Carbon monoxide adaptation and utilisation in *Cupriavidus necator*

Charles Wickham-Smith, MRes

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2019

Declaration

Unless otherwise acknowledged, the work presented in this thesis is my own. No part has been submitted for another degree in the University of Nottingham or any other institute of learning.

Charles Wickham-Smith

September 2019

Abstract

The C1 gas carbon monoxide (CO) is abundantly present in industrial waste gases, which can be used as a feedstock for the biological production of industrially significant chemicals and fuels. Fermentation with such CO-rich gases may form a vital part of the strategy to reduce the levels of greenhouse gases in the Earth's atmosphere, which are damaging to the environment. For efficient bacterial growth and to increase product yield in the presence of high CO concentrations, resistance to the gas and its utilisation are required. The aerobic C1 chassis *Cupriavidus necator* H16 is able to utilise CO_2 + H_2, although cannot use CO as a sole source of carbon and energy and is inhibited by the gas. To increase the tolerance of *C. necator* to CO, the organism was continually subcultured in the presence of CO. Ten individual cultures were evolved in this manner and, after 15 subcultures, displayed a clear growth advantage when compared to the wild type strain. Following whole genome sequencing of single isolates from these cultures, a number of unique mutations were found, including a single point mutation upstream of a cytochrome bd unbiquinol oxidase operon (*cydA2B2*), which was present in all ten isolates. When a subset of these mutations was engineered into the genome of the original parental strain, only the upstream *cydA2B2* mutation enabled the fast growing phenotype in the presence of CO. Characterisation of this mutation and the *cydA2B2* operon through gene expression analysis, mutation and overexpression suggested that cytochrome bd production is upregulated in the evolved isolates, resulting in increased CO tolerance.

Utilisation of CO requires CO dehydrogenases (CODHs), which are found in various chemolithoautorophic organisms, which, in contrast to *C. necator*, harbour functional CODH-encoding cox gene clusters. To equip *C. necator* with the required set of genes, transfer of the *Oligotropha carboxidovorans* pHCG3 megaplasmid was investigated, as it contains one of the best characterised cox gene cluster and a quorum sensing (QS) system assumed to play a role in regulating conjugal transfer. To optimise pHCG3 transfer, the putative QS repressor TraM was inactivated, enabling constitutive N-acyl-L-homoserine lactone production, boosting the conjugative functions of pHCG3. Transfer into plasmid-less strains of *Agrobacterium tumefaciens* was achieved, enabling AHL production in this organism, but not the ability to utilise CO or CO_2 + H_2. A novel species of *Pseudomonas* was accidentally

isolated whilst carrying out conjugation experiments, which stably maintained this pHCG3 derivative, although the plasmid appeared to be non functional. pHCG3 could not be successfully transferred into *C. necator*, presumably due to species specific barriers such as restriction systems, plasmids instability, host immunity or plasmid incompatibility.

Acknowledgements

I would like to sincerely thank my primary supervisor Dr Klaus Winzer for his support, whose encouragement guided me throughout this PhD project. I would additionally like to thank my co-supervisor Dr Naglis Malys for his help and direction and my bench supervisor Dr Swathi Alagesan for her patient assistance at the beginning of this project. I would also like to thank all the staff in the CBS and SBRC for their help, in particular Dr Nigel Halliday who helped me with crucial sample analysis. Without funding for this work, it would not have been possible so thanks to the University of Nottingham's BBSRC Doctoral Training Partnership (DTP).

Thanks to everyone in B67 and B52 for their help and making life in the lab highly enjoyable. Jon Humpherys also deserves good thanks for his advice and pub presence, as well as Raj, Alex and Sarah for their support from the start. Luke, Oran, Max, James, Chris, Stef, Stan, the Techfest and ArcTanGent gangs, thanks for all the good times and the sweet music. Also thanks to all the GMs for being great and supportive.

A special thank you to Beth for being so understanding and inspirational and to my Mum, Dad, brother Henry, sister Georgie, Nanna, Aunty Manda and rest of my family for all the support and encouragement.

Table of contents

1: Introduction....24
1.1. Tackling the problem of greenhouse gas pollution....24
1.1.1. C1 gas....24
1.1.2. Fermentation and product formation....26
1.1.2.1. Chemolithoautotrophs....26
1.2. *Cupriavidus necator* H16....27
1.2.1. Genome features....28
1.2.2. Metabolism....28
1.2.3. Industrial career and product formation....29
1.2.4. Enhancing the use of waste gas feedstocks for *C. necator* H16....31
1.3. Carbon monoxide....31
1.3.1. CO-oxidising bacteria....32
1.3.2. Tolerance to CO....33
1.3.3. *Oligotropha carboxidovorans*....35
1.3.3.1 Genome features....35
1.3.3.2. Cox genes....37
1.4. Megaplasmids....39
1.4.1. RepABC-family plasmids....39
1.4.1.1. Replication....39
1.4.1.2. Evolution....40
1.4.1.3. Conjugal transfer systems....40
1.4.1.4. Quorum sensing....41
1.4.2. Horizontal gene transfer....43
1.4.2.1. Megaplasmid transfer in α-proteobacteria....43
1.4.2.2. Megaplasmid transfer in other proteobacteria....44
1.5. Evolution and mutant selection....45
1.5.1. Selection of industrial organisms....45
1.5.1.1. Induced mutagenesis....45
1.5.1.2. Identification of mutations....46
1.5.1.3. Characterisation of mutations....47
1.6. Genetic tools for *Cupriavidus necator*....47

1.6.1. Transformation....48
1.6.1.1. Classic Allelic exchange....48
1.6.1.2. CRISPR / Cas 9....49
1.6.1.3. Electroporation....49
1.6.4. Expression vectors....49
1.7. Aims of this study....50
2. Materials and Methods....52
2.1. Bacterial strains, plasmids and oligonucleotides....52
2.2. Microbiological methods....62
2.2.1. Preparation of Media....62
2.2.2. Bacterial cultivation....63
2.2.2.1. Growth at atmospheric pressure....63
2.2.2.2. Growth under 2 bar pressure....64
2.2.3. Storage of bacterial cultures....64
2.2.5. Transformation via conjugation....65
2.2.5.1. *C. necator* H16....65
2.2.5.2. *Oligotropha carboxidovorans* OM5....65
2.2.5.3. Transfer of megaplasmid pHCG3....66
2.2.6. Strain evolution for *C. necator* H16....66
2.3. Molecular biology methods....67
2.3.1. Isolation of DNA from bacterial strains....67
2.3.1.1. Isolation of plasmid DNA from *E. coli*....67
2.3.1.2. Isolation of genomic DNA from *C. necator, O. carboxidovorans, A. tume-faciens* and *P.* cws....67
2.3.2. Analysis and manipulation of DNA....67
2.3.2.1. DNA quantification and purity analysis....67
2.3.2.2. Polymerase chain reaction (PCR) for amplification of DNA....68
2.3.2.3. Agarose gel electrophoresis....68
2.3.2.4. DNA purification....68
2.3.2.5. Restriction digest of DNA....69
2.3.2.6. Ligation of DNA....69
2.3.2.7. Other DNA assembly methods....69
2.3.2.8. DNA sequencing....69
2.4. Phenotypic analysis....69

2.4.1. Optical density measurements 69
2.4.2. Measurement of *rfp* reporter genes (for mRFP1) expression 70
2.4.3. Liquid chromatography mass spectrometry (LC-MS) analysis 70
2.6. *In silico* work and tools.....71
2.6.1. Analysis of raw numerical data.....71
2.6.2. Nucleotide and protein sequences.....71
2.6.3. Genome sequencing analysis 72
2.6.3.1. Quality of genome sequences 72
2.6.3.2. Mapping of Reads.....72
6.6.3.3. Single nucleotide variant calling.....73
Chapter 3: Adaptation of *Cupriavidus necator* to carbon monoxide.....75
3.1. Introduction.....75
3.1.1. Aims of this study 76
3.2. Results.....78
3.2.1. Growth profiles with increasing CO concentrations.....78
3.2.2. Continuous subculturing of *C. necator* with CO 79
3.2.2.1. Growth comparison after one month of subculturing 79
3.2.2.2. Growth comparison after 8 months of subculturing 80
3.2.2.3. Whole genome sequencing 81
3.2.3. Repeated subculturing with ten individual *C. necator* culture lineages 84
3.2.3.1 Growth profiles of evolved cultures Evo1 to Evo10 in the presence of 50% CO.....84
3.2.3.2. Growth profiles of defined isolates obtained from cultures Evo1 to Evo10 in the presence of 50% CO / 50% air 86
3.2.3.3. Whole genome sequencing 87
3.2.3.4. Analysis of mutations 88
3.2.4. Replication of the evolved CO resistant phenotype through targeted mutation of strain H16 90
3.2.4.1. Knock-out/knock-in plasmid construction.....92
3.2.4.2. Generation of a knock-out/knock-in mutant library 95
3.2.4.3. Growth comparison of generated single knock-in mutants with wild type *C. necator*.....101
3.2.4.4. Growth comparison of generated double and triple knock-in mutants with wild type *C. necator*.....103

3.2.5. Analysis of cytochrome *bd* ubiquinol oxidase in *C. necator* 106
3.2.5.1. RFP expression analysis of *cydA2B2* intergenic region in *C. necator* 106
3.2.5.2. RFP expression analysis of the *cydA2B2* intergenic region in *E. coli* 108
3.2.5.3. Overexpression of cytochrome *bd* 110
3.2.5.4. Knock-out mutant generation 112
3.2.5.5. Growth comparison of the cytochrome *bd* knock-out with wild-type *C. necator* and SNP knock-in strain *cydA2B2* SNP 112
3.3. Discussion 115
Chapter 4: Conjugative transfer of the megaplasmid pHCG3 121
4.1. Introduction 121
4.1.1. Aims of this study 123
4.2. Results 124
4.2.1. Analysis of *cox* genes in *C. necator* H16 124
4.2.1.1. Identification of putative *C. necator cox* gene clusters 124
4.2.1.2. Sequence alignments 125
4.2.2. Modifying pHCG3 to allow enhanced transfer and efficient detection 128
4.2.2.1. Plasmid construction for overexpression of AHL synthase genes 129
4.2.2.2. Detection of AHLs produced by *E. coli* pBBR1_*traI* and *E. coli* pBBR1_*solI* using a biosensor 129
4.2.2.3. Mass spectrometry analysis of *traI* and *solI* 131
4.2.2.4. Plasmid construction for a *traM* knockout 132
4.2.2.5. Attempted knock-out of *traM* in *O. carboxidovorans* 133
4.2.2.6. Plasmid construction for *traM* truncation 135
4.2.2.7. Truncation of *traM* in *O. carboxidovorans* 136
4.2.2.8. Mass spectrometry analysis of AHLs produced by OM5_*traM* 138
4.2.3. Investigation of pHCG3 transfer into *A. tumefaciens* 140
4.2.3.1. Conjugation with *A. tumefaciens* 140
4.2.3.2. PCR confirmation of plasmid transfer into *A. tumefaciens* and absence of the conjugal *O. carboxidovorans* donor. 141
4.2.3.3. Whole genome sequencing 144
4.2.3.4. Monitoring of autotrophic growth with CO and CO_2 145
4.2.3.5. Mass spectrometry analysis of AHLs produced by A676 147

4.2.4. Investigation of pHCG3_*traM* transfer into a *Pseudomonas* species.....148
4.2.4.1. Conjugation with *P.* cws and other species 149
4.2.4.2. PCR screening.....150
4.2.4.3. Whole genome sequencing 151
4.2.4.4. Phylogeny of *Pseudomonas* cws.....152
4.2.4.5. Monitoring of autotrophic growth 154
4.2.4.6. Mass spectrometry analysis of *P.* cws supernatants 155
4.2.5. Investigation of pHCG3 transfer in *C. necator*.....156
4.2.5.1. Conjugation with *C. necator*.....156
4.2.5.2. Analysis of *repABC* 156
4.2.5.3. Analysis of pMOL28 replication unit.....158
4.2.6. Analysis of pLO3 functionality 159
4.2.7. Characterisation of putative CO transcription factors 159
4.2.7.1. Construction of CO responsive RFP reporters.....160
4.2.7.2. *coxM*-RFP reporter expression in *E. coli* DH5α.....161
4.2.7.3. *coxM*-RFP reporter expression in in *C. necator* H16 162
4.2.7.4. *coxM*-RFP reporter expression in in *O. carboxidovorans* 163
4.3. Discussion.....165
Chapter 5: General discussion 172
5.1. Background.....172
5.2. Key findings and future prospects 173
5.2.1. Genotypic and phenotypic analysis of carbon monoxide resistant *C. necator* 173
5.2.2. Characterisation of selected mutations 174
5.2.3. Characterisation of cytochrome *bd* ubiquinol oxidase expression 174
5.2.4. Elucidation of the *O. carboxidovorans* quorum sensing systems.....175
5.2.5. Analysis of the conjugative capacity of the megaplasmid pHCG3 and expression of encoded functions.....177
5.2.6. Analysis of CO regulation 178
5.3. Concluding statements.....179
Chapter 6: Appendix.....180
6.1. Read coverage distribution for pHCG3_*traM* containing strains.....180
.....181
.....183

.....184
6.2. 16S *rRNA* sequences.....186
6.2.1. 16S *rRNA* and *rpoB* from GenBank.....186
Biblography.....191

List of figures

Figure 1.1. CO_2 emissions by sector for selected regions. 24
Figure 1.2. *C. necator* growth modes. 29
Figure 1.3. Plasmid map of *O. carboxidovorans* OM5 megaplasmid pHCG3. 36
Figure 1.4. Organisation of the complete *cox* gene cluster of *O. carboxidovorans* OM5 38
Figure 1.5. LuxRI-type quorum sensing system as exemplified by the *A. tumefaciens* TraRI system. 42
Figure 3.1. *C. necator* H16 growth with increasing concentrations of CO. 78
Figure 3.2. Growth of *C. necator* Sub and its parent H16 strain in the presence of 50% CO. 80
Figure 3.3. Growth of *C. necator* Evo and its parent H16 strain in the presence of 50% CO/50 air. 81
Figure 3.4. Growth of evolved *C. necator* cultures Evo1 to Evo10 on CO compared to the wild type. 85
Figure 3.5. Growth of evolved *C. necator* isolates E1 to E10 on CO compared to wild type. *C. necator*. 86
Figure 3.6. Promoter prediction for G→A mutation upstream of *cydA2B2*. 88
Figure 3.7. Generation of knock-out / knock-in workflow. 91
Figure 3.8. Generation of knock-out and knock-in plasmids exemplified for *cydA2B2* upstream region. 92
Figure 3.9. Plasmids used for the generation of knock-out/knock-in mutant library (A). 93
Figure 3.10. Plasmids used for the generation of knock-out/knock-in mutant library (B). 94
Figure 3.11. Key steps in the generation of defined deletions and associated PCR screening approaches exemplified for pLO3_*cydA2B2* SNP KO. 96
Figure 3.12. PCR confirmation of *cydA2B2* upstream region knock-out (KO) and knock-in (KI). 98
Figure 3.13. PCR confirmation of *phcR* knock-out (KO) and knock-in (KI). 99
Figure 3.14. PCR confirmation of *phcA* upstream knock-out (KO) and knock-in (KI). 99

Figure 3.15. PCR confirmation of RNA polymerase sigma-54 factor gene 5744 knock-out (KO) and knock-in (KI). 100
Fgure 3.16. Figure 3.16. PCR confirmation of RNA polymerase sigma-54 factor gene 4574 knock-out and knock-in. 100
Figure 3.17. Single knock-in mutant growth compared to wild type in the presence of 50% CO/ 50% air. 101
Figure 3.18. Figure 3.18. Single knock-in mutant growth compared to wild type in the presence of 50% N_2/ 50% air. 102
Figure 3.19. Double and triple knock-in mutant growth compared to the *cydA2B2* SNP strain in the presence of 50% CO/50% air. 104
Figure 3.20. Double and triple knock-in mutant growth compared to the *cydA2B2* SNP strain in the presence of 50% N_2/50% air. 105
Figure 3.21. *cydA2B2* transcriptional reporter plasmid. 107
Figure 3.22. RFP expression of *C. necator* p006_*cyd* US, *C. necator* p006_*cyd* SNP and *C. necator* p006E in 50% CO / 50% air and 50% N_2/50% air. 108
Figure 3.23. RFP expression of *E. coli* p006_*cyd* US, *E. coli* p006_*cyd* SNP and *E. coli* p006E in 50% CO / 50% air and 50% N_2 / 50% air. 109
Figure 3.24. Construction of the cytochrome *bd* ubiquinol oxidase operon overexpression vector pEH006_*cydA2B2*. 110
Figure 3.25. Growth of *C. necator* p006_*cyd* Exp and *C. necator* p006 with arabinose induction in the presence of 50% CO/50% air. 111
Figure 3.26. Growth of the *cydA2B2* knock-out mutant compared to the H16 wild type and *cydA2B2* SNP strains in the presence of 50%CO/ 50% air. 113
Figure 3.27. Growth of the *cydA2B2* knock-out mutant compared to the H16 wild type and *cydA2B2* SNP strains in the presence of 50% N_2/50% air. 114
Figure 4.1. Comparison of cox gene clusters identified in *O. carboxidovorans* OM5 and *C. necator* H16. 124
Figure 4.2. CoxL sequence alignment. 126
Figure 4.3. CoxM sequence alignment. 127
Figure 4.4. CoxS sequence alignment. 127
Figure 4.5. Plasmid maps of AHL synthase overexpression vectors pBBR1_traI (right) and pBBR1_solII (left). 129
Figure 4.6 Map of the *N*-acyl homoserine lactone sensor plasmid pSB403. 130

Figure 4.7. Bioluminescent response of *E. coli* DH5α pSB403 when T-streaked against strains overexpressing putative AHLs synthase genes *traI* and *solI*. 130
Figure 4.8. Relative amounts of AHLs produced by E. *coli* expressing *traI* and *solI*. 131
Figure 4.9. Map of pLO3S_*traM*: a suicide plasmid for the replacement of *traM* with a spectinomycin cassette. 132
Figure 4.10. PCR confirmation of single crossover (A) and failure of second crossover (B) of pLO3S_*traM* in *O. carboxidovorans.* 134
Figure 4.11. Cloning of *traMt* into pLO3. 135
Figure 4.12. PCR confirmation of pLO3_*traMt* single crossover. 136
Figure 4.13. Sequencing of pLO3_*traMt* insertion region. 137
Figure 4.14. Mass spectrometry identification of AHLs produced by OM5_*traM*. ..139
Figure 4.15. PCR confirmation of pHCG3_*traM* presence in A676/A677. 141
Figure 4.16. PCR confirmation of A676/A677 transconjugant chromosomal DNA. 142
Figure 4.17. PCR analysis to confirm absence of contaminating *O. carboxidovorans* DNA in isolated transconjugant A676/A676 chromosomal DNA. 142
Figure 4.18. PCR confirmation of multiple regions of pHCG3_traM in A676/A677. 143
Figure 4.19. Growth of A676 pHCG3_*traM* in comparison to the parental strain A676 with either CO or CO_2 + H_2. 146
Figure 4.20. Mass spectrometry detection of AHLs produced by A676 pHCG3_*traM*. 147
Figure 4.21. Contamination observed on selective plates for initial conjugations between OM5 pHCG3_*traM* and *A. tumefaciens* A676. 148
Figure 4.22. PCR confirmation of pHCG3_*traM* presence in *P.* cws transconjugants. 150
Figure 4.23. PCR amplification of multiple regions of pHCG3_*traM* confirming presence in *P.* cws. 151
Figure 4.24. Neighbour-joining tree based on *rpoD* sequences from different *Pseudomonas* species 153
Figure 4.25. Growth of *Pseudomonas* cws pHCG3_*traM*_A growth in comparison to the parental strain with either CO or CO_2 + H_2. 155
Figure 4.26. The pHCG3 *repABC* operon cloned into pLO3. 157

Figure 4.27. pMTL72111_E1 plasmid and PCR confirmation of its presence in *C. necator*. 158
Figure 4.28. Reporter plasmids to test CO-responsiveness of *coxM expression*. 160
Figure 4.29. Expression of *cox*-rfp *reporter* plasmids in *E. coli* DH5α. 162
Figure 4.30. Expression of *cox*-rfp reporter plasmids in *C. necator* H16.....163
Figure 4.31. Expression of *cox*-rfp reporter plasmids in *O. carboxidovorans* OM5. 164

List of tables

Table 2.1. Bacterial strains used in this study. 52
Table 2.2. Plasmids used in this study. 54
Table 2.3. Oligonucleotides used in this study. 55
Table 3.1. Comparison of mutations across three isolated strains of *C. necator* evolved in the presence of CO. 83
Table 3.2. Growth rates of evolved cultures and evolved single isolates 85
Table 3.3. Comparison of mutations across all ten isolated strains of *C. necator* evolved with CO. 89
Table 4.4. PCR screening guide to mutations generated in *C. necator* 98
Table 4.1. Conjugation frequency of pHCG3_*traM* for *A. tumefaciens* A676 and A677 141
Table 4.2. Illumina sequencing reads of A676/677 pHCG3_*traM* mapped to the wild type *O. carboxidovorans* OM5 and *A. tumefaciens* A676 genomes. 144
Table 4.3. Conjugation frequency of pHCG3_*traM* for *P.* cws and other selected proteobacteria 149
Table 4.4. Illumina sequencing reads of *P.* cws pHCG3_*traM* mapped to wild type *O. carboxidovorans* and *A. tumefaciens* genomes. 152
Table 4.5. Similarities of the 16S *rRNA* and *rpoB* sequences of *Pseudomonas* strains to *P.* cws 153
Table 4.6. Average nucleotide identity and 16S *rRNA* sequence identity of *P.* cws to other *Pseudomonas* species 154

Abbreviations

Mg	Microgram
Ml	Microliter
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
Bp	Base pair
CFU	Colony forming unit
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic acid
g	gravity
GC	Gas chromatography
gDNA	Genomic DNA
InDel	Insertion/deletion
LB	Lysogeny Broth
LHA	Left homology Arm
mg	Milligram
ml	Millilitre
mm	Millimetre
mM	Millimolar
NADH	Nicotinamide adenine dinucleotide
NCBI	National Centre for Biotechnology information
NGS	Next Generation Sequencing
nm	Nanometres
OD	Optical density
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
RHA	Right homology arm
RNA	Ribonucleic acid
RPM	Rotations per minute
SNP	Small Nucleotide Variant
TAE	Tris acetate EDTA

1: Introduction

1.1. Tackling the problem of greenhouse gas pollution

1.1.1. C1 gas

The accelerated release of greenhouse gases into the environment through the burning of fossil fuel has been shown to increase global temperatures (Butler 2018, Marcott *et al.* 2013, Dentener *et al.* 2007, Hansen *et al.* 2000). C1 greenhouse gases, including carbon dioxide (CO_2) and methane (CH_4) are likely to contribute significantly to this problem due to their absorbance of infrared radiation emitted by the earth, with CO_2 accounting for the largest proportion of globally emitted greenhouse gases (>90%), followed by CH_4 (US Environmental protection agency 2016). Global CO_2 emissions as of 2016 were at 32.31 gigatonnes of equivalent carbon dioxide ($GtCO_2$) (Fig. 1.1.) having risen by 40% since 2000 and could reach 40.2 $GtCO_2$ by 2030. The C1 gas carbon monoxide (CO) is not considered a main greenhouse gas but has indirect radiative effects, as it can easily be oxidised to CO_2 and readily reacts with hydroxyl radicals (•OH), preventing the reduction of greenhouse gas (The Intergovernmental Panel on Climate Change 2018).

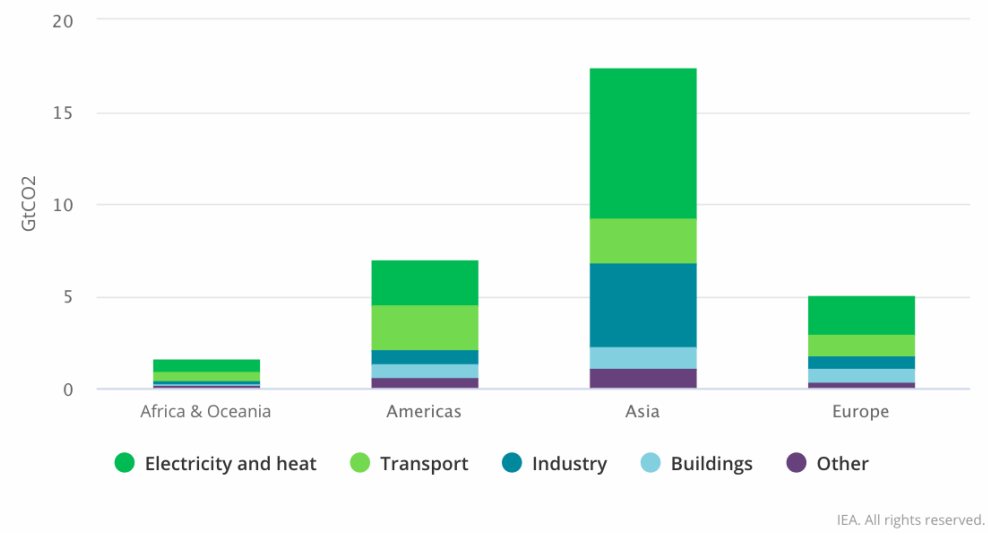


Figure 1.1. CO_2 emissions by sector for selected regions.
Global emissions of CO_2 in 2016 from electricity and heat, transport, industry, building and other sources. Asia producing the most at 17.4 $GtCO_2$, followed by the Americas producing 7 $GtCO_2$, Europe at 5.85 and Africa & Oceania at 1.43. (Figure taken from: IEA, World Energy Outlook 2016).

Global warming has been linked to various detrimental environmental phenomena such as glacial and arctic sea ice retreat, severe drought, increased hurricane intensity and flooding (Fell *et al.* 2017, Karl & Trenberth 2003; Solomon *et al.* 2009). These destructive effects are further exacerbated by the decrease of natural carbon sinks including the ocean, through acidification and the rainforests, through deforestation (Hoegh-Guldberg *et al.* 2017, Geider *et al.* 2001, Hoegh-Guldberg *et al.* 2007, Karl & Trenberth 2003). Rising concerns over these significant environmental issues and the global over reliance on fossil fuel has led to increased interest in the redirection of C1 gas and the use of renewable biofuels (International Energy Agency 2015). The urgent need to tackle global warming was affirmed by the 2015 Paris Agreement where 55 countries pledged to lower GHG emissions to ensure their peak is reached in the shortest possible time (United Nations Framework Convention on Climate Change 2015). To achieve this goal, developing and implementing sustainable strategies to redirect C1 gases from their source is essential.

The application of carbon capture and storage technology is a possible strategy to reduce carbon emissions from coal-burning power plants, although the usage of these technologies on a large scale is highly costly (Blomen *et al.* 2009). As an alternative, C1 gases can be used as a feedstock, to obtain valuable chemicals and fuels and this may form part of a strategy to further reduce greenhouse gas emissions in the future. CO is currently used as a reducing agent to produce methanol, acetic acid and oxo alcohols (Kirk 1998), CH_4 can be used as a precursor for chemicals such as ammonia or methanol (Schnürer & Nordberg 2008, Hess *et al.* 1993) and CO_2 can be used to produce urea, propylene carbonate and salicylic acid (Jajesniak *et al.* 2104). Additionally, C1 gases can be used as a substitute to biomass feedstocks for a microbial host to produce a variety of fuels and platform chemicals. Feedstocks such as first-generation sugars obtained from crops can be fermented to produce biofuels; however, their production potentially puts pressure on food supplies. Second-generation sugars from lignocellulosic material in waste plant matter can alternatively be used but can only be produced from hemicellulose and cellulose through enzymatic break down following physical and/or chemical pre-treatment, with the lignin component being wasted (Gomez *et al*. 2008, Naik *et al*. 2010). C1 feedstocks may represent a cheaper and less wasteful alternative for microbial fermentation with the aid of bacteria able to utilise CO, CO_2 or CH_4 as a carbon source for growth and energy.

