-603.
308. Dees, M. W. and Wanner, L. A. (2012) In Search of Better Management of Potato Common Scab. *Potato Res*, 55, 249-268.
309. Lapwood, D. H., Wellings, L. W. and Rosser, W. R. (1970) The control of common scab of potatoes by irrigation. *Ann. Appl. Biol.*, 66, 397-405.

310. Davis, J. R., McMaster, G.M., Callihan, R.H., Nissley, F.H., Pavek, J. J. (1976) Influence of Soil Moisture and Fungicide Treatments on Common Scab and Mineral Content of Potatoes. *Phytopathology*, 66, 228-233.
311. Larkin, R. P., et al. (2011) Effects of different potato cropping system approaches and water management on soilborne diseases and soil microbial communities. *Phytopathology*, 101, 58-67.
312. Lambert, D. H. and Manzer, F. E. (1991) Relationship of Calcium to Potato Scab. *Phytopathology*, 81, 632-636.
313. Waksman, S. A. (1922) The influence of soil reaction upon the growth of Actinomycetes causing potato scab. *Soil Science*, 14, 61-80.
314. Loria, R., et al. (1997) Plant Pathogenicity in the Genus *Streptomyces*. *Plant Dis.*, 81, 836-846.
315. Haynes, K. G., et al. (2010) Common Scab Trials of Potato Varieties and Advanced Selections at Three US Locations. *Am J Potato Res*, 87, 261-276.
316. Al-Mughrabi, K. I., et al. (2016) Management of common scab of potato in the field using biopesticides, fungicides, soil additives, or soil fumigants. *Biocontrol Sci. Technol.*, 26, 125-135.
317. Santos-Cervantes, M. E., et al. (2017) Characterization, Pathogenicity and Chemical Control of *Streptomyces acidiscabies* Associated to Potato Common Scab. *Am J Potato Res*, 94, 14-25.
318. Tomihama, T., et al. (2016) Rice Bran Amendment Suppresses Potato Common Scab by Increasing Antagonistic Bacterial Community Levels in the Rhizosphere. *Phytopathology*, 106, 719-728.

319. Sarikhani, E., et al. (2017) The effect of peat and iron supplements on the severity of potato common scab and bacterial community in tuberosphere soil. *FEMS Microbiol. Ecol.*, 93, fiw206.
320. Thompson, H. K., et al. (2014) Foliar treatments of 2,4-dichlorophenoxyacetic acid for control of common scab in potato have beneficial effects on powdery scab control. *ScientificWorldJournal*, 2014, 947167.
321. Sarwar, A., et al. (2018) Biological Control of Potato Common Scab With Rare Isatropolone C Compound Produced by Plant Growth Promoting Streptomyces A1RT. *Front Microbiol*, 9, 1126.
322. Lin, C., et al. (2018) Biological control of potato common scab by *Bacillus amyloliquefaciens* Ba01. *PLoS One*, 13, e0196520.
323. Prevost, K., et al. (2006) Effect of chitosan and a biocontrol streptomycete on field and potato tuber bacterial communities. *BioControl*, 51, 533-546.
324. Stefanato, F. L., et al. (2019) Pan-genome analysis identifies intersecting roles for *Pseudomonas* specialized metabolites in potato pathogen inhibition. *bioRxiv*, doi: 10.1101/783258783258.
325. Alanjary, M., Steinke, K. and Ziemert, N. (2019) AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res.*, 47, W276-W282.
326. Richter, M. and Rossello-Mora, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A*, 106, 19126-19131.
327. Peix, A., Ramirez-Bahena, M. H. and Velazquez, E. (2018) The current status on the taxonomy of *Pseudomonas* revisited: An update. *Infect Genet Evol*, 57, 106-116.

328. Blin, K., et al. (2019) antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.*, 47, 81-87.
329. Altschul, S. F., et al. (1990) Basic local alignment search tool. *J. Mol. Biol.*, 215, 403-410.
330. Laville, J., et al. (1998) Characterization of the hcnABC gene cluster encoding hydrogen cyanide synthase and anaerobic regulation by ANR in the strictly aerobic biocontrol agent *Pseudomonas fluorescens* CHA0. *J. Bacteriol.*, 180, 3187-3196.
331. Meyer, J. M., et al. (2008) Siderotyping of fluorescent *Pseudomonas*: molecular mass determination by mass spectrometry as a powerful pyoverdine siderotyping method. *BioMetals*, 21, 259-271.
332. Budzikiewicz, H., et al. (2007) Characterization of the chromophores of pyoverdins and related siderophores by electrospray tandem mass spectrometry. *BioMetals*, 20, 135-144.
333. Malone, J. G., et al. (2010) YfiBNR mediates cyclic di-GMP dependent small colony variant formation and persistence in *Pseudomonas aeruginosa*. *PLoS Path.*, 6, e1000804.
334. Bryan, G., Garza, D. and Hartl, D. (1990) Insertion and excision of the transposable element mariner in *Drosophila*. *Genetics*, 125, 103-114.
335. De Vrieze, M., et al. (2020) Linking Comparative Genomics of Nine Potato-Associated *Pseudomonas* Isolates With Their Differing Biocontrol Potential Against Late Blight. *Front. Microbiol.*, 11, 857.
336. Drenkard, E., et al. (2018) Replication of the Ordered, Nonredundant Library of *Pseudomonas aeruginosa* strain PA14 Transposon Insertion Mutants. *J. Vis. Exp.*, doi: 10.3791/57298e57298.

337. Moffat, A. D., et al. (2021) A biofoundry workflow for the identification of genetic determinants of microbial growth inhibition. *Synth Biol (Oxf)*, 6, ysab004.
338. Horii, T., Ogawa, T. and Ogawa, H. (1980) Organization of the recA gene of Escherichia coli. *Proc Natl Acad Sci U S A*, 77, 313-317.
339. Blankenfeldt, W., et al. (2000) The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). *EMBO J.*, 19, 6652-6663.
340. Muzio, F. M., et al. (2020) 7-hydroxytropolone is the main metabolite responsible for the fungal antagonism of Pseudomonas donghuensis strain SVBP6. *Environ. Microbiol.*, 22, 2550-2563.
341. Krzyzanowska, D. M., et al. (2016) When Genome-Based Approach Meets the "Old but Good": Revealing Genes Involved in the Antibacterial Activity of Pseudomonas sp P482 against Soft Rot Pathogens. *Front. Microbiol.*, 7, 782.
342. Choi, K. H. and Schweizer, H. P. (2006) mini-Tn7 insertion in bacteria with single attTn7 sites: example Pseudomonas aeruginosa. *Nat Protoc*, 1, 153-161.
343. Heeb, S., et al. (2000) Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria. *Mol. Plant-Microbe Interact.*, 13, 232-237.
344. Kautsar, S. A., et al. (2020) MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.*, 48, D454-D458.
345. Jiang, Z., et al. (2016) 7-Hydroxytropolone produced and utilized as an iron-scavenger by Pseudomonas donghuensis. *BioMetals*, 29, 817-826.

346. Yu, X., et al. (2014) The two-component regulators GacS and GacA positively regulate a nonfluorescent siderophore through the Gac/Rsm signaling cascade in high-siderophore-yielding *Pseudomonas* sp. strain HYS. *J. Bacteriol.*, 196, 3259-3270.
347. Chen, M., Wang, P. and Xie, Z. (2018) A Complex Mechanism Involving LysR and TetR/AcrR That Regulates Iron Scavenger Biosynthesis in *Pseudomonas donghuensis* HYS. *J. Bacteriol.*, 200, e00087-00018.
348. Guo, H., Roman, D. and Beemelmans, C. (2019) Tropolone natural products. *Nat Prod Rep*, 36, 1137-1155.
349. Bentley, R. (2008) A fresh look at natural tropolonoids. *Nat Prod Rep*, 25, 118-138.
350. Duan, Y., et al. (2020) Bacterial Tropone Natural Products and Derivatives: Overview of their Biosynthesis, Bioactivities, Ecological Role and Biotechnological Potential. *ChemBioChem*, 21, 2384-2407.
351. Cai, X., et al. (2017) Structure and Biosynthesis of Isatropolones, Bioactive Amine-Scavenging Fluorescent Natural Products from *Streptomyces* Go66. *Angew Chem Int Ed Engl*, 56, 4945-4949.
352. Palleroni, N. J., et al. (1978) Production of a novel red pigment, rubrolone, by *Streptomyces echinoruber* sp. nov. I. Taxonomy, fermentation and partial purification. *J Antibiot (Tokyo)*, 31, 1218-1225.
353. Porsby, C. H., et al. (2011) Resistance and tolerance to tropodithietic acid, an antimicrobial in aquaculture, is hard to select. *Antimicrob. Agents Chemother.*, 55, 1332-1337.
354. Wilson, M. Z., et al. (2016) Mode of action and resistance studies unveil new roles for tropodithietic acid as an anticancer agent and the gamma-glutamyl cycle as a proton sink. *Proc Natl Acad Sci U S A*, 113, 1630-1635.

355. Beyersmann, P. G., et al. (2017) Dual function of tropodithietic acid as antibiotic and signaling molecule in global gene regulation of the probiotic bacterium *Phaeobacter inhibens*. *Sci Rep*, 7, 730.
356. Seyedsayamdost, M. R., et al. (2011) The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis*. *Nat Chem*, 3, 331-335.
357. Trust, T. J. (1975) Antibacterial activity of tropolone. *Antimicrob. Agents Chemother.*, 7, 500-506.
358. Morita, Y., et al. (2003) Biological activity of tropolone. *Biol. Pharm. Bull.*, 26, 1487-1490.
359. Allen, N. E., et al. (1982) 7-Hydroxytropolone: an inhibitor of aminoglycoside-2"-O-adenylyltransferase. *Antimicrob. Agents Chemother.*, 22, 824-831.
360. Teufel, R., et al. (2011) Studies on the mechanism of ring hydrolysis in phenylacetate degradation: a metabolic branching point. *J. Biol. Chem.*, 286, 11021-11034.
361. Teufel, R., Friedrich, T. and Fuchs, G. (2012) An oxygenase that forms and deoxygenates toxic epoxide. *Nature*, 483, 359-U161.
362. Duan, Y., et al. (2021) A Flavoprotein Dioxygenase Steers Bacterial Tropone Biosynthesis via Coenzyme A-Ester Oxygenolysis and Ring Epoxidation. *J. Am. Chem. Soc.*, 143, 10413–10421.
363. Azegami, K., et al. (1985) Tropolone as a Root Growth-Inhibitor Produced by a Plant Pathogenic *Pseudomonas* sp. Causing Seedling Blight of Rice. *Japanese Journal of Phytopathology*, 51, 315-317.

364. Agaras, B. C., Iriarte, A. and Valverde, C. F. (2018) Genomic insights into the broad antifungal activity, plant-probiotic properties, and their regulation, in *Pseudomonas donghuensis* strain SVBP6. *PLoS One*, 13, e0194088.
365. Gao, J., et al. (2015) *Pseudomonas donghuensis* sp. nov., exhibiting high-yields of siderophore. *Antonie Van Leeuwenhoek*, 107, 83-94.
366. Tao, X., et al. (2020) *Pseudomonas* species isolated via high-throughput screening significantly protect cotton plants against verticillium wilt. *AMB Express*, 10, 193.
367. Matuszewska, M., et al. (2021) The carbon source-dependent pattern of antimicrobial activity and gene expression in *Pseudomonas donghuensis* P482. *Sci Rep*, 11, 10994.
368. Chen, X. F., et al. (2018) Biosynthesis of Tropolones in *Streptomyces* spp.: Interweaving Biosynthesis and Degradation of Phenylacetic Acid and Hydroxylations on the Tropone Ring. *Appl. Environ. Microbiol.*, 84, e00349-00318.
369. Medema, M. H., Takano, E. and Breitling, R. (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol. Biol. Evol.*, 30, 1218-1223.
370. Wang, M. C., et al. (2016) Indole-3-Acetic Acid Produced by *Burkholderia heleaia* Acts as a Phenylacetic Acid Antagonist to Disrupt Tropolone Biosynthesis in *Burkholderia plantarii* (vol 6, 22596, pg 2016). *Sci Rep-Uk*, 6, Article number: 22596.
371. Azegami, K., Nishiyama, K. and Kato, H. (1988) Effect of Iron Limitation on "Pseudomonas plantarii" Growth and Tropolone and Protein Production. *Appl. Environ. Microbiol.*, 54, 844-847.
372. Takeshita, H., et al. (1987) Preparation of Polyacetoxytropolones and Polyhydroxytropolones by Acetolysis and Hydrolysis of Halotroponoids by Acetyl Trifluoroacetate with Exhaustive Displacement of Halogens on the Tropone Ring -

Predominant Formation of Reductive Acetolysates from Fully-Substituted Tropones. *B Chem Soc Jpn*, 60, 4325-4333.

373. Schwyn, B. and Neilands, J. B. (1987) Universal Chemical-Assay for the Detection and Determination of Siderophores. *Anal. Biochem.*, 160, 47-56.

374. Kirst, H. A., et al. (1982) Synthesis and characterization of a novel inhibitor of an aminoglycoside-inactivating enzyme. *J Antibiot (Tokyo)*, 35, 1651-1657.

375. Blin, K., et al. (2013) antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.*, 41, W204-212.

376. Fernandez, C., Diaz, E. and Garcia, J. L. (2014) Insights on the regulation of the phenylacetate degradation pathway from *Escherichia coli*. *Environ Microbiol Rep*, 6, 239-250.

377. Weber, T., et al. (2015) antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.*, 43, W237-W243.

378. Almagro Armenteros, J. J., et al. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.*, 37, 420-423.

379. Das, D., et al. (2010) The structure of the first representative of Pfam family PF09836 reveals a two-domain organization and suggests involvement in transcriptional regulation. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 66, 1174-1181.

380. Kenney, G. E., et al. (2018) The biosynthesis of methanobactin. *Science*, 359, 1411-1416.

381. Ting, C. P., et al. (2019) Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science*, 365, 280.

382. Kohler, C., et al. (2012) Extracytoplasmic function (ECF) sigma factor sigma(F) is involved in *Caulobacter crescentus* response to heavy metal stress. *BMC Microbiol.*, 12, 210.
383. Sarkisova, S. A., et al. (2014) A *Pseudomonas aeruginosa* EF-Hand Protein, EfhP (PA4107), Modulates Stress Responses and Virulence at High Calcium Concentration. *Plos One*, 9, e98985.
384. Clark, I. C., et al. (2014) Chlorate reduction in *Shewanella* algae ACDC is a recently acquired metabolism characterized by gene loss, suboptimal regulation and oxidative stress. *Mol. Microbiol.*, 94, 107-125.
385. Halary, S., et al. (2013) EGN: a wizard for construction of gene and genome similarity networks. *BMC Evol. Biol.*, 13, 146.
386. Kenney, G. E. and Rosenzweig, A. C. (2013) Genome mining for methanobactins. *BMC Biol.*, 11, 17.
387. Geer, L. Y., et al. (2002) CDART: protein homology by domain architecture. *Genome Res.*, 12, 1619-1623.
388. Gerlt, J. A., et al. (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim Biophys Acta*, 1854, 1019-1037.
389. Alvarez-Martinez, C. E., Baldini, R. L. and Gomes, S. L. (2006) A *Caulobacter crescentus* extracytoplasmic function sigma factor mediating the response to oxidative stress in stationary phase. *J. Bacteriol.*, 188, 1835-1846.
390. Dammeyer, T., et al. (2011) Efficient production of soluble recombinant single chain Fv fragments by a *Pseudomonas putida* strain KT2440 cell factory. *Microbial cell factories*, 10, 11.

391. Wittgens, A., et al. (2011) Growth independent rhamnolipid production from glucose using the non-pathogenic *Pseudomonas putida* KT2440. *Microbial cell factories*, 10, Article number: 80.
392. Nickel, P. I. and de Lorenzo, V. (2014) Robustness of *Pseudomonas putida* KT2440 as a host for ethanol biosynthesis. *New Biotechnol*, 31, 562-571.
393. Heeb, S., Blumer, C. and Haas, D. (2002) Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0. *J. Bacteriol.*, 184, 1046-1056.
394. Forsberg, E. M., et al. (2018) Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS Online. *Nat Protoc*, 13, 633-651.
395. Maertens, L., et al. (2021) Environmental Conditions Modulate the Transcriptomic Response of Both *Caulobacter crescentus* Morphotypes to Cu Stress. *Microorganisms*, 9, 1116.
396. Ozaki, K., et al. (2002) Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat. Genet.*, 32, 650-654.
397. San, J. E., et al. (2019) Current Affairs of Microbial Genome-Wide Association Studies: Approaches, Bottlenecks and Analytical Pitfalls. *Front Microbiol*, 10, 3119.
398. Power, R. A., Parkhill, J. and de Oliveira, T. (2017) Microbial genome-wide association studies: lessons from human GWAS. *Nat. Rev. Genet.*, 18, 41-50.
399. Earle, S. G., et al. (2016) Identifying lineage effects when controlling for population structure improves power in bacterial association studies. *Nat Microbiol*, 1, Article number: 16041.

400. Amos, G. C. A., et al. (2017) Comparative transcriptomics as a guide to natural product discovery and biosynthetic gene cluster functionality. *P. Natl. Acad. Sci. USA*, 114, E11121-E11130.
401. Petriccione, M., et al. (2017) Occurrence of copper-resistant *Pseudomonas syringae* pv. *syringae* strains isolated from rain and kiwifruit orchards also infected by *P. s.* pv. *actinidiae*. *Eur. J. Plant Pathol.*, 149, 953-968.
402. Bentley, S. D., et al. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, 417, 141-147.
403. Kamoun, S., et al. (1998) Resistance of *Nicotiana benthamiana* to *Phytophthora infestans* is mediated by the recognition of the elicitor protein INF1. *Plant Cell*, 10, 1413-1426.
404. Scott, T. A., et al. (2017) An L-threonine transaldolase is required for L-threo-beta-hydroxy-alpha-amino acid assembly during obafluorin biosynthesis. *Nat Commun*, 8, 15935.
405. Choi, K. H., et al. (2005) A Tn7-based broad-range bacterial cloning and expression system. *Nat. Methods*, 2, 443-448.
406. Loudon, B. C., Haarmann, D. and Lynne, A. M. (2011) Use of Blue Agar CAS Assay for Siderophore Detection. *J Microbiol Biol Educ*, 12, 51-53.
407. Madeira, F., et al. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.*, 47, W636-W641.
408. Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, 42, W320-324.

409. Miller, M. A., Pfeiffer, W., Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees, presented at: Proceedings of the Gateway Computing Environments Workshop (GCE). New Orleans, LA, 14 November 2010.
410. Letunic, I. and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.*, 47, W256-W259.
411. Darling, A. C., et al. (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.*, 14, 1394-1403.
412. Bosi, E., et al. (2015) MeDuSa: a multi-draft based scaffold. *Bioinformatics*, 31, 2443-2451.
413. Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30, 2068-2069.
414. Navarro-Munoz, J. C., et al. (2020) A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.*, 16, 60-68.
415. Winsor, G. L., et al. (2016) Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.*, 44, D646-653.

Appendix 1: Moffat, A. D., *et al.* (2021) A biofoundry workflow for the identification of genetic determinants of microbial growth inhibition. *Synth. Biol. (Oxf)*, 6, ysab004.



Synthetic Biology, 2021, 6(1): ysab004

doi: 10.1093/synbio/ysab004

Advance Access Publication Date: 28 January 2021
Research Article

A biofoundry workflow for the identification of genetic determinants of microbial growth inhibition

Alaster D. Moffat ¹, Adam Elliston², Nicola J. Patron ²,
Andrew W. Truman ^{1,*}, and Jose A. Carrasco Lopez ^{2,*}

¹Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, UK and

²Department of Engineering Biology, Earlham Institute, Norwich Research Park, Norwich, UK

*Corresponding authors: E-mails: jose.carrasco-lopez@earlham.ac.uk and andrew.truman@jic.ac.uk

Abstract

Biofoundries integrate high-throughput software and hardware platforms with synthetic biology approaches to enable the design, execution and analyses of large-scale experiments. The unique and powerful combination of laboratory infrastructure and expertise in molecular biology and automation programming, provide flexible resources for a wide range of workflows and research areas. Here, we demonstrate the applicability of biofoundries to molecular microbiology, describing the development and application of automated workflows to identify the genetic basis of growth inhibition of the plant pathogen *Streptomyces scabies* by a *Pseudomonas* strain isolated from a potato field. Combining transposon mutagenesis with automated high-throughput antagonistic assays, the workflow accelerated the screening of 2880 mutants to correlate growth inhibition with a biosynthetic gene cluster within 2 weeks.

Key words: biofoundry; automated testing workflow; *Pseudomonas*; biocontrol; *Streptomyces scabies*

1. Introduction

Biofoundries are integrated automation platforms for the design, construction and characterization of organisms for biotechnology applications and research. They combine experimental approaches, such as mathematical and statistical modeling (i.e. design of experiments, DoE), computer-aided design (CAD) and analytical software, together with automated experimental platforms to enable high-throughput automated workflows. Over the last decade, many research institutions have established biofoundries to accelerate both fundamental and applied research (1). Previous to this, the use of automation in research labs was sporadic with uptake mainly limited to service laboratories performing repetitive workflows, such as diagnostics or genotyping.

The era of synthetic biology has applied engineering principles, such as standardization and predictive models to inform experimental design (2). This has enabled experiments to be scaled, making automated workflows not only attractive but

necessary. Additional benefits of automation include improvements in the accuracy and reproducibility of experimentation (1, 3). A further tenet of engineering prescribed by synthetic biology is the progression of design-build-test-learn (DBTL) cycles, around which biofoundry workflows are typically centered. In the 'design' phase of the cycles, CAD tools are often used to assemble complex experimental designs, e.g. Ref. (4–6) and statistical techniques, such as DoE, are used to identify critical influencing factors and define the scale of the design space (7). Following this, automated workflows are used to assemble or 'build' the designs, typically from DNA. In the 'test' phase, specific analytic measurements are taken using techniques relevant to the phenotype or characteristic of interest. Finally, data are analyzed in the 'learn' phase and conclusions are used to inform the design phase of the subsequent cycle (1).

While the DBTL has clear applications in metabolic engineering (8), the applicability of biofoundries to the wider bioscience research community is not always obvious. The significant costs of establishing and maintaining biofoundries require an

Submitted: 7 July 2020; Received (in revised form): 4 January 2021; Accepted: 12 January 2021

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from <https://academic.oup.com/synbio/article/6/1/ysab004/6122745> by guest on 13 September 2021

active user base. Barriers identified for offsetting the cost of newly established biofoundries include a limited awareness of the capabilities among potential users (1). The basic software and hardware infrastructure of biofoundries, combined with the associated expertise in large-scale experimental design can, however, be applied to a wide range of molecular and microbial workflows enabling biofoundries to become central research infrastructures.

Here, we demonstrate the broader applicability of biofoundries to established molecular and microbial methods through the development and application of automated workflows for the rapid and scalable creation and phenotypic screening of mutant libraries. The ability to identify single-gene determinants of inhibitory activity in random mutagenesis studies is limited only by library size, which presents a significant limitation in manual screening efforts. In this example, we describe workflows for the identification of genes associated with the inhibition of microbial growth. This leads to the identification of a biosynthetic gene cluster (BGC) in the genome of *Pseudomonas* sp. Ps652, recovered from a commercial potato field in Norfolk (UK). This isolate had previously been shown to display strong *in vitro* activity against *Streptomyces scabies* 87-22, a commercially relevant potato pathogen (9, 10). The determinants of *Pseudomonas* sp. Ps652 inhibition of *S. scabies* 87-22 were not evident from the Ps652 genome sequence or rational gene deletions and, given the large zones of inhibition observed on plates, it was not possible to use a standard 96 solid pin replicator to assay transposon mutants on the same plate alongside the pathogen. To overcome this constraint, we developed biofoundry workflows to generate and assay a library of random transposon mutants. This enabled the identification of a BGC associated with the production of a compound able to inhibit the growth of *S. scabies* 87-22. There is substantial interest in the role of microbial antagonism in the health of humans (11), animals (12) and plants (13), but it can be difficult to elucidate the genetic determinants of this antagonism, which can be complex and manifold. The workflow described here provides a template for the automated identification of these genes.

2. Materials and methods

2.1 Bacterial strains and media

Pseudomonas sp. Ps652ΔHCN, a cyanide null mutant of an environmental isolate from a commercial potato field in Norfolk, UK (GenBank Accession GCA_902497775.1) was used as the parent strain for transposon mutagenesis. *Pseudomonas* sp. Ps652ΔHCN was stored at -80°C in 925 μL L medium with 75 μL DMSO. *Streptomyces coelicolor* M145 and *S. scabies* 87-22 were stored as spore stocks that were grown to sporulation on potato dextrose agar, harvested in 20% glycerol and stored at -80°C . The same spore stock was used for the entire workflow.

All media used in this study were adjusted to pH 7.2. Rye Sucrose Agar (RSA): 60 g/l rye grains, 20 g/l sucrose, 20 g/l Agar was prepared by soaking rye grains in a 1/40 dilution of 10% chlorite solution for 4 min before rinsing with reverse osmosis purified water and germinating overnight. Germinated rye grains were then ground at high speed for 2 min with a hand blender, before being transferred to a water bath for 3 h at 50°C in Milli-Q water. This was then filtered through a sieve and muslin cloth before addition of sucrose and agar and sterilization by autoclaving. Malt Extract-Yeast Extract Maltose Medium (MYM): 10 g/l malt extract, 4 g/l yeast extract, 4 g/l maltose, 18 g/l bacteriological agar (Sigma-Aldrich, UK), made up to 1 l with tap

water. Lennox Broth and Lennox Agar (L): 10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, 1 g/l D-glucose, 10 g/l Formedium agar (Lennox Agar), made up to 1 l with Milli-Q water. Lysogeny Broth (LB): 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, made up to 1 l with Milli-Q water. Potato Dextrose Agar: 41 g/l potato dextrose agar (Formedium PDA0102S), made up to 1 l with Milli-Q water.

2.2 Production and selection of a *Pseudomonas* mutant library

Pseudomonas sp. Ps652ΔHCN was transformed by electroporation in 300 mM sucrose with the mariner transposon-containing plasmid pALMAR3 (14) and 100 μL was plated directly in single-well Nunc™ OmniTray™ plates (242811, Thermo Scientific, UK) containing L agar + 25 $\mu\text{g}/\text{ml}$ tetracycline with a plating density that enabled automated colony picking. Plates were incubated overnight at 28°C for 16 h and resulting colonies were selected using the Hamilton Microlab STARplus (Hamilton, Bonaduz AG, Switzerland) equipped with one 96-channel pipetting head and eight individual single channels capable of liquid level detection, a light table and camera for colony picking, and grippers for plate movement steps (Supplementary Data S1). Methods were programmed using the Venus Three software (version 4.4.07740, Hamilton). Details of scripts, deck layout and associated Hamilton libraries are provided in Supplementary Data S2.