1.1.2. Fermentation and product formation

A viable feedstock for CO and CO_2 utilising bacteria is synthesis gas (syngas), which can be produced by gasification, whereby biomass, coal or natural gas are reacted at extremely high temperatures (>800^{o}C) with a low level of oxygen or steam (Halvorsen *et al.* 2015). Crucially, syngas can be produced from renewable sources such as lignocellulose, which is otherwise difficult to process (Rauch *et al.* 2014), but also municipal waste (Zheng *et al.* 2018). Syngas is comprised of a mixture containing hydrogen (25-30%) and inorganic CO (30-60%), CO_2 (5-15%) and CH_4 (0-5%) (National Energy Technology Laboratory 2018), which can be used as a raw material in the Fischer–Tropsch process, whereby the gas mixture is converted into liquid fuel. Another source of CO and CO_2 are industrial waste gases such as steel mill off-gases, which can have a composition similar to that of syngas (Schittkowski *et al.* 2018). The use of such gas mixtures as the feedstock for fermentation has gained recent attention, with new research aiming to improve CO and H_2 solubility to maximise gas conversion efficiency (Liu *et al.* 2019). Use of industrial off-gases for the anaerobic fermentative production of ethanol and other products is currently pioneered at industrial scale by the company LanzaTech (Li *et al.* 2019).

1.1.2.1. Chemolithoautotrophs

Many microorganisms are well suited to grow using syngas as they naturally contain dedicated metabolic pathways that enable them to utilise CO and CO_2 + H_2 or CH_4. Oxidation of these substrates releases chemical energy, which these organisms can capture in the form of ATP generated through either oxidative (electron transport) or substrate-level phosphorylation (Yu & Chen 2019, McKinlay & Harwood 2010, Hanson & Hanson 1996, Meyer & Schlegel 1983, Peck 1968). Bacteria possessing the ability to utilise CO_2 as their major carbon source for growth are referred to as autotrophs (Schlegel 1975). Some autotrophs are incapable of growing in organic media and may be referred to as obligate autotrophs. Those that utilise inorganic substrates to obtain the reducing equivalents needed for CO_2 fixation are widely known as chemolithoautotrophs.

Methanotrophic bacteria are able use CH_4 as their sole carbon and energy source by overcoming the C-H bond through oxygen-dependent enzymes called methane monooxygenases, which form methanol, which subsequently is converted to

formaldehyde (Clomburg *et al.* 2017). This enters the bacterial metabolism through the serine and/or the ribulose monophosphate (RuMP) cycles and can be converted into biopolymers, single cell protein (SCP), vitamins and antibiotics (Dürre & Eikmanns 2015, Fei *et al.* 2014). However, these organisms are not well suited to grow using syngas as a substrate, since CH_4 comprises such a low proportion of the gas. In contrast, the anaerobic, soil and sludge-dwelling acetogenic bacteria can utilise syngas for the production of bioethanol via the Wood-Ljungdahl (WL) pathway, in which CO or CO_2 + H_2 is converted to acetyl-CoA and further to acetate and, sometimes, ethanol (Wilkins & Atiyeh 2011). Syngas has recently been used as a feedstock for fermentation with acetogens such as *Clostridium carboxidivorans* to produce butanol and hexanol (Philips *et al.* 2015), *Clostridium ragsdalei* and *Alkalibaculum bacchi* to enhance ethanol production (Lui *et al.* 2012; Sun *et al.* 2018), *Clostridium ljungdahlii* to produce n-butanol (Sandoval-Espinola *et al.* 2017) and, most notably, *Clostridium autoethanogenum* to produce ethanol (Norma *et al.* 2018). Additionally *Acetobacterium woodii* has been shown to produce high levels of acetate from CO_2 + H_2 (Straub *et al.* 2014). CO_2 can be metabolised by other pathways such as the 3-hydroxypropionate/4-hydroxybutyrate cycle in the hyperthermophilic archaeon *Pyrococcus furiosus,* which has been engineered to produce the platform chemical 3-hydroxypropionic acid (3-HP) from the fixed carbon (Keller *et al.* 2013). Alternatively, CO_2 can be fixed using the widely distributed Calvin-Benson-Bassham (CBB) cycle by organisms such as *Cupriavidus necator* H16 to produce poly [(R)-3-hydroxybutyrate] (PHB) (Ishizaki *et al.* 2001). Recent research has focussed on optimising the use of syngas as a feedstock for *C. necator,* through the implementtation of CO utilisation to increase levels of PHB production (Heinrich *et al.* 2018).

1.2. *Cupriavidus necator* H16

The Gram-negative chemolithoautotrophic bacterium *C. necator* H16 (formerly *Ralstonia eutropha* or *Alcaligenes eutrophus*) belongs to the β-subclass of the *Proteobacteria* and was first isolated in the early 1960s (Schlegel *et al.* 1961). The organism lives ubiquitously in freshwater and soil biotopes and is able to grow concomitantly, or alternatively using inorganic (lithoautotrophically) and organic substrates (heterotrophically) as an energy source (Fig. 1.1) (Pohlmann et al. 2006, Volodina *et al.* 2016).

1.2.1. Genome features

C. necator H16 harbours a genome consisting of three circular replicons: chromosome 1 (4,052,032 bp), chromosome 2 (2,912,490 bp) and megaplasmid pHG1 (452,156 bp), with a total size of 7,416,678 bp (Pohlmann *et al.* 2006). This differs from the closely related and more recently sequenced N-1 strain, which harbours an extra megaplasmid (Poehlein *et al.* 2011). These replicons have been confirmed by pulse-field electrophoresis (Schwartz & Friedrich 2001). Two *cbb* operons (required for autotrophic fixation of CO_2) are present in *C. necator* H16, one of which maps to the megaplasmid pHG1, the other to chromosome 2. Genes for two [NiFe] hydrogenases responsible for oxidising H_2 are located on pHG1. Gene clusters coding for fructose, glucose, 2-ketogluconate and glucosaminate are found on chromosome 2 (Pohlmann *et al.* 2006). Putative replication origins and proteins have been found in the *C. necator* genome with gene distribution analysis revealing that key replication functions are encoded on chromosome 1. GC skew analysis has revealed a *dnaA* ortholog with an adjacent consensus DnaA-binding site on this chromosome. Chromosome 2 contains a putative RepA binding site in the 5'-flanking region of *repA,* with adjacent *parA* and *parB* orthologs, which are additionally present on pHG1 (Pohlmann *et al.* 2006). Since ParAB proteins are responsible for plasmid partitioning and *C. necator* RepA displays similarity to its homologue on the closely related *Cupriavidus metallidurans* megaplasmid pMOL28, chromosome 2 is thought to have megaplasmid ancestry (Fricke *et al.* 2009).

1.2.2. Metabolism

C. necator H16 can only catabolise a limited number of carbohydrate derivatives including fructose, gluconate and *N*-acetylglucosamine. However, a large number of other organic substrates including aromatic compounds, fatty acids, alcohols, tricarboxylic acid cycle intermediates and oleic acid derivatives can be used for growth (Brigham *et al.* 2010, Schwartz *et al.* 2003, Park *et al.* 2011, Johnson *et al.* 1971). Depending on growth conditions, short chain fatty acids and hydroxylated fatty acids such as 3-hydroxypropionic acid (3-HP) can either be incorporated into PHAs or used for growth (Yu & Si 2004; Andreeßen & Steinbüchel 2010, Kukui *et al.* 2009). *C. necator* H16 can oxidise transiently available H_2 through its two energy-conserving hydrogenases (Schwartz & Friedrich 2003), in contrast to the N-1 strain, which lacks

hydrogenase genes (Poehlein *et al.* 2011). It can switch to anaerobic respiration in the absence of O_2 and can also fix CO_2 using the CBB cycle (Fig. 1.2) (Bowien & Kusian 2002). When carbon is abundant and factors such as O_2, bound nitrogen or phosphate are limiting, *C. necator* H16 is able to retain organic carbon in specialised storage granules as PHB and is able to form up to 61 g/liter of PHB in 48 hours (70% or more of total cell weight) (Steinbüchel & Füchtenbusch 1998, Ishizaki *et al.* 2001). PHB and PHAs (related polyhydroxyalkanoates) are of industrial interest as possible alternatives to petroleum-based polymer materials due to their biodegradability and renewability. These properties make *C. necator* H16 a promising candidate for the sustainable production of bio products from industrial waste gas (Heinrich *et al.* 2017, Brigham 2019, Brigham *et al.* 2012).

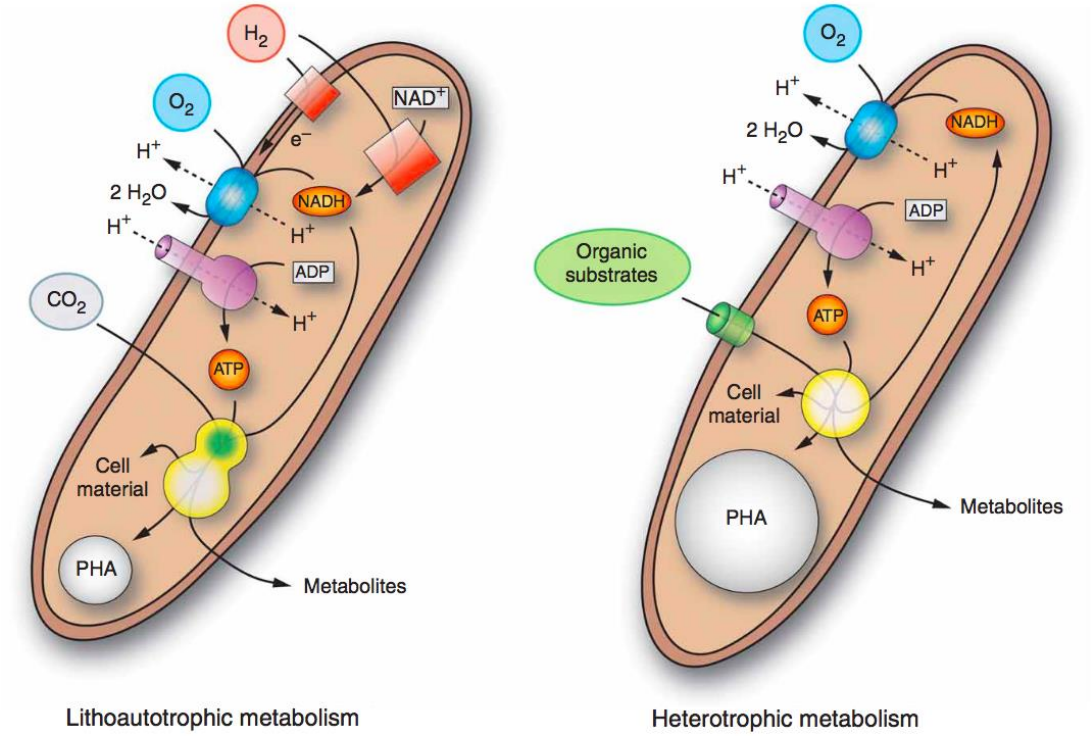


Figure 1.2. *C. necator* growth modes.

Graphic representation showing the main features of lithoautotrophic and heterotrophic metabolism; the latter employs central metabolic processes (yellow circle); in addition to this, the former uses the CBB cycle (yellow / green circles). Poly[(*R*)-3-hydroxybutyrate] (PHA) storage granules shown in grey (Figure taken from Pohlmann *et al.* 2006).

1.2.3. Industrial career and product formation

Initially, *C. necator* H16 was considered as a source of single cell protein in the 1960s due to the rising demand for animal feed (Schlegel & Lafferty 1965). Additionally,

research was carried out to assess the viability of *C. necator* as a life support system on long-term space missions, to utilise CO_2 whilst producing O_2 (Schlegel & Lafferty 1971). However, its potential to produce large amounts of PHB drew focus towards thermoplastic polyester production. Due to the frailty of PHB, the organism was grown under phosphate limiting conditions with glucose and propionic acid, leading to the production of mixed polymers (Byrom 1987). These related polyhydroxyalkanoates (PHAs) displayed improved properties and were used to manufacture bottles and films (Byrom 1987, Chen 2009). However, production of PHA was subsequently performed by *E. coli* strains expressing the PHA biosynthesis genes from *C. necator* H16 due to their faster growth (Slater *et al.* 1988). The array of products synthesised by *C. necator* H16 has expanded in recent years with low molecular mass compounds, alcohols, novel biopolymers of 3-mercaptoalkanoates and cyanophycin being produced (Wübbeler & Steinbüchel 2014, Li *et al.* 2012, Frommeyer *et al.* 2016). Knowledge of the proteome and transcriptome provided by the published genome of *C. necator* H16 has elucidated key metabolic pathways, enabling the targeting of specific genes to enhance its product formation further (Pohlmann *et al.* 2006).

Current studies have focussed on producing platform chemicals through knock out and overexpression of targeted metabolic pathway genes in *C. necator* H16 using CO_2 and H_2 as the sole sources of carbon and reducing equivalents, respectively. *C. necator* H16 has been engineered to increase production of fatty acid-derived, diesel-range methyl ketones, through overexpression of TesA thioesterase, the acyl coenzyme A oxidase gene, *fadB* and *fadM* and deletion of two putative β-oxidation operons (Müller *et al.* 2013). Optimisation of acetoin production has also been achieved in *C. necator* H16, through controlled expression of acetolactate dehydrogenase and acetolactate decarboxylase and deletion of a PHB synthase gene (*phaC1*) (Windhorst & Gescher 2019). Additionally, *C. necator* can be engineered to produce different chemicals electrochemically, such as isobutanol, 3-methyl butanol and C13 or C15 methyl ketones, isopropanol and the terpene α-humulene (Li *et al.* 2012, Muller *et al.* 2013, Torella *et al.* 2015, Krieg *et al.* 2108). Other studies have focussed on using sucrose for PHA production in engineered strains of *C. necator* (Park *et al.* 2015), and a combination of sugars and different precursor substrates for 2-methylcitric acid production (Ewering *et al.* 2006).

Recent studies have explored the potential of *C. necator* H16 to produce 3-HP, a platform chemical with great industrial potential as a precursor for the production of acrylic acid, acrylamide, propandiol and 3-HP polymers (Arenas *et al.* 2019, Hanko *et al.* 2017, Andreeßen & Steinbüchel 2010). *C. necator* H16 naturally metabolises 3-HP; however, three (methyl) malonate semialdehyde dehydrogenase genes (*mmsA1*, *mmsA2* and *mmsA3)* were identified in its genome and deactivated, preventing growth with 3-HP (Arenas *et al.* 2019). Additionally, biosensor tools have been developed for the improved production of itaconic acid in *C. necator* H16 (Hanko *et al.* 2018).

1.2.4. Enhancing the use of waste gas feedstocks for *C. necator* H16

For chemolithoautotrophic growth with waste gases, feedstock optimisation for *C. necator* H16 it is essential to boost the efficiency of growth and boost product formation. The growth rate, cell density and gas uptake rate of *C. necator* H16 can be enhanced during gas fermentation with CO_2 through increased pressure (Yu & Munasinghe 2018).

Syngas is a promising energy rich feedstock, which could be utilised by *C. necator* H16 to produce the bio-products described. However the gas contains high concentrations of CO (30-60%), which the wild type is unable to metabolise (Heinrich *et al.* 2018). *C. necator* H16 can be genetically engineered to oxidise CO to CO_2 by introducing and expressing a functional carbon monoxide dehydrogenase (CODH), which under specific conditions enhanced PHB synthesis (Heinrich *et al.* 2017). There have also been claims that CO oxidation was achieved by anchoring an *E. coli*-produced anaerobic (*A. woodii*) CODH on its cell surface (Hyeon *et al.* 2015). Syngas therefore, has potential as a feedstock for *C. necator* H16 for the production of valuable polymers, chemicals and fuels (Cook & Schlegel 1978, Müller *et al.* 2013, Lu & Yu 2017). However CO significantly inhibits the growth and respiration of *C. necator* H16, which is likely due to its toxicity towards the electron transport chain (Cypionka and Meyer 1982).

1.3. Carbon monoxide

CO participates in a range of processes, both biotic and abiotic including controlling the oxidative capacity of the troposphere (Khalil & Rasmussen 1988, Zhang *et al.* 2001), cell signalling during immunological stress (Verma *et al.* 1993, Zakhary *et al.*

1997, Xue *et al.* 2000), bacterial and eukaryotic haem turnover and microbial biosynthesis and energetics (Meyer & Schlegel 1983, Cypionka *et al.* 1984). CO is highly toxic to many organisms. However, certain microorganisms have evolved to oxidise the gas, some of which are capable of using the gas as a sole carbon and energy source (King & Weber 2007). This is achieved through the utilisation of the *cox* genes able to form CODH, which oxidises CO. CO oxidation is considered a thermodynamically favourable process, which results in CO_2 formation and a pair of reducing equivalents (or molecular hydrogen). This exergonic reaction is given as the equation:

$$CO + H_2O \rightleftharpoons CO_2 + 2H^+ + 2e^-$$

The CO-catalysed reaction is the biological equivalent of the water gas shift reaction ($CO + H_2O \rightleftharpoons CO_2 + H_2$; $\Delta G^{0/}$= 20 kJ/mol (Oelgeschläger & Rother 2008).

CO oxidising bacteria inhabiting soil play an important role in removing CO directly from the atmosphere, consuming ~10% of annual CO emissions (Khalil *et al.* 1999, King 1999). As CO regulates the primary tropospheric chemical oxidant; the hydroxyl radical, these bacteria therefore alter the tropospheric chemistry though this activity (Crutzen & Gidel 1983, Lu & Khalil 1993, Bergamaschi *et al.* 2000). Chemolithotrophs such a *Carboxydothermus hydrogenformans* and species of anoxic archaea were among the first organisms that evolved to utilise CO for growth (Klenk et al. 1997, Sokolova et al. 2004). Anoxic hydrothermal vent ecosystems containing dissolved CO (in concentrations around 100 nM) may have given rise to these bacterial communities (Lilley et al. 1982).

1.3.1. CO-oxidising bacteria

Aerobic bacteria capable of CO oxidation can be categorised into carboxydovores and carboxydotrophs (King & Weber 2007). The former are able to couple CO oxidation with O_2 reduction but are incapable of using CO as a sole source of carbon and energy (King 2003). These organisms span a large range of bacterial and archaeal genera and can oxidise CO at low, atmospheric concentrations but require an organic source of carbon for growth (King 2003, Cunliffe 2013, King 2015, Weber & King 2011). Many carboxydovores lack a complete CBB cycle, rendering them incapable of fixing carbon

from CO (King & Weber 2007). In contrast, carboxydotrophs can grow chemolithoautotrophically with CO as a sole source of carbon and energy and can tolerate the gas at high concentrations. Past studies involving *Oligotropha carboxidovorans* have elucidated the role of form I carbon monoxide dehydrogenase (CODH) in the facilitation of aerobic CO oxidation (Rovaletti *et al.* 2019, Kreß *et al.* 2014, Resch *et al.* 2005, Kraut & Meyer 1988, Meyer *et al.* 2000). Most carboxydotrophs are thought to harbour only form I CODH, with exceptions such as *Bradyrizobium japonicum* USDA 110 which uses a form II CODH (King & Weber 2007) and *Streptomyces thermoautotrophicus,* which is thought to contain form I and form II (MacKellar *et al.* 2016). The large subunit (CoxL) gene for form II CODH is phylogenetically close to form I, however the active site motifs of the two forms differ distinctly (King 2003). Some carboxydovores are thought to use form I CODHs, however others contain form II enzymes, which have been shown to oxidise CO at a slower rate than form I enzymes (Lorite *et al.* 2000).

Until recently, the physiological role of CO oxidation in carboxyovores was thought to be to aid mixotrophic growth (King 2003, King & Weber 2007). However, there is now evidence to support the claim that CO enables carboxydovores to survive carbon limitation as a supplementary source of energy (Islam *et al.* 2019, Greening *et al.* 2015, Ji *et al.* 2017). Several different bacteria all upregulate CODH when carbon is limited and CO has been shown to be rapidly oxidised in ecosystems containing low organic carbon (Patrauchan *et al.* 2012, Berney & Cook 2010, Muthusamy *et al.* 2017, King *et al.* 2008, King & Weber 2008). Genetic, proteomic and biochemical analyses of form I CODH from *Mycobacterium smegmatis* has revealed that the enzyme is strongly induced by organic carbon starvation, enhancing cell survival, but is not used for CO metabolism (Coredo *et al.* 2019). Therefore, carboxydovores may use CO as an additional energy source, yielding NADH and thus electrons to be fed into the respiratory chain.

1.3.2. Tolerance to CO

As well as their inability to grow using CO as a sole source of carbon and energy, carboxydovores are also characterised by their inability to grow at elevated concentrations of CO (King & Weber 2007). In some bacteria, a failure to fix CO through the lack of a CBB pathway may cause this phenotype, however others appear

to not be able to overcome the inhibitory toxic effects posed by raised CO concentrations. Inhibition of growth has been reported when CO concentrations rise above 1% in carboxydovores such as *Burkholderia fungorum* LB400, *Mesorhizobium sp*., *Stappia sp*. and *Stenotrophomonas maltophila sp*. (King 2003). Aside from the CO-oxidisers, many aerobic bacteria such as *E. coli* and *C. necator* show high sensitivity when cultured with CO, with their respiration becoming inhibited, suggesting that components of the electron transport chain are targeted in these organisms (Shen *et al.* 2014, Cypionka & Meyer 1982, Gee & Brown 1981), as will be detailed below. However, growth and respiration of *Pseudomonas denitrificans* and *Pseudomonas stutzeri* were unaffected by high concentrations of CO, suggesting that these bacteria may possess CO-insensitive terminal oxidases.

CO growth inhibition has also been documented in anaerobic carboxydotrophs such *Butyribacterium methylotrophicum,* which is completely inhibited in the presence of 16% CO (Shen *et al.* 1999). However, this organism is able to adapt to higher CO concentrations through successive subculturing. CO growth adaptation over time periods spanning months has been observed in several different carboxydotrophs (Lynd *et al.* 1982, O'Brien *et al.* 1984, Oelgeschläger & Rother 2008, Krajmalnik-Brown *et al.* 2017). In contrast, certain aerobic carboxydotrophs such as *Bradyrhizobium japoncium* USDA 110, *Oligotropha carboxidovorans* and *Rhodocyclus gelatinosus* have all been shown to tolerate high concentrations of CO (Lorite *et al.* 2000, Cypionka & Meyer 1982).

The CO-oxidising capacity of *O. carboxidovorans* has been well documented but does not influence its CO insensitivity (Meyer & Schlegel 1979, Cypionka & Meyer 1982). *O. carboxidovorans* as well as *P. carboxydohydrogena, P. carboxydoflava* (now *Hydrogenophaga pseudoflava*), *P. compransoris* (now *Zavarzinia compransoris*) and *Alcaligenes carboxydus* (now *Carbophilus carboxidus*) are aerobic carboxydotrophs, which contain respiratory chain components including *b*-, *c*- and *a*-type cytochromes (Meyer *et al.*1993, An & Kim 1997, Zavarzin *et al.* 1977, Meyer & Schlegel 1978). These bacteria contain CO-insensitive cytochromes b_{561} and b_{563}, which have previously been studied in *O. carboxidovorans* (Wilcoxen *et al.* 2011, Cypionka & Meyer 1983). When CO is present, b_{563} functions as an alternative terminal oxidase to cytochrome *a*, which is inhibited by CO (Cypionka & Meyer 1986). Due to the low

redox potential of cytochrome b_{563} and its high affinity for O_2, the cytochrome is permanently oxidised and untargeted by CO, which readily binds to reduced cytochromes (Meyer *et al.* 1986). Other bacterial cytochromes appear to be unaffected by CO such as *E. coli* cytochrome *bd-I,* which has also been described as CO-releasing molecule (CORM)-insensitive (Jesse *et al.* 2013).

1.3.3. *Oligotropha carboxidovorans*

O. carboxidovorans is a soil dwelling Gram-negative chemolithoautotrophic α-proteobacterium [formerly *Pseudomonas carboxydovorans*] (Meyer & Schlegel 1978). Like *C. necator,* is it able to grow chemolithoautotrophically with CO_2 and H_2 and heterotrophically in aerobic conditions, however it can also grow using CO as a sole source of carbon and energy (Schlegel 1979).

1.3.3.1 Genome features

The genome of *O. carboxidovorans* OM5 contains a 3,595-kbp chromosome and a low-copy-number megaplasmid pHCG3 (133 kb) (Fig. 1.2) as well as a second 167-kb megaplasmid (pOC167). The latter was originally thought to be part of the main chromosome; however, discovery of the presence of the replicon *repABC,* efflux pumps and components of a type IV secretion system indicated that it was a megaplasmid (Volland et al. 2011). The megaplasmid pHCG3 harbours various gene clusters, which enable the chemolithoautotrophic utilisation of H_2, CO_2 and CO (Fig. 1.3). The function of these genes has been demonstrated through a series of plasmid-cured and deletion mutants in *O. carboxidovorans* (Meyer 1985, Kraut & Meyer 1988). pHCG3 also carries a *tra/trb* gene cluster (25.01 kb) required for conjugal plasmid transfer (Fuhrmann et al. 2003). The chromosome sequence has been shown to encode putative enzymes involved in energy metabolism, biosynthesis of amino acids/fatty acids, nucleotide metabolism and enzymes allowing heterotrophic growth with organic acids, such as pyruvate, acetate, succinate, formate and lactate (Paul *et al.* 2008).

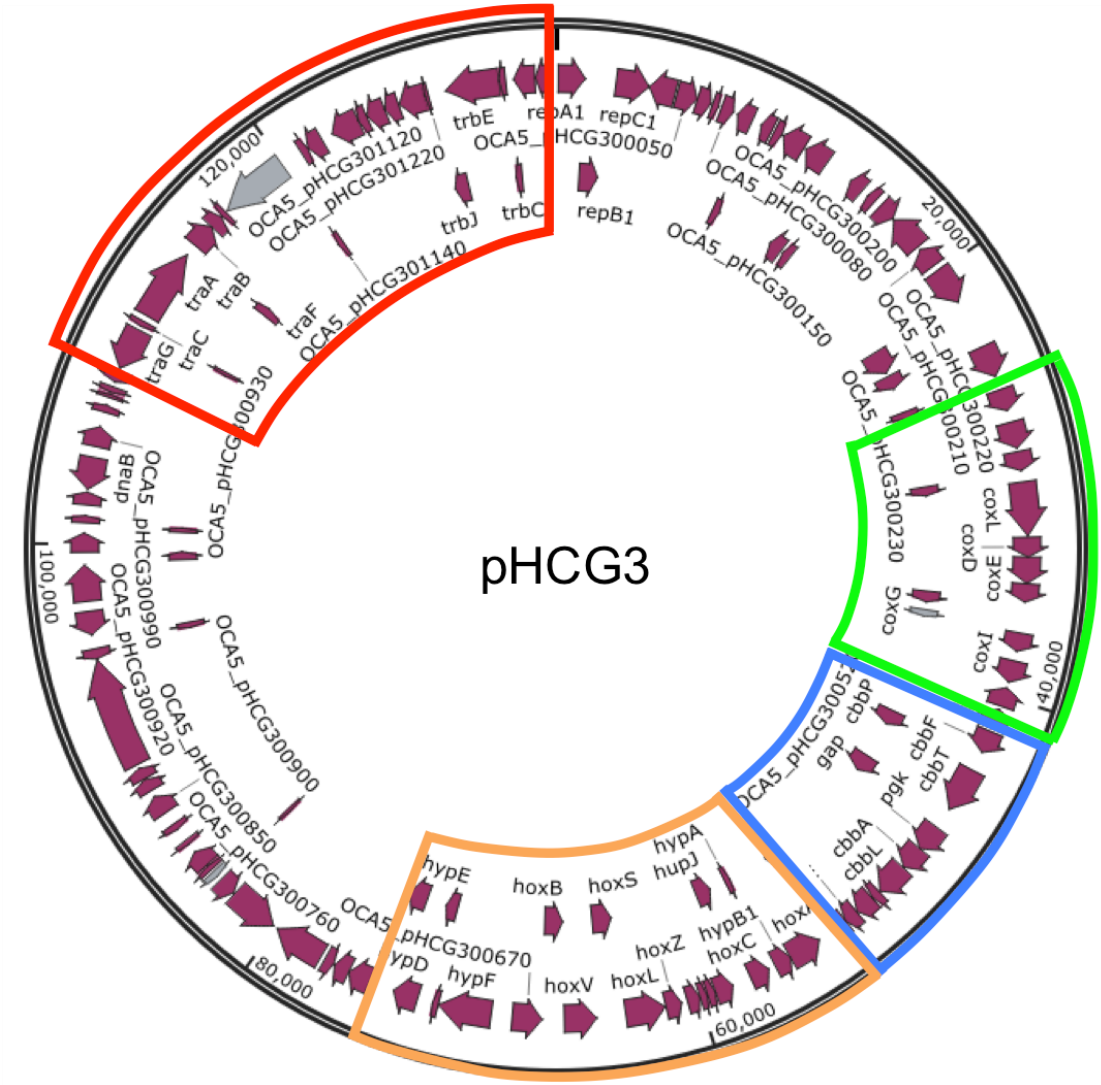


Figure 1.3. Plasmid map of *O. carboxidovorans* OM5 megaplasmid pHCG3.