The detection and selection of the maximum number of viable colonies were enabled by a modified picking method using the EasypickII software (Version 1.0.2, Hamilton). A maximum number of 500 colonies from each plate were selected to generate the library. Criteria for colony picking were established to avoid cross-contamination and aberrant phenotypes. Colonies were defined with a minimal circularity factor of 0.05 and an area of 0.4–15 mm^2 . The liquid transfer and aspiration process were initially hampered by the presence of air bubbles formed by successive pipetting that hampered the autosensing of liquid levels, leading to inaccurate transfers. To overcome this, the pipette position was altered during aspiration and ejection of media. Suitable colonies were transferred using disposable 300 μL tips into liquid media (200 μL L + 25 $\mu\text{g}/\text{ml}$ tetracycline) in 96-well microplates (4ti-0117, Brooks Life Science, UK) (Supplementary Data S2). Cultures were grown at 28°C , 250 rpm for 16 h. 15 μL DMSO was added to each well manually using a multichannel pipette, and plates were stored at -80°C until required.

2.3 Automated plating of *S. coelicolor* M145

Spores of *S. coelicolor* M145 were resuspended into Milli-Q water at a concentration of 35 μL spore stock per 80 ml H_2O . A Hamilton Microlab STARplus, as described above, was used to dispense 20 μL of spore suspension into the center of each well of 120 24-well plates (Greiner Bio-One, 662102) containing 1 ml RSA per well. Spores were distributed using the on-deck shaker at 1500 rpm for 5 s. The plates were dried for 1 h before use in automated bioassays. Supplementary Data S3 provides the link to the automated spore plating method scripts, deck layout and associated Hamilton libraries.

2.4 Automated bioassays against *S. coelicolor* M145

The library of *Pseudomonas* sp. Ps652ΔHCN mutants were spiked onto the *S. coelicolor* M145 plates using a Hamilton Microlab STARplus as described above. Pipette tips were used to transfer droplets by submersion without aspiration followed by a brief touch of the agar surface using a liquid level setting of 0 mm.

Plates were incubated at 28°C for 5 days before photographic imaging (Canon Ixus 175, Canon, Tokyo, Japan). [Supplementary Data S4](#) provides the link to the bioassay method scripts, deck layout and associated Hamilton libraries.

2.5 Growth curves

Mutants of interest were streaked by hand to single colonies and grown overnight in 10 ml LB. Growth curves were performed by manually inoculating 5 µl of overnight culture diluted to an optical density at 600 nm (OD₆₀₀) = 0.01 into 150 µl of LB medium in 96-well plates. Plates were incubated at 30°C with shaking at 200 rpm for 48 h with absorbance readings taken every 30 min. Data were acquired on a SPECTROstar Nano UV/Vis microplate reader (BMG Labtech, Ortenberg, Germany), and processed using Microsoft Excel and R (15). Three replicates were used per mutant.

2.6 Cross-streak assays

Cross-streak assays against *S. scabies* 87-22 were performed manually on MYM tap agar. *Streptomyces* spores were streaked in two parallel lines on the plate using a sterile toothpick, which were then 'cross streaked' at a 90° angle with overnight cultures of *Pseudomonas* sp. Ps652ΔHCN mutants of interest. All mutants were assayed in triplicate. Plates were incubated for 5 days at 28°C before photographic imaging (Canon Ixus 175, Canon, Tokyo, Japan).

2.7 Insertion site sequencing

Sites of transposon insertion were determined by polymerase chain reaction (PCR) as described in Ref. (16). DNA flanking the insertion sites was enriched in two rounds of amplification using primers specific to the ends of the mariner transposon element (Arb1b 5'-GGCCAGCGAGCTAACGAGACNNNGATAT-3' and Arb-PCR 5'-CGCAAACCAACCCCTTGGCAG-3') with an initial denaturation at 95°C for 3 min followed by 30 cycles of 95°C for 15 s, 38°C for 30 s and 72°C for 90 s, followed by a final extension of 72°C for 3 min. Products were purified, eluted in 30 µl sterile water and used as template for a second round of PCR with degenerate primers that anneal to chromosomal sequences flanking the transposon (Arb1 5'-GGCCAGCGAGCTAACGAGAC-3' and Almar3-seq 5'-ACATATCCATCGCGTCCGCC-3') with an initial denaturation of 95°C for 3 min followed by 30 cycles of 95°C for 30 s, 56°C for 30 s, and 72°C for 90 s, followed by a final extension of 72°C for 3 min. Amplicons were purified using a QIAquick Spin PCR Purification Kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions and sequenced by Sanger sequencing (Eurofins Genomics, Luxembourg City, Luxembourg) using the Almar3-seq primer (5'-ACATATCCATCGCGTCCGCC-3').

3. Results

3.1 Automated production of a *Pseudomonas* sp. Ps652ΔHCN mutant library

Prior analysis of the *Pseudomonas* sp. Ps652 genome determined that it features no BGCs for known antimicrobial compounds (10), with the exception of a hydrogen cyanide (HCN) BGC (17). However, in preliminary experiments (unpublished data), we determined that HCN production in *Pseudomonas* sp. Ps652 only partly explained inhibition of *S. scabies* growth and that a non-volatile compound was contributing to inhibition. Therefore, by using a cyanide null mutant, Ps652ΔHCN, as the parent strain

for the generation of a transposon mutant library, we predicted the screen would identify new genes relevant to the inhibitory phenotype.

The primary bottleneck in the production of microbial mutant libraries is the selection of colonies and their subsequent curation. Manual colony-picking and subsequent screening of the resultant library is a tedious and time-consuming task with a consequential risk of human error and repetitive strain injury. We, therefore, automated the selection, culture and screening of a library of *Pseudomonas* sp. Ps652ΔHCN mutants to which the mariner transposon-containing plasmid pALMAR3 (14) had been delivered. Mariner type-transposons consist of a transposase gene flanked by inverted tandem repeats that insert their DNA cargo into the target genome at TA sites (18); the Ps652 genome contains 228 924 TA sites meaning that insertions of the transposase cargo DNA occur approximately randomly across the genome. The use of automation enabled the identification and selection of 2880 colonies (30 × 96-well plates) in 11 h with the added benefits of colony traceability (a digital record that links the original position of colonies to plate position for verification) and walk-away time (Figure 1A, [Supplementary Data S2](#)). The workflow was optimized to avoid colonies with aberrant morphologies and to prevent cross-contamination from neighboring colonies, facilitating scalability. The selection of 2880 mutants was based on the number of accessory coding sequences typically found in environmental *Pseudomonas* strains (19), however, the experiment is easily scalable.

3.2 Automated identification of *Pseudomonas* sp. Ps652 mutants unable to inhibit growth of *S. coelicolor* M145

Streptomyces coelicolor M145 was used in automated bioassays as a non-pathogenic proxy for *S. scabies* 87-22, which we were unable to work with in the biofoundry due to plant health license limitations. *Streptomyces scabies* and *S. coelicolor* both belong to clade II of a genome-based phylogeny for the *Streptomyces* genus (20), and preliminary work indicated that both were inhibited by Ps652ΔHCN (data not shown). Given the large zones of inhibition observed in pilot experiments, it was not possible to assay the mutant library on plates containing pathogen inoculum using a standard 96-pin replicator. We, therefore, developed an automated assay to assess the ability of 2880 *Pseudomonas* sp. Ps652ΔHCN mutants to affect the growth of *S. coelicolor* M145 using 24-well plates. This automated workflow consists of two steps. Firstly, the uniform spread of *Streptomyces* spores to each well (Figure 1B; [Supplementary Data S3](#)) and, secondly, inoculation with *Pseudomonas* sp. Ps652ΔHCN (Figure 1B; [Supplementary Data S4](#)). The automated assay required the optimization of (i) pipetting settings to minimize the time taken for the transfer of spore solution and plate inoculation, (ii) inoculation position (spiking) settings to enable the use of different volumes of agar in 24-well plates, (iii) lid-gripping and plate-gripping positions for accurate and rapid transfer of plates to stacks, and (iv) the speed (rpm) and time (s) for which plates were shaken to achieve a reproducibly homogenous spread of spores in every individual well. We also noted that it was essential to allow the spore solution to dry on the plates for 60 min before spiking the mutant library, to prevent Ps652 mutants from colonizing the entire well.

The optimized bioassay was performed across 120 24-well plates and took ~10 h to set up. Following incubation at 28°C for 5 days, the plates were imaged to identify mutants unable to inhibit *S. coelicolor* M145 growth (Figure 1C). Mutants of interest were defined as those in which *S. coelicolor* M145 grew immediately

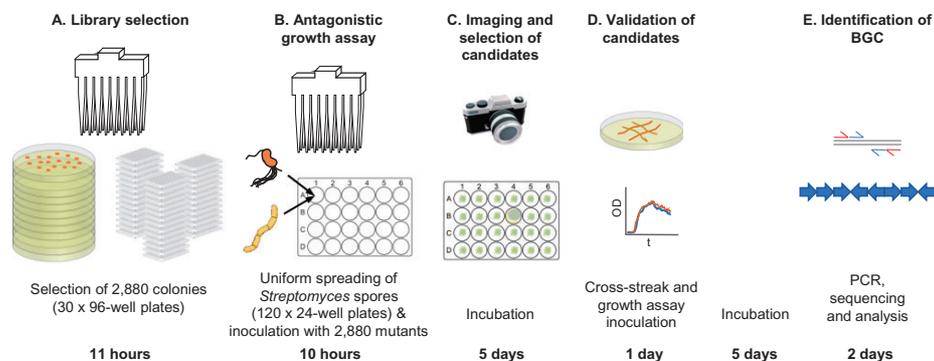


Figure 1. A Biofoundry workflow for the automated screening of microbial interactions. A mutant library of *Pseudomonas* sp. Ps652ΔHCN was created by automated selection of 2880 colonies (A). These were used in an automated bioassay with *S. coelicolor* M145 (B) and image analyzed after 5 days of incubation (C). Candidate strains were validated in growth assays and against the target pathogen, *S. scabiei*, in a licensed laboratory (D) before the insertion site was identified by PCR and sequencing (E).

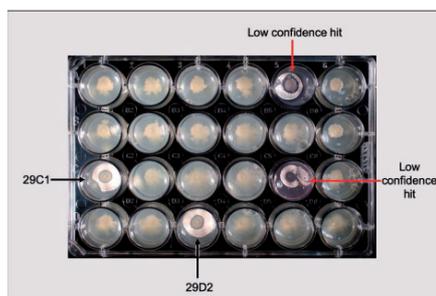


Figure 2. Representative image of a 24-well plate showing both low-confidence and high-confidence candidates, including 29C1 and 29D2, which were selected for further analysis.

adjacent to the pseudomonad. High-confidence mutants were those in which the growth of the mutants themselves was uncompromised. Low-confidence mutants were defined as those in which *Pseudomonas* sp. Ps652ΔHCN growth appeared abnormal or in which a minimal zone of inhibition was visible (Figure 2). We identified 23 mutants of interest, which were further investigated in low-throughput screens.

3.3 Validation of candidates in cross-streak assays with *S. scabiei*

In order to validate the high-throughput screen, the 23 mutants identified in the high-throughput screen against *S. coelicolor* M145 were streaked to single colonies and used to inoculate cross-streak assays (see Section 2) with the target strain, *S. scabiei* 87-22 (Figure 1D). Five mutant strains, 14G7, 24D9, 28E3, 29C1 and 29D2 were also unable to inhibit the growth of *S. scabiei* 87-22 (Figure 3A). To rule out mutants that simply caused a Ps652 growth defect, growth curves were performed in liquid medium. Under these conditions, mutant 28E3 displayed a significant growth defect when compared with the parent

strain and was excluded from further consideration (Figure 3B). Sequencing revealed that 28E3 had a transposon insertion in the *rfaA* gene, which encodes RmlA, a thymidyltransferase involved in the biosynthesis of the cell wall component deoxythymidine diphosphate-L-rhamnose (21), highlighting the potential for off-target hits and false positives in this type of screen. All 22 remaining initial hits retained normal growth characteristics, including 14G7, 24D9, 29C1 and 29D2 (Figure 3B). Sites of transposon insertion for all 23 mutants are available in Supplementary Data S5.

3.4 Sequencing of transposon insertion sites identifies a candidate gene cluster

The sites of transposon insertion were determined by PCR, and the four Ps652 mutants that were unable to inhibit the growth of both *S. coelicolor* M145 and *S. scabiei* 87-22 all harbored transposon insertions within the same 8 kbp chromosomal region (Figure 4A). This region had not been identified as a BGC via antiSMASH 5.0 (22) or PRISM (23) analysis of the Ps652 genome, but did match a putative BGC that had been identified in *Pseudomonas donghuensis* SVBP6 in parallel with our experimental work. This was associated with antifungal activity and 7-hydroxytropolone biosynthesis in *P. donghuensis* (24). Comparative sequence analyses using MultiGeneBlast (25) indicated that the BGCs in *Pseudomonas* sp. Ps652 and *P. donghuensis* SVBP6 are highly similar, where they share the same genetic architecture (Figure 4B), as well as high identity/coverage scores for all genes (Supplementary Data S6).

4. Discussion

High-throughput screening is an emerging approach in many scientific fields and can be a useful tool in investigating microbial interactions (26–29). However, many researchers lack access to automation equipment or to programming expertise to enable existing equipment to be applied to their experimental workflow. As a result, high-throughput approaches often bypassed in favor of simple or low-throughput methodologies. This is particularly pertinent in microbial interactions, where the ability to scale bioassays holds immense promise but is

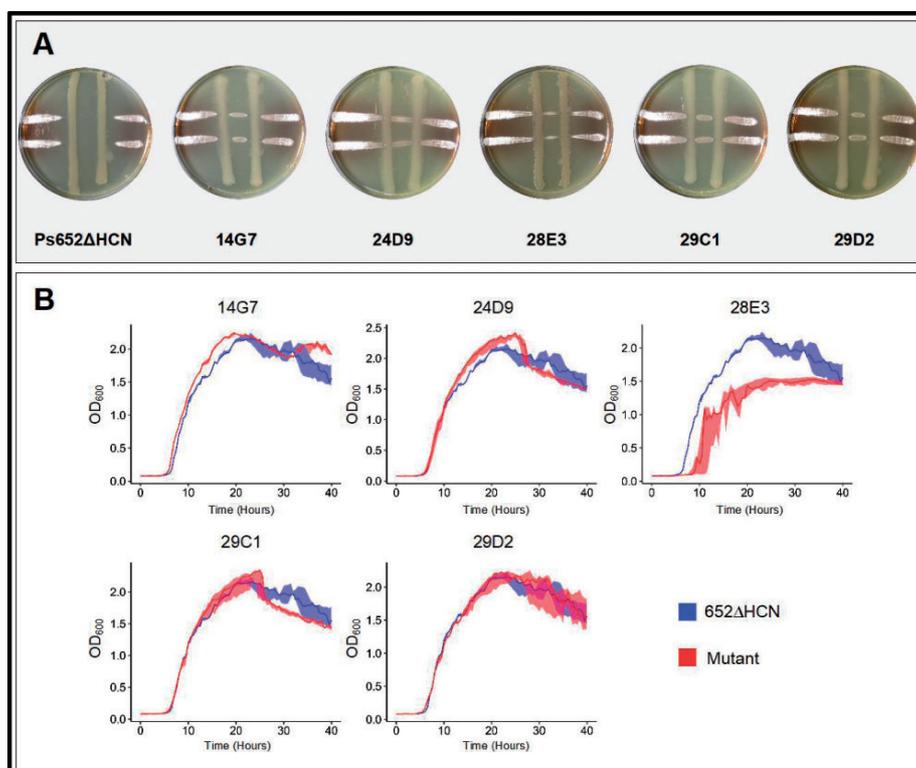


Figure 3. (A) Cross-streak assays of five *Pseudomonas* sp. Ps652ΔHCN mutants (vertical lines) against *S. scabies* 87-22 (horizontal lines). Images are representative of 3 replicates after incubation for five days. (B) Growth curves of *Pseudomonas* sp. Ps652ΔHCN (control) and five transposon mutants. Solid lines represent the median at each time point and the shaded area represent the maximum and minimum values; $n = 3$.

impractical without the application of automation (27). In this study, we exemplified the relevance to biofoundries to the study of microbial interactions in the soil interactions, by conducting an automated screen that identified a BGC implicated in growth inhibition in just a few days of laboratory time. Such microbial interactions are of interest to agriculture, particularly the identification and characterization of plant-associated or rhizospheric soil bacteria capable of inhibiting the growth of specific plant pathogens (13, 30).

Advances in bioinformatic approaches to natural product discovery can be applied to such problems, aiding discovery of novel antimicrobial compounds in plant-associated bacteria (31). However, many existing bioinformatic approaches rely on homology to characteristic enzymes, such as polyketide synthases and non-ribosomal peptide synthetases, which represent roughly 65% of microbial metabolites used commercially (32). At present, it is likely that such approaches may be less reliable in identifying novel BGCs that display unusual and unprecedented genetic architectures, or that lack unifying features shared across the whole molecular class. Similarly, currently available bioinformatics tools can be of limited use in

uncovering specialized metabolites where the biosynthesis is performed by a single protein, as has been shown previously for a volatile compound that is involved in biocontrol, 1-undecene (33). Consequently, novel compounds with societally relevant activities are still being discovered through bioactivity guided screening methods (34). As such, interesting phenotypes from select bacterial isolates still warrant investigation through screening against pathogens of interest.

To develop biofoundry workflows for studying microbial interactions, we focused on a *Pseudomonas* strain for which previous investigations had not identified BGCs of known antimicrobials (10, 24). The utility of our automated workflow was initially validated by growth curves and cross-streak assays (Figure 3), and subsequently by the identification of a BGC highly likely to produce the tropolonoid antimicrobial 7-hydroxytropolone. In parallel with our work, this BGC was identified through a transposon mutant screen of *P. donghuensis* SVBP6 (24) (Figure 4B). In that study, the authors associated the BGC with inhibitory activity of fungi and oomycetes, but not *Bacillus subtilis* or *Escherichia coli*. Our screen identified this BGC as being implicated in the inhibition of the filamentous Gram-positive

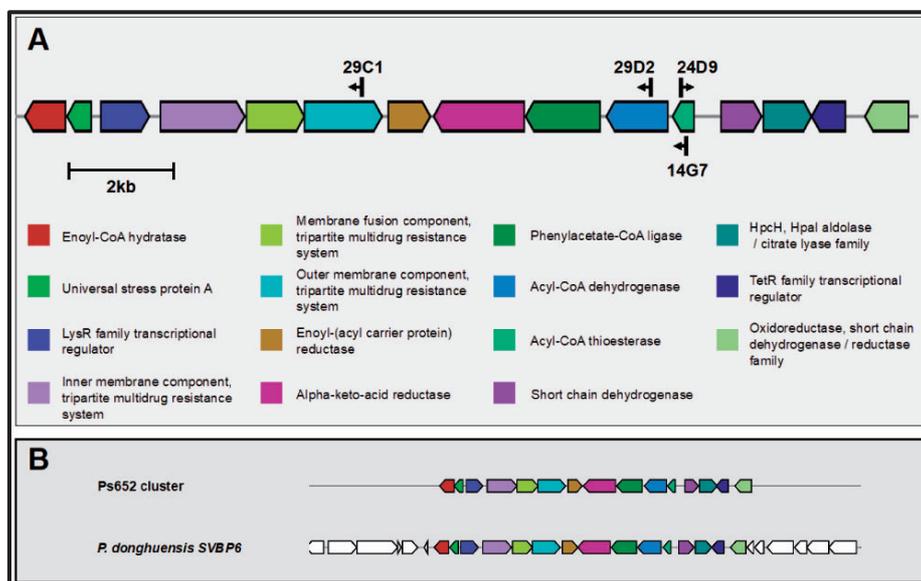


Figure 4. (A) Schematic of the *Pseudomonas* sp. Ps652 BGC indicating the location of transposon insertions (black arrows) and the conserved domains of the proteins encoded in this BGC. (B) Comparison of the cluster identified in *Pseudomonas* sp. Ps652 to a recently reported cluster linked to 7-hydroxytropone production in *P. donghuensis* SVBP6.

bacterium *S. scabies* 87-22, a commercially important potato pathogen (35). Lacking the natural product class-defining proteins and being considerably different to published non-*Pseudomonas* tropolone BGCs (36), this cluster was not readily identified by genome mining tools.

The development of the workflow identified a number of important considerations and advantages of automated processes. The screening of bacterial libraries often relies on 96-solid pin replicators (37); however, some phenotypes are affected by the genotype or phenotype of the surrounding colonies on the plate. In the case of *Pseudomonas* sp. Ps652, mutant colonies that had lost the ability to produce a diffusible natural product were undetectable due to the production of the inhibitory natural product by neighboring colonies. Conducting the assays in individual wells of 24-well plates, an approach only practicable with automation, overcame this complication. Automating the assay also enabled the uniform distribution of *Streptomyces* spores, as well as the exact positioning and depth of the *Pseudomonas* inoculum. This used the liquid level detection feature and conductive tips to avoid the challenges of inconsistent agar depth. An additional advantage of automation is the ability to screen very large libraries. This is particularly important for single-gene determinants of phenotypes. Finally, the selection of false positives can be a potent obfuscating factor. These can increase due to inconsistent assay set-up as well as human errors or bias in visual screening. The ability to accurately and consistently inoculate assays and to program defined parameters for the automated identification of colony size and shape aided the robustness and reproducibility of the assay. Through the application of biofoundry workflows and approaches to

experimentation, we were able to rapidly identify and prioritize a manageable number of high-confidence candidates for further characterization using manual approaches, ultimately implicating a BGC proposed to produce 7-hydroxytropone in the biological control of the important potato pathogen *S. scabies* 87-22.

Most biofoundries are already engaged in the reconstruction and optimization of biosynthetic pathways to enable natural product biosynthesis. This study highlights their potential in the upstream experimental processes of gene discovery. In the workflow described here, the DBTL cycle consists of the selection of appropriate tools for the production of libraries, the construction and arraying of such libraries, phenotypic screens and genotyping of candidates, and, finally, data analysis to provide new knowledge of gene function. The outcomes may lead to either iterative cycles in which either library complexity or the robustness of the phenotypic screen are improved, or excitingly, link directly through the provision of novel genes to interconnected DBTL cycles in which the aim is biosynthesis. To aid this, critical comparisons of library production methods and the integration of high-throughput computer-aided quantitative phenotyping into biofoundry platforms would be highly beneficial.

SUPPLEMENTARY DATA

Supplementary Data are available at SYN BIO Online.

Acknowledgments

We thank James Titchmarch (Hamilton Company) for the original Hamilton picking method, which was adapted by

A.E. We are grateful to Jacob Malone (John Innes Centre) for providing the pALMAR3 mariner transposon plasmid, and to David Widdick (John Innes Centre) for advice on transposon mutagenesis screening.

Funding

The authors would like to acknowledge support from the UK Research and Innovation (UKRI) Biotechnology and Biological Sciences Research Council (BBSRC) Core Capability Grant BB/CCG1720/1 for the Earlham Institute and the National Capability BBS/E/T/000PR9815. Additional funding was provided via the BBSRC MfN Institute Strategic Programme grant [BBS/E/J/000PR9790] for the John Innes Centre, a Norwich Research Park BBSRC Doctoral Training Partnership grant [BB/M011216/1] for A.D.M. and a Royal Society University Research Fellowship for A.W.T.

Data availability

All relevant data (scripts packages for the automated colony picking, spore spreading and automated inoculation of *Pseudomonas* sp. Ps652ΔHCN mutant library) are available via the GitHub repository of the Earlham Institute Biofoundry (<https://github.com/eibiofoundry/A-biofoundry-workflow-for-the-identification-of-genetic-determinants-of-microbial-growth-inhibition>). All videos depicting the automated workflows are available via the YouTube channel of the Earlham Institute Biofoundry (Automated colony picking: <https://youtu.be/cLQZDdMr0l0>, automated distribution of *S. coelicolor* M145 spores on 24-well plates <https://youtu.be/3TEFXuciNu8>, and automated inoculation of 24-well plates containing *S. coelicolor* M145 with the *Pseudomonas* sp. Ps652ΔHCN mutant library: <https://youtu.be/nlIDy0Vf5sw>).

References

- Hillson,N., Caddick,M., Cai,Y., Carrasco,J.A., Chang,M.W., Curach,N.C., Bell,D.J., Le Feuvre,R., Friedman,D.C., Fu,X. et al. (2019) Building a global alliance of biofoundries. *Nat. Commun.*, 10, 2040.
- Jessop-Fabre,M.M. and Sonnenschein,N. (2019) Improving reproducibility in synthetic biology. *Front. Bioeng. Biotechnol.*, 7, 18.
- Chao,R., Mishra,S., Si,T. and Zhao,H. (2017) Engineering biological systems using automated biofoundries. *Metab. Eng.*, 42, 98–108.
- Czar,M.J., Cai,Y. and Peccoud,J. (2009) Writing DNA with genoCAD™. *Nucleic Acids Res.*, 37, W40–W47.
- Hillson,N.J., Rosengarten,R.D. and Keasling,J.D. (2012) j5 DNA assembly design automation software. *ACS Synth. Biol.*, 1, 14–21.
- Nielsen,A.A.K., Der,B.S., Shin,J., Vaidyanathan,P., Paralanov,V., Strychalski,E.A., Ross,D., Densmore,D. and Voigt,C.A. (2016) Genetic circuit design automation. *Science*, 352, aac7341.
- Braniff,N. and Ingalls,B. (2018) New opportunities for optimal design of dynamic experiments in systems and synthetic biology. *Curr. Opin. Syst. Biol.*, 9, 42–48.
- Carbonell,P., Jervis,A.J., Robinson,C.J., Yan,C., Dunstan,M., Swainston,N., Vinaixa,M., Hollywood,K.A., Currin,A., Rattray,N.J.W. et al. (2018) An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun. Biol.*, 1, 66.
- Seipke,R.F. and Loria,R. (2008) *Streptomyces scabies* 87-22 possesses a functional tomatinase. *J. Bacteriol.*, 190, 7684–7692.
- Stefanato,F.L., Trippel,C., Uszkoreit,S., Ferrafiat,L., Grenga,L., Dickens,R., Kelly,N., Kingdon,A.D.H., Ambrosetti,L., Findlay,K.C., et al. (2019) Pan-genome analysis identifies intersecting roles for *Pseudomonas* specialized metabolites in potato pathogen inhibition. *bioRxiv*, doi: 10.1101/783258.
- Waldman,A.J. and Balskus,E.P. (2018) The human microbiota, infectious disease, and global health: challenges and opportunities. *ACS Infect. Dis.*, 4, 14–26.
- Mueller,U.G. and Sachs,J.L. (2015) Engineering microbiomes to improve plant and animal health. *Trends Microbiol.*, 23, 606–617.
- Kim,Y.C., Leveau,J., McSpadden Gardener,B.B., Pierson,E.A., Pierson,L.S. 3rd and Ryu, C.M. (2011) The multifactorial basis for plant health promotion by plant-associated bacteria. *Appl. Environ. Microbiol.*, 77, 1548–1555.
- Malone,J.G., Jaeger,T., Spangler,C., Ritz,D., Spang,A., Arriemerlou,C., Kaever,V., Landmann,R. and Jenal,U. (2010) YfiBNR mediates cyclic di-GMP dependent small colony variant formation and persistence in *Pseudomonas aeruginosa*. *PLoS Pathog.*, 6, e1000804.
- R Core Team. (2019) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- O'Toole,G.A. and Kolter,R. (1998) Initiation of biofilm formation in *Pseudomonas fluorescens* wcs365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Mol. Microbiol.*, 28, 449–461.
- Pessi,G. and Haas,D. (2000) Transcriptional control of the hydrogen cyanide biosynthetic genes hcnABC by the anaerobic regulator ANR and the Quorum-sensing regulators LasR and RhIR in *Pseudomonas aeruginosa*. *J. Bacteriol.*, 182, 6940–6949.
- Lampe,D.J., Grant,T.E. and Robertson,H.M. (1998) Factors affecting transposition of the Himar1 mariner transposon in vitro. *Genetics*, 149, 179–187.
- De Vrieze,M., Varadarajan,A.R., Schneeberger,K., Bailly,A., Rohr,R.P., Ahrens,C.H. and Weisskopf,L. (2020) Linking comparative genomics of nine potato-associated *Pseudomonas* isolates with their differing biocontrol potential against late blight. *Front. Microbiol.*, 11, 857.
- McDonald,B.R. and Currie,C.R. (2017) Lateral gene transfer dynamics in the ancient bacterial genus *Streptomyces*. *mBio.*, 8, e00644-17.
- Blankenfeldt,W., Asuncion,M., Lam,J.S. et al. (2000) The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). *EMBO J.*, 19, 6652–6663.
- Blin,K., Shaw,S., Steinke,K., Villebro,R., Ziemert,N., Lee,S.Y., Medema,M.H. and Weber,T. (2019) AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.*, 47, W81–W87.
- Skininder,M.A., Merwin,N.J., Johnston,C.W. and Magarvey,N.A. (2017) PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.*, 45, W49–W54.
- Muzio,F.M., Agaras,B.C., Masi,M., Tuzi,A., Evidente,A. and Valverde,C. (2020) 7-hydroxytropolone is the main metabolite responsible for the fungal antagonism of *Pseudomonas donghuensis* strain SVBP6. *Environ. Microbiol.*, 22, 2550–2563.