Labelled features of pHCG3 with purple arrows indicating gene open reading fames. The LuxR family quorum sensing system and *tra/trb* gene cluster highlighted in red, the *cox* genes cluster highlighted in green, the *cbb* gene cluster highlighted in blue and the *hox* gene cluster highlighted in orange. Diagram adapted from GenBank file published by Volland *et al.* 2011.

The pHCG3-encoded *hox* gene cluster (23.55 kb), is thought to be involved with the hydrogenotrophic utilisation and sensing of H_2. The large (HoxL) and small (HoxS) subunit of the [NiFeS]-hydrogenase are encoded in the *hoxSLZ* subcluster (Santiago & Meyer 1997). HoxZ is a cytochrome, which potentially serves as an electron acceptor and membrane anchor of the hydrogenase. The *hyp* genes within the cluster show similarity to the corresponding genes in *C. necator*, which encode a hydrogenase-maturation system (Dernedde *et el.* 1996). Additionally the *hoxB* and *hoxC* encoded

proteins have sequence similarity to their corresponding *C. necator* homologues, which are subunits of a sensing NiFeS-hydrogenase (Kleihues *et al.* 2000). The *cbb* gene cluster (13.33 kb) enables *O. carboxidovorans* to fix CO_2 via the Calvin-Benson-Bassham cycle (Fuhrmann *et al.* 2003). The genes *cbbL* and *cbbS* encode the subunits for the key enzyme for the cycle, ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). The cluster also contains the gene for CbbX, which has been shown to be essential for autotrophic growth in *C. necator* (Bowien & Kusian 2002) and has been reported to bind ATP and activate Rubisco in *Rhodobacter sphaeroides* (Mueller-Cajar 2012). The megaplasmid also contains a *cox* gene cluster (14.54 kb), enabling oxidation of CO (Fuhrmann *et al.* 2003).

1.3.3.2. Cox genes

The *cox* gene cluster comprises 12 genes, *coxBCMSLDEFGHIK* (Fig. 1.4), which function to utilise CO and are transcribed specifically in *O. carboxidovorans* when CO is present (Santiago *et al.* 1999). The key enzyme involved in the carboxydotrophic utilisation of CO in *O. carboxidovorans* is CODH, which has been characterised through X-ray spectroscopy and by crystallography at an atomic level of resolution (Gnida *et al.* 2003, Dobbek *et al.* 2002). The structural proteins of CO dehydrogenase form a heterohexameric enzyme, which includes the medium subunit CoxM (flavoprotein), the small subunit CoxS ([2Fe-2S]-protein) and the large subunit CoxL ([$CuSMoO_2$]-protein) (Schübel *et al.* 1995, Dobbek *et al.* 2002 & Gnida *et al.* 2003). [$CuSMoO_2$] is the bimetallic centre of the active site that must undergo posttranslational synthesis and maturation. This is carried out by the enzymes encoded by *coxDEF* and *coxI* (Meyer et al. 2000, Kaufmann *et al.* 2018). In *O. carboxidovorans*, CoxD has been characterised for its function in the biogenesis of the [$CuSMoO_2$] cluster (Pelzmann *et al.* 2009). CoxD acts as a MoxR-like AAA+ ATPase chaperone, which unfolds CoxL to allow its [MoO_3] centre to be modified through the addition of sulphur and copper. CoxD facilitates this process by interacting with von Willebrand factor A (VWA)-domain containing proteins to form a chaperone system (Pelzmann *et al.* 2009). CoxE harbours a VWA domain and is thought to be associated with CoxD when carrying out metal cofactor insertion in CoxL (Fuhrmann et al. 2003).

CoxC and CoxH are putative transcriptional regulators, which have been proposed to sense CO. They are remarkably similar to each other and carry an MHYT sensor

domain consisting of transmembrane segments as well as a putative LytTR DNA-binding domain. The MHYT sensor domain is characterised by conserved histidine and methionine residues also found in other species (Fuhrmann *et al.* 2003, Santiago *et al.* 1999, Galperin *et al.* 2001 & Nikolskaya & Galperin 2002). *coxC and coxH* both flank the CO dehydrogenase operon (*coxMSLDEFG),* and the encoded proteins are thought to transcriptionally regulate the autotrophic metabolism of CO (Santiago *et al.* 1999). However, in other proteobacteria regulation of CO metabolism relies on RcoM, which has been described in *Burkholderia xenovorans* and is thought to regulate (anaerobic) *cox* and *coo* expression (Kerby *et al.* 2008). CoxC and CoxH remain uncharacterised in *O. carboxidovorans* but have been described as putative transmembrane proteins along with CoxB, CoxK and CoxG (Fuhrmann *et al.* 2003). CoxG has been proposed to anchor CO dehydrogenase in the cytoplasmic membrane.



Figure 1.4. Organisation of the complete *cox* gene cluster of *O. carboxidovorans* OM5
Structural genes (*coxS, M & L)* coloured in red (figure adapted from King & Weber (2007).

CoxM, S, L, D, E, F and G from the *O. carboxidovorans* cluster have been expressed in *C. necator* via an inducible plasmid-borne system, enabling oxidation of CO to CO_2 (Heinrich *et al.* 2018). However, expressing these genes did not provide *C. necator* with adequate means to utilise CO as a sole carbon and energy source. Seemingly, the organism was unable to channel electrons produced from CO oxidation into the respiratory chain, so additional H_2 was required for chemolithoautotrophic growth. These electrons were thought to be accepted by cytochrome b_{561} (Cypionka & Meyer 1983), although more recent work suggests that initially the *O. carboxidovorans* quinone pool accepts the CODH-derived reducing equivalents (Wilcoxen *et al.* 2011). Cytochrome b_{561} encoded by *hoxZ* on pHCG3 may therefore be required for CO to be used as a sole source of carbon and energy. The transfer of megaplasmid pHCG3 into *C. necator* via conjugation with *O. carboxidovorans* may solve this issue. Transfer of pHCG3 is potentially possible due to the presence of a conjugal transfer system encoded by the *tra/trb* genes and will be explored at length in Chapter 4.

1.4. Megaplasmids

Genome segmentation is common to a large number of bacteria, which maintain low-copy replicons, classified as megaplasmids in addition to their main chromosome. Megaplasmids can be up to 1.7 Mb in size, many of which have impart beneficial phenotypes to their host such as antimicrobial resistance, virulence or the ability to use a novel catabolic pathway (Finan *et al.* 2001, diCenzo & Finan 2017).

1.4.1. RepABC-family plasmids

In addition to *O. carboxidovorans,* numerous bacteria of the order *Rhizobiale*s contain megaplasmids with a *repABC-type* replication system (Pinto *et al.* 2012). The system enables acquisition, maintenance and vertical transfer of large amounts of genetic information, conferring a wide variety of functions to its host.

1.4.1.1. Replication

The *repABC* operon is responsible for the replication and segregation of low copy-number plasmids and is widely believed to exist exclusively in α-proteobacteria (Pinto *et al.* 2012). Most of these plasmids employ a Type I segregation system with Walker-type ATPase (RepA, ParA superfamily) and *parS* binding protein (RepB, ParB superfamily), *parS* representing a centromere-like site (Cevallos *et al.* 2008). These proteins play a key role in plasmid partitioning whereby ParB acts as a mediator binding to ParA and *parS,* whilst ParA segregate the plasmids via filament formation (Chai & Winans 2005, Ebersbach & Gerdes 2005). Induced mutagenesis of *repA* or *repB* has been shown to substantially decrease the stability of *repABC* plasmids (Tabata *et al.* 1989, Bartosik *et al.* 1998, Ramirez-Romero *et al.* 2000). Recent findings have revealed that plasmid segregation is aided through the specific binding of RepB to centromere-like *parS* elements in *Rhizobium leguminosarum bv. trifolii* TA1 (Koper *et al.* 2016). In the same organism, RepA has been shown to specifically bind to the *repABC* operator sequence to autorepress the *repABC* operon (Żebracki *et al.* 2015). RepA autorepression has also been documented in *A. tumefaciens* and *Rhizobium etli* (Pappas & Winans 2003, Ramirez-Romero *et al.* 2001).

RepC is essential for plasmid replication and is solely capable of performing autonomous replication in *A. tumefaciens* (Chai & Winans 2005). Additionally, it has been proposed as a replication initiator protein due to its capacity as a limiting

replication factor (Pinto *et al.* 2011, Ramirez-Romero *et al.* 2000, Tabata *et al.* 1989, Bartosik *et al.* 2001). In *A. tumefaciens* and *Rhizobium etli*, the origin of replication is thought to lie within *repC* and its expression is controlled by a small antisense RNA encoded in the region between *repB* and *repC* (Ramirez-Romero *et al.* 2000, Cervanes-Rivera *et al.* 2011, Venkova-Canova *et al.* 2004). In *O. carboxidovorans* the origin of replication is thought to lie in the *repB-repC* intergenic region (Fuhrmann *et al.* 2003).

1.4.1.2. Evolution

RepA and RepB coevolve with one another, exclusive of RepC, presumably to generate different incompatibility properties. Therefore, some plasmids harboring *repABC* can be stably maintained after transfer in a range of different strains due to a lack of plasmid competition (Wetzel *et al.* 2015). The presence of multiple *repABC* replicons in certain bacteria is thought to be a result of this process, with one example, *Rhizobium etli* CFN 42, maintaining up to six different replicons (González *et al.* 2006). On the other hand, RepA, RepB, RepC, and *cis* acting sites upstream (*inc*α) and downstream of *repC* (*inc*β / *parS*) have all been shown to be involved with plasmid incompatibility, causing plasmid competition in different α-proteobacteria (Ramírez-Romero *et al.* 2000, Cervantes-Rivera *et al.* 2011, Pérez-Oseguera & Cevallos 2013). Recently, a series of modular vectors was generated that harbor 11 different *repABC*-based regions, three of which were shown to be compatible with one another and were stably maintained in *Sinorhizobium meliloti*, whereas others were shown to be incompatible (Döhlemann *et al.* 2017). Despite their proximity and sharing of promoters and *tra* boxes allowing TraR-mediated transcription (as described below), RepA, RepB and RepC have been shown to evolve independently from the quorum-sensing (QS) regulatory systems that are responsible for plasmid transfer (Wetzel *et al.* 2015).

1.4.1.3. Conjugal transfer systems

repABC plasmids contain transfer systems, which can be split into at least four classes (Giusti *et al.* 2012). *O. carboxidovorans* houses a Class I transfer system on pHCG3 encoded by the *tra* and *trb* operons, which is used by the majority of *repABC* plasmids (Chen *et al.* 2002, Wetzel *et al.* 2015). These operons contain all the genes essential for conjugal transfer, as first exemplified for the Ti plasmid of *Agrobacterium tumefaciens* (Li *et al.* 1999, Cook *et al.* 1998). The Class I system consists of a chimeric IncQ- and IncP- region, which comprises the DNA transfer and replication (Dtr) function and the

relaxosome complex (*oriT*) encoded by the *traAFBH* and *traCDG* operons (Cook & Farrand 1992, Farrand *et al.* 1996). Three other classes of transfer system are found in *repABC* plasmids with the Class II system also being well characterised (Giusti *et al.* 2012). This system is composed of an IncQ-like *oriT* region and a Dtr system similar to that of the Class I system, however the *traF*, *traB,* and *traH* genes are absent (Chen *et al.* 2002). The mating pair formation (Mpf) apparatus, required for the translocation of *repABC* plasmids from donors to recipients is encoded by the *trb* operon consisting of the genes *traI* and *trbB, -C, -D, -E, -J, -K, -L, -F, -G* and *–I* (Li *et al.* 1998). The transcription of the Class I transfer genes is regulated by LuxRI-family quorum sensing systems.

1.4.1.4. Quorum sensing

Cell-to-cell communication via the secretion of small diffusible signal molecules (often termed autoinducers) into the local environment is known as quorum sensing (QS) (Fig. 1.5). Bacteria are thought to use such systems to regulate population-wide responses in a cell density-dependent manner through coordinated gene expression. For Gram-negative bacteria, QS was first discovered in *Vibrio fischeri* (now *Photobacterium fischeri*), a marine symbiont of squid which develops a bioluminescent phenotype when high signal molecule concentrations and, thus, cell densities are reached in the host's light organ, but has since been shown to control expression of virulence factors, biofilm formation and many other traits in various other bacteria (Parsek & Greenberg 2000, Chhabra *et al.* 2005).

The LuxRI-family QS system has been well characterized in *Agrobacterium tumefaciens,* and consists of the transcriptional activator TraR and a population-dependent QS signal produced by TraI; the autoinducer *N*-acyl-homoserine lactone (AHL) (Piper *et al.* 1993, Zhang *et al.* 1993). This system enables conjugative transfer of the crown gall tumour-inducing (Ti) megaplasmid in *A. tumefaciens* (Hwang *et al.* 1995, Piper & Farrand 2000). To initiate this process, TraR binds the AHL, produced by the AHL synthase TraI. This dimerises and stabilises this transcriptional activator, which initiates transcription of the *trb* operon (Zhu & Winans 1999, Chai *et al.* 2001,

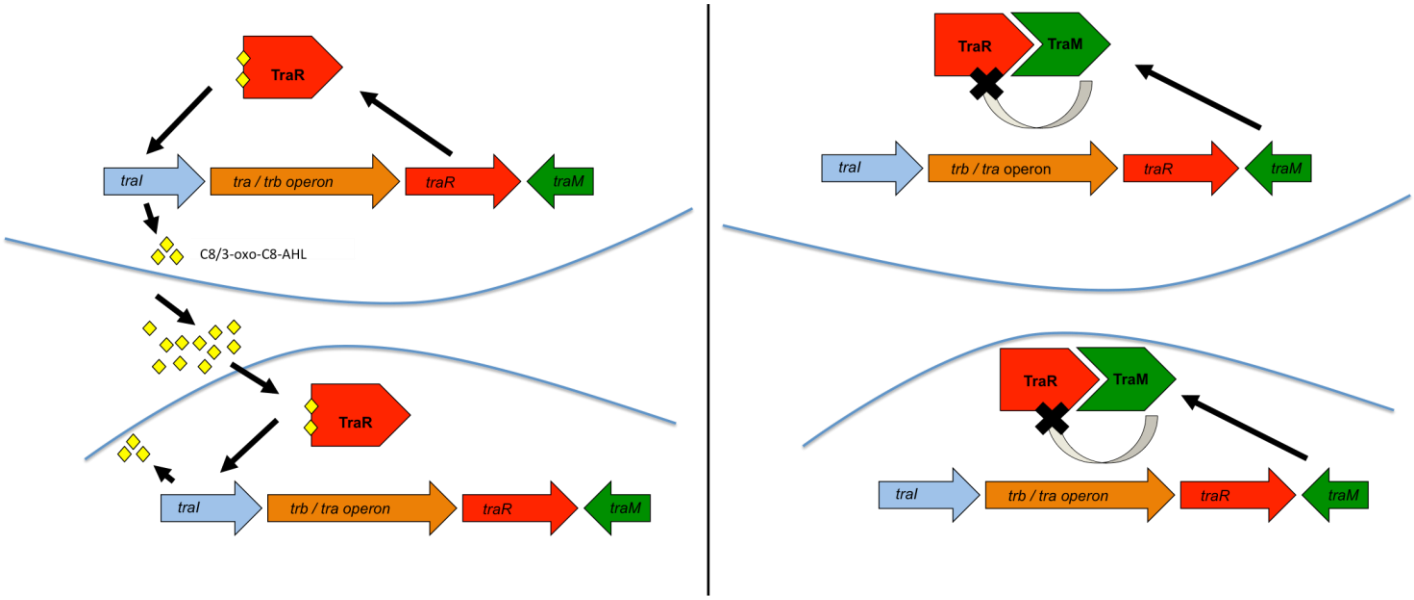


Figure 1.5. LuxRI-type quorum sensing system as exemplified by the *A. tumefaciens* TraRI system.

Overview of the TraRI-quorum sensing system used by *A. tumefaciens* and, presumably, *O. carboxidovorans.* TraR (red) and TraM (green) proteins control *traI* expression and hence activation of *trb* genes. AHLs (yellow) produced by TraI diffuse across the cell membrane (blue line) from one cell into the surrounding extracellular matrix and from there across the cell membrane of another cell when a sufficiently high concentration is reached (left). When TraM is produced, it binds to TraR and thus prevents transcription of *traI* and the *trb/tra* operon (right).

Qin *et al*. 2000, Fuqua & Winans 1996, He *et al.* 2003). The *A. tumefaciens* TraRI system is one of a few that are also controlled by the antiactivator TraM. This protein prevents induction of the signaling process by binding TraR, thus inhibiting transcription of *tra* and *trb* operons (Fuqua & Winans 1994, Fuqua *et al.* 1995, Hwang *et al.* 1995, Danino *et al.* 2003, Luo *et al.* 2000). When *traR* expression is induced, the increase in TraR protein overcomes the suppression of TraM, permitting the QS response (Hwang *et al.* 1995, Su *et al.* 2008).

The organisation of the *tra*, *trb*, *rep* and QS genes, which are conserved in the Class I group (Moriguchi et al. 2001), can be separated into two categories; Group I and Group II (Wetzel *et al.* 2015). The *traI/trb* operon is adjacent and divergently oriented to *repABC* in both groups. In Group I plasmids such as the Ti plasmid, the *tra* operons, *traR*, *traM* and the cis-acting *oriT* are separated from the *traI*/*trb* genes by at least 60 kb. By contrast in Group II plasmids such as pHCG3, the *tra* locus neighbours the *trb* locus and *traM* and *traR* are found in-between these regions (Wetzel et al. 2015).

Phylogenetic analysis of TraI, TraR and TraM from various α-proteobacteria reveals coevolution whereby the protein set from Group I plasmids forms Clade I and the protein set from most Group II plasmids form Clade II (Wetzel *et al.* 2015). Within Group II, the proteins from *O. carboxidovorans* and *Nitrobacter hamburgensis* form a distinct third clade. The *tra/trb* cluster of *O. carboxidovorans* begins at the plasmid replication origin oriV, and its products are believed to mobilise pHCG3 through conjugal transfer (Fuhrmann et al. 2003). The cluster exhibits a similar arrangement and high identity to the corresponding genes of the Ti plasmid of *A. tumefaciens* and shares many of the same functions. When cultured in the presence of syngas, *O. carboxidovorans* upregulates the genes for TrbL, TrbI and the autoinducer synthase TraI (Paul *et al.* 2010). This suggests that pHCG3 is being actively transferred via conjugation when transcribing the *cox*, *cbb* and *hox* genes for utilization of CO and CO + H_2. Conjugation of pHCG3 has not previously been demonstrated; however, megaplasmid transfer has been documented for several proteobacteria as a form of horizontal gene transfer (HGT) (Danino *et al.* 2003, Blanca-Ordóñez 2010, Pérez-Mendoza 2005).

1.4.2. Horizontal gene transfer

HGT enables the movement of genetic information between organisms and is a key evolutionary mechanism within bacterial communities (Syvanen 2012, Keeling 2009). HGT between bacteria can provide recipients of genetic information with a phenotypic benefit but can also be an inefficient process (Diaz-Ricci & Hernández 2000, Park & Zhang 2012).

1.4.2.1. Megaplasmid transfer in α-proteobacteria

HGT via the conjugal transfer of megaplasmids has been documented in several α-proteobacteria, in experiments that demonstrate the acquisition of a particular encoded phenotype. The native pM2 megaplasmid of *Rhizobium meliloti*, which enables nodulation and extracellular polysaccharide production for host infection, was equipped with selectable antibiotic resistance genes (Pretorius-Güth *et al.* 1990). Cellulose acetate filter paper mating enabled the marked megaplasmid to be conjugated into mutant *R. meliloti* strains restoring the nodulation and infection phenotypes. More recently, the conjugal transfer of the megaplasmids from *Sinorhizobium meliloti* has

been demonstrated, whereby RctA regulates pSymA and pSymB transfer relies on pSymA functionality (Blanca-Ordóñez 2010, Pérez-Mendoza *et al.* 2005).

Interspecies megaplasmid transfer has also been demonstrated between different α-proteobacteria demonstrating the host range of *repABC* plasmids. Conjugation of pRL1JI from *Rhizobium leguminosarum* bv. *viciae* into a plasmidless strain of *A. tumefaciens* has previously been demonstrated (Danino *et al.* 2003). Additionally, the pSym1 megaplasmid from *Rhizobium leguminosarum* strain 1001, which contains genes for root nodulation, can be conjugated into *A. tumefaciens* and different species of *Rhizobium* which enables induction of root nodules in pea plants (Hooykaas *et al.*, 1981, Hooykaas *et al.*, 1982). Megaplasmids which are not naturally mobilised have been modified to allow conjugal transfer, such as the pRme41*b* megaplasmid from *Rhizobium meliloti* 41, which was equipped with the mobilization mob region of RP4 and conjugated into two other species of *Rhizobium* and a plasmidless *A. tumefaciens* strain, enabling nodulation (Kondorosi *et al.* 1982). A similarly mobilised version of the pSym megaplasmid from *R. meliloti* 2011, which carries nodulation genes, has been conjugated into *A. tumefaciens* to induce a root nodulation in a variety of host plants (Truchet *et al.* 1984).

1.4.2.2. Megaplasmid transfer in other proteobacteria

Megaplasmid transfer has also been demonstrated in other proteobacteria, which do not harbour *repABC* plasmids such as the β-proteobacterium *Cupriavidus metallidurans* CH34 (Diels *et al.* 1989, Mergeay *et al.* 1985). The native megaplasmids pMOL28 and pMOL30, which confer resistance to various heavy metals, can be conjugated into plasmid-free recipients and selected for their resistant phenotype (Diels *et al.* 1989). Conjugation of these plasmids can be enhanced by the presence of IncP1 plasmids such as RP4 or pUZ8, which function as helper plasmids to stimulate the mobilisation functions of the native plasmids (Mergeay *et al.* 1985). Additionally, interspecies transfer of the megaplasmid pMOL85 from the closely related *Cupriavidus campinensis* DS185 into *C. metallidurans* has been shown (Diels *et al.* 1989). Interspecies megaplasmid transfer has also been reported in γ-proteobacteria, whereby pMPPla107 from *Pseudomonas syringae* pv. *lachrymans* 107 was conjugated into several different *Pseudomonas* strains including *Pseudomonas stutzeri* DBL332, which induced a range of phenotypic changes including decreased antibiotic resistance and

biofilm formation but increased motility (Romanchuk *et al.* 2014, Dougherty *et al.* 2014).

1.5. Evolution and mutant selection

Different environmental changes cause bacteria to continuously evolve through the generation of heritable mutations, which favour survival of the best-adapted organism (Lenski 2017). This central dogma, first proposed by Darwin, can be exploited under controlled conditions with a limited number of variables, known as adaptive laboratory evolution (ALE).

1.5.1. Selection of industrial organisms

1.5.1.1. Induced mutagenesis

ALE of organisms is an effective method to generate improved strains for the production of biofuels and chemicals (Portnoy *et al.* 2012, Debabov 2015). ALE involves propagation of a microbial strain over an extended time period to allow for numerous rounds of generation, with a specific selective pressure driving the process (Winkler *et al.* 2013). Genetic changes for this process can be generated randomly by a chemical mutagen or error prone PCR (Landt *et al.* 1990, Füzik *et al.* 2014). Such an approach can be more effective by performing long cultivations with a continuous culture to allow for a greater number of generations (Debabov 2015). However, such experiments can also be performed in batch culture via continuous subculturing (Lee *et al.* 2016). For instance, continuous cultivation has been used to with *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine to generate greater butanol yields in *Clostridium pasteurianum* (Malaviya *et al.* 2012). Various ALE experiments have been successfully performed on *Saccharomyces cerevisiae* to improve the efficiency of xylose and arabinose assimilation (Sonderegger *et al.* 2003, Kuyper *et al.* 2005, Shen *et al.* 2012, Garcia *et al.* 2010).

ALE can also be used to generate production strains that are more resistant to inhibitory compounds they are cultured with. Mutant strains of *C. pasteurianum* showing high tolerance to crude glycerol can be generated through random mutagenesis mediated by ethane methyl sulfonate (Jensen *et al.* 2012). Such strains exhibited greater 1,3-PDO and butanol production as a result. Similarly, adaptive evolution can be applied to *E. coli* strains to increase resistance to ethanol or isobutanol (Cao *et al.* 2017, Smith *et al.*

2011, Atsumi *et al.* 2010). Obtaining strains though ALE which are highly tolerant to CO is essential for syngas fermentation. For instance, the tolerance of *Butyribacterium methylotrophicum* can be dramatically improved through successive transfer in media containing increasing concentrations of CO (Shen *et al.* 1999).

1.5.1.2. Identification of mutations

Following ALE, genomic analysis of selected variant strains can be performed to target specific genes that have been mutated, which may lead to the desired phenotype. Next generation sequencing techniques such as Illumina, PacBio or Oxford Nanopore can be employed to identify single nucleotide polymorphisms, insertions, deletions or mismatches that may occur at mutational hotspots across the whole genome of an organism. These platforms sequence millions of DNA fragments in parallel, which can be assembled into longer contig files available for analysis (Behjati & Tarpet 2013). Recent studies have implemented whole genome sequencing in combination with ALE using *E. coli* to elucidate genes and pathways involved with metabolism of different substrates such as glucose, xylose, glycerol, sucrose or acetate (Sandberg *et al.* 2017, Long *et al.* 2017). However, strains generated using ALE often accumulate multiple mutations, therefore genetic determination of the organism's phenotype is not straightforward. However, ALE can be performed with multiple independent experiments, examining the most frequently occurring mutations to help filter causal mutation candidates (La Croix *et al.* 2015).

Other methods can also be used in combination with ALE and whole genome sequencing to accurately locate mutations. For example, the hyperthermophilic archaeon *Thermococcus onnurineus* NA1 was adapted to enhance H_2 productivity through repeated serial transfers in serum bottles containing CO (Lee *et al.* 2016). Following adaptation, enhanced strains were analysed simultaneously by Illumina Hi-seq2000 and SMRT DNA sequencing as well as RNA-seq to identify unique mutations upregulated at a particular serial transfer. Alternatively, genome-scale analysis can be performed by microarray to assess differences in the transcriptome, an approach that was used to identify regulatory and metabolic mechanisms, which influence ethanol tolerance in *E. coli* (Goodarzi *et al.* 2010, Horinouchi *et al.* 2010). These approaches can identify sets of genes, which are involved in generating the desired phenotype;

however, they do not specify the exact changes individual mutations can have on an organism.