25. Medema, M.H., Takano, E. and Breitling, R. (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol. Biol. Evol.*, 30, 1218–1223.
26. Geng, H., Bruhn, J.B., Nielsen, K.F., Gram, L. and Belas, R. (2008) Genetic dissection of tropodithetic acid biosynthesis by marine roseobacters. *Appl. Environ. Microbiol.*, 74, 1535–1545.
27. Wang, R., Gallant, E. and Seyedsayamdost, M.R. (2016) Investigation of the genetics and biochemistry of roseobactin production in the *Roseobacter* clade bacterium *Phaeobacter inhibens*. *mBio.*, 7, e02118.
28. Santos Kron, A., Zengerer, V., Bieri, M., Dreyfuss, V., Sostizzo, T., Schmid, M., Lutz, M., Remus-Emsermann, M.N.P. and Pelludat, C. (2020) *Pseudomonas orientalis* F9 pyoverdine, safracin, and phenazine mutants remain effective antagonists against *Erwinia amylovora* in apple flowers. *Appl. Environ. Microbiol.*, 86, e02620–19.
29. Agaras, B.C., Iriarte, A. and Valverde, C.F. (2018) Genomic insights into the broad antifungal activity, plant-probiotic properties, and their regulation, in *Pseudomonas donghuensis* strain SVBP6. *PLoS One*, 13, e0194088.
30. Philippot, L., Raaijmakers, J.M., Lemanceau, P. and van der Putten, W.H. (2013) Going back to the roots: the microbial ecology of the rhizosphere. *Nat. Rev. Microbiol.*, 11, 789–799.
31. Paterson, J., Jahanshah, G., Li, Y., Wang, Q., Mehnaz, S. and Gross, H. (2017) The contribution of genome mining strategies to the understanding of active principles of PGPR strains. *FEMS Microbiol. Ecol.*, 93, fiw249.
32. Katz, L. and Baltz, R.H. (2016) Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.*, 43, 155–176.
33. Rui, Z., Li, X., Zhu, X., Liu, J., Domigan, B., Barr, I., Cate, J.H.D. and Zhang, W. (2014) Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *Proc. Natl. Acad. Sci. U S A*, 111, 18237–18242.
34. Ling, L.L., Schneider, T., Peoples, A.J., Spoering, A.L., Engels, I., Conlon, B.P., Mueller, A., Schäberle, T.F., Hughes, D.E., Epstein, S. et al. (2015) A new antibiotic kills pathogens without detectable resistance. *Nature*, 517, 455–459.
35. Lerat, S., Simao-Beanoir, A.M. and Beaulieu, C. (2009) Genetic and physiological determinants of *Streptomyces scabies* pathogenicity. *Mol. Plant Pathol.*, 10, 579–585.
36. Chen, X., Xu, M., Lü, J., Xu, J., Wang, Y., Lin, S., Deng, Z. and Tao, M. (2018) Biosynthesis of tropolones in *Streptomyces* spp.: Interweaving biosynthesis and degradation of phenylacetic acid and hydroxylations on the tropone ring. *Appl. Environ. Microbiol.*, 84, e00349–18.
37. Drenkard, E., Hibbler, R.M., Gutu, D.A., Eaton, A.D., Silverio, A.L., Ausubel, F.M., Hurley, B.P. and Yonker, L.M. (2018) Replication of the ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *J. Vis. Exp.*, 4, 57298.

Appendix 2: Moffat, A. D., *et al.* (2021) A user guide for the identification of new RiPP biosynthetic gene clusters using a RiPPER-based workflow. *Methods Mol. Biol.*, 2296, 227-247.



Chapter 14

A User Guide for the Identification of New RiPP Biosynthetic Gene Clusters Using a RiPPER-Based Workflow

Alaster D. Moffat, Javier Santos-Aberturas, Govind Chandra, and Andrew W. Truman

Abstract

In recent years, genome mining has become a powerful strategy for the discovery of new specialized metabolites from microorganisms. However, the discovery of new groups of ribosomally synthesized and post-translationally modified peptides (RiPPs) by employing the currently available genome mining tools has proven challenging due to their inherent biases towards previously known RiPP families. In this chapter we provide detailed guidelines on using RiPPER, a recently developed RiPP-oriented genome mining tool conceived for the exploration of genomic database diversity in a flexible manner, thus allowing the discovery of truly new RiPP chemistry. In addition, using TfuA proteins of Alphaproteobacteria as an example, we present a complete workflow which integrates the functionalities of RiPPER with existing bioinformatic tools into a complete genome mining strategy. This includes some key updates to RiPPER (updated to version 1.1), which substantially simplify implementing this workflow.

Key words Biosynthetic gene cluster, RiPP, Genome mining, Bioinformatics, Antibiotics, Natural products, Specialized metabolites, Peptides

1 Introduction

Microbial specialized metabolites (also known as secondary metabolites or natural products) constitute an essential source of bioactive molecules for both pharmaceutical and agrochemical industries. In the last 15 years, traditional bioactivity-based discovery methods have found an invaluable sidekick in “genome mining” approaches that leverage affordable next-generation genome sequencing [1, 2]. These strategies are focused on the bioinformatic exploration of growing genomic databases for the identification of biosynthetic gene clusters (BGCs) responsible for making novel bioactive compounds, thus avoiding re-discovery of already known molecules.

Standard genome mining approaches have proven to be very effective for the identification of BGCs containing large modular

biosynthetic systems, as in the case of polyketide synthases (PKSs) or non-ribosomal peptide synthetases (NRPSs), as well as smaller BGCs containing characteristic enzymes (e.g., terpene synthases or cyclodipeptide synthases) [3–5]. However, it is much more challenging to use standard genome mining methods to comprehensively identify BGCs for ribosomally synthesized and post-translationally modified peptides (RiPPs) [6], one of the major classes of specialized metabolites. This is a consequence of: (a) the unrivalled chemical diversity of their biosynthetic pathways, and (b) a lack of unifying features that could be employed as general rules for their detection.

The biosynthesis of RiPPs is characterized by the use of a genetically encoded precursor peptide as a substrate for a series of chemical modifications catalyzed by clustered tailoring enzymes (Fig. 1). These are made to a core peptide region of the precursor peptide, which is usually at its C-terminus. The unmodified leader peptide region is then cleaved as a late-stage step in the pathway. Due to the high variability of their sequences (specific for each pathway) and their short length (usually <120 amino acids, AAs), the genes encoding these precursor peptides are sometimes not annotated in databases, so are not ideal starting points for the discovery of new RiPP BGCs. In addition, the accompanying tailoring enzymes are also specific for different RiPP subclasses or families.

Different approaches have been recently developed for the discovery of new RiPP BGCs by genome mining in a user-friendly manner [3, 4, 7–12]. However, most are designed for the identification of BGCs belonging to specific RiPP families, searching for defined tailoring enzymes and considering certain pre-established sequence features for the correct identification of the precursor peptides. Although such specialized genome mining tools allow the reliable identification of RiPP BGCs, their inherent biases constrain the user's freedom to explore truly new RiPP metabolic landscapes, thus limiting the likelihood of finding strikingly new RiPP chemistry.

We therefore developed RiPPER (*RiPP Precursor Peptide Enhanced Recognition*) as a RiPP genome mining tool to help overcome these limitations. This was achieved by providing the user with a highly customizable tool to explore genomic databases in a strongly bias-reduced manner [13]. RiPPER functions as a command line tool that employs its own RiPP-specific functionality alongside a series of pre-existing tools to retrieve and annotate precursor peptide candidates encoded within the same genomic locus as a user-defined tailoring enzyme.

RiPPER first uses RODEO2 [7] to retrieve a genomic region associated with a given tailoring enzyme as a GenBank file. RODEO2 was itself designed for the identification of RiPP BGCs, but it is used here to retrieve accessions and to generate an

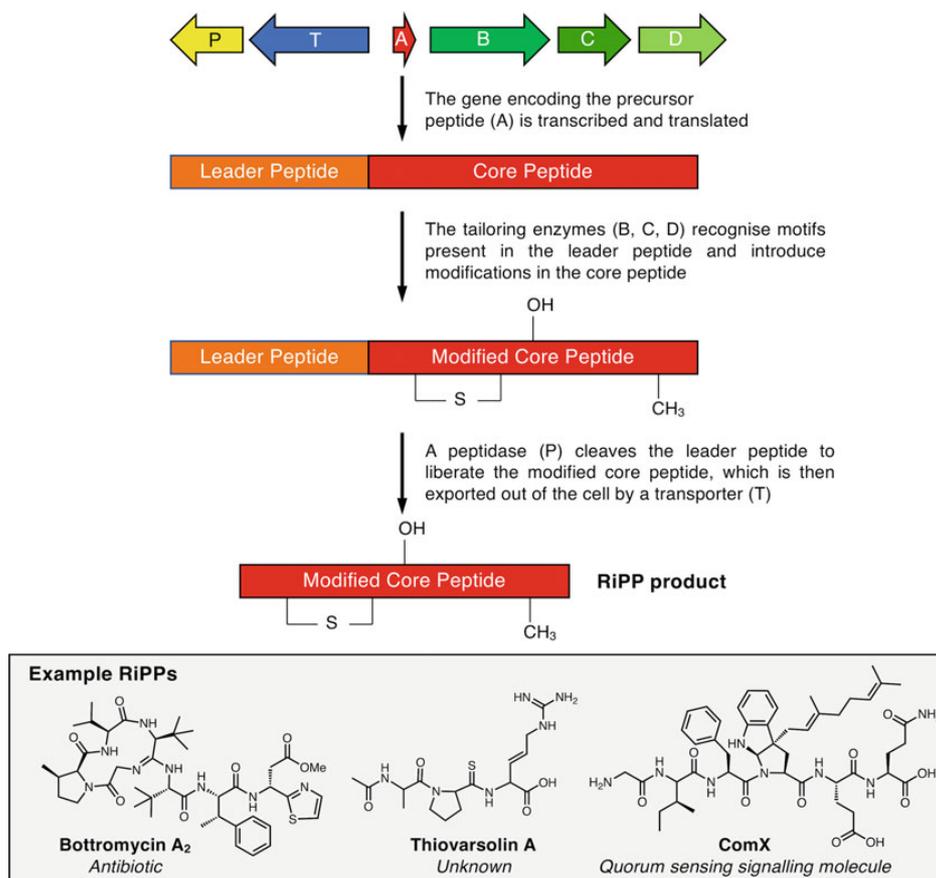


Fig. 1 Overview of RiPP biosynthesis

html output for the genomic locus. RiPPER then re-annotates open reading frames (ORFs) encoding short peptides (20–120 AAs) in the putative BGC using Prodigal-short, a modified version of the established gene finding tool, Prodigal [14] (Fig. 2). This is much more effective at identifying true coding regions compared to a crude 6-frame translation approach. RiPPER specifically searches for short ORFs in intergenic regions and re-annotates short peptides (<120 AAs) already annotated in the retrieved GenBank file. This yields a modified GenBank file that can be viewed in Artemis [15] (Fig. 2). Here, the gene encoding the “bait” tailoring protein is colored green, and the short peptides are color coded on a white-red scale for their Prodigal-short score, where Prodigal-short assigns a score to these ORFs based on numerous genetic factors. The top three scoring peptides are retained for downstream analysis, as are any additional ORFs that are over a defined score

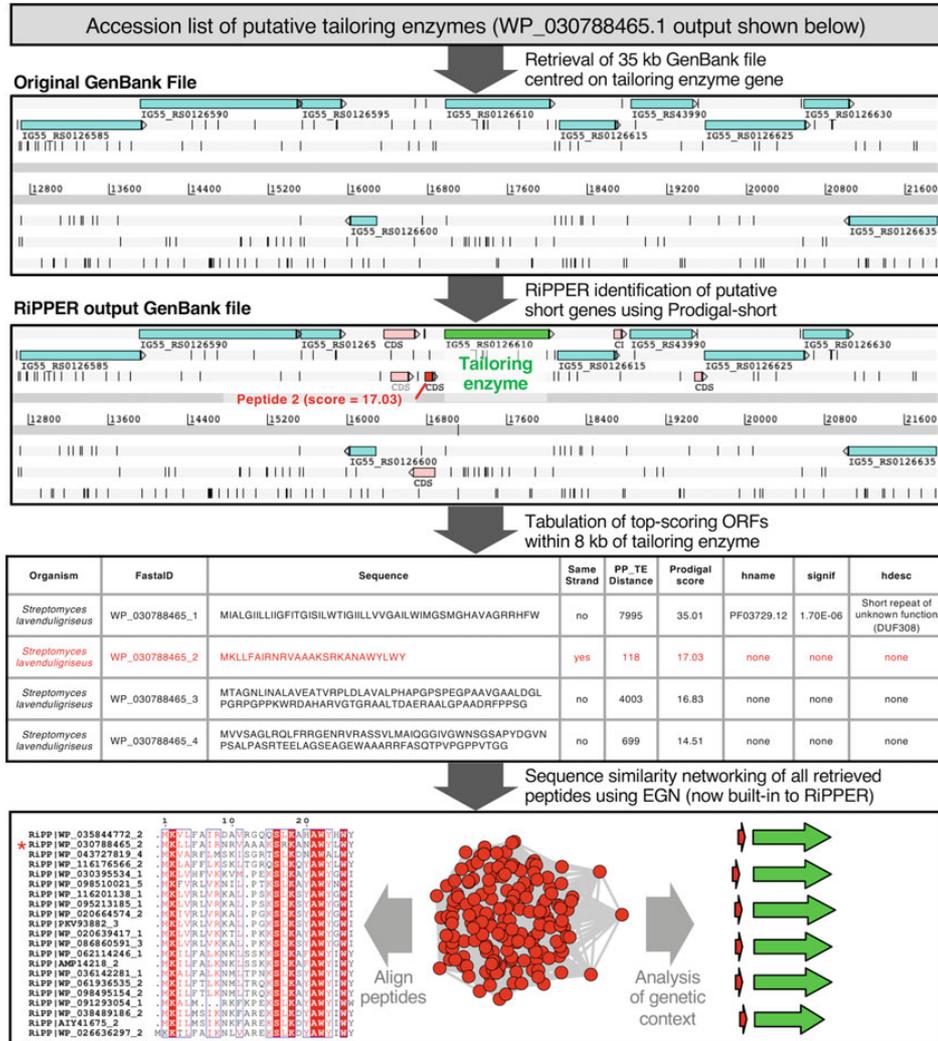


Fig. 2 Overview of RiPPER 1.1 outputs using an input of accessions of cytochrome P450s with homology to the P450 in the tryptorubin BGC [18]. Dark grey arrows indicate automated steps in the RiPPER tool. The GenBank files are visualized in Artemis [15], and the tryptorubin-like precursor peptide is highlighted in red in the table and starred in the alignment

threshold (default = 7.5). This RiPPER output includes tables of numerous peptide parameters, such as distance and orientation of gene in relation to the bait gene, as well as a conserved domain search against both the Pfam database [16] and a NCBI database with conserved domains specifically for RiPP precursor peptides [17] (Fig. 2).

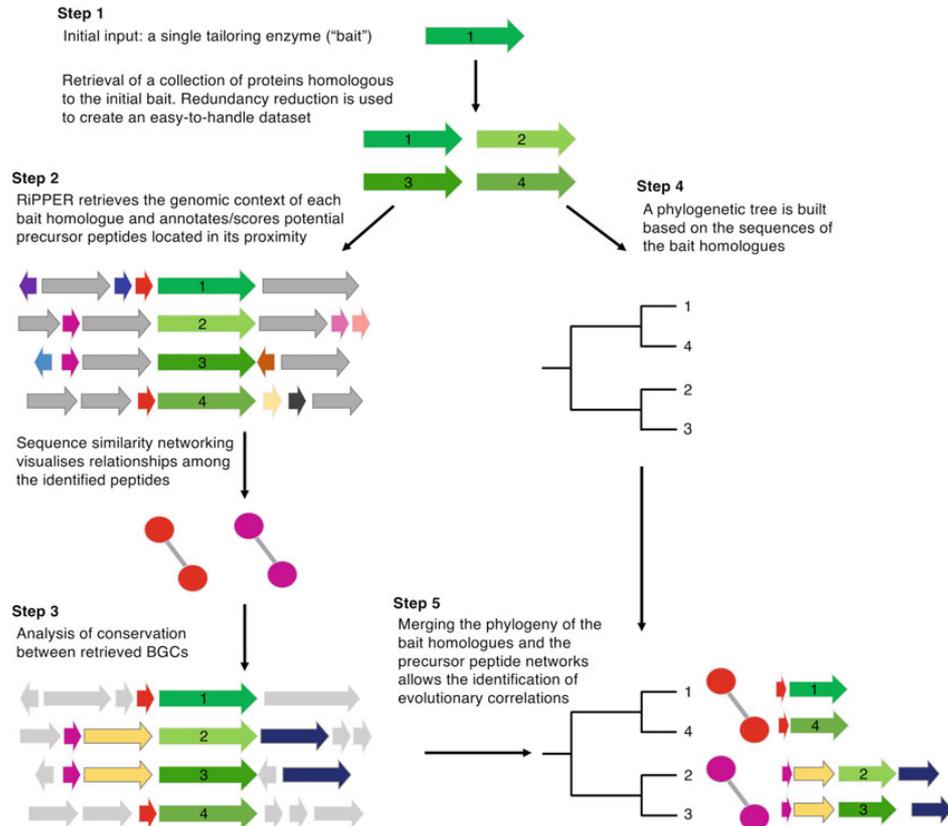


Fig. 3 The biological logic of the RiPPER 1.1 mining workflow described in this chapter

In this chapter, we will describe a workflow that combines RiPPER with a series of pre-existing bioinformatic functionalities, allowing identification of phylogenetic correlations between networks of homologous precursor peptide candidates and groups of homologous tailoring enzymes (Fig. 3). The RiPPER workflow requires more commitment from the user than web server-based genome mining tools, but its advantages are substantial. RiPPER facilitates the exploration of the RiPP diversity landscapes in a very flexible way as it allows the user to focus on any tailoring enzyme (well-characterized or suspected) as starting bait. We believe this is critical for the discovery of truly new RiPP chemistry. Therefore, instead of looking for predefined RiPP classes across a given genome, the RiPPER workflow described here allows exploration of the whole diversity of RiPP BGCs containing a given tailoring enzyme.

We have now updated RiPPER in numerous ways to generate RiPPER 1.1. This includes the integration of sequence similarity networking by the EGN (Evolutionary gene and genome network) tool [19], which allows the user to rapidly identify groups of related precursor peptides. In the workflow we describe, the output is then assessed for the co-occurrence of tailoring enzymes with co-evolving candidate precursor peptides. Further data processing is then used to dissect this diversity into subgroups by comparing BGC architecture. We have found this approach invaluable for the identification of large families of novel RiPP BGCs, where the existence of conserved precursor peptides that belong to evolutionarily conserved gene clusters highlights the likelihood of new RiPP families.

2 Materials

1. A PC, Macintosh or Linux computer with administrator privileges and an internet connection. The example reported here was carried out using a Mac with macOS Catalina (10.15), a 1.8 GHz (two cores) processor and 8 GB RAM. If a Windows PC is used, this requires Windows 10 Pro, Education or Enterprise for the installation of Docker. This requires a 64-bit processor and a minimum of 4 GB RAM. Hyper-V and Windows Container Features must be enabled. Approximately 14 GB hard drive is required for the storage of Docker images.
2. Installed versions of Docker (<https://www.docker.com/products/docker-desktop>), Cytoscape (<https://cytoscape.org>), Microsoft Excel, and a text editor capable of displaying both Unix and DOS line endings correctly (we recommend Sublime Text or jEdit, which is free to download). The software versions we use in the steps below are Docker Desktop 2.2.0.3, Cytoscape v3.7.2, Excel 16.16.17, and Sublime Text 3.0. On a Windows PC the user must specify using Linux containers rather than using Windows containers upon installation of Docker. This is done by not clicking the box that says “use Windows containers instead of Linux containers”.
3. Tools to generate and visualize phylogenetic trees. In the steps below, we use iTOL [20] (<https://itol.embl.de>) and CIPRES Science Gateway [21] (<https://www.phylo.org>), which both require free online accounts.

3 Methods

The RiPPER-based genome mining strategy we will describe here involves several data processing steps (Fig. 3). The initial input consists of the accession of a single protein that is suspected or known to be involved in RiPP biosynthesis. In the method presented here, we use AAB17515.1, which is a TfuA domain protein from *Rhizobium leguminosarum* bv. *trifolii*. TfuA domain proteins function with YcaO domain proteins to catalyze thioamidation of peptide backbones [13, 22]. This was believed to be a very rare modification, but we had previously shown that TfuA domain proteins are widespread in novel RiPP BGCs in Actinobacteria [13]. In the steps below, as an example of how to operate with the RiPPER-based genome mining workflow, we provide a detailed summary of how to search for potential RiPP BGCs in Alphaproteobacteria using a TfuA domain protein.

This input accession can be used in a number of ways to retrieve related proteins, such as BLAST analysis to identify related proteins. Here, we will instead use the Conserved Domain Architecture Retrieval Tool (CDART) [23] to download all proteins with the same conserved domain. CDART can additionally filter by taxonomic classification if required (for example, if the user wants to focus on a particular area of biodiversity). As the number of retrieved proteins can potentially be enormous and difficult to handle, a trimming process is normally performed to reduce the redundancy of protein set. Here, the Enzyme Function Initiative-Enzyme Similarity (EFI-EST) tool is used [24]. A list of accession numbers corresponding to the proteins “baits” is now ready to be used with the RiPPER tool.

RiPPER uses this accession list to retrieve the genomic context of each of those proteins in a flexible manner, and annotates and scores the potential precursor peptides existing within a given genomic window around the bait homologue following a series of customisable parameters. Following RiPPER analysis, we will have a list of baits associated with three or more top-scoring co-occurring short peptides, as well as a GenBank files relating to each putative BGC.

The ultimate aim of the RiPPER-based strategy described here is to identify groups of related precursor peptides that have undergone co-evolution alongside groups of homologous bait proteins. The identification of conserved short peptides associated with conserved genomic loci suggests that potential precursor peptides are genuine, and potentially acted on by the bait proteins. It is possible that short peptides with other functions also associate with conserved BGCs, such as transcription factors and chaperones like PqqD [25]. Here, the conserved domain search for the retrieved short peptides assists the user. We have also recently updated

RiPPER to automatically carry out sequence similarity networking analysis of the resulting short peptides by incorporating the EGN tool with predefined settings for short peptide networking. This provides a series of peptide networks, which should each contain closely related peptides, and this can all be visualized in Cytoscape [26]. In the workflow described below, a phylogenetic tree of the bait proteins is created. The association of networks of precursor peptides alongside defined branches of the phylogenetic tree is a strong indication of the evolution of a family of related RiPP BGCs and suggests that the precursor peptide and tailoring enzyme are co-evolving. We also describe how the architecture of the candidate BGCs can be explored using MultiGeneBlast [27], providing additional insight into their conservation and diversity. The software and bioinformatic tools associated with this workflow are summarized in Fig. 4.

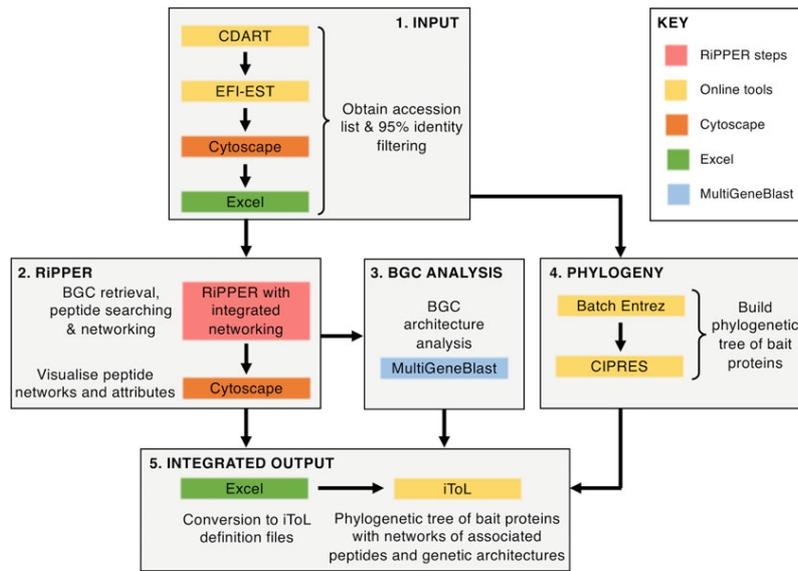


Fig. 4 Suggested workflow for identification of RiPP gene clusters and precursor peptides from genomes using bait proteins in RiPPER 1.1. Tools are shown in colored boxes, with arrows indicating the order of steps described in this chapter. Some steps are facilitated by the use of Excel formulas and/or usage of the command line; the formulas/code we use are presented in the text. Note, the presented workflow is suggested but is not the only possible means of conducting the analysis

3.1 Obtain an Input List of Protein Accessions

3.1.1 CDART

1. Begin by selecting a desired bait protein, which you suspect may be involved in RiPP biosynthesis. In the example here, we use a TfuA domain protein, AAB17515.1, from the alphaproteobacterium *Rhizobium leguminosarum* bv. *trifolii*. Similar proteins can be found in other organisms using CDART from NCBI (<https://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi>).
2. Within the CDART interface, use the “Filter your results” option (NCBI Taxonomy Tree setting) to limit the search to your organism(s) of interest and then select “Lookup sequences in Entrez”. Here, we have chosen to focus on proteins from Alphaproteobacteria.
3. Download (“Send to” option) as an Accession List, and you can simply proceed to Section 3.2 if desired. However, we recommend the use of EFI-EST to reduce the dataset size by reducing bait protein redundancy. This is useful as it reduces the likelihood of a set of very closely related proteins from a highly sequenced species dominating the RiPPER output (*see Note 1*).