1.5.1.3. Characterisation of mutations

Following ALE and identification of target genes, reverse engineering of specific mutations can provide more accurate information explaining how genetic changes affect the phenotype of an organism (Mundhada *et al.* 2017, Mohamed *et al.* 2019). This method has been successfully used to elucidate individual genes involved with enhanced sucrose utilisation in *E. coli*, whereby specific mutations were "reverse engineered" individually and in parallel into the starting strain (Mohamed *et al.* 2019). The growth rates of these adapted strains were measured following reverse engineering to assess their fitness in comparison the starting strain. Mutations in *E. coli* can be reverse engineered using counter selection with an antibiotic marker and *sacB* (Mundhada *et al.* 2017) or using the more modern CRISPR/Cas 9 technique (Jiang *et al.* 2015).

ALE will be used in this study to generate *C. necator* strains that are adapted to aerobic growth in the presence of CO, since this gas significantly inhibits its growth and respiration (Cypionka & Meyer 1982). Genetic changes that have occurred to the evolved strains can be identified by Illumina whole genome sequencing and can then be engineered into the original wild type strain to observe if the same phenotypic changes occur. This type of genome editing may involve the altering of single base pairs in the *C. necator* genome, which will require specific tools to be designed *in-silico.*

1.6. Genetic tools for *Cupriavidus necator*

Extensive research has been carried out on *C. necator* since its discovery more than 50 years ago, during which a variety of tools have been developed for the genetic manipulation of the organism. The viability of *C. necator* as a production strain has increased since the publication of its annotated genome, allowing identification of specific genes and operons for omics studies (Raberg *et al.* 2018). Tools for transformation, chromosomal modification and an array of plasmids are available for *C. necator* and other, more sophisticated tools are still under development (Nybo *et al.* 2015, Gruber *et al.* 2015).

1.6.1. Transformation

Standard transformation techniques can be used for *C. necator* such as electroporation or the transfer of plasmids via conjugation. Initial mating experiments using the donor *E. coli* S17-1 strain proved biparental mating was possible using a broad host range vector with an origin of transfer (*oriT*) and RK2 mobilisation machinery (Friedrich *et al.* 1981, Ditta *et al.* 1985). *C. necator* also naturally possess resistance to gentamicin, enabling improved selection for transformants following conjugation.

1.6.1.1. Classic Allelic exchange

For deletion and integration of target genes and DNA fragments in the genome of *C. enactor, sacB* counter selection can be used (Quandt *et al.* 1993). The suicide vector pLO3, derived from the plasmid pBR322 is a suitable vector for this process, carrying the *B. subtilis sacB* gene obtained from pUCD 800 (Lenz & Friedrich 1998, Gay *et al.* 1985). Additionally pLO3 carries the RP4 origin of transfer and a tetracycline resistance cassette, enabling selection following integration of plasmid into the host genome. Knock-out / knock-in vectors can be designed with homology arms flanking the site of interest, thus enabling single crossover of pLO3 via homologous recombination upon transfer into *C necator* via a conjugal donor strain of *E. coli* such as S17-1 (Lenz & Friedrich 1998). The origin of replication from pBR322 is non-functional in *C. necator;* therefore, upon transfer of pLO3, the plasmid will either be lost or integrate into the chromosome via allelic exchange if homology arms are present. The exact mechanisms of pLO3-mediated gene integration and knockout are described in detail in a different section (see 3.2.4.2).

The pLO3 suicide vector has been used previously to generate multiple genetic modifications in *C. necator*, including successive in-frame deletions to generate a strain incapable of metabolising 3-HP (Arenas *et al.* 2019), knockouts of the HPr kinase/phosphorylase gene *hprK* (Kraußе *et al.* 2009) and PHB depolymerase gene *phaZa1* (Juengert *et al.* 2018, Handrich *et al.* 2000). The plasmid was used extensively in this study to create knock-out / knock-in vectors for the creation of a mutant library in *C. necator*.

1.6.1.2. CRISPR / Cas 9

Genome editing can also be performed using an electroporation-based type II CRISPR / Cas9 technique (Xiong *et al.* 2018). The plasmids containing this system are based on pBBR1MCS-2 (Kovach *et al.* 1995), which optimises *cas9* expression via a pBAD promoter. The system operates by locating unique sequences then generating targeted double strand breaks, which are repaired through homologous recombination of homology arms cloned into the vector.

1.6.1.3. Electroporation

Methods for transforming *C. necator* via electroporation have previously been developed (Aneja *et al.* 2009). As opposed to transformation via conjugation with *E. coli*, this method is much faster, giving recombinant colonies in a shorter time scale. In addition, plasmids for improved electroporation efficiency in *C. necator* have been developed (Solaiman *et al.* 2010). Recently, a considerably more efficient electroporation protocol has been developed for *C. necator* (Tee *et al.* 2017). This approach optimised different variables such as cell growth, cell number, transformation buffer, $CaCl_2$ incubation and electroporation voltage and gave a transformation efficiency 10^3 times higher than previously documented.

1.6.4. Expression vectors

The broad host range plasmid pBBR1, which replicates in a range of bacteria was used to construct the expression vector pBBR1 MCS (Kovach *et al.* 1994, Kovach *et al.* 1995). For the overexpression of target genes, derivatives of this vector have been used in *C. necator* (Voss & Steinbüchel 2006), with the mobilisation efficiency and stability of pBBR1-derived vectors improved through the introduction of RP4 and RSF1010 based mobilisation sequences (Gruber *et al.* 2014). Recently, *cox* genes from *O. carboxidovorans* were expressed via an inducible promoter on a pBBR1 MCS derivate enabling the utilisation of CO in *C. necator* (Heinrich *et al.* 2018). Additionally, derivatives of pBBR1MCS have been used to express an *efe* gene enabling production of bioethylene (Ding *et al.* 2019), the *ispS* gene to produce isoprene (Alagesan *et al.* 2018) and the *glpK* gene to improve the growth of *C. necator* on glycerol (Aboulnaga *et al.* 2018). A derivative of this vector pBBR1MCS-2-PphaC-eyfp-c1 (Pfeiffer & Jendrossek 2011) was used in this study to overexpress AHL synthase genes from pHCG3 in *E. coli* to elucidate the type of signal molecules produced (see 4.2.1.1).

Previously, the activity of the inducible P_{BAD}, P_{tet}, $P_{xyls/PM}$, P_{lacUV5} promoters has been monitored to optimise gene expression (Bi *et al.* 2013), however the range of available, inducible promoters and ribosome binding sites for *C. necator* has now been expanded (Alagesan *et al.* 2018). Recently, 42 constitutive promoters have been shown to function in the organism, such as the IPTG-inducible P_{j5} promoter, which displays greater strength than previously used inducible promoters (Johnson *et al.* 2018, Gruber *et al.* 2016). Components of an arabinose inducible system have been cloned into the vector pEH006, which contains a red fluorescent protein (RFP) gene under control of the P_{BAD} promoter. This system enables the detection of a fluorescent signal if gene expression is induced. Derivatives of this plasmid have been used to elucidate native production of 3-HP and itaconic acid in *C. necator* (Hanko *et al.* 2017, Hanko *et al.* 2018). Derivatives of this vector were also used in this study to characterise putative transcription factors responsive to CO from megaplasmid pHCG3 to observe whether they are functional in *C. necator*.

1.7. Aims of this study

The primary objectives of this study were to develop a strain of *C. necator*, which exhibited increased tolerance to CO through adaptive evolution and to enable CO oxidation through the conjugative transfer of pHCG3. At present, *C. necator* H16 is significantly inhibited by high concentrations of CO, which are present in feedstocks such as syngas, although a closely related strain (*R. eutropha* B5786) has been claimed to display CO resistance at lower concentrations (Volova *et al.* 2002). Therefore, increased resistance to the gas is a crucial step if syngas or CO is to be used as a substrate for *C. necator* for the production of valuable platform chemicals. Whether *C. necator* can adapt to high concentrations of CO remains unclear and the genetic basis for CO sensitivity is still not fully understood. Recent advances published during the middle of this project have enabled the utilisation of CO in *C. necator* via *cox* gene expression. However, this gave a low efficiency of CO oxidation and required the addition of reducing equivalents such as hydrogen (Heinrich *et al.* 2018), suggesting absence of crucial components that allow utilisation of CO-derived reducing equivalents. For efficient utilisation of CO as a sole source of carbon and energy, conjugation of pHCG3, which harbours complete CO, CO_2 and H_2 metabolism pathways, may offer a possible solution and help to identify any missing or limiting

factors. It is, however, unclear whether initiation of *cox* gene expression from native transcription factor binding sites on pHCG3 is possible in *C. necator*, since oxidation of CO was not achieved following integration of the entire *O. carboxidovorans cox* gene cluster into its genome (Alagesan & Malys, University of Nottingham, unpublished) and hence this point also needed to be addressed.

Specific objectives of this study were:

I. To develop and apply a protocol for the isolation and phenotypic characterisation of *C. necator* H16 mutants evolved in the presence of CO to obtain a profile of specific mutations that may lead to increased CO tolerance.

II. To characterise specific obtained mutations in wild type *C. necator* H16 to elucidate which confer CO adaptation, then further characterise genes and gene regions in which these mutations are found to generate a defined, more resistant strain.

III. To explore the conjugative capacity of pHCG3 and to attempt its transfer into *A. tumefaciens* and *C. necator*. This will involve analysis of the LuxRI-family quorum sensing system of pHCG3 to elucidate AHL signal molecules that are produced and to exploit this knowledge to achieve conjugative transfer.

IV. To characterise putative CO-responsive transcription factors from pHCG3 to reveal whether they are functional in *C. necator* H16 and thus able to drive *cox* gene expression.

2. Materials and Methods

2.1. Bacterial strains, plasmids and oligonucleotides

All bacterial strains, plasmids and oligonucleotides used or generated in this study are listed in tables 2.1, 2.2 and 2.3, respectively.

Table 2.1. Bacterial strains used in this study.

Strains	**Description**	**Source**
Cupriavidus necator H16 (DSM 428)	Wild type strain	German Collection of Microorganisms and Cell Cultures GmbH (DSMZ)
C. necator cydA2B2 SNP KO	H16 derivative with deletion upstream of *cydA2B2*	This study
C. necator cydA2B2 SNP	H16 derivative with SNP generated upstream of *cydA2B2*	This study
C. necator phcR KO	H16 derivative with phcR gene deleted	This study
C. necator phcR 12bp	H16 derivative with 12 bp (4 amino acid) deletion in *phcR*	This study
C. necator phcA SNP KO	H16 derivative with upstream *phcA* region deleted	This study
C. necator phcA SNP	H16 derivative with SNP generated upstream of *phcA*	This study
C. necator RNApol 5744 KO	H16 derivative with part of the open reading frame of the RNA polymerase sigma-54 factor gene deleted	This study
C. necator RNApol 5744	H16 derivative with SNP generated in the RNA polymerase sigma-54 factor gene (position 405744)	This study
C. necator RNApol 4574 KO	H16 derivative with part of the open reading frame of the RNA polymerase sigma-54 factor gene deleted	This study
C. necator RNApol 4574	H16 derivative with SNP generated in the RNA polymerase sigma-54 factor gene (position 404574)	This study
C. necator cyd + *phcR* KO	H16 derivative with SNP generated upstream of *cydA2B2* and *phcR* gene deleted	This study
C. necator cyd + *phcR* 12bp	H16 derivative with SNP generated upstream of *cydA2B2* and 12 bp (4 amino acid) deletion in *phcR* gene	This study
C. necator cyd + *phcA* SNP KO	H16 derivative with SNP generated upstream of *cydA2B2* and upstream *phcA* region promoter region deleted	This study
C. necator cyd + *phcA* SNP	H16 derivative with SNP generated upstream of *cydA2B2* and *phcA*	This study

C. necator cyd + *phcA* SNP + RNApol 4574 KO	H16 derivative with SNP generated upstream of *cydA2B2* and *phcA* with part of the open reading frame of the RNA polymerase sigma-54 factor gene	This study
C. necator cyd + *phcA* SNP + RNApol 4574	H16 derivative with SNP generated upstream of *cydA2B2* and *phcA* and SNP in the RNA polymerase sigma-54 factor gene (position 404574)	This study
C. necator cyd + RNApol 5744 KO	H16 derivative with SNP generated upstream of *cydA2B2* with part of the open reading frame of the RNA polymerase sigma-54 factor gene	This study
C. necator cyd + RNApol 5744	H16 derivative with SNP generated upstream of *cydA2B2* and in the RNA polymerase sigma-54 factor gene (position 405744)	This study
C. necator cyd + RNApol 4574 KO	H16 derivative with SNP generated upstream of *cydA2B2* with part of the open reading frame of the RNA polymerase sigma-54 factor gene	This study
C. necator cyd + RNApol 4574	H16 derivative with SNP generated upstream of *cydA2B2* and in the RNA polymerase sigma-54 factor gene (position 405744)	This study
C. necator Δ*cydA2B2*	H16 derivative containing knock out of *cydA2B2*	This study
C. necator Sub	Isolated *C. necator* H16 derivative displaying greater CO tolerance	This study
C. necator EvoA	Isolated CO resistant *C. necator* H16 derivative	This study
C. necator EvoB	Isolated CO resistant *C. necator* H16 derivative	This study
C. necator EvoC	Isolated CO resistant *C. necator* H16 derivative	This study
C. necator E1-E10	Isolated CO resistant *C. necator* H16 derivatives	This study
Oligotropha carboxidovorans OM5 (DSM 1227)	Wild type OM5 strain carrying megaplasmid pHCG3; chemolitho-autotrophic, can grow on CO	DSMZ
O. carboxidovorans_ traM	OM5 strain carrying deactivated *traM* gene on pHCG3 megaplasmid; tetracycline resistant	This study
E. coli DH5a	Used for cloning. DH5α F- φ80lacZΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 hsdR17*(rk-, mk+) *phoA supE44 thi*-1 *gyrA96 relA1* λ-	Grant *et al.* 1990
E. coli S17-1	Conjugal donor strain: *recA thi pro hsdR*$^-$ *hsdM*$^+$ RP4::2-Tc::Mu::Km Tn7 *λpir*, Tp^R Sm^R	Simon *et al.* 1983
E. coli JM109 pSB403	DH5α containing the pSB403 bioluminescent reporter plasmid	Winson *et al.* 1998
A. tumefaciens A676	Plasmidless strain of *A. tumefaciens*; spontaneous rifampicin-resistant derivative of strain C58.00	Danino *et al.* 2003

A. tumefaciens A677	Plasmidless strain of *A. tumefaciens*; spontaneous streptomycin-resistant derivative of strain C58.00	Danino *et al.* 2003
A. tumefaciens A676 pHCG3_*traM*	*A. tumefaciens* transconjugant containing pHCG3_*traM*	This study
A. tumefaciens A677 pHCG3_*traM*	*A. tumefaciens* transconjugant containing pHCG3_*traM*	This study
Pseudomonas strain ws	*Pseudomonas* isolate found as contaminant during conjugation experiments	This study
P. cws pHCG3_*traM*_A	*Pseudomonas* ws transconjugant containing pHCG3_*traM*; isolate A	This study
P. cws pHCG3_*traM*_B	*Pseudomonas* ws transconjugant transformed with pHCG3_*traM*; isolate B	This study
Pseudomonas azotoformas (DSM 18862)	Wild type strain for conjugation control	DSMZ
Pseudomonas lactis (DSM 29167)	Wild type strain for conjugation control	DSMZ
Pseudomonas paralactis (DSM 29164)	Wild type strain for conjugation control	DSMZ

Table 2.2. Plasmids used in this study.

Plasmid	**Description**	**Source**
pLO3	Suicide plasmid containing tetracycline resistance marker and s*acB* gene	Lenz & Friedrich 1998
pBBR1MCS-2-PphaC-eyfp-c1	Overexpression vector containing *tac* promoter with *eyfp* gene	Pfeiffer & Jendrossek 2011
pLO3S_*traM*	Plasmid for deletion of *traM* and replacement with spectinomycin cassette	This study
pLO3_*traMt*	Integration plasmid containing a truncated form of *traM*	This study
pLO3_*cydA2B2* SNP KO	Deletion plasmid for upstream region of cytochrome *bd* ubiquinol oxidase *(cydA2B2)*	This study
pLO3_*cydA2B2* SNP KI	Knock-in plasmid to generate a defined SNP within the *cydA2B2* upstream region	This study
pLO3_Δ*cydA2B2*	Plasmid for in-frame deletion of *cydA2B2*	
pLO3_ *phcR* 12bp	Integration plasmid for inserting *phcR* with a defined 12 bp deletion into a larger *phcR* knock-out	This study
pLO3_Δ*phcR*	Deletion plasmid for in frame *phcR* knockout	This study
pLO3_ *phcA* US SNP KO	Deletion plasmid for the upstream region of *phcA*	This study
pLO3_ *phcA* US SNP KI	Integration plasmid for the upstream region of *phcA* containing novel SNP	This study
pLO3_ *phcA* IF SNP KO	Deletion plasmid for part of the open reading frame of *phcA*	This study

pLO3_ *phcA* IF SNP KI	Integration plasmid to generate a defined SNP within the *phcA* open reading frame	This study
pLO3_ RNApol 5744 KO	Deletion plasmid for part of the open reading frame of the RNA polymerase sigma factor-54 gene (position 405744)	This study
pLO3_ RNApol 5744 KI	Knock-in plasmid to generate a defined SNP within the RNA polymerase sigma factor-54 gene containing novel SNP (position 405744)	This study
pLO3_ RNApol 4574 KO	Deletion plasmid for an in-frame section of the RNA polymerase sigma factor-54 gene (position 404574)	This study
pLO3_ RNApol 4574 KI	Knock-in plasmid to generate a defined SNP within RNA polymerase sigma factor-54 gene containing novel SNP (position 404574)	This study
pBBR1_*traI*	pBBR1MCS-2-PphaC-eyfp-c1-based overexpression vector containing *traI* downstream of promoter	This study
pBBR1_*solI*	pBBR1MCS-2-PphaC-eyfp-c1-based overexpression vector containing *solI* downstream of promoter	This study
pEH006	RFP expressing plasmid with an arabinose-inducible system	Erik Hanko
pEH006_TFBS1	RFP expressing plasmid with upstream intergenic region of *coxM*	This study
pEH006_TFBS2	RFP expressing plasmid with the region from the start of *coxC* to the start of *coxM* replacing the arabinose inducible promoter	This study
pEH006_TFBS3	RFP expressing plasmid with the intergenic region downstream of *coxB* to the start of *coxM* replacing the arabinose inducible promoter	This study
pEH006_TFBS4	RFP expressing plasmid with the region upstream of *coxB* to the start of *coxM* replacing the arabinose inducible promoter	This study
pEH006_*cyd* US	RFP expressing plasmid containing upstream intergenic region of *cydA2B2* replacing the arabinose inducible promoter	This study
pEH006_*cyd* US SNP	RFP expressing plasmid containing upstream intergenic region of cytochrome *bd* with altered SNP replacing the arabinose inducible promoter	This study
pEH006_*cyd* Exp	pEH006 with RFP gene replaced with *cydA2B2* downstream of the arabinose-inducible promoter	This study

Table 2.3. Oligonucleotides used in this study.

Oligonucleotide	**Purpose in this study**	**Sequence (5' to 3')**
PHA_F2	Confirmation of successful ligations into pLO3	AGTGGAACGAAAACTCACGTTAAGGG
PHA_R2	Confirmation of successful ligations into pLO3	ACACGGTGCCTGACTGCG
CytoD_LHA F	To amplify left homology arm of *cydA2B2* upstream region for cloning into pLO3	GATCCTTTTAATTCGAGCTCGGTACCCGGGAGCATCGGCAACATGACGG
CytoD_LHA R	To amplify left homology arm of *cydA2B2* upstream region for cloning into pLO3	GAGGAGATTGCGCGGGATCACAGTC

CytoD_RHA F	To amplify right homology arm of *cydA2B2* upstream region for cloning into pLO3	GACTGTGATCCCG CGCAATCTCCTCGC GCCTAGCCTGAC
CytoD_RHA R	To amplify right homology arm of *cydA2B2* upstream region for cloning into pLO3	AATTAGCTTGCAT GCCTGCAGGTCAT GCAGCAGCTGCCA G
CytoD SNP F	To amplify right homology arm containing SNP in *cydA2B2* upstream region	ACGCTTGTCCATG GTATCCTGTCCCGC AATTGACCC
CytoD SNP R	To amplify left homology arm containing SNP in *cydA2B2* upstream region	GGACAGGATACCA TGGACAAGCGTAA GGACT
CytoD SNP det F	Confirmation of single and double crossover of pLO3_*cydA2B2* KO and KI	TGCAAGACGACCG TACTTTCGG
CytoD SNP det R	Confirmation of single and double crossover of pLO3_*cydA2B2* KO and KI	GCGTGCTTCCAGG AACGC
CytoD seq F	Confirmation of single crossover of pLO3_*cydA2B2* KO and KI	GGTGATGATCCAG GAAACCTTCAAGG G
CytoD seq R	Confirmation of single crossover of pLO3_*cydA2B2* KO and KI	CGATCAGCACCAT CCACAGCG
HisKin_LHA F	To amplify left homology arm of *phcR* 12 bp deletion region for cloning into pLO3	CACCTAGATCCTTT TAATTCGAGCTCC ATTCGGTTCGGTCA TCGGTCA
HisKin_LHA R	To amplify left homology arm of *phcR* 12 bp deletion region for cloning into pLO3	CACCGCGAGCGCA CCCTGCGCCAGCA CCGGCTGATGG
HisKin_RHA F	To amplify right homology arm of *phcR* 12 bp deletion region for cloning into pLO3	CGCAGGGTGCGCT CGCGGTGGCGCTC GGCGGCGCCTGCC AGCACCTCG
HisKin_RHA R	To amplify right homology arm of *phcR* 12 bp deletion region for cloning into pLO3	GCAGGTTTAAACA GTCGACTCTAGAC CGACACCCTGGTG CTG
HisKin Del det F	Confirmation of single and double crossover of pLO3_*phcR* 12bp	CTCATGCGCCAGG AAGGC
HisKin Del det R	Confirmation of single and double crossover of pLO3_*phcR* 12bp	CTTCCGCATCCTGG AAAAGCC
HisKin seq SC F	Confirmation of single crossover of pLO3_*phcR* 12 bp and pLO3_Δ*phcR*	GCTGGAAATACTT GACTGCAGTCGC
HisKin seq SC R	Confirmation of single crossover of pLO3_*phcR* 12 bp and pLO3_Δ*phcR*	GCGCTCGATTCGTC GATCCG
PhcR Ko LHA F	To amplify left homology arm of *phcR* for cloning into pLO3	TCCTTTTAATTCGA GCTCGGTACCCGG GGGCTGCACCAGG TACTCGG

PhcR Ko LHA R	To amplify left homology arm of *phcR* for cloning into pLO3	AAGCAGGCACAGTCGAAAGCATAGGAGTGACCG
PhcR Ko RHA F	To amplify left homology arm of *phcR* for cloning into pLO3	GCTTTCGACTGTGCCTGCTTCGCTTCGC
PhcR Ko RHA R	To amplify left homology arm of *phcR* for cloning into pLO3	GCAGGTTTAAACAGTCGACTCTAGAGTCTACACGTTCATGAACGAGCG
PhcR seq F	Confirmation of single and double crossover of pLO3_Δ*phcR*	GCACGGCACAGGATGGC
PhcR seq R	Confirmation of single and double crossover of pLO3_Δ*phcR*	GCTGAACCTGGTCATCTTCGC
PhcA IF LHA F	To amplify left homology arm of *phcA* in frame region for cloning into pLO3	CACCTAGATCCTTTTAATTCGAGCTCCGGGAGGGAAAGAAAAGTACCG
PhcA IF LHA R	To amplify left homology arm of *phcA* in frame region for cloning into pLO3	GTGCATGTTGCCCTGCAAGACCTGCAGGC
PhcA IF RHA F	To amplify right homology arm of *phcA* in frame region for cloning into pLO3	GCAGGTCTTGCAGGGCAACATGCACTCGTACATCCG
PhcA IF RHA R	To amplify right homology arm of *phcA* in frame region for cloning into pLO3	GCAGGTTTAAACAGTCGACTCTAGAGCTTCCGCAGTAAAGGAAAACGC
PhcA IF SNP F	To amplify right homology arm containing SNP in of *phcA* in frame region	ATGTCCTTGGTCCGCGCATTGACGGTGAAGATCATGTCG
PhcA IF SNP R	To amplify right homology arm containing SNP in of *phcA* in frame region	CGTCAATGCGCGGACCAAGGACATGGCCTGCG
PhcA IF detect F	Confirmation of single and double crossover of pLO3_*phcA* IF KO and KI	AATCCCATTCGTGGGAAACGG
PhcA IF detect R	Confirmation of single and double crossover of pLO3_*phcA* IF KO and KI	CCACAGACCCAAACGATAAGCC
PhcA IF SNP detect F	Confirmation of single and double crossover of pLO3_*phcA* IF KO and KI	GGCAGTCCGACAAACTCGC
PhcA IF 3 seq SC F	Confirmation of single crossover of pLO3_*phcA* IF KO and KI	CGCGAGGTGGAATCGTAGGC
PhcA IF 3 seq SC R	Confirmation of single crossover of pLO3_*phcA* IF KO and KI	AGCTGGTACGACAACGAGTGGG

PhcA US SNP LHA F	To amplify left homology arm of *phcA* upstream region for cloning into pLO3	GGATCTTCACCTA GATCCTTTTAATTC GAGCTCCGAGCAT GGTGTTCAGTTCCG
PhcA US SNP LHA R	To amplify left homology arm of *phcA* upstream region for cloning into pLO3	GGATGTACGAGTG CATGTTGCCGAAG CTTTTAGTGATCTT CACCG
PhcA US SNP RHA F	To amplify right homology arm of *phcA* upstream region for cloning into pLO3	TCGGTGAAGATCA CTAAAAGCTTCGG CAACATGCACTCG TACATCCG
PhcA US SNP RHA R	To amplify right homology arm of *phcA* upstream region for cloning into pLO3	GCAGGTTTAAACA GTCGACTCTAGAG CTTCCGCAGTAAA GGAAAACGC
PhcA US SNP F	To amplify right homology arm of *phcA* upstream region containing SNP for cloning into pLO3	TTTCAACGGCGTCT TACTAGTGTAAGA GGAAAAAACAAGC TCGG
PhcA US SNP R	To amplify left homology arm of *phcA* upstream region containing SNP for cloning into pLO3	TTTTTTCCTCTTAC ACTAGTAAGACGC CGTTGAAATCTGA TGTGC
PhcA US SNP detect F	Confirmation of single and double crossover of pLO3_*phcA* US KO and KI	GCGGATCGTCAAA GATTTCACGC
PhcA US SNP detect R	Confirmation of single and double crossover of pLO3_*phcA* US KO and KI	GCAGGCGGAAATG TTCAGGC
PhcA US 4 seq R	Confirmation of single and double crossover of pLO3_*phcA* US KO and KI	AGCTGGTACGACA ACGAGTGGG
RNApol 5744 LHA F	To amplify left homology arm of RNA polymerase 5744 region for cloning into pLO3	GGATCTTCACCTA GATCCTTTTAATTC GAGCTCGCCTCGG CATCACATCCG
RNApol 5744 LHA R	To amplify left homology arm of RNA polymerase 5744 region for cloning into pLO3	ATGCAGCGGCCAC TGACGGAAAACCC TCTGC
RNApol 5744 RHA F	To amplify right homology arm of RNA polymerase 5744 region for cloning into pLO3	TTTCCGTCAGTGGC CGCTGCATCACAT GCG
RNApol 5744 RHA R	To amplify right homology arm of RNA polymerase 5744 region for cloning into pLO3	TGCAGGTTTAAAC AGTCGACTCTAGA GAAGCCTTCCGTC GAAGCC
RNApol 5744 SNP F	To amplify right homology arm of RNA polymerase 5744 region containing SNP for cloning into pLO3	TGGGAAAGGCGGA GCGGTAGCGACGG TTTCATGCG