3.1.2 EFI-EST

1. Use the sequence similarity networking tools at EFI-EST (<https://efi.igb.illinois.edu/efi-est/>) to reduce the size of the dataset, where the accession list is inputted via the Accession IDs input option. We generally use a 95% identity reduced dataset, where any sequences more than 95% identical are collapsed to a single representative sequence and accession. In this example, we used the default settings for the initial submission. EFI-EST will then require you to finalize the Sequence Similarity Networks (SSNs) by selecting size and alignment score cutoffs.
2. It is advised within EFI-EST to select an alignment score corresponding to 35% similarity. In this example, we selected 120 AAs as the lower end cutoff, with no upper limit, and an alignment score threshold of 22.
3. Once the SSN is finalized (*see Note 2*), download the 95% identity network and open it using Cytoscape.
4. Export the associated attribute table as a .csv file, and open this using Excel. Find the column containing the list of accession IDs. In our experience, these are in a column called “Query IDs” or “Description” and include multiple IDs if the node in the network represents multiple proteins that are >95% identical.
5. To obtain a list that includes one accession from each node, we do the following: copy this column to column A of a new Excel sheet, and extract the first accession from each row using the following Excel command in a different column of that new

sheet, where “|” represents the symbol that follows the first accession. In some examples, this is instead a space (“ ”, *see Note 3*):

```
= IF (ISERR(FIND("|",A2)),A2,LEFT(A2,FIND("|",A2)-1))
```

6. Apply this to all rows, thereby generating a new column with only the one accession per row. Now copy this column and paste it into a new .txt file, generating a shortened accession list in which entries share less than 95% identity to use as input for RiPPER and phylogenetic analysis. Note that Section 3.4 (phylogenetic analysis of bait proteins) can take significant time to compute and only requires this input accession list, so can be started prior to Sections 3.2 and 3.3 if the user prefers.

3.2 Short Peptide Searching and Networking Using RiPPER 1.1

3.2.1 BGC Retrieval and Short Peptide Searching

1. Online details on how to use RiPPER 1.1 are found at <https://github.com/streptomyces/ripper>. We highly recommend that it is run from a Docker container, which includes all dependencies and can be run from Linux/Mac/Windows systems. Docker must be installed and running before RiPPER 1.1 can be used (<https://www.docker.com/get-started>). To install RiPPER 1.1, open the command line interface (terminal in Mac, command prompt in Windows) with Docker running and use the command:

```
docker pull streptomyces/ripdock
```

2. Following installation, run the container using the following command, where your input accession list is stored in “filepath”. Do not change /home/mnt as this refers to a directory in the container and is required for the scripts to run:

```
# Example usage on Linux / Mac
```

```
docker run -it -v /filepath:/home/mnt streptomyces/ripdock
```

```
# Example usage on MS Windows.
```

```
docker run -it -v C:/filepath:/home/mnt streptomyces/ripdock
```

3. Now you can run RiPPER 1.1 on your accession list, to search for short peptides co-occurring with your bait protein and obtain all RiPPER output files (*see* summary below). RiPPER 1.1 will retrieve short peptides based on a set of parameters that have sensible defaults for RiPP analysis (Fig. 2): using Prodigal-short, peptides between 20 AA (minPPlen) and 120 AA (maxPPlen) in length are searched for within a 17.5-kb window (flanklen) either side of the bait protein, where a score boost of 5 (sameStrandReward) is included if the peptide is encoded on the same strand as the bait protein, which is a common feature in RiPP BGCs. Peptides within a window of 8 kb (maxDistFromTE) either side of the bait protein are

considered for the output peptide list, and three are retrieved (fastaOutputLimit) based on the Prodigal-short score ranking. Additional peptides are retrieved if they have a score above 7.5 (prodigalScoreThresh). We use these defaults for the analysis described here, but these parameters can be modified (*see Note 4*). Assuming your list of bait proteins is stored in a file called “Accession_IDs.txt”, use the following command to run RiPPER:

```
./ripper_run.sh /home/mnt/Accession_IDs.txt
```

4. RiPPER 1.1 may take several hours to run, due to the time required to download GenBank files from NCBI sequentially based on the accession ID list, which requires a stable internet connection (*see Note 5*).
5. A number of output files and folders will be generated:
 - orgnamegbk:** A folder containing GenBank files for all retrieved gene clusters. These files are named by the host organism and the input accession ID.
 - out.txt:** This is a tab-delimited file that reports peptide data for all peptides retrieved by RiPPER (as defined by the analysis parameters above) for all gene clusters. This includes various associated data for the peptides, including sequence, species, Prodigal-short score, Pfam domain, and distance from the bait protein.
 - distant.txt:** This is a tab-delimited file that reports peptide data for all peptides retrieved by RiPPER via a precursor peptide Hidden Markov Model (HMM) search across the full size of the retrieved gene cluster (i.e., ignores maxDistFromTE), and only reports peptides that were not identified in the original RiPPER search. Peptide identifiers are unique from out.txt, so these text files can be combined for downstream analysis. This is useful for retrieving peptides from large RiPP BGCs (e.g., thiopeptides).
 - out.faa:** A fasta file for peptides reported in out.txt that is formatted as an input file for similarity networking using EGN. An equivalent distant.faa file is also provided.
 - rodeohtml:** A folder containing RODEO [8] html output files for all retrieved gene clusters. Additional RODEO data is provided in a separate rodout folder. The data stored in rodout can be further analyzed via approaches described on the RODEO website (<http://ripp.rodeo/advanced.html>).
 - pna:** A folder containing the output of EGN-based sequence similarity networking, which includes files for visualization in Cytoscape (*see below*).
 - Networks:** A folder containing subfolders (e.g., “Network1”, “Network2”) that contain GenBank files associated

with each peptide network identified by EGN networking (*see* below).

6. Following the RiPPER run, exit the Docker container by typing “exit”.

3.2.2 Short Peptide Networking

1. It is possible to utilize the output of RiPPER directly, for example, by looking at the .gbk files stored in orgnamegbk, but in earlier studies [13] we found it very helpful to generate sequence similarity networks (SSNs) of the output short peptides using EGN [19]. This step identifies families of related short peptide sequences, which can highlight the existence of conserved RiPP-like BGCs [13]. Other SSN software may be used, but we find that the customisable nature of EGN works best with the short sequences generated in this analysis.
2. To significantly simplify the workflow, we have now improved RiPPER to incorporate an automated EGN-based networking analysis. The default parameters we have defined within RiPPER are that peptides are networked if they have over 40% identity and there is over 40% sequence coverage, with a minimum alignment region of 14 AA. These are relatively permissive networking settings, but our testing indicates they are highly effective at grouping families of related precursor peptides. An advanced user may wish to use EGN themselves to assess alternative settings [19]. The EGN output is provided in folder “pna.” Within this folder the network file is found in /GENENET_10.40.40.0.0/CYTOSCAPE as a file titled “cc_1.to.X.txt”, which in our example is “cc_1.to.78.txt”.
3. This should be imported into Cytoscape to visualize the series of networks generated with the short peptides found by RiPPER (Fig. 5). Open Cytoscape and import the file “cc_1.to.X.txt” by dragging it into the **network** panel on the left.
4. Following this, the Cytoscape attributes file outputted by RiPPER (“cytoattrib.txt” in “pna” folder) can be imported into Cytoscape by dragging into the **table** panel along the bottom. Mapping attributes to the network means that the nodes are associated with the information provided in the RiPPER output, such as peptide sequence, Prodigal-short score, and conserved domain. This is therefore a suitable interface for assessing the RiPPER output.
5. The network can now be colored by using the “Style” pane. Under the “Fill Color” option, select “Column: Color” and “Mapping type: Passthrough Mapping”. The networks will now be color-coded the same as in a phylogenetic tree that is color-coded in Section 3.5 (Fig. 5). Node formatting can also be used to highlight numerous other peptide features (*see* **Note 6**).

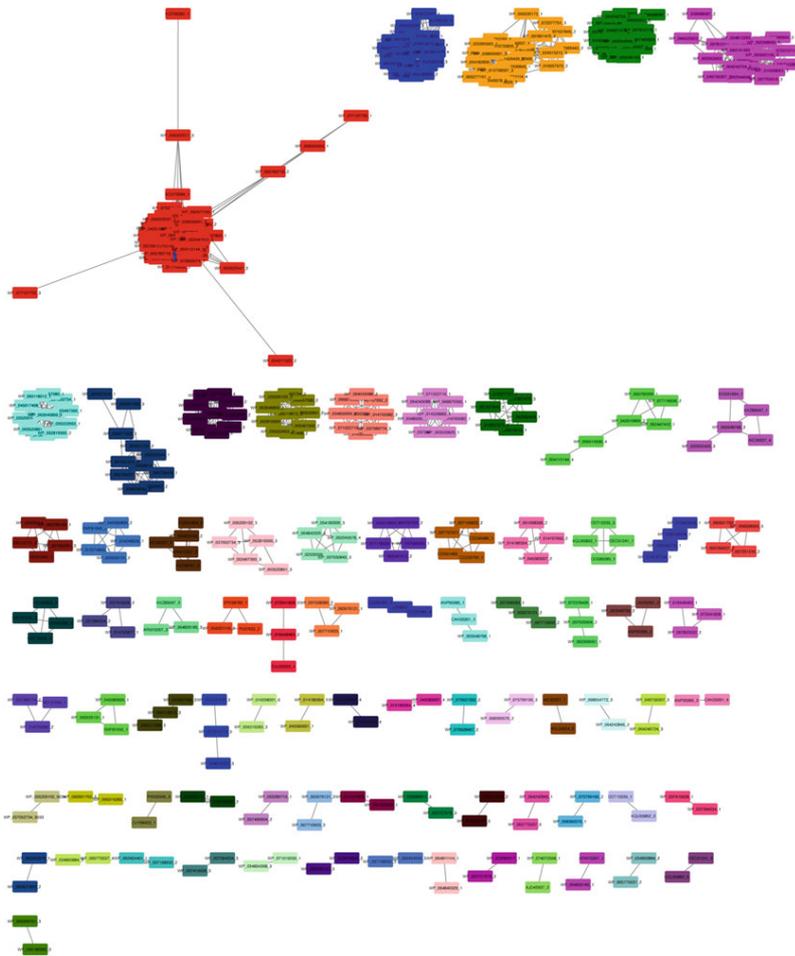


Fig. 5 The short peptide network output when using alphaproteobacterial TfuA proteins as the input is described in this chapter. The network file is found in the /GENENET_10.40.40.0.0/CYTOSCAPE/ folder and then recolored using the “cytoattrib.txt” file

3.3.1 MultiGeneBlast (MGB)

1. MGB can be very useful to use on genomic loci associated with each network of short peptides. To achieve this, the GenBank files outputted by RiPPER are divided into different networks identified in Section 3.2. RiPPER now does this automatically for GenBank files associated with the top 30 networks, which are saved in folders for each network in /Networks/NetworkX (e.g., “Network1”, “Network2”).

2. MGB for Windows is currently only available as 32-bit software, which is incompatible with modern Windows operating systems. We have therefore generated a Docker image for MGB which contains the 64-bit Linux version of MGB, which we recommend is used by both PC and Mac users. This can be used as follows. Start Docker desktop, open a new terminal window, and type the command:

```
docker pull streptomyces/multigeneblast
```

3. Start the Docker container as below, where “filepath” relates to the path to the “Networks” folder (or equivalent if the user moves the “NetworkX” folders):

```
docker run -it -v /filepath:/home/workstreptomyces/multigeneblast
```

4. MGB requires a custom database to be searched against. Therefore, make a database for each relevant network with the following command (DBnetwork_X = database name, Network_folder = folder with desired GenBank files):

```
makedb DBnetwork_X Network_folder
```

Example:

```
makedb DBnetwork_2 Network2
```

Where the files for creating the “DBnetwork_2” database are GenBank format files from the RiPPER output, stored in the directory /Networks/Network2.

5. You can then run MGB using the following commands (“DBnetwork_X” = database previously generated; Network_folder/example.gbk = query file; “mgbout” = destination folder name for MGB results). This assumes the query file is in a “NetworkX” folder, so specify a different path if this is not the case:

```
multigeneblast -db DBnetwork_X -in Network_folder/example.gbk -from 0 -to 35000 -out mgbout
```

Example:

```
multigeneblast -db DBnetwork_2 -in Network2/WP_003184680.1.gbk -from 0 -to 35000 -out Network2_WP_003184680_MGB
```

6. The folder “Network2_ WP_003184680_MGB” will be created, which contains .svg files that visualize each GenBank file compared to your query GenBank file, and a combined .svg file with all gene clusters compared to the query (Fig. 6). .xhtml files are also outputted that can be viewed in a web browser. This provides an interactive color-coded visualization of all gene clusters compared to the query file. The user must ensure the output folder name is changed in the command when running another search, or the folder will be overwritten.
7. Following MGB analysis, exit the Docker container by typing “exit”.

3.3.2 Peptide Alignment

1. It can be informative to directly assess the sequence conservation of peptides that belong to a given network (Fig. 6). This step can help identify potential outliers within a Network. Network-specific FASTA files are generated automatically by EGN running within RiPPER (saved in /pna/GENE-NET_10.40.40.0.0/FASTA/) and are named “p_ccX.faa”, where X is the network number.
2. These can therefore be used directly as the input for multiple sequence alignment using MUSCLE (<https://www.ebi.ac.uk/Tools/msa/muscle/>).

3.4 Building Phylogenetic Tree of Bait Proteins

Building a phylogenetic tree of your selected bait proteins (TfuA domain proteins in Alphaproteobacteria in this example) can help illustrate the diversity of these proteins in your dataset, and aids downstream identification of evolutionarily associated BGCs.

3.4.1 Batch Entrez

1. The first step is to retrieve sequences for all the proteins in your accession list. We do this using the Batch Entrez tool from NCBI (<https://www.ncbi.nlm.nih.gov/sites/batchentrez>). Simply upload the 95% identity filtered list that was generated in Section 3.1, and click “retrieve”, making sure to select the “protein” option from the dropdown menu. This results in a list of these proteins on the NCBI website. Simply click the “Send to” option, select “File”, and change format to “FASTA”, then “Create file.” This will download the full list in FASTA format with amino acid sequences for each accession ID in the list. We will use this as an input to create the phylogenetic tree.

3.4.2 CIPRES

1. We use the CIPRES Science Gateway (<https://www.phylo.org>) to create a phylogenetic tree of our sequences. Users will need to create a free account in order to use the service.
2. Create a new project and upload your FASTA file from the above step in the “Data” subfolder. In the “Tasks” subfolder, create a new task with your FASTA file as the input, and use the

“Muscle” tool to perform an alignment. Ensure that sufficient time is allocated via the maximum run time setting, then save and run the analysis. We use default alignment parameters here.

3. Once the alignment has run, download the “output.fasta” file, and upload this in the “Data” subfolder as before.
4. Then create a new task to build the phylogenetic tree, selecting the “output.fasta” file as input. We find the “RAxML-HPC2 on XSEDE” tool works well, which we use with default parameters in this example, ensuring that sufficient time is allocated via the maximum run time setting and that “Protein” is set for Data Type.
5. Once this has run, download the “RAxML_bestTree” output file. This can then be imported directly into Interactive tree of Life (iTOL, <https://itol.embl.de>) [20].

3.5 Visualize the Integrated Output

3.5.1 Mapping Short Peptide Networks to Protein Clades

1. In order to map short peptide networks to protein clades in iTOL, it is necessary to make use of iTOL definition files. A template is available at <https://itol.embl.de/help.cgi> as a file “colors_styles_template.txt”, which should be downloaded. To convert the data into the right format to import into iTOL requires processing using Microsoft Excel and a text editor, such as Sublime Text.

Start by opening the network attribute table .csv file from Section 3.2. Cluster 1 peptides (cc1 in “Network” column) occur in rows 2 to X (2–67 in our example), all cluster 2 peptides (cc2) in rows $X + 1$ to Y (68–91 in our example), and so on. Save as an Excel Workbook (e.g., “iTOL_definitions.xlsx”) and create a new sheet with the following columns defined: “Accession” (column A), “Full Accession” (B), “Color type” (C), “Color” (D), “Style” (E), “Network” (F). Copy and paste columns “Accession”, “Color”, and “Network” from the original sheet into the relevant columns. The “Full Accession” column is required to generate an accession code format that matches the format in the tree. The following Excel command is therefore used in cell B2 and then applied to all rows in the column:

```
= A2&" .1"
```

“Color type” has multiple options, but we find “branch” works well; simply type “branch” in the top row for the “Color type” column, and use the fill down option to apply this to all rows. For “Style”, simply type “normal” into the top row and use the fill down option to apply it to all rows.

2. Once this is done, columns B–E for the relevant networks can be pasted into the bottom of the iTOL-supplied “colors_styles_template.txt” file below “DATA”. It will also be necessary

to change the separator value in the template file to TAB; do this by typing “#” before the “Space” and “Comma” options, and remove the “#” before “Tab”; failing to do so will render iTOL unable to locate the nodes in the tree to assign colors to.

3.5.2 Final Output

1. The above steps provide a large amount of data that simply originates from a single input bait protein: (i) 35 kb GenBank files centered on putative BGCs with short peptides scored and visualized; (ii) tabulated data for short peptides and their associated attributes, including conserved domains; (iii) Peptide networks that highlight conserved sets of peptides across multiple BGCs; (iv) MultiGeneBlast comparison of BGCs for specified networks; (v) A phylogenetic tree of non-redundant homologues of the original bait protein. This data can now be integrated and visualized.
2. Firstly, to recolor the tree based on peptide networks, the iTOL definition file(s) generated above are simply dragged into a web browser while the phylogenetic tree is open in iTOL. The branches of the bait proteins that have nearby short peptides belonging to a given network will be automatically recolored (Fig. 7). Some networks may map to monophyletic groups, while others may not. The recolored tree can be exported as a .png file.
3. It can be useful to investigate the MultiGeneBlast output to view the surrounding genes to help explain this. Graphics generated in Section 3.2 (Cytoscape-generated network images) and Section 3.3 (MultiGeneBlast outputs and peptide sequence alignments) can then be visualized alongside the colored tree using software such as Microsoft PowerPoint or Adobe Photoshop (Fig. 7).

Where closely related bait proteins are associated with similar short peptides, and the surrounding gene architecture is highly conserved, it highlights the possibility that this represents a genuine RiPP biosynthetic gene cluster that is conserved across multiple species (Fig. 7).

4 Notes

1. Bait proteins from heavily sequenced species, such as widely studied clinical pathogens, can potentially distort the output, as proteins with almost identical sequences are very likely to retrieve the same precursor peptide and BGC via RiPPER, and will also be over-represented in phylogenetic analyses. This can lead to an overestimation of the natural prevalence of the resulting RiPP BGC family. These problems are solved

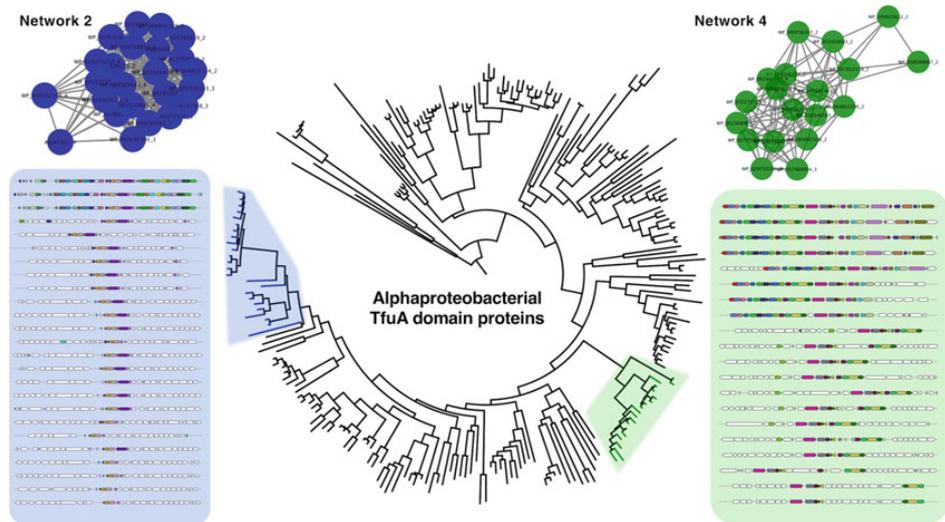


Fig. 7 Final RiPPER workflow output combining bait protein phylogenetics, precursor peptide identification, and genetic architecture investigation. The final output shows a phylogenetic tree of supplied bait proteins across a user-defined phylogenetic group, onto which are mapped networks of similar putative RiPP precursor peptides identified by RiPPER. MGB can then inform the user whether the surrounding gene architecture is conserved, potentially supporting the hypothesis that these are real RiPP biosynthetic gene clusters

by using an identity cutoff to reduce redundancy. We use 95% identity to filter any proteins that are almost identical, but lower cutoffs can be used, especially if the input protein list is very large. EFI-EST provides the option to filter based on numerous percentage identity cutoffs. The user is warned that the use of a lower cutoff (e.g., 70% identity) will likely mean that RiPP diversity is reduced in the output.

2. In our experience, there can sometimes be a disparity between the number of NCBI accessions and the number of proteins that are used for network analysis in EFI-EST. This is due to issues with matching NCBI accessions with UniProt IDs. This can be overcome by using the FASTA input option in EFI-EST instead. *See* Section 3.4.1 for instructions on how to generate a FASTA file from an accession list via Batch Entrez.
3. Excel formulas provided are tailored towards the example presented. Most will work as presented with data from the RefSeq database. A simple modification to the presented formula = IF (ISERR(FIND("|",A2)),A2,LEFT(A2,FIND("|",A2)-1)) would be to change the "|" character to whichever character is used to separate the accession IDs in the Cytoscape output "Query IDs" column.

4. The RiPPER parameter file is called “local.conf”. If you would like to modify this, once the Docker container is running, copy it into the working directory (i.e., “filepath”) on your computer using the following command:

```
cp local.conf /home/mnt
```

Modify and save this locally stored file using a text editor and then use the following command to copy it back into the working directory for RiPPER:

```
cp /home/mnt/local.conf /home/work
```

Following this step, we recommend testing a small subset of input accessions to ensure parameters have been properly modified.

5. Occasionally, RiPPER may fail to retrieve a sequence for a given accession, so an error will be listed on the command line, and RiPPER will then proceed to the next accession. This can occur if the sequence is corrupted, if there is an error with the accession ID (for example, a nucleotide accession is used, an accession ID has changed, or there is no nucleotide accession associated with the protein), if a lapse in internet connection occurs or if there is a problem with the NCBI server. At present, RiPPER does not attempt to re-run these at the end of the analysis, and they will be absent from your dataset. However, this will be addressed in a future version of RiPPER.
6. In our experience, Cytoscape visualization is a powerful tool to assess the properties of RiPPER-identified peptides. For example, as an alternative to coloring by network, color-coding and/or node size can be used to assess Prodigal score (using Continuous Mapping to visualize low to high scores), conserved domains (“hdesc” using Discrete Mapping), or genus (Discrete Mapping).

References

1. Katz L, Baltz RH (2016) Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol* 43:155–176
2. Genilloud O (2019) Natural products discovery and potential for new antibiotics. *Curr Opin Microbiol* 51:81–87
3. Medema MH, Blin K, Cimermancic P et al (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* 39:W339–W346
4. Blin K, Shaw S, Steinke K et al (2019) antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87
5. Skinnider MA, Merwin NJ, Johnston CW et al (2017) PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res* 45:W49–W54
6. Montalbán-López M, Scott TA, Ramesh S et al (2021) New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* 38: 130–239 <https://doi.org/10.1039/d0np00027b>
7. Tietz JI, Schwalen CJ, Patel PS et al (2017) A new genome-mining tool redefines the lasso

- peptide biosynthetic landscape. *Nat Chem Biol* 13:470–478
8. Schwalen CJ, Hudson GA, Kille B et al (2018) Bioinformatic expansion and discovery of thiopeptide antibiotics. *J Am Chem Soc* 140:9494–9501
 9. Li J, Qu X, He X et al (2012) ThioFinder: a web-based tool for the identification of thiopeptide gene clusters in DNA sequences. *PLoS One* 7:e45878
 10. Agrawal P, Khater S, Gupta M et al (2017) RiPPMiner: a bioinformatics resource for deciphering chemical structures of RiPPs based on prediction of cleavage and cross-links. *Nucleic Acids Res* 45:W80–W88
 11. Skinnider MA, Johnston CW, Edgar RE et al (2016) Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining. *Proc Natl Acad Sci U S A* 113:E6343–E6351
 12. Mohimani H, Kersten RD, Liu W-T et al (2014) Automated genome mining of ribosomal peptide natural products. *ACS Chem Biol* 9:1545–1551
 13. Santos-Aberturas J, Chandra G, Frattaruolo L et al (2019) Uncovering the unexplored diversity of thioamidated ribosomal peptides in Actinobacteria using the RiPPER genome mining tool. *Nucleic Acids Res* 47:4624–4637
 14. Hyatt D, Chen G-L, LoCascio PF et al (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119
 15. Carver T, Harris SR, Berriman M et al (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28:464–469
 16. Finn RD, Coggill P, Ry E et al (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285
 17. Haft DH, DiCuccio M, Badretdin A et al (2018) RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860
 18. Reisberg SH, Gao Y, Walker AS et al (2020) Total synthesis reveals atypical atropisomerism in a small-molecule natural product, tryptorubin A. *Science* 367:458–463
 19. Halary S, McInerney JO, Lopez P et al (2013) EGN: a wizard for construction of gene and genome similarity networks. *BMC Evol Biol* 13:146
 20. Letunic I, Bork P (2019) Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259
 21. Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES science gateway for inference of large phylogenetic trees. In: 2010 gateway computing environments workshop (GCE). IEEE, New Orleans, LA, pp 1–8
 22. Mahanta N, Liu A, Dong S et al (2018) Enzymatic reconstitution of ribosomal peptide backbone thioamidation. *Proc Natl Acad Sci U S A* 115:3030–3035
 23. Geer LY, Domrachev M, Lipman DJ et al (2002) CDART: protein homology by domain architecture. *Genome Res* 12:1619–1623
 24. Gerlt JA, Bouvier JT, Davidson DB et al (2015) Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks. *Biochim Biophys Acta* 1854:1019–1037
 25. Latham JA, Iavarone AT, Barr I et al (2015) PqqD is a novel peptide chaperone that forms a ternary complex with the radical S-adenosylmethionine protein PqqE in the pyrroloquinoline quinone biosynthetic pathway. *J Biol Chem* 290:12908–12918
 26. Shannon P, Markiel A, Ozier O et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504
 27. Medema MH, Takano E, Breitling R (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* 30:1218–1223
 28. Navarro-Muñoz JC, Selem-Mojica N, Mullenney MW et al (2020) A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol* 16:60–68