RNApol 5744 SNP R	To amplify left homology arm of RNA polymerase 5744 region containing SNP for cloning into pLO3	ATGAAACCGTCGC TACCGCTCCGCCTT TCCCAGCATCTGG
RNApol 5744 5 seqF	Confirmation of single and double crossover of pLO3_RNA polymerase 5744 KO and KI	GCCGGACTCACCC TTTGCG
RNApol 5744 5 seqR	Confirmation of single and double crossover of pLO3_RNA polymerase 5744 KO and KI	GCCAGAAACGCCT GAACGC
RNApol 4574 LHA F	To amplify left homology arm of RNA polymerase 4574 region for cloning into pLO3	GGATCTTCACCTA GATCCTTTTAATTC GAGCTCCCAGGAT TCCCTCTCTTTGTA TGGC
RNApol 4574 LHA R	To amplify left homology arm of RNA polymerase 4574 region for cloning into pLO3	CATCCAGCAGCCT TGATCAAGCAACT GATAG
RNApol 4574 RHA F	To amplify right homology arm of RNA polymerase 4574 region for cloning into pLO3	TTGATCAAGGCTG CTGGATGTTCTTGA TCAG
RNApol 4574 RHA R	To amplify right homology arm of RNA polymerase 4574 region for cloning into pLO3	TGCAGGTTTAAAC AGTCGACTCTAGA GCTGACGACGACA GCTATGGC
RNApol 4574 SNP F	To amplify right homology arm of RNA polymerase 4574 region containing SNP for cloning into pLO3	GGGAGATGGTTTA CTCGTGTAAACCC AGTGTATCGG
RNApol 4574 SNP R	To amplify left homology arm of RNA polymerase 4574 region containing SNP for cloning into pLO3	ACTGGGTTTACAC GAGTAAACCATCT CCCGGGTGACG
RNApol SNP detect F	Confirmation of single and double crossover of pLO3_RNA polymerase 4574 and pLO3_RNA polymerase 5744 KO and KI	GTCGGAAAGGGGA TTCCTCGG
RNApol SNP detect R	Confirmation of single and double crossover of pLO3_RNA polymerase 4574 and pLO3_RNA polymerase 5744 KO and KI	GCCTACATCATCA GCGAAGGC
RNApol 4574 seqF	Confirmation of single and double crossover of pLO3_RNA polymerase 4574 and pLO3_RNA polymerase 5744 KO and KI	AGATACATGGTGA TGCAATGGTGCC

RNApol 4574 seqR	Confirmation of single and double crossover of pLO3_RNA polymerase 4574 and pLO3_RNA polymerase 5744 KO and KI	CTCGGTCAACCTG CAGAGCG
CytoD inter US F	To amplify *cydA2B2* upstream intergenic region for cloning into pEH006	ATATGACGTCGCG CCACCCCGGCTTCC G
CytoD inter US R	To amplify *cydA2B2* upstream intergenic region for cloning into pEH006	ATATATCATATGC GATGTGCTCTCCGT GGGGTGGC
CytoD Exp F	To amplify *cydA2B2* for cloning into pEH006	AATTCAAAAGATCTTTT AAGAAGGAGATATACA ATGTACGGACTCACAG CC
CytoD Exp R	To amplify *cydA2B2* for cloning into pEH007	TGGAGATCCTTACTCGA GTTTGTCAGTGGTACCC CTCCCC
Cbd op del LHA F	To amplify left homology arm for the in frame knock out of *cydA2B2* for cloning into pLO3	GGATCTTCACCTA GATCCTTTTAATTC GAGCTGCACAAGC TGGCCATGGG
Cbd op del LHA R	To amplify left homology arm for the in frame knock out of *cydA2B2* for cloning into pLO3	CCTGGCCATCAGT GGCGTACATCGAT GTGCTCTCCG
Cbd op del RHA F	To amplify right homology arm for the in frame knock out of *cydA2B2* for cloning into pLO3	AGAGCACATCGAT GTACGCCACTGAT GGCCAGGATCCC
Cbd op del RHA R	To amplify right homology arm for the in frame knock out of *cydA2B2* f3r cloning into pLO3	GCAGGTTTAAACA GTCGACTCACGTC GGTGATCCAGCG
TraI_F	To amplify *traI* for cloning into pBBR1MCS-2	GATATACATATGC GAGCAATAGCAAT CACAGC
TraI_R	To amplify *traI* for cloning into pBBR1MCS-2	GATGCTAGCTGAC TACGCGGCCGAGC GGC
SolI_F	To amplify *solI* for cloning into pBBR1MCS-2	GATATACATATGA TCAGCGTCATCCG CGGA
SolI_R	To amplify *solI* for cloning into pBBR1MCS-2	CATGGATCCAATC ATTGTAGTCGAGC CGGTATGA
TraMt_F	To amplify truncated version of *traM* for cloning into pLO3	ATTCGAGCTCTAA GTTTTCAAGAAAG AA
TraMt_R	To amplify truncation version of *traM* for cloning into pLO3	TAAGTCTAGAACG CTCGATCAGTGTTG C
1Int_coxCM_F	To amplify upstream intergenic region of *coxM*	ATATGACGTCGCC AGTAATGCTGACG TATTTCG

2CoxC_AatII_F	To amplify region from the start of *coxC* to the start of *coxM*	ATATGACGTCATG CCCATCACACACG ATCCG
3Int_coxB_F	To amplify region from the start of *coxC* to the start of *coxM*	ATATGACGTCCGC GGGTAACAATGGT GACGAG
4Int_coxB_F	To amplify region upstream of *coxB* to the start of *coxM*	ATATGACGTCCAA AAGCCTTTGATGCT GCTTTGCG
Int_coxCM_R	To amplify all CO transcription factor binding sites upstream of *coxM*	ATATATCATATGG ACCTCCCCAGGTT ACCTGCG
40 seqF	Confirmation of pHCG3_*traM* presence in putative transconjugants	CGGCCTTGGTCAA GTTGACCC
40 seqR	Confirmation of pHCG3_*traM* presence in putative transconjugants	CCGCCAGGCTTTG AGAATGCG
220 seqF	Confirmation of pHCG3_*traM* presence in putative transconjugants	GACTCTCTATCGAT TTGTTGCATGGG
220 seqR	Confirmation of pHCG3_*traM* presence in putative transconjugants	CATCACGAGCACA AAGTCAATTTCGG
880 seq F	Confirmation of pHCG3_*traM* presence in putative transconjugants	CATCGCTAGCCCG CTGTACCC
880 seq R	Confirmation of pHCG3_*traM* presence in putative transconjugants	TGCCCCATTCCGAC GTACCG
1000 seqF	Confirmation of pHCG3_*traM* presence in putative transconjugants	CGCCAAGACGCAG AGGTTTGGG
1000 seqR	Confirmation of pHCG3_*traM* presence in putative transconjugants	CGTTCGCGCAGCG AATAGAGCC
HoxA F	Confirmation of pHCG3_*traM* presence in putative transconjugants	ATCATCGTCGGGA GCATCCC
HoxA R	Confirmation of pHCG3_*traM* presence in putative transconjugants	ATGGCCGATACTC CGCTTCC
TraG F	Confirmation of pHCG3_*traM* presence in putative transconjugants	TCTCCTGACAGGTC TTCCAAATCCC
TraG R	Confirmation of pHCG3_*traM* presence in putative transconjugants	CAGTGGATGACGA TTAACGAGGC
FabG F	Confirmation of *A. tumefaciens* circular chromosome presence	ATGCAGCGTTTTAC CAACAGAAC
FabG R	Confirmation of *A. tumefaciens* circular chromosome presence	TCAGGAAAACAGG ACGCC
FtsA F	Confirmation of *A. tumefaciens* circular chromosome presence	ATGAATTTCGATGT CCGCCACC

FstA R	Confirmation of *A. tumefaciens* circular chromosome presence	TGACCGTCGACAG CCTGATCGTTA
RecF F	Confirmation of *O. carboxidovorans* circular chromosome presence	ATGATCCCCTCCCG CATCCT
RecF R	Confirmation of *O. carboxidovorans* circular chromosome presence	TCACCCGCGCGGC GTGAT
Rbfa F	Confirmation of *O. carboxidovorans* circular chromosome presence	ATGCCCCGTCATTC ACGCGG
Rbfa R	Confirmation of *O. carboxidovorans* circular chromosome presence	TCATGCATTGTCGT CCGAATCGG
Tet R	Confirmation of single crossover of pLO3_*traMt*	CTTCCATTCAGGTC GAGGTG
TraM seqF	Confirmation of single crossover of pLO3_*traMt* and pLO3S_*traM*	GGGCCTGTGACGC ATGTCCAGGT
TraM seqR	Confirmation of single crossover of pLO3_*traMt* and pLO3S_*traM*	CTCATCTAATGCG ACGCCTGAATAT
TraM HA seqF	Confirmation of single and double crossover of and pLO3S_*traM*	ACGTTAGCGAGCT CGCACCC
TraM HA seqR	Confirmation of single and double crossover of and pLO3S_*traM*	CCTTTGGTGTCCAG GTCTTCGC

2.2. Microbiological methods

2.2.1. Preparation of Media

Escherichia coli, *Agrobacterium tumefaciens*, *Pseudomonas* strain cws, *Pseudomonas azotoformans* (DSM 18862), *Pseudomonas lactis* (DSM 29167), *Pseudomonas paralactis* (DSM 29164) and *Cupriavidus necator* were routinely grown in lysogeny broth (LB) medium, consisting of 10 g/l tryptone, 5 g/l yeast extract, 5 g/l sodium chloride at pH 7.2 with the addition of 15 g/l agar (Bacto-agar, BD) if for solid agar medium. For selection of plasmids bearing strains, 15 µg/ml tetracycline and 150 µg/ml spectinomycin were added to the medium, if required. *E. coli* was additionally grown in M9 minimal medium for analysis of supernatants for *N* – acyl-L-homoserine lactone (AHL) presence.

O. carboxidovorans OM5 and (for selection purposes and AHL analysis) *A. tumefaciens* and *P.* cws were grown in mineral salts medium (Schlegel et al., 1961),

which contained (per litre) 4.5 g sodium phosphate dibasic, 1.5 g potassium phosphate monobasic, 1.0 g ammonium chloride, 0.2 g magnesium sulphate, 0.02 g calcium chloride, 1.2 mg iron (III) citrate and 1 ml TS2 trace elements (sodium molybdate 900 mg/l, boric acid 300 mg/l, cobalt (II) chloride 200 mg/l, zinc sulphate 100 mg/l, manganese (II) chloride 30 mg/l, sodium selenite 20 mg, nickel (II) chloride 20.0 mg/l, copper (II) chloride 10 mg/l) with the addition of 15 g/l agar (Bacto-agar, BD) if for solid agar medium. For growth on solid media and for precultures 0.2% (w/v) of sodium pyruvate was added to the medium for *O. carboxidovorans* OM5 and 0.2% (w/v) of glucose was added for *A. tumefaciens* and *P.* cws.

For precultures, selection of transconjugants and growth in the presence of CO, *C. necator* was grown in a different mineral salts medium, containing (per litre) 4.5 g sodium phosphate dibasic, 1.5 g potassium phosphate monobasic, 1.0 g ammonium chloride, 0.2 g magnesium sulphate, 0.02 g calcium chloride, 1.2 mg iron (III) citrate and 1 ml SL7 solution (hydrochloric acid (25%, v/v) 1 ml /l, zinc chloride 70 mg/l, manganese (II) chloride 100 mg/l, boric acid 60 mg/l, cobalt (II) chloride 200 mg/l, copper (II) chloride 20 mg/l, nickel (II) chloride 20 mg/l, sodium molybdate 40 mg/l), with the addition of 15 g/l agar (Bacto-agar, BD) if for solid agar medium. For growth on solid media and for precultures, 0.4% (w/v) of fructose was added to the medium. For counter selection during conjugation with *E. coli* S17-1, 10 μg/ml gentamicin was added to agar plates.

2.2.2. Bacterial cultivation

2.2.2.1. Growth at atmospheric pressure

E. coli was grown aerobically in LB or M9 medium at 37 °C and liquid cultures shaking at 200 revolutions per minute (RPM) in 50 ml Falcon tubes. *C. necator, O. carboxidovorans, A. tumefaciens* and *P.* cws were grown in LB or minimal medium at 30°C and liquid cultures shaking at 200 revolutions per minute (RPM) in 50 ml Falcon tubes or 250 ml baffled Erlenmeyer flasks if larger volumes were required. *E. coli* was taken from bead stocks (Microbank^{TM}) and put directly into 5 ml liquid media for growth to use for either pre-cultures or for plasmid preparations. *C. necator, O. carboxidovorans, A. tumefaciens* and *P.* cws were first streaked onto agar plates from bead stocks, and growing cells used to inoculate 5 ml liquid media for pre-cultures.

2.2.2.2. Growth under 2 bar pressure

For aerobic growth under 2 bar pressure, *C. necator* was grown with 20 ml minimal medium (with fructose for strains growing in the presence of CO) in 150 ml serum bottles, which were sealed with butyl rubber stoppers and aluminium clamps. Different concentrations of CO, N_2, CO_2 and H_2 were injected into the serum bottles through a hypodermic needle and 0.2 µm nitrocellulose filter using a commercial gas exchange system (GR Instruments). Gas mixtures were injected in addition to the existing 1 bar atmospheric pressure to give a total of 2 bar acting upon growing cultures.

2.2.3. Storage of bacterial cultures

E. coli, C. necator, O. carboxidovorans, A. tumefaciens and *P.* cws stocks were made by adding 800 µl of cell culture to a storage bead vial (Microbank^{TM}, Pro-lab), which was inverted 5-10 times for attachment. The top liquid layer was removed and the vials were stored at -80 °C. CO-evolved *C. necator* strains were stored as a 50% (v/v) glycerol stock in 2 ml cryogenic vials at -80 °C.

2.2.4. Preparation and transformation of chemically competent *E. coli*

Desired *E. coli* strains (DH5α or S17-1) were grown overnight in 5 ml LB liquid medium in 50 ml Falcon tubes. 1 ml of the overnight medium was inoculated into 100 ml LB liquid medium in 250 ml shaking flasks and grown aerobically at 37°C shaking at 200 RPM. OD_{600} measurements were taken hourly until the OD_{600} reached values between 0.7-0.8. The culture was then cooled on ice for 20 minutes and subsequently centrifuged using a chilled rotor at 3000xg for 15 minutes. The supernatant was removed and the pellet was suspended in 100 ml ice-cold 0.1 M $CaCl_2$. Cells were centrifuged at 3000xg for 15 minutes and the pellet resuspended in 50 ml of ice-cold 0.1 M $CaCl_2$ with 15% (v/v) glycerol and cells. Cells were centrifuged again at 3000xg for 15 minutes and the final pellet was resuspended in 2 ml ice-cold 0.1 M $CaCl_2$ with 15% (v/v) glycerol and aliquoted into pre-chilled Eppendorf tubes (50 µl each) and frozen at -80°C.

For transformations, aliquots were thawed on ice and 1-5 µl of desired plasmid (50-200 ng/µl) was added, mixed and incubated on ice for 1 hour. Cells were heat shocked at 42°C for 2 minutes then left on ice for 1 minute and 950 µl S.O.C medium (Invitrogen™) was added to the cell mixture which was transferred to a sterile 1.5 ml

reaction tube. The cell mixture was then incubated at 37 °C shaking at 200 RPM for 1 hour. After recovery, 100 µl of the mixture was plated on LB agar with antibiotic selection and incubated overnight at 37 °C. Colonies found were screened for the presence of the plasmid.

2.2.5. Transformation via conjugation

2.2.5.1. *C. necator* H16

Transformation of *C. necator* was performed via conjugation with strains of *E. coli* S17-1, containing the desired plasmid. Wild type *C. necator* H16 and *E. coli* S17-1 were cultured in 5 ml LB medium (+ appropriate antibiotic for S17-1 depending on the plasmid used, see Table 2.4) to an OD_{600} of around 3. 1 ml of each culture was centrifuged at 5000 RPM for four minutes. The supernatants were discarded and the cell pellets resuspended in 1 ml LB medium and centrifuged at 5000 RPM for four minutes. The supernatants were discarded and the cell pellet combined by resuspending in 1 ml LB medium and centrifuging at 5000 RPM for four minutes. The supernatant was discarded and the cell pellet was resuspended in 50 ml LB medium that was spotted onto an LB agar plate. Following incubation for 6 hours at 30°C, the cell spot was resuspended in 200 ml PBS and plated onto F-MM + gen 10 µg/ml + (appropriate antibiotic for selection, see Table 2.4). Following incubation at 30°C for at least 48 hours, colonies were restreaked on F-MM selective plates. Following incubation at 30°C for 18 hours, restreaked colonies were screened by PCR for plasmid presence. For detection of single-crossover of pLO3-based suicide plasmids, colonies were screened using two primer pairs for the confirmation of right homology arm (RHA) or left homology arm (LHA) crossover. A full schematic of this screening process is shown in section 3.2.4.2. (Fig. 3.11), which describes the generation of a mutant library in *C. necator*.

2.2.5.2. *Oligotropha carboxidovorans* OM5

Transformation of *O. carboxidovorans* was performed via conjugation with strains of *E. coli* S17-1, containing the desired pLO3-based suicide plasmid. Wild type *O. carboxidovorans* was cultured in P-MM to and OD_{600} of around 0.8 and *E. coli* S17-1 was cultured in 5 ml LB medium (+ appropriate antibiotic for S17-1 depending on the plasmid used, see Table 2.4) to an OD_{600} of around 3 and diluted back to 0.8. The same washing steps described in 2.2.5.1 were performed but using agar plates containing

75% pyruvate (0.2%) minimal medium (P-MM) and 25% LB (P-MM/LB)**,** instead of LB. Following incubation for 18 hours at 30 ^{o}C, the cell spot was resuspended in 200 ml PBS and plated onto P-MM + (appropriate antibiotic for selection, see Table 2.4). Following incubation at 30^{o}C for at least 72 hours, antibiotic-resistant colonies were restreaked onto P-MM selective plates and incubated for 48 hours. Colony PCR was then performed to screen for single crossover events (for further detail see 4.2.2.5. and 4.2.2.7).

2.2.5.3. Transfer of megaplasmid pHCG3

For pHCG3 transfer, *O. carboxidovorans_traM* (containing modified version of pHCG3, pHCG3_*traM*) was used as a conjugation donor for recipient strains of *A. tumefaciens* and *Pseudomonas* species, as well as wild type *C. necator* H16 and *E. coli* DH5α. The methods for mating and conjugation frequency determination were taken from Danino *et al.* (2003). Agar plate-grown cells of the donor and recipient were resuspended separately in 3 ml P-MM/LB to a density of approximately 10^{9} cells ml^{-1}. Equal volumes of each suspension were mixed and 50 µl of this mixture were plated on a 0.2 µm nitrocellulose filter on P-MM/LB agar plates. Following incubation for 18 hours, the mating spot was resuspended from the filter in 0.9 ml PBS and transferred to minimal medium (MM) agar plates with an appropriate carbon source + tetracycline at 15 µg/ml (+ appropriate counter-selective antibiotic). In addition, serial dilutions of the recipient and donor resuspensions were made and plated out on P-MM/LB agar plates without tetracycline. Colony numbers were counted in order to calculate the number transconjugants per 10^{8} recipients, following screening by PCR using a pHCG3-specific primer pair (see 4.2.3.1. and 4.2.4.1. for PCR analysis).

2.2.6. Strain evolution for *C. necator* H16

For strain evolution of *C. necator* H16 as described in section 3.2.3.1, a single colony was used to inoculate 5 ml F-MM and grown to an OD_{600} of around 3. This culture was used to inoculate ten separate serum bottles containing 20 ml F-MM under an atmosphere of 50% CO/50% air (v/v), prepared as previously described (2.2.2.2). The cultures were shaken in at 200 RPM in a 30^{o}C incubator for at least 40 hours to an OD_{600} of around 2. This culture was used to inoculate 20 ml F-MM to an OD_{600} of 0.01 in a fresh serum bottle and the incubation cycle repeated. The cultures were monitored 18 hours after each inoculation to observe whether the OD_{600} had dramatically

increased. After this length of time wild type cultures would yield an OD_{600} of around 0.1, whereas evolved cultures yielded an OD_{600} of around 1, which was clearly distinguishable. Growth profiling was performed following observation of the fast growing phenotype, which is described in Chapter 3 (3.2.3.1 and 3.2.3.2). The purity of these cultures was checked taking a 1 ml sample of OD_{600} cells and making serial dilutions, plating out 10^{-5} on F-MM agar plates. Single colonies were screened by colony PCR using primers for the *hsdRN* gene (H16_A0006). Samples were taken every three subcultures and preserved in case they needed reanalysing or following cultures became contaminated.

2.3. Molecular biology methods

2.3.1. Isolation of DNA from bacterial strains

2.3.1.1. Isolation of plasmid DNA from *E. coli*

Plasmid DNA was isolated from 5 ml overnight *E. coli* cultures grown in LB medium (+ appropriate antibiotic for S17-1, depending on the plasmid used) using the GenElute Miniprep Kit (Sigma-Aldrich) following the manufacturer's instructions. Isolated plasmids were stored at -20 °C.

2.3.1.2. Isolation of genomic DNA from *C. necator, O. carboxidovorans, A. tumefaciens* and *P.* cws

Genomic DNA was isolated from 5 ml overnight minimal medium cultures using the Sigma-Aldrich GenElute Bacterial Genomic DNA kit following the manufacturer's instructions, with the exception of DNA elution being performed with 50 µl deionised water. Isolated gDNA was stored at -20 °C.

2.3.2. Analysis and manipulation of DNA

2.3.2.1. DNA quantification and purity analysis

A NanoDrop Lite (Thermo Scientific) was used to analyse isolated DNA, measuring the absorption at 260 nm, with purity estimations made using the 260/280 nm absorption ratio. DNA of decent quality had an A260/A280 ratio of 1.8-2.0, with an absorption of 1 representing 50 µg of dsDNA/ml. A Qubit Fluorometric quantification (Thermo Scientific) was used to analyse concentrations and purity of genomic DNA, which was sent for whole genome sequencing.

2.3.2.2. Polymerase chain reaction (PCR) for amplification of DNA

Screening PCRs for the analysis of colonies, plasmids or mutants was carried out using DreamTaq (Thermo Scientific) in 10 µl reactions. Each reaction contained 5 µl DreamTaq master mix, 0.5 µl 10 mM forward primer, 0.5 µl 10 mM reverse primer, 3 µl deionised water and 1 µl of DNA template at a concentration of between 1-100 ng/µl. The thermal cycling settings for DreamTaq PCRs were 95 °C for 1 minute (colony PCR samples were heated at 99°C in the thermocycler for 10 minutes in 20 ml deionised water prior) followed by 30 cycles of 95 °C for 10 seconds, 50 - 65 °C (depending on primer annealing temperature) for 30 seconds and 72 °C for 1 minute per kb of DNA amplified and a final extension of 72 °C for 10 minutes.

DNA used for cloning or sent for Sanger sequencing was amplified using Phusion polymerase (New England BioLabs) in 25 µl reactions. Each reaction contained 12.5 µl Phusion master mix, 1.25 µl 10 mM forward primer, 1.25 µl 10 mM reverse primer, 9 µl deionised water and 1 µl DNA at a concentration of between 1-50 ng/µl. The thermal cycling settings for Phusion PCRs was 98 °C for 2 minutes followed by 30 cycles of 98 °C for 10 seconds, 60-72 °C (depending on primer annealing temperature) for 30 seconds, 72 °C at 30 seconds per kb and a final extension at 72 °C for 10 minutes. PCR amplifications were carried using a TProfessional Trio cycler (Biometra).

2.3.2.3. Agarose gel electrophoresis

DNA was separated using agarose gel electrophoresis at 90 V for 1–1.5 hours in 1x TAE running buffer (40 mM Tris, 1 mM EDTA and 0.1% (v/v) glacial acetic acid) using a compact gel tank (Biometra) and a PowerPac 300 (Bio-Rad). Agarose gel was routinely made to 1% agarose (w/v) with 1x TAE buffer and 10 µl per 100 ml SYBR Safe (Thermo Scientific). DNA was visualised with a Gel Dock XR+ System (BioRad). Band sizes were determined using a 2-log DNA ladder (New England BioLabs).

2.3.2.4. DNA purification

DNA from agarose gels was purified using Monarch DNA Gel Extraction Kit (New England BioLabs) following the manufacturer's instructions. DNA from PCR reactions was purified using Monarch PCR & DNA Cleanup Kit (New England BioLabs) following the manufacturer's instructions.

2.3.2.5. Restriction digest of DNA

Cleavage of DNA was carried out in 20 µl reactions using 1 µl of various CutSmart restriction enzymes (New England BioLabs) following the manufacturer's instructions, with the exception of longer incubations (18 hours) with NdeI.

2.3.2.6. Ligation of DNA

DNA ligations were carried out in 20 µl reactions using 1 µl T4 DNA ligase (New England Biolabs), 2 µl 10x ligation buffer, 14 µl of purified DNA fragments and 3 µl cut vector. Reactions were incubated for 1 hour at room temperature.

2.3.2.7. Other DNA assembly methods

Plasmid DNA was additionally assembled using NEBuilder HiFi DNA assembly (New England Biolabs) following the manufacturer's instructions. This method permits the combination of multiple (up to six) DNA fragments irrespective of end compatibility, due to the presence of homologous overlapping sequences a the 5' and 3' ends of the different fragments and the vector. 10 µl reactions consisted of 5 µl HiFi master mix, 2.5 µl DNA fragments, 2.5 µl vector. Fragments for the assembly were created from NEBuilder HiFi DNA assembly tool (http://nebuilder.neb.com).

2.3.2.8. DNA sequencing

Sanger sequencing to confirm correct DNA assembly or confirming chromosomal mutations was carried out by Eurofins Genomics (Munich) with samples prepared as per the service's instructions. Next generation whole genome sequencing was carried out by MicrobesNG (University of Birmingham) using their standard and enhanced services. For the former, genomic DNA samples were prepared as described in 2.3.1.2. and the quality of DNA checked by gel electrophoresis and Qubit analysis. For the latter, the appropriate strain was cultured in 2-4 replicates in 50 ml LB medium in 250 ml baffled Erlenmeyer flasks, and centrifuged at 13000 RPM in 50 ml Falcon tubes. The cell pellet was resuspended in cryopresevant liquid and shipped to MicrobesNG in a 2 ml storage bead vial (Microbank, Pro-lab) for genomic extraction.

2.4. Phenotypic analysis

2.4.1. Optical density measurements

For *E. coli, C. necator, O. carboxidovorans, A. tumefaciens* and *P.* cws, optical density at 600 nm (OD_{600}) was used as a measure of cell growth, which was

determined using a 1-cm gap plastic cuvette (Sigma-Aldrich) and a Thermo Scientific Biomate 3 spectrophotometer. Sterile liquid medium was used as a blank and samples with an OD_{600} above 0.8 were diluted 1:10 and re-measured.

2.4.2. Measurement of *rfp* reporter genes (for mRFP1) expression

For *E. coli*, *C. necator* and *O. carboxidovorans,* promoter activity was measured using a red fluorescence protein-containing biosensor plasmid and an Infinite® M1000 PRO micro plate reader (Tecan, Switzerland). 180 µl of each sample was pipetted into each well of a black 96-well plate (Greiner Bio One International, Germany; Cat. no 655090) and sterile liquid medium was used as a blank in each row used. Fluorescence was measured using 585 nm as excitation wavelength and an emission wavelength of 620 nm. The gain factor for fluorescence was set manually to 80%. The OD_{600} of each sample, along with an MM blank and plasmid-less control were measured using the micro plate reader to normalise fluorescence by optical density. The normalised fluorescence values of the MM blank and plasmid-less control were taken away from each measured sample value to give normalised absolute fluorescence.

2.4.3. Liquid chromatography mass spectrometry (LC-MS) analysis

For *E. coli*, *O. carboxidovorans*, *A. tumefaciens* and *P.* cws, AHL production was determined by LC-MS using 0.2 µm nitrocellulose filer-sterilised supernatant samples grown to mid exponential phase. 1 ml of filter-sterilised supernatants were spiked with a solution of internal standard (10µl of 10µM d9-C5-AHL). These samples were then solvent extracted three times with 0.5ml aliquots of acidified ethyl acetate 0.1% (v/v) acetic acid in ethyl acetate, and dried under vacuum (Speedvac (Thermo-Scientific^{TM}) centrifugal evaporator at ambient temperature). Dried extracts were stored at -20°C. Prior to analysis, each sample extract was reconstituted in 50 µl of MeOH + 0.1% (v/v) formic acid. The injection volume was 5 µl.

Chromatography was achieved using a Shimadzu series 10AD VP LC system. The column oven was maintained at 50°C. The HPLC Column used was a Phenomenex Gemini C18 column (3.0 µm, 100 x 3.0 mm) with an appropriate guard column. Mobile phase A was 0.1% (v/v) formic acid in water, and mobile phase B 0.1% (v/v) formic acid in methanol. The flow rate throughout the chromatographic separation was 450 µL/min. The binary gradient initially began at 10% B for 1.0 min, increased linearly to

50% B over 0.5 min, then to 99% B over 4.0 min. The composition remained at 99% B for 1.5 min, decreased to 10% B over 0.1 min, and stayed at this composition for 2.9 min. Total run time per sample was 10 min.

The MS system used was an Applied Biosystems Qtrap 4000 hybrid triple-quadrupole linear ion trap mass spectrometer equipped with an electrospray ionisation (ESI) interface. Instrument control, data collection and analysis were conducted using Analyst software. Source parameters were set as: curtain gas: 20.0, ion source potential: 5000 V, temperature: 450 °C, nebulizer gas: 20.0, and auxiliary gas: 15.0. LC-ESI-MS/MS methods were used for the analysis of AHLs in bacterial supernatant extract samples. Analysis was conducted with the MS operating in MRM (multiple reaction monitoring) mode, screening the LC eluent for specific AHLs (unsubstituted, 3-oxo and 3-OH AHLs with even chain lengths from C4-C14).

The work described above was carried out by Dr Nigel Halliday, University of Nottingham.

2.6. *In silico* work and tools

2.6.1. Analysis of raw numerical data

Raw numerical data was analysed using Microsoft Excel 2013 (Microsoft). Graphical data and statistical analysis was presented and calculated using GraphPad Prism 7 (GraphPad software). Statistically significant differences were determined via paired T-test.

2.6.2. Nucleotide and protein sequences

SnapGene (GSL biotech, available at: www.snapgene.com) was used for plasmid construction and visualisation of DNA. Protein and nucleotide sequences were acquired from the National Center for Biotechnology Information (NCBI) and sequences compared using the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.* 1990, Zhang & Madden 1997). Protein alignments were generated using Clustal Omega (Sievers & Higgins 2013) and edited in Jalview (Waterhouse *et al.* 2009).

2.6.3. Genome sequencing analysis

2.6.3.1. Quality of genome sequences

Sequencing of genomic DNA was performed at MicrobesNG (University of Birmingham, UK), where their in house software was used to trim sequencing reads and assess their quality. The 'Standard Whole Genome Service" included Illumina next-generation sequencing with a guaranteed minimum coverage of 30x. For newly isolated (*P.* cws and derivatives) or highly modified strains (*A. tumefaciens* A676 and derivatives) the 'Enhanced Genome Service' was used which combined the short reads obtained by the Illumina-based standard service with the long reads generated by Oxford Nanopore technology.

2.6.3.2. Mapping of Reads

The published genomes of *C. necator* H16, *A. tumefaciens*, *O. carboxidovorans* and *P.* cws was used as a template to map paired end reads to (Pohlmann *et al.* 2006, Slater *et al.* 2008, Volland *et al.* 2011) by using the software CLC Genomics work bench 9 (Qiagen). The following parameters set by the read mapping analysis were used:

Masking mode = no masking
Mismatch cost = 2
Cost of insertions and deletions = linear gap cost
Insertion cost = 3
Deletion cost = 3
Insertion open cost = 6
Insertion extend cost = 1
Deletion open cost = 6
Deletion extend cost = 6
Length fraction = 0.5
Similarity fraction = 0.8
Global alignment = No
Auto-detect paired distances = yes
Non-specific match handling = map randomly

6.6.3.3. Single nucleotide variant calling

The mapped reads were used to call single nucleotide variants (SNVs) by using CLC Genomics Workbench 9 (Qiagen). The following parameters were set and tested by Dr Alex Grosse (University of Nottingham) and used for the Mapping and Variants workflow:

Ploidy = 1
Ignore positions with coverage above = 100,000
Restrict calling to target regions = not set
Ignore broken pairs = yes
Ignore non-specific matches = reads
Minimum coverage = 10
Minimum count = 2
Minimum frequency (%) = 40.0
Base quality filter = yes
Neighbourhood radius = 5
Minimum central quality = 20
Minimum neighbourhood quality = 15
Read direction filter = no
Relative read direction filter = yes
Significance (%) = 1.0
Read position filter = no
Remove pyro-error variants = yes
In homopolymer regions with minimum length = 5
With frequency below = 0.7

Chapter 3: Adaptation of *Cupriavidus necator* to carbon monoxide.

3.1. Introduction

C. necator H16 is a promising candidate for the biological conversion of CO_2 from industrial waste gas into valuable bioproducts, including polymers, chemicals and fuels (Cook & Schlegel 1978, Müller *et al.* 2013, Lu & Yu 2017). Recent studies are exploring the potential of synthesis gas (syngas) as an alternative feedstock for gas fermentation (Drzyzga *et al.* 2015, Steinbüchel *et al.* 2017). This cheap and energy-rich gas contains CO_2 and hydrogen but is mostly comprised of carbon monoxide.

C. necator can be engineered to oxidise CO to CO_2 (Hyeon *et al.* 2015, Heinrich *et al.* 2018) and its membrane-bound [NiFe] hydrogenases show low affinity towards CO and appear to be unaffected by its presence (Van der Linden *et al.* 2004, Bernard *et al.* 1974, Schneider *et al.* 1979). Additionally the cytochromes of *C. necator* show low affinity for CO (Bernard *et al.* 1974). Despite this, at high concentrations (90%), CO has been shown to drastically inhibit the growth of *C. necator* (Cypionka and Meyer 1982). Syngas typically contains 30-60% CO (National Energy Technology Laboratory 2018), a range in which *C. necator* growth is likely to be significantly inhibited.

Numerous aerobic bacteria are highly sensitive to carbon monoxide, their growth and respiration becoming inhibited when cultured with the gas (Cypionka and Meyer 1982). Isolated aerobic CO oxidisers such as *Burkholderia fungorum* LB400, *Mesorhizobium sp*. strain SH 15003, *Stappia sp*. strain CV902-700 & CV812-530 and *Stenotrophomonas maltophila sp*. strain LMG-957 are all inhibited by and are unable to adapt to CO concentrations higher than 1% following incubations lasting several months (King 2003). However, other CO oxidisers can tolerate higher CO concentrations including *Xanthobacter sp.* strain INA43/2-2 (20%), *Bradyrhizobium japoncium* USDA 110 (50%), *Oligotropha carboxidovorans* and *Rhodocyclus gelatinosus* (90%) (Lorite *et al.* 2000, King 2003, Cypionka and Meyer 1982). *Ralstonia eutropha* B5786, a close relative of *C. necator* H16, exhibits CO tolerance but its growth is inhibited by CO concentrations above 20% (Volova *et al.* 2001).

The anaerobic CO utilising bacteria can tolerate the gas at high concentrations and utilise it for growth and production of industrially relevant products such as butyrate, acetate, butanol and methane (Vega *et al.* 1989, Grethlein *et al.* 1991, Luo *et al.* 2013). High tolerance to CO has been documented in carboxidotrophs such as *Butyribacterium methylotrophicum* (Lynd *et al.* 1982) and in various species of *Carboxydocella* (Sokolova *et al.* 2002, Sokolova *et al.* 2005, Slepova *et al.* 2006). Adaptation to growth in the presence of CO has been observed in several different carboxyodotrophs, a process which can take a period of weeks to months (Lynd *et al.* 1982, O'Brien *et al.* 1984, Oelgeschläger & Rother 2008, Krajmalnik-Brown *et al.* 2017). *B. methylotrophicum* is only able to grow with CO following adaptation by successive transfer with increasing CO concentrations (Shen *et al.* 1999). Growth of wild type *B. methylotrophicum* is completely inhibited in presence of 16% CO but is uninhibited in its CO-adapted form due to the altered regulation of ferredoxin-NAD oxidoreductase (Shen *et al.* 1999).

For *C. necator* to be efficiently cultured with syngas or any alternative feedstock containing high concentrations of CO, adaptation to the toxicity of CO is essential. Industrial organisms used are typically produced through evolutionary selection (Debabov 2015, Lütke-Eversloh & Bahl, 2011). Whether tolerance of *C. necator* to CO can be improved through adaptive evolution has yet to be investigated. Through repeated exposure to CO over time, a selective pressure will be provided, reshaping the genome to give genetically adapted strains of *C. necator* a growth advantage.

3.1.1. Aims of this study

The overall purpose of this study was to isolate and characterise mutant strains of *C. necator* that have adapted to growth with CO. Specific objectives were:

- Developing a CO adaptation regime based on the continuous subculturing of *C. necator* in batch culture on minimal media containing an additional carbon source of fructose.
- Identification and characterisation of potential CO-adapted strains by comparing their growth kinetics (growth rates, final optical densities, length of lag phase) to the original wild type "mother culture".

- Further characterisation of evolved candidate strains by whole genome sequencing to identify mutational hotspots and determine whether mutations have occurred in any genes of interest, which may be involved in CO regulation or metabolism.
- To re-introduce identified mutations in the original wild type background to establish if the adapted growth phenotype could be recreated and to generate a fully defined CO-resistant strain.

3.2. Results

3.2.1. Growth profiles with increasing CO concentrations

To determine the effect CO has on growth, *C. necator* was grown with increasing concentrations of CO (0%, 12%, 25%, 50%) together with air in the headspace. Pre-cultures were grown in 5 ml fructose minimal media (F-MM) overnights and grown to OD_{600}= 2. This was injected into serum bottles under 2-bar pressure, containing 20 ml F-MM (0.4%). The bottle headspaces comprised of 50% air with either 50% CO, 25% CO and 25% N_2 or 12.5% CO and 37.5% N_2 or 0% CO, 50% N_2. The experiment was initiated with a starting OD_{600} of 0.05.

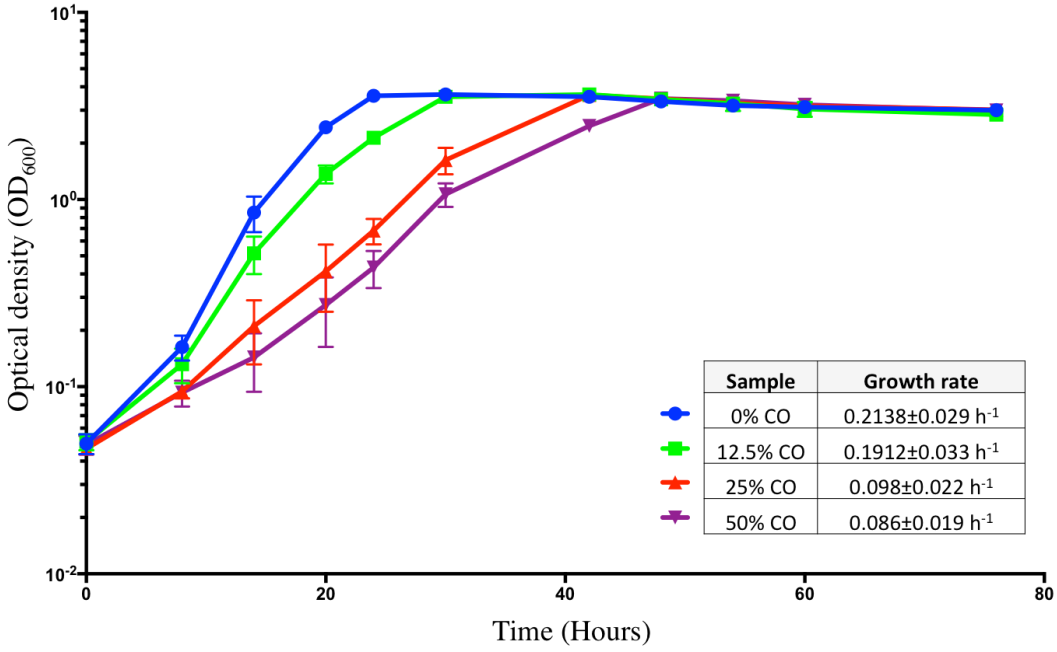


Figure 3.1. *C. necator* H16 growth with increasing concentrations of CO.

C. necator was grown at 30 °C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% N_2 / 50% air (blue circles), 12.5% CO / 37.5% N_2 / 50% air (green squares), 25% CO / 25% N_2 / 50% air (red triangles) and 50% CO / 50% air (purple inverted triangles), respectively (all bottles with 2 bar total pressure). Error bars represent the standard deviation of the mean for three independent replicates. Standard deviation (SD) between growth rates for each growth condition is given.

The growth curves followed a time course of 76 hours, starting with exponential growth and transitioning to stationary phase between 20 and 50 hours (Fig. 3.1). With 0% CO *C. necator* reached stationary phase after 24 hours but with 12.5% CO, the lag phase was increased, with cultures taking 30 hours to reach stationary phase. The growth rate decreased but this was not statistically significant (p = 0.428). This effect was seen to

an even greater extent at 25% and 50% CO, with *C. necator* taking 42 hours and 50 hours, respectively, to reach stationary phase and both conditions significantly reducing the growth rate ($p = 0.005$ and $p = 0.003$, respectively) (Fig. 3.1). This demonstrates that CO inhibits the growth of *C. necator* but can be tolerated to some degree at relatively high concentrations.

3.2.2. Continuous subculturing of *C. necator* with CO

The ideal conditions for evolving *C. necator* during continuous aerobic subculturing would involve an artificial syngas mixture containing $CO_2 + H_2$ and high concentrations of CO, since this gas mixture would be more industrially relevant and because hydrogenases and respiratory chain components are main targets of CO inhibition (Schuchmann *et al.* 2018, Cypionka and Meyer 1982). This approach, however, poses certain limitations including the potentially explosive nature of the gas mixture. It would also be more experimentally demanding given the increased time *C. necator* takes to grow chemolithoautotrophically and potential difficulties in reproducibly generating such complex gas mixtures. Importantly, the hydrogenases of *C. necator* have been shown to be highly CO resistant (Schneider et al, 1979, Van der Linden *et al.* 2004), hence further hydrogenase adaptation seems unlikely, so there would be little benefit for employing the above conditions. Hence, fructose was used as a source of carbon and energy. Since the decrease in growth rate of *C. necator* between 25% CO and 50% CO was not significant ($p > 0.05$), it was assumed that no further significant decrease would be seen above 50%. Therefore, to observe whether *C. necator* can adapt to growth in the presence of high CO concentrations, wild type cells were maintained in serum bottles with 20 ml F-MM and 50% CO / 50% air through continuous subculturing. Subculturing was carried out by allowing cultures to grow to an OD_{600} of approximately 3, then diluting them back to an OD_{600} of 0.01 every 3 days. The culture was plated out on F-MM agar every two weeks to assess the purity of the culture.

3.2.2.1. Growth comparison after one month of subculturing

Following one month of subculturing the selective pressure of the CO was removed, by plating out the culture on F-MM agar. Single colonies were screened by PCR using primers for the *hsdRN* gene (H16_A0006) to confirm the presence of *C. necator*. Three separate colonies were picked together with three of the unevolved wild type and grown in 5 ml F-MM to an OD_{600} of approximately 2. These were diluted into serum bottles

containing 20 ml F-MM and 50% CO and 50% air (v/v), to obtain a starting OD_{600} of 0.05.

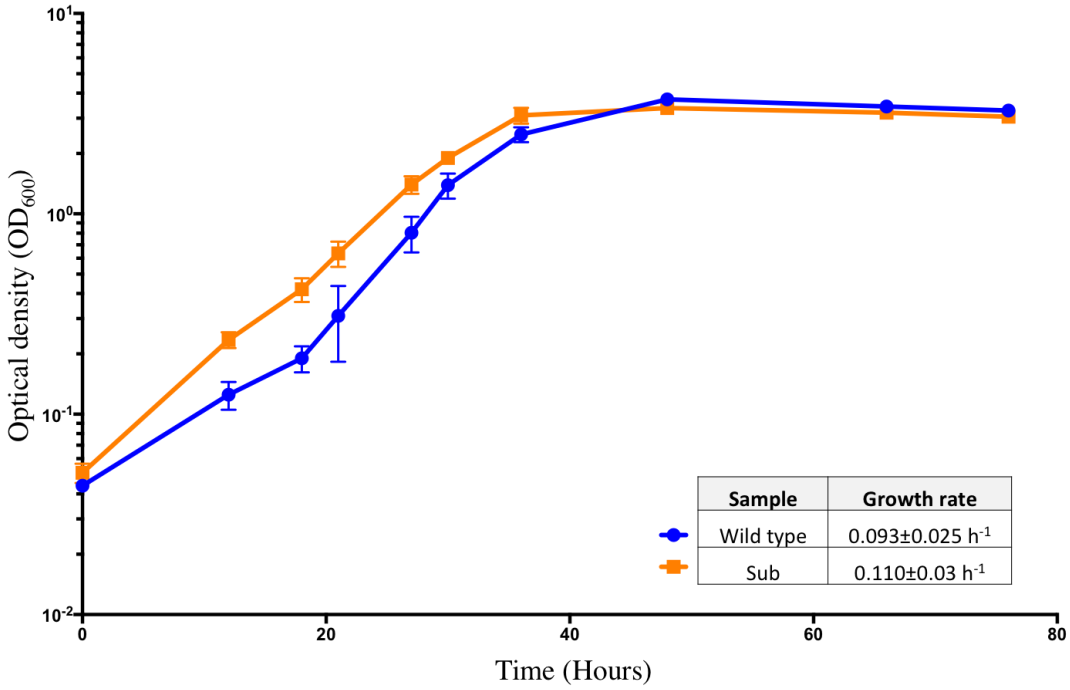


Figure 3.2. Growth of *C. necator* Sub and its parent H16 strain in the presence of 50% CO.
C. necator Sub (orange squares,) the parent H16 wild type (blue triangles) in 20 ml F-MM under an atmosphere of 50% CO/50% air (v/v; 2 bar total pressure). Single colonies were used for inoculation. Error bars represent the standard deviation of the mean for three independent replicates.

The growth curves followed a time course of 76 hours with both the wild type and subcultured strain (*C. necator* Sub) reaching peak OD_{600} after 48 hours (Fig. 3.2). *C. necator* Sub did show a slight increase in growth rate (0.110±0.025 h^{-1}) in comparison to the wild type (0.093±0.030 h^{-1}), however this was not statistically significant (p = 0.603).

3.2.2.2. Growth comparison after 8 months of subculturing

Since *C. necator* only displayed a slight improvement in CO resistance after 1 month, the period of subculturing was extended to eight months. After 8 months, the selective pressure of the CO was removed, by plating out the culture on F-MM agar. Single colonies were confirmed to be *C. necator* H16 by colony PCR. Three separate colonies were picked together with three of the unevolved wild type and grown in 5 ml F-MM to an OD_{600} of approximately 2. These were diluted into serum bottles containing 20 ml F-MM and 50% CO and 50% air (v/v), to obtain a starting OD_{600} of 0.05. If taken as technical replicates, the three isolates from the evolved strain (termed *C. necator*

Evo) displayed a significantly higher growth rate (0.234±0.009 h^{-1}; p = 0.0017) than the wild type strain (0.075±0.004 h^{-1}), each taking 30 hours to reach peak OD_{600} whereas the wild type strain took 48 hours (Fig. 3.3). However, the isolates were later found to differ from each other genotypically (3.2.2.3.; Table 1) and were therefore considered different strains. This indicated that the evolved culture had a clear growth advantage in the presence of CO when compared to wild type and was no longer genetically uniform.

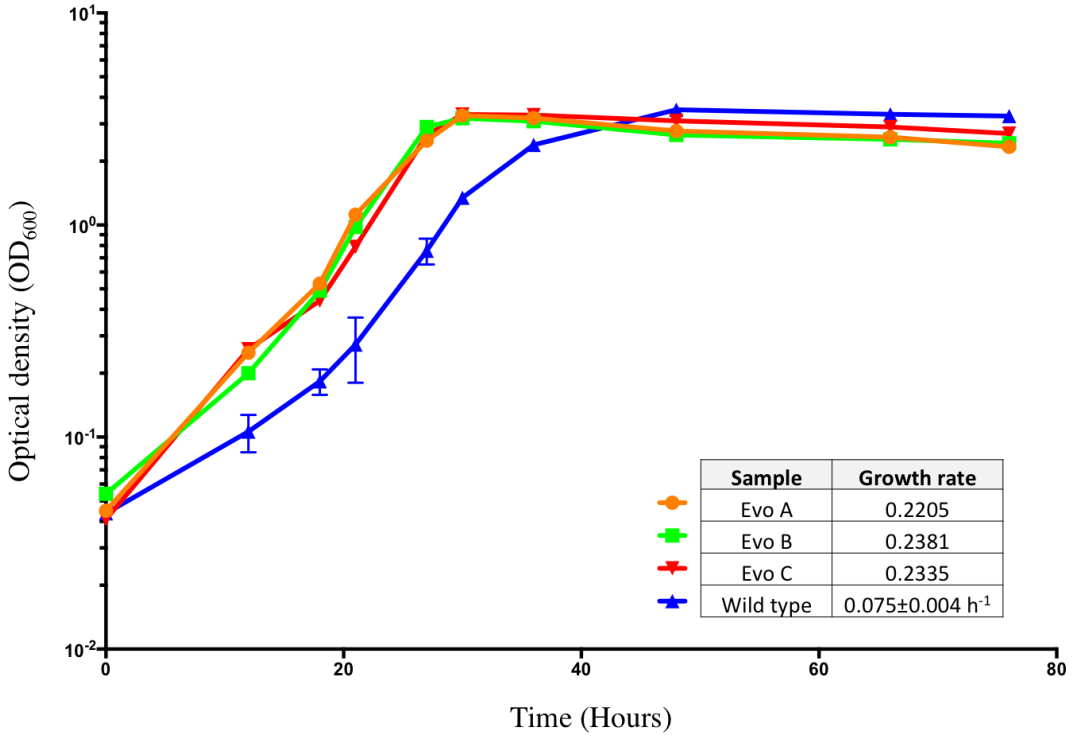


Figure 3.3. Growth of *C. necator* Evo and its parent H16 strain in the presence of 50% CO/50 air. Three distinct colonies obtained from the Evo culture were used for inoculation: *C. necator* Evo A (orange circles), *C. necator* Evo B (green squares) and *C. necator* Evo C (red inverted triangles). These three isolates and the parent H16 wild type (blue triangles) were grown in 20 ml F-MM under an atmosphere of 50% CO/50% air (v/v; 2 bar total pressure). The inset shows growth using a logarithmic scale. Error bars represent the standard deviation of the mean for three independent experiments.

3.2.2.3. Whole genome sequencing

Illumina next generation sequencing was performed following growth comparisons between the three evolved isolates and wild type cultures. This was carried out to compare the genomes of the evolved strain to the to the original wild type genome and identify mutations that have occurred which may confer tolerance to CO.

The genomic DNA of *C. necator* Evo isolates Evo A, Evo B, Evo C and the H16 parent strain was isolated following pre-culturing (3.2.2.1) and sent for next generation Illumina sequencing as described in Materials & Methods (2.6.3). The obtained raw data were processed using the software CLC Genomics Workbench 11 by, first performing read mapping of all samples onto the published *C. necator* H16 reference genome. The 'Resequencing Analysis' function was used to compare variants between the evolved strains and the parental H16 strain (Pohlmann *et al.* 2006), by using the 'Compare Variants' and 'Compare Sample Variant Tracks' functions. Any mutations found in the isolates and parental strain but absent in the published genome were discarded as these were assumed to be the common starting mutations present in the ancestral strain.

This identified several unique SNPs that were only present in the evolved strains (Table 3.1). A number of these were common between the three individual isolates including a point mutation (G→A) upstream of the cytochrome *bd* ubiquinol oxidase operon, a point mutation (G→C) in the iron transporter *feoA* and an insertion (C) in a pseudogene. A 12 base pair deletion in *phcR*, a point mutation (A→G) in the gene for cytochrome *c* biogenesis protein, ResB and a point mutation (C→G) in an intergenic region were detected in two of the three strains (Table 3.1). Various other mutations appeared only once in two different isolates (2 & 3), including two point mutations in the TonB-dependent receptor gene (C→G & G→T), a point mutation (A→C) in cobalamin biosynthesis gene, encoding CobW, and a point mutation (G→A) in the β-ketoacyl-ACP reductase gene.

Table 3.1. Comparison of mutations across three isolated strains of *C. necator* evolved in the presence of CO.

H16_A = chromosome 1 and H16_B = chromosome 2.

Locus	Mutation	Amino Acid change	EvoA	EvoB	EvoC	Region	Gene region/encoded protein
H16_B1461	G → A	_	Yes	Yes	Yes	1638520	Intergenic region upstream of cytochrome ubiquinol oxidase operon
H16_B0085	G → C	alanine → glycine	Yes	Yes	Yes	98474	Iron transporter FeoA
H16_A2071	insertion: C	frame shift	Yes	Yes	Yes	2253191	Pseudogene
H16_A3118	Deletion: GCGCTCGCGGTG	_	/	Yes	Yes	3376140	Hybrid sensor histidine kinase/response regulator PhcR
H16_A3450	A → G	serine → proline	Yes	/	Yes	3730714	Cytochrome C biogenesis protein ResB
H16_A3140	C → G	_	Yes	Yes	/	3400456	Intergenic
H16_B2109	G → T	glutamine → lysine	/	Yes	/	125022	TonB-dependent receptor
H16_B2109	G → C	asparagine → lysine	/	Yes	/	125377	TonB-dependent receptor
H16_A0195	A → C	histidine → glutamine	/	Yes	/	208932	Cobalamin biosynthesis CobW
H16_B0167	G → T	_	/	Yes	/	1522049	Intergenic
H16_A1439	G → A	glycine → serine	/	Yes	/	1560730	β–ketoacyl-ACP reductase
H16_A1847	A → G	_	/	Yes	/	2003025	Intergenic
H16_A1846	T →G	isoleucine → serine	/	/	Yes	2001162	Hypothetical protein
H16_A3734	T → G	isoleucine → serine	/	/	Yes	4034982	Hypothetical protein

3.2.3. Repeated subculturing with ten individual *C. necator* culture lineages

The subculturing experiment described in 3.2.2 did not allow for the accurate estimation of when the *C. necator* evolved the fast growing phenotype. The experiment required repetition with a larger sample size to observe which of the identified regions mutated again and also to observe any potential mutations in new genes. Closer growth monitoring was required to determine the time at which the cells evolved increased CO tolerance and whether this occurred gradually or relatively quickly. The second subculturing experiment started with a single colony of *C. necator* H16, which was cultured in 5 ml F-MM. The genomic DNA of this culture was extracted and ten separate pre-cultures in F-MM were created. These were used to inoculate ten serum bottles containing 20 ml F-MM and 50% CO/50% air (v/v) to an OD_{600} of 0.05 to create ten separate evolving lineages of *C. necator* exposed to CO stress. Following growth for 48 hours, after reaching an OD_{600} of about 2, each culture was propagated by inoculating a fresh serum bottle under the same conditions. This continued until a significant visible change in OD appeared within 12 hours of re-inoculation, which would signify that the cells had adapted to CO. This was observed after the 15^{th} round of re-inoculation, at which point the cultures, now designated Evo1 to Evo10, were stored at -80^{o}C with 50% glycerol and inoculated into fresh serum bottles for further analysis.

3.2.3.1 Growth profiles of evolved cultures Evo1 to Evo10 in the presence of 50% CO

To confirm that the fast growing *C. necator* cultures Evo1 to Evo10 had adapted to CO, the growth of all ten cell lineages was compared to the original wild type culture (Fig. 3.4). Three wild type colonies were pre-cultured in 5ml F-MM and split into three serum bottles under an atmosphere of 50% CO/50% air concurrently with the ten evolved cultures and allowed to grow to OD=~2. All serum bottle pre-cultures were then used to inoculate fresh bottles containing 20 ml F-MM and 50% CO / 50% air (v/v) to a starting OD_{600} of 0.05. The growth curves followed a time course of 90 hours, with the evolved cultures displaying higher growth rates than the wild type (Table 3.2). All evolved cultures reached their highest OD_{600} after 25-30 hours and the wild type after 52 hours (Fig. 3.4).

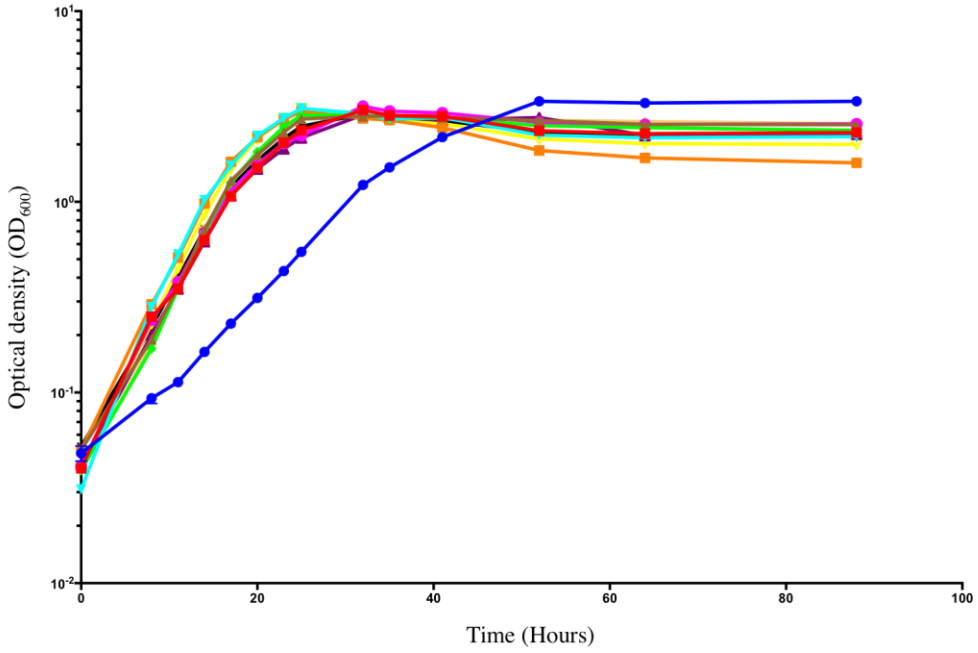


Figure 3.4. Growth of evolved *C. necator* cultures Evo1 to Evo10 on CO compared to the wild type.

C. necator Evo1 (red squares), Evo2 (brown triangles), Evo3 (cyan inverted triangles), Evo4 (green diamonds), Evo5 (pink circles), Evo6 (orange squares), Evo7 (purple triangles), Evo8 (yellow inverted triangles), Evo9 (orange diamonds), Evo10 (black stars) and wild type H16 (blue circles) were grown in F-MM under an atmosphere of 50% CO/50% air (v/v; 2 bar total pressure). For wild type H16, error bars represent the standard deviation of the mean for three independent replicates. Single replicates were used for the evolved cultures.

Table 3.2. Growth rates of evolved cultures and evolved single isolates

Lineage	Evo cultures	Evo isolates
1	0.186	0.201
2	0.213	0.1952
3	0.179	0.177
4	0.214	0.206
5	0.185	0.174
6	0.229	0.251
7	0.188	0.182
8	0.199	0.173
9	0.208	0.188
10	0.183	0.171
Average	0.198±0.017	0.192±0.024
Wild type	0.113±0.034	0.097±0.019

3.2.3.2. Growth profiles of defined isolates obtained from cultures Evo1 to Evo10 in the presence of 50% CO / 50% air

The evolved lineages are potentially comprised of mixed populations of cells, whereby different subpopulations may have displayed differing rates of growth. To address this possibility, the evolved cultures were plated out on F-MM agar plates and single colonies were isolated to represent defined strains designated E1 to E10, according to their origin (Evo1 to Evo10). Pre-cultures were made from these colonies alongside wild type colonies in 5 ml F-MM. These were then subcultured in 20 ml F-MM in serum bottles with 50% CO/50% air (v/v) and grown to an OD_{600} of about 2. Using these cultures as an inoculum under the same conditions, the same growth comparison previously described for total cultures Evo1 to 10 was carried out with similar results, with the experiment starting with an OD_{600} of 0.05.

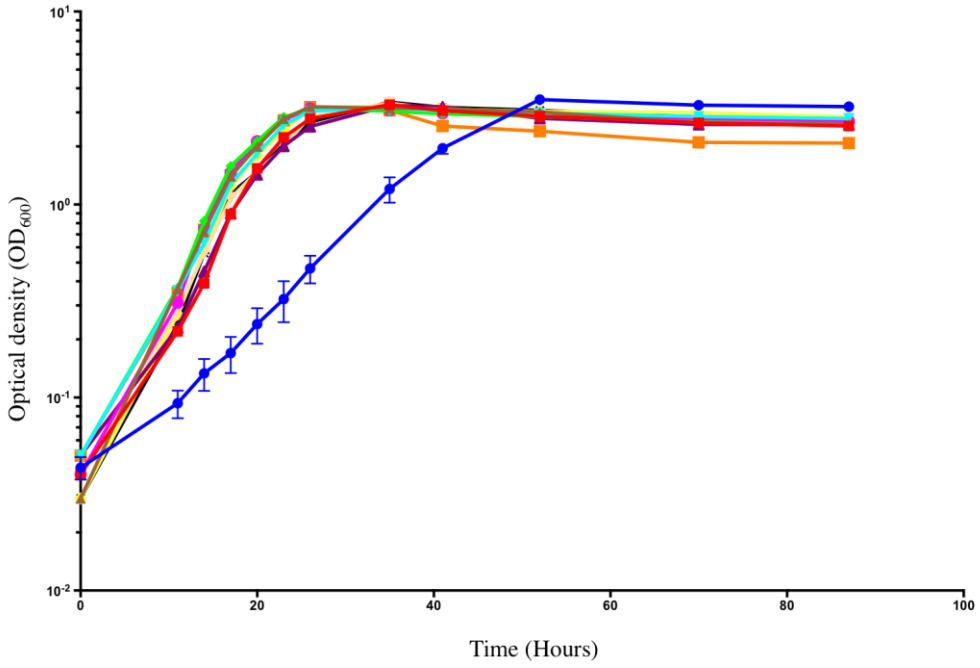


Figure 3.5. Growth of evolved *C. necator* isolates E1 to E10 on CO compared to wild type. *C. necator*.

Evo1 (red squares), Evo2 (brown triangles), Evo3 (cyan inverted triangles), Evo4 (green diamonds), Evo5 (pink circles), Evo6 (orange squares), Evo7 (purple triangles), Evo8 (yellow inverted triangles), Evo9 (orange diamonds), Evo10 (black stars) displayed faster growth than the wild type (blue circles) when grown in F-MM under an atmosphere of 50% CO/50% air (v/v; 2 bar total pressure). For wild type H16, error bars represent the standard deviation of the mean for three independent replicates. Single replicates were used for the evolved isolates.

The growth curves followed a time course of 90 hours, with the evolved strains reaching peak OD_{600} after 25-30 hours and the wild type controls after 53 hours (Fig. 3.5). The

similarity of E1 to E10 growth curves to those obtained for the respective Evo 1 to Evo10 parent cultures suggested that the evolved cultures are either dominated by the obtained isolates or comprised of a mixed population of strains with very similar fitness. There was a statistically significant difference between the growth rates of the evolved strains and the wild type replicates ($p < 0.01$ for all strains; Table 3.2). This indicated that the evolved strains displayed a clear growth advantage over the wild type strain. Ideally this experiment would be performed with technical replicates so that statistical significance could be calculated for each of the evolved isolates.

3.2.3.3. Whole genome sequencing

Illumina next generation sequencing was performed as previously described (3.2.2.3) following growth comparisons between the evolved isolates E1 to E10 and wild type cultures. For this, chromosomal DNA was isolated from each evolved isolated strain as well as the initial, unevolved parent strain culture to be certain that identified differences in the former, when compared to the published H16 genome, were not already present in the latter. Analysis of the obtained sequences using CLC Genomics Workbench 11 was carried out as described above. This identified several unique mutations that were only present in the evolved strains E1 to E10, several of which were also present in the evolved Iso1 to Iso3 strains analysed in 3.2.2.3. Interestingly, a single point mutation upstream of the cytochrome *bd* ubiquinol oxidase operon (G→A) was detected in all ten isolates (Table 3.3). A 12 bp deletion in the Hybrid sensor histidine kinase/response regulator, PhcR was found in 5 of the isolates (Iso2, E3, E4, E5 and E9) (Table 3.3). The deletion removed the coding sequence for the peptide sequence Ala-Leu-Ala-Val from this gene. Two point mutations (both A→C) were found within the PhcS coding sequence and upstream of it, respectively, in two separate isolates (E6 and E8). The SNP within the PhcA gene changed the coding sequence from leucine to arginine (Table 3.3). Two point mutations were found within the RNA polymerase sigma factor 54 (*rpoN*) gene. A G→T transversion converted a serine codon into a stop codon and was found in two of the isolates (E6 and E7) and a T→G transversion changed a glutamine codon into a proline codon and was only seen on one occasion (Iso1). An A →T transversion was found in a sigma factor 54-dependent Fis family transcriptional regulator gene, converting isoleucine to phenylalanine (E4).

Various other mutations were identified across the ten isolates, four of which were located within 1.5 kb of each other, in the intergenic regions upstream and downstream of a transposase gene (H16_A3141). The upstream mutations (deletion: G and insertion: T) occurred within 4 base pairs of each other. The deletion occurs twice (E8 and E9) and the insertion once (E5). The downstream mutations (A→G and G→A) occurred within 3 base pairs of each other, both occurring twice in the same isolates (E8 and E10). A lysine to arginine mutation (A→G) in a membrane protein was seen in two strains (E3 and E10) and a frame shift mutation (deletion of a G) occurred in a putative pseudogene in one strain (E7).

3.2.3.4. Analysis of mutations

From the mutations described, several appear to be the most likely candidates that may contribute to CO resistance. Cytochrome *bd* ubiquinol oxidase II (encoded by *cydA2B2*) is one of two cytochrome *bd* terminal oxidases (the other encoded by *cydA1B1*, H16_B1177) of the *C. necator* electron transport chain, which may be resistant to CO. The mutation upstream of *cydA2B2* (G→A) may alter expression of the *cydA2B2* gene.

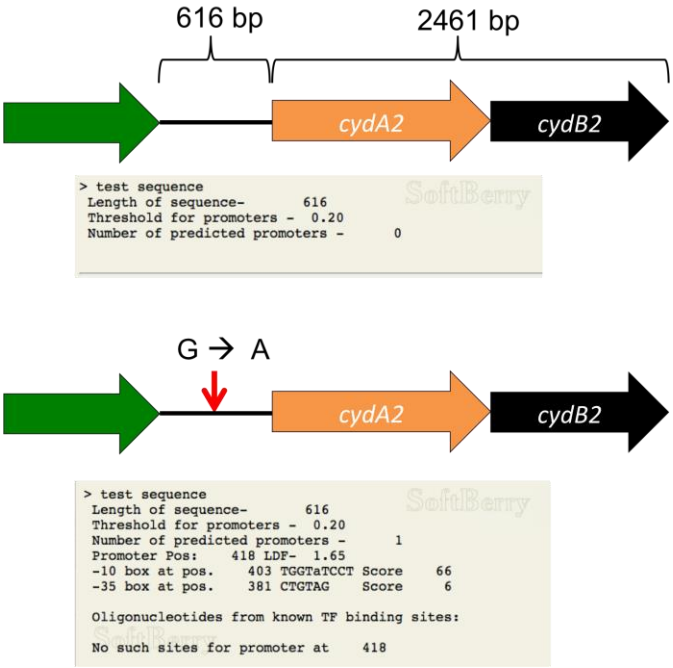


Figure 3.6. Promoter prediction for G→A mutation upstream of *cydA2B2*.

The BPROM programme (http://www.softberry.com) does not predict a promoter upstream of *cydA2B2* in the wild type sequence (top). Following the G→A change, a promoter is predicted with an extended -10 box (TGGTATCCT) score of 66 and a -35 box (CTGTAG) score of 6 (bottom). Linear discriminant function (LDF) score was 1.65. This is a measure of the promoter strength with a promoter is considered active with an LDF score of at least 0.2 (Hücker *et al.* 2017).

Table 3.3. Comparison of mutations across all ten isolated strains of *C. necator* evolved with CO.

H16_A = chromosome 1 and H16_B = chromosome 2.

Locus	Mutation	Amino acid change	Iso 1	Iso 2	Iso 3	Iso 4	Iso 5	Iso 6	Iso 7	Iso 8	Iso 9	Iso 10	Region	Gene/encoded protein
H16_B1461	G -> A	_	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	1638520	Intergenic region upstream of *cydA2B2*
H16_A3118	Deletion: GCG CTCGCGGTG	Ala-Leu-Ala-Val sequence deleted	_	Yes	Yes	Yes	Yes	/	/	/	Yes	/	3376140	Hybrid sensor histidine kinase/response regulator PhcR
H16_A3140	Deletion: G	_	/	/	/	/	/	/	/	Yes	Yes	/	3399139	Intergenic
	Insertion: T	_	/	/	/	/	Yes	/	/	/	/	/	3399143	Intergenic
H16_A3141	A -> G	_	/	/	/	/	/	/	/	Yes	/	Yes	3400457	Intergenic
	G -> A	_	/	/	/	/	/	/	/	Yes	/	Yes	3400460	Intergenic
H16_A3144	A -> C	Leucine → Arginine	/	/	/	/	/	/	/	Yes	/	/	3403057	LysR family transcriptional regulator PhcA
	A -> C	_	/	/	/	/		Yes		/	/	/	3403900	Intergenic region upstream of PhcA gene
H16_A0387	G -> T	Serine → Stop codon	/	/	/	/	/	Yes	Yes	/	/	/	404574	RNA polymerase sigma-54 factor
	T -> G	Glutamine → Proline	Yes	/	/	/	/	/	/	/	/	/	405744	RNA polymerase sigma-54 factor
H16_B0573	A -> G	Lysine → Arginine	/	/	Yes	/	/	/	/	/	/	Yes	645528	Membrane Protein
H16_A0146	Deletion: G	Glycine frame shift	/	/	/	/	/	/	Yes	/	/	/	159883	Pseudogene
H16_pHG19	A -> T	Isoleucine → phenylalanine	/	/	/	Yes	/	/	/	/	/	/	16093	Sigma-54-dependant Fis family transcriptional regulator

To test this idea the BPROM software was used which has been designed to predict bacterial sigma-70 promoter sequences (www.softberry.com). Indeed, BPROM predicted the generation of a promoter sequence when this single base is changed, supporting this hypothesis (Fig 3.6). PhcA and PhcR are the transcriptional activator and regulator of the Phc system. This functions as a putative quorum sensing system, where PhcR regulates activity of PhcA (Garg *et al.* 2002). This system could be influenced by CO to improve the growth of *C. necator* when these genes contain the mutations described in Table 3.3 (12 bp deletion in PhcR, A→C in frame/upstream of PhcA). Genes under the control of RNA polymerase sigma-54 factor could also be affected by the presence of CO when either of the mutations in this gene are present (G→T or T→G).

These six mutations were therefore chosen for characterisation in *C. necator* (Table 3.4). Before embarking on a detailed analysis of the indicated loci and their potential involvement in CO resistance, it was necessary to confirm the existence of the proposed mutations by independent means. This was achieved through amplification of the gene regions carrying these SNPs and Indels by PCR and subsequent Sanger sequencing which confirmed their presence in all cases/all cases tested. Each mutation was confirmed by carrying out PCR using the primers shown below (3.2.4.2, Table 3.4) on an isolate harbouring a specific mutation, with Sanger sequencing being performed on each PCR fragment. A mutant name was given for the strain that would be generated from each mutation. A shortened version of the locus tag was given for each mutation found in the RNA polymerase sigma-54 factor gene to distinguish them. For the *phcA* mutations, IF and US were used to indicate whether the mutation was located within the open reading frame or upstream of it, respectively (Table 3.4). The other mutations found across the ten isolates were deemed unlikely to be contributing significantly towards carbon monoxide resistance due to their location and observed frequency.

3.2.4. Replication of the evolved CO resistant phenotype through targeted mutation of strain H16

To test whether the six chosen mutations increase CO tolerance in *C. necator*, it was decided to construct a series of knock-out / knock-in plasmids to artificially recreate the individual mutations in the wild type genome (Fig. 3.9 & Fig. 3.10). Growth of the resulting mutant strains was then compared to the H16 wild type they were generated

in. To achieve this, the following strategy was employed. First, a deletion of 200 bp was introduced into the H16 genome at the position of the respective mutation. This deletion would then be repaired with a sequence containing the given mutation but otherwise identical to the wild type (Fig. 3.7). Deletions and integrations were screened using the same primer pair for each mutation site (Table 3.5).

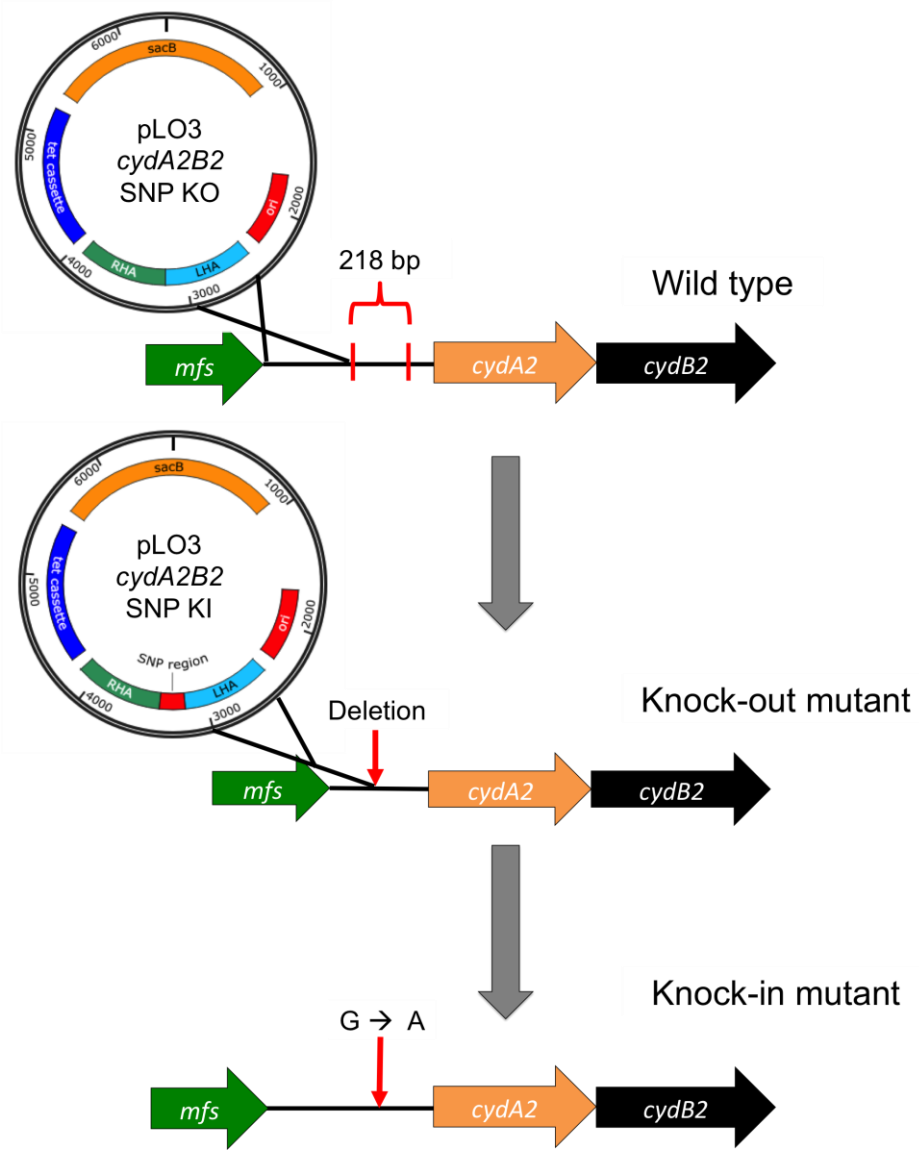


Figure 3.7. Generation of knock-out / knock-in workflow.

The workflow was the same for all mutants and is exemplified here for the *cydA2B2* upstream mutation. The left (LHA) and right homology arms (RHA) of pLO3_*cydA2B2* SNP KO recombine with the corresponding chromosomal regions (top), resulting first in a single and then a double crossover, following tet selection and *sacB* counter selection, respectively. As a result 218 bp upstream of *cydA2B2* are deleted. Following conjugation, the knock-in plasmid pLO3_*cydA2B2* SNP KI is introduced into the knock-out strain and either the left or right homology arm recombine next to the deletion site (middle). This resulting single and subsequent double crossover reintroduce the deleted region into the chromosome with the altered base pair (G to A on the top strand) (bottom).

3.2.4.1. Knock-out/knock-in plasmid construction

Homology arms (750-850 bp in length) upstream and downstream of the six deletion sites were amplified from the *C. necator* genome with primers containing overlapping regions, as described in the Method (2.4.1).

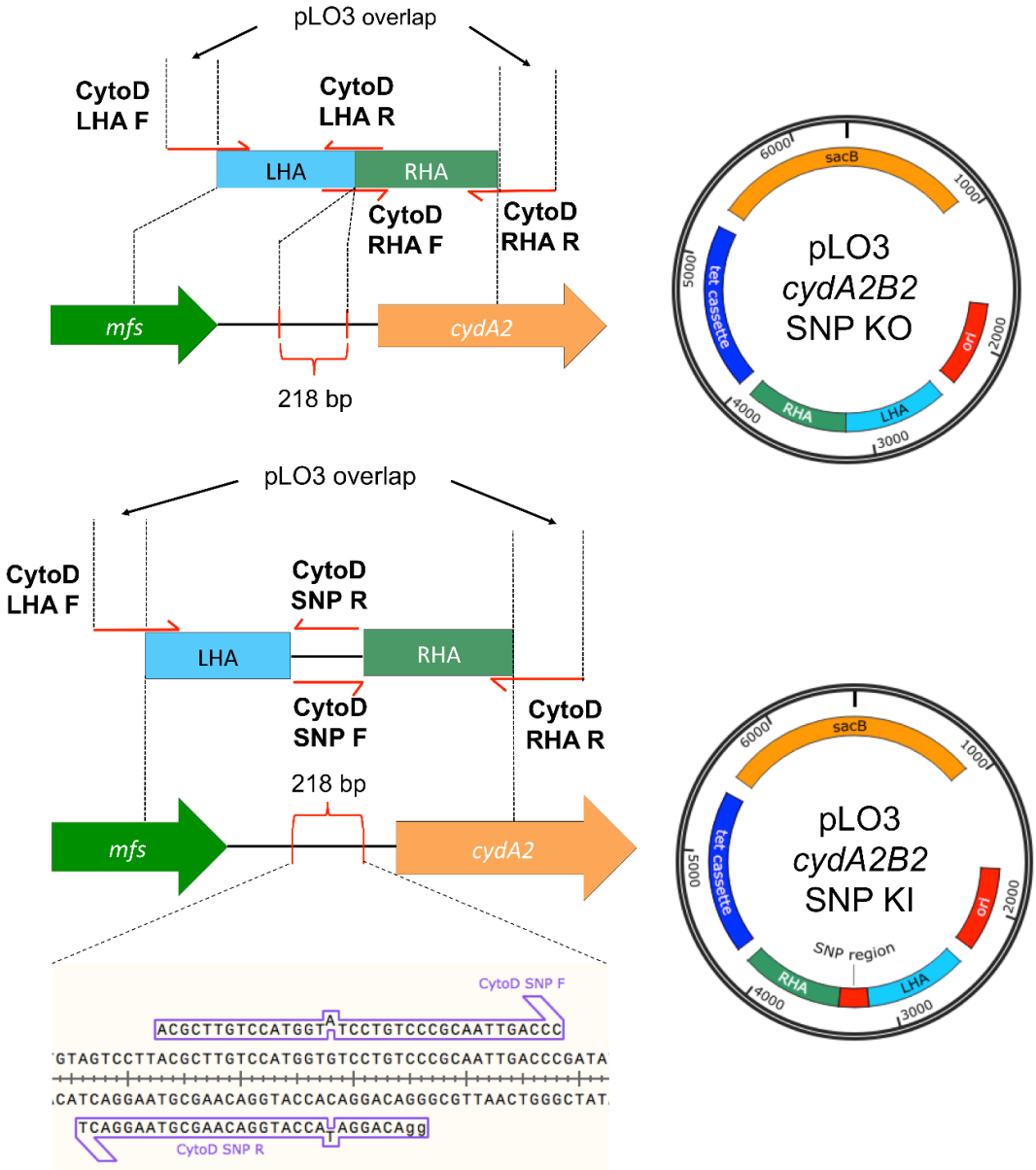


Figure 3.8. Generation of knock-out and knock-in plasmids exemplified for *cydA2B2* upstream region.

Primer pairs used to generate LHA and RHA either side of the 218 bp deletion site for insertion into pLO3 to create the knock-out plasmid pLO3_*cydA2B2* SNP KO (top). Primer pairs used to generate LHA and RHA which include the altered SNP region (bottom). CytoD SNP F (containing A instead of G) and CytoD RHA R used to amplify right homology arm and CytoD LHA F and CytoD SNP R (containing T instead of C) used to amplify left homology arm, which are assembled together with pLO3 by Hifi DNA assembly to create pLO3_*cydA2B2* SNP (Fig. 3.7).

The left homology arm (LHA) forward primer contained base pairs overlapping with pLO3 at the chosen restriction site and the LHA reverse primer contained base pairs overlapping with right homology arm (RHA) forward primer. The RHA forward primer contained base pairs overlapping with the LHA reverse primer and the RHA reverse primer contained base pairs overlapping with pLO3 at the second restriction site.

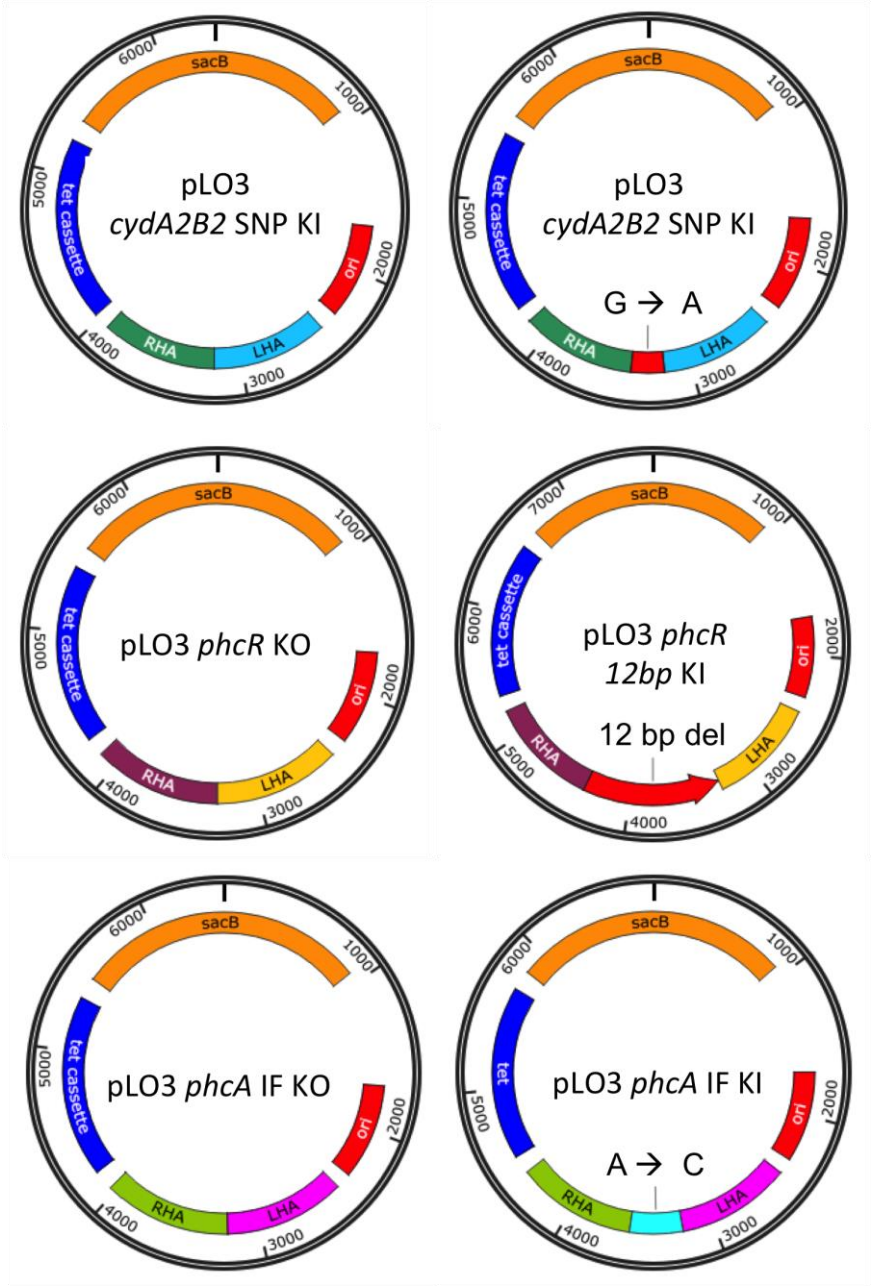


Figure 3.9. Plasmids used for the generation of knock-out/knock-in mutant library (A).
Knock-out and knock-in plasmids constructed for *cydA2B2* upstream SNP generation (top), *phcR* 12 bp deletion generation (middle) and *phcA* open reading frame SNP generation (bottom). All knock-in plasmids contained the same homology arms as their corresponding knock-out plasmid but also contained the deletion region of DNA with the desired mutation.

The LHA and RHA PCR fragments could therefore be used in a HiFi-DNA assembly reaction with restriction-digested pLO3, to construct the respective knock-out plasmid. The LHA reverse and RHA forward primers for all knock-in plasmids contained the altered single nucleotide or deletion to generate the desired mutation but otherwise covered the entire native sequence (Fig. 3.8). The generated knock-out and knock-in plasmids are shown in Fig. 9 and Fig.10.

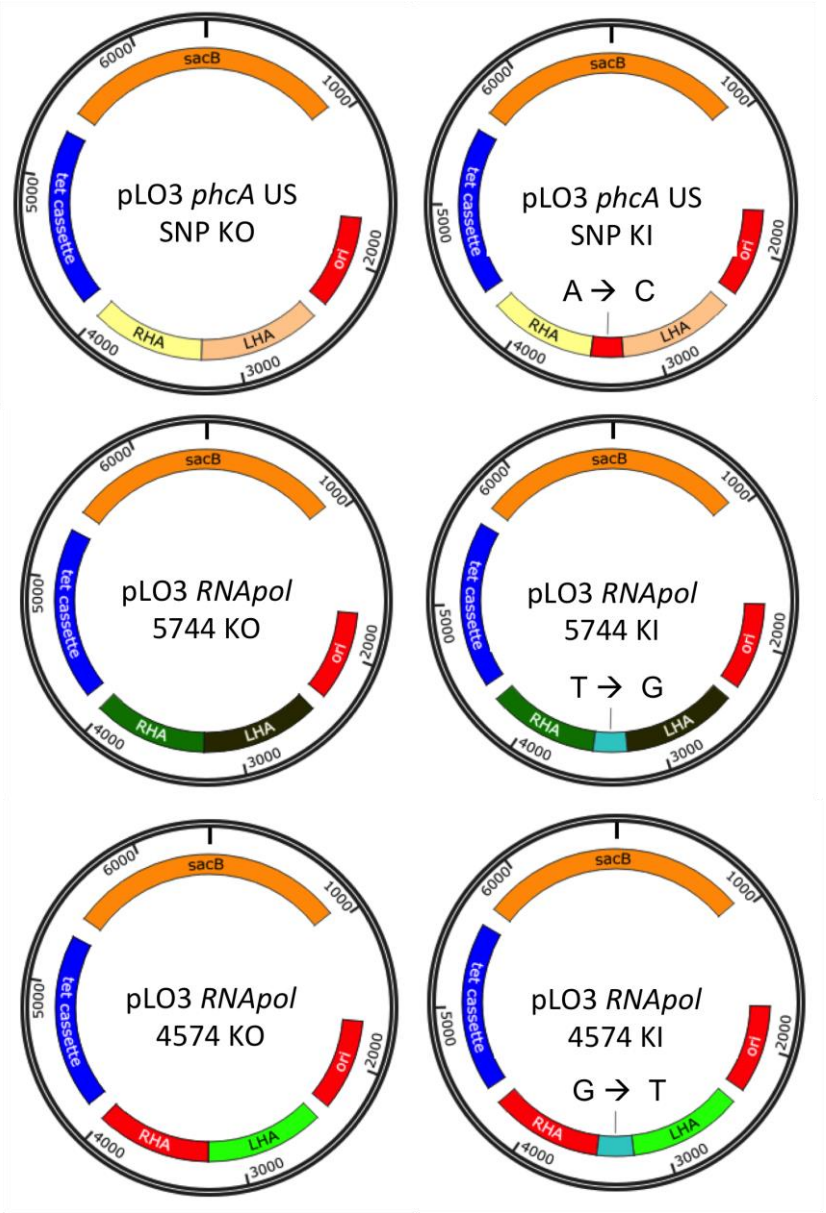


Figure 3.10. Plasmids used for the generation of knock-out/knock-in mutant library (B). Knock-out and knock-in plasmids constructed for *phcA* upstream SNP generation (top), RNA polymerase sigma-54 gene 5744 open reading frame SNP generation (middle) and RNA polymerase sigma-54 gene 4574 open reading frame SNP generation (bottom). All knock-in plasmids contained the same homology arms as their corresponding knock-out plasmid but also contained the deletion region of DNA with the desired mutation.

3.2.4.2. Generation of a knock-out/knock-in mutant library

All generated plasmids were transformed into *E. coli* DH5α for propagation, miniprepped and transformed into the *E. coli* conjugation donor strain S17-1. Conjugation was performed between S17-1 and the original *C. necator* H16 parent strain. Potential transconjugants were selected for on F-MM agar plates containing both gentamicin and tetracycline for presence of the suicide plasmid and against *E. coli* growth. After 3-4 days, single colonies were picked and re-streaked twice on selective plates and single crossover of all plasmids was confirmed by colony PCR. For this, two primer pairs were used to verify whether the LHA or RHA had initiated the single crossover integration of the suicide plasmid into the chromosome (Fig. 3.11).

Confirmed single crossover mutants were grown in 5 ml LB medium supplemented with gentamicin and tetracycline and allowed to grow to an OD_{600} of about 3. 2 µl of this culture was used to inoculate 5 ml LSLB with 15% sucrose, which was allowed to grow to an OD_{600} of about 3. Serial dilutions of this culture were made with 10^{-5} and 10^{-6} dilutions being plated onto LSLB 15% sucrose agar plates. These steps allowed for the selection of cells in which the second crossover event of the integrated construct had occurred (Fig. 3.11).

The *sacB* gene on pLO3 drives this process by encoding the enzyme levansucrase, which catalyses formation of toxic fructose polymers (levans), when sucrose is present. This creates a selective pressure and thus supports selection of recombinant cells which have lost *sacB* through either loss of the plasmid backbone through double crossover of the second homology arm with the corresponding sequence on the chromosome (see Materials and Methods 2.4.1) or the entire original plasmid through reversal of the single crossover. This will result in recombinant cells, which either have reverted to the wild type or generate the desired mutation.

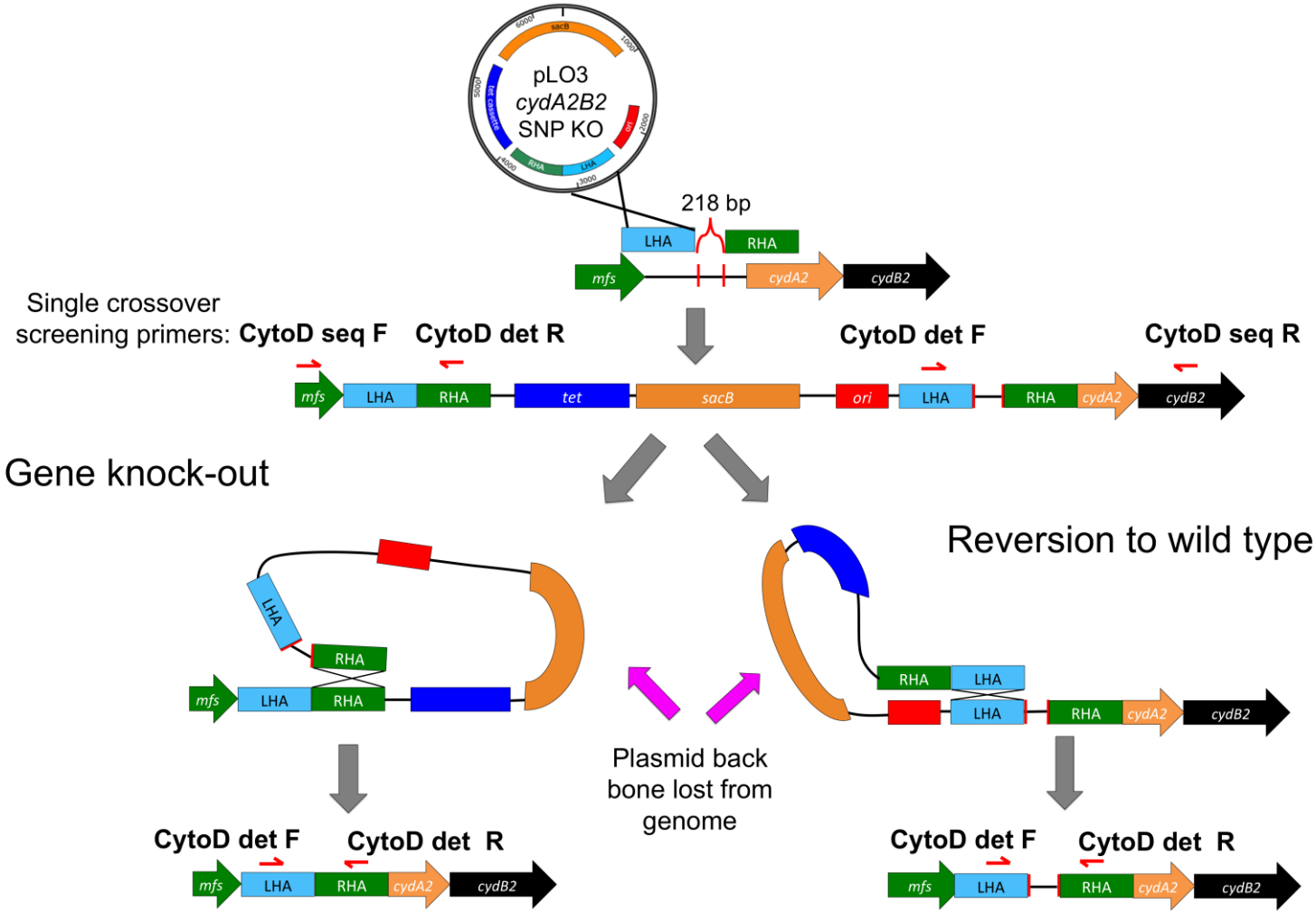


Figure 3.11. Key steps in the generation of defined deletions and associated PCR screening approaches exemplified for pLO3_*cydA2B2* SNP KO.

pLO3_*cydA2B2* SNP KO initiates a single crossover, either via the left homology arm (LHA) as shown or the right homology arm (RHA), to allow its integration next to the desired site for deletion, which is screened for by two primer pairs (top). Insertion will take place upstream or downstream of the intended deletion site, respectively, dependent on whether single crossover events involved the LHA or RHA. For LHA crossovers, a smaller PCR product will be generated with CytoD seq F and Cyto D det R primers in comparison to the wild type and CytoD det F and CytoD seq R primers will produce a wild type band. For an RHA crossover, the former pair will produce a wild type band and the latter pair a smaller band. In the shown example, the second crossover event will produce a deletion for an RHA crossover, yielding a smaller PCR product with CytoD det F and CytoD det R (bottom left), or will revert to the wild type for a LHA crossover (bottom right).

After 2 days, single colonies were transferred onto LB plates containing gentamicin and gentamicin/tetracycline respectively to screen for loss of pLO3 from the genome. Colonies not able to grow on tetracycline should no longer contain pLO3, since the second crossover event has occurred and the plasmid backbone has been lost from the genome (Fig. 11). These colonies were screened for the loss of the desired sequence by PCR (Figs. 3.12, 3.13, 3.14, 3.15 and 3.16; top panel) using the relevant primers (Table 3.4.). During generation of *C. necator phcA* IF, multiple colonies were screened, originating from different single crossover mutants which all grew on both selective plates following patching (data not shown). This indicated the suicide plasmid could not be lost from the genome, perhaps due to a lack of sacB expression or because the gene is essential to *C. necator* and cannot be knocked out.

This library of knock-out strains was generated with the intention of performing further screening in the presence of CO, to observe whether the loss of these regions influenced CO tolerance. However due to time constraints, these experiments were not performed. For all confirmed knock out mutants, conjugation was performed with the corresponding knock-in plasmid and the mutant selection process described was repeated (Figs. 3.12, 3.13, 3.14, 3.15 and 3.16; bottom panel). This generated knock-in mutant strains each containing the desired mutation. Knock-out and knock-in PCR fragments were purified and confirmed by Sanger sequencing. This way, successful introduction of the chosen mutations could be confirmed. The generated strains were designated *C. necator cydA2B2*_SNP, *C. necator phcR*_12bp, *C. necator phcA*_US, *C. necator* RNApol 5744 and *C. necator* RNApol 4574 for mutations in or upstream of *cydA2B2, phcR, phcA* and *RNA polymerase sigma 54 factor*, respectively. This knock-out/knock-in method was repeated for one of the generated SNP-bearing mutants, the *C. necator cydA2B2* SNP strain, to generate strains carrying a combination (one or two additional mutations) of some of the above gene variants (screening data not shown). These strains were created to assess whether particular SNP profiles (based on the evolved *C. necator* isolate strains; Table 3.3) generated greater CO tolerance.

Table 4.4. PCR screening guide to mutations generated in *C. necator*

Mutation	Gene	Mutant Name	Screening primers	Knock-out PCR length	Knock-in PCR length
G -> A	Intergenic region upstream *cydA2B2*	*C. necator cydA2B2* SNP	CytoD SNP det F & CytoD SNP det R	476 bp	695 bp
Deletion: GCGCTC GCGGTG	*phcR*	*C. necator phcR* 12bp	PhcR seq F and PhcR seq R	565 bp	1650 bp
A -> C	*phcA*	*C. necator phcA* IF SNP	PhcA IF SNP F & PhcA IF SNP R	352 bp	717 bp
A -> C	Intergenic region upstream of *phcA*	*C. necator phcA* US SNP	PhcA US SNP F & PhcA US SNP R	238 bp	465 bp
T -> G	*RNA polymerase sigma-54 factor*	*C. necator* RNApol 5744	RNApol SNP det F & RNApol SNP det R	287 bp	532 bp
G -> T	*RNA polymerase sigma-54 factor*	*C. necator* RNApol 4574	RNApol SNP det F & RNApol SNP det R	233 bp	492 bp

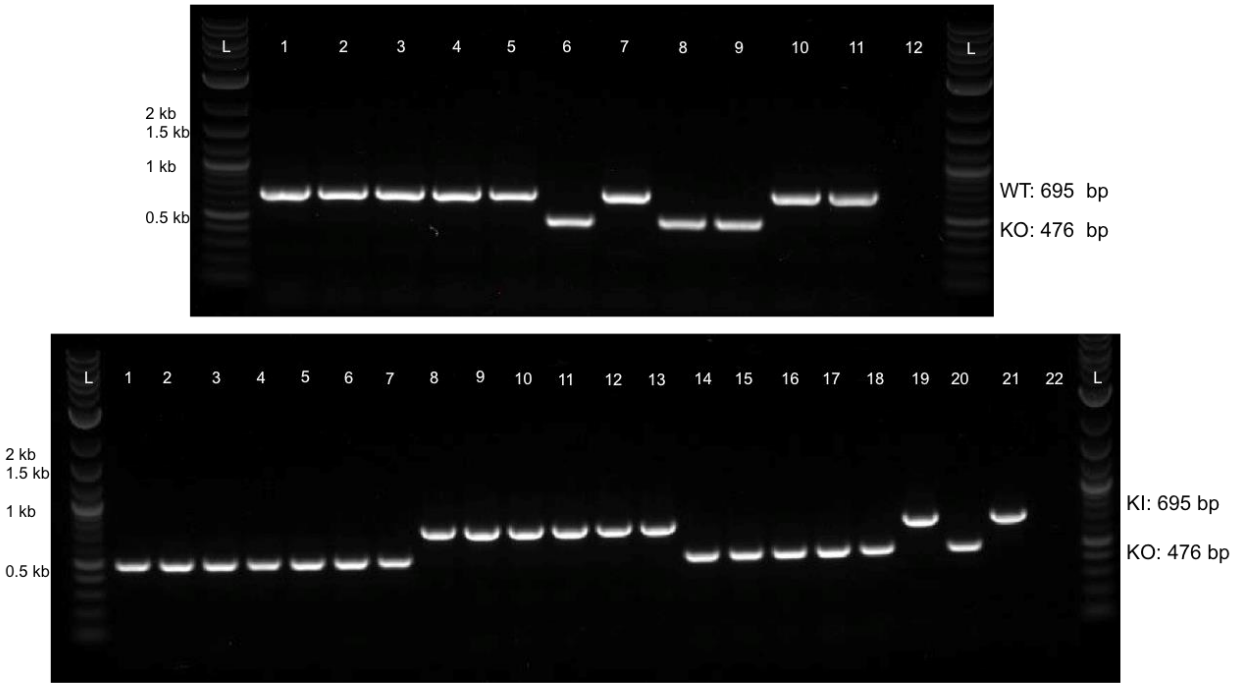


Figure 3.12. PCR confirmation of *cydA2B2* upstream region knock-out (KO) and knock-in (KI). PCR amplification of *C. necator cydA2B2* upstream region using the primer pair CytoD SNP det F and CytoD SNP det R in wild type (lane 11) and mutant candidates (lanes 1-10) to confirm knock-out (top). Candidates in lanes 6, 8 and 9 displayed the knock-out genotype. Lane 12 contains a blank PCR reaction. PCR amplification of the same region in a knock-out mutant (lane 20), wild type (lane 21) and derived knock-in mutant candidates (lanes 1-19) to confirm knock-in of regions containing the altered nucleotide (bottom). Candidates in lanes 8-13 and 19 displayed the knock-in genotype. Lane 22 contains a blank PCR reaction. The size of PCR products was estimated following 1% agarose gel electrophoresis using a 2-log ladder (L).

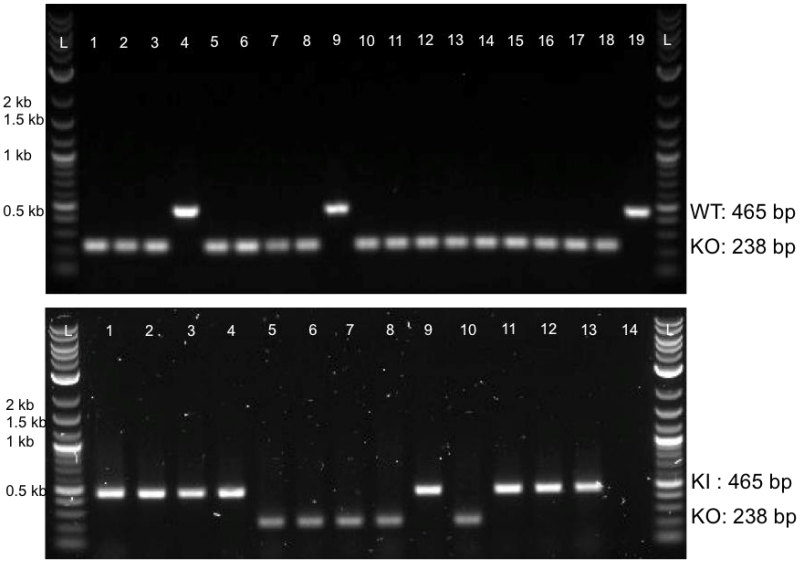


Figure 3.13. PCR confirmation of *phcR* knock-out (KO) and knock-in (KI).

PCR amplification of *C. necator phcR* using the primer pair PhcR seq F and PhcR seq R in wild type (lane 22) and mutant candidates (lane 1-21) to confirm knock-out (top). Candidates 2, 3, 9, 10, 12, 14, 18, 19 and 21 display knock-out genotype. Lane 23 contains a blank PCR reaction. PCR amplification using the same primer pair in a knock-out mutant (lane 8) and mutant candidates (lanes 1-7) to confirm knock-in containing the 12 bp deletion (bottom). Candidates 5 and 6 display knock-in genotype. Lane 9 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose electrophoresis using a 2-log ladder (L).

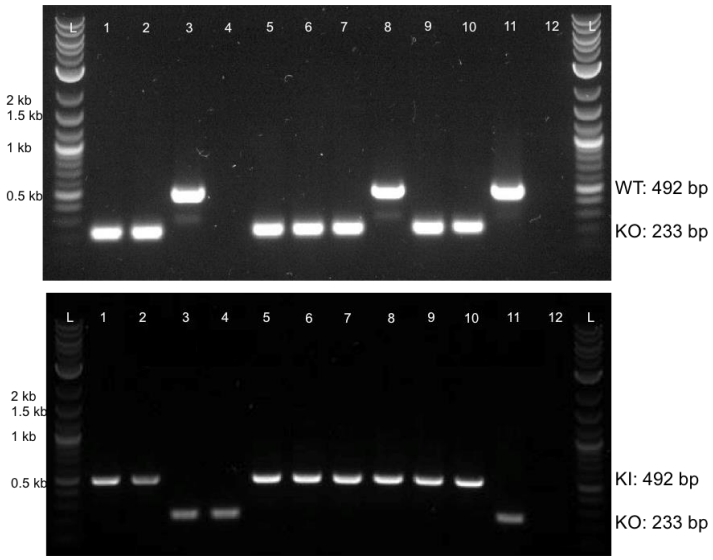


Figure 3.14. PCR confirmation of *phcA* upstream knock-out (KO) and knock-in (KI).

PCR amplification of *C. necator phcA* using the primer pair PhcA US SNP F and PhcA US SNP R upstream region in wild type (lane 19) and mutant candidates (lanes 1-18) to confirm knock-out (top). Candidates in lanes 1-3, 5-8 and 10-18 display the knock-out genotype. PCR using the same primer pair in wild type (lane 13) and mutant candidates (lanes 1-12) to confirm knock-in containing the altered nucleotide (bottom). Candidates 1-4, 9, 11 and 12 display the knock- in genotype. Lane 14 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose electrophoresis using a 2-log ladder (L).

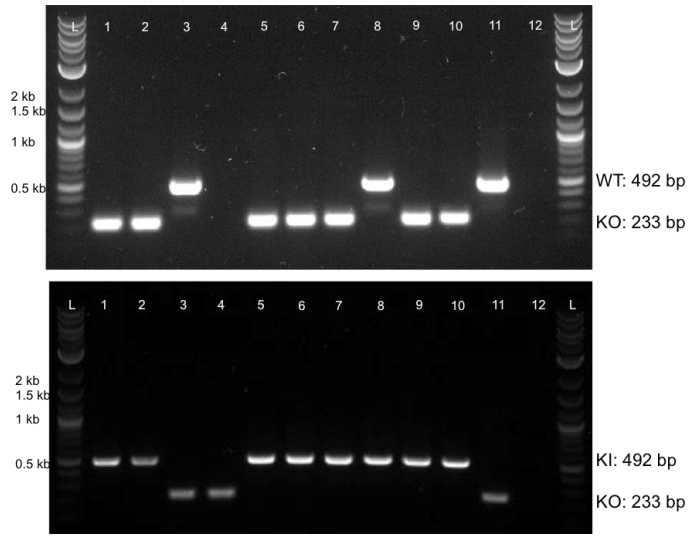
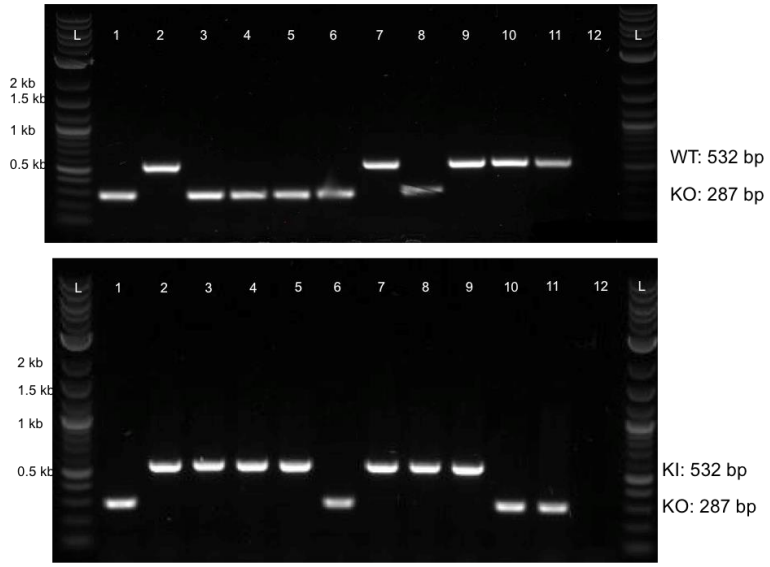


Figure 3.15. PCR confirmation of RNA polymerase sigma-54 factor gene 5744 knock-out (KO) and knock-in (KI).

PCR amplification of *C. necator* RNApol 5744 using the primer pair RNApol SNP det F and RNApol SNP det R in wild type (lane 11) and mutant candidates (lanes 1-10) to confirm knock-out (top). Lane 12 contains a blank PCR reaction. Candidates in lanes 1, 3-6 and 8 display the knock-out genotype. PCR using the same primer pair in knock-out mutant (lane 11) and mutant candidates (lanes 1-10) to confirm knock-in containing the altered nucleotide (bottom). Candidates 2-5 and 7-9 display the knock-in genotype. Lane 12 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose electrophoresis using a 2-log ladder (L).



Fgure 3.16. Figure 3.16. PCR confirmation of RNA polymerase sigma-54 factor gene 4574 knock-out and knock-in.

PCR amplification of *C. necator* RNApol 4574 using the primer pair RNApol SNP det F and RNApol SNP det R in wild type (lane 11) and mutant candidates (lanes 1-10) to confirm knock-out (top). Candidates in lanes 1, 2, 5-7, 9 and 10 display the knock-out genotype. Lane 12 contains a blank PCR reaction. PCR using the same primer pair in a knock-out mutant (lane 11), wild type (lane 10) and mutant candidates (lanes 1-9) to confirm knock-in containing the altered nucleotide (bottom). Candidates in lanes 1, 2 and 5-9 display knock-in genotype. Lane 12 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose electrophoresis using a 2-log ladder (L)

3.2.4.3. Growth comparison of generated single knock-in mutants with wild type *C. necator*

Having generated mutant strains carrying defined genetic changes, it was necessary to test which of these improved growth in the presence of CO. For each single knock-in mutant, three colonies were grown in 5 ml F-MM alongside wild type colonies. These were then subcultured in 20 ml F-MM in serum bottles with 50% CO and 50% air or 50% N_2 and 50% air. The latter condition was chosen as a control, to verify that a given mutation was not generally increasing growth, but that this effect was specific for the presence of CO. The experiment started with an inoculation OD_{600} of 0.05.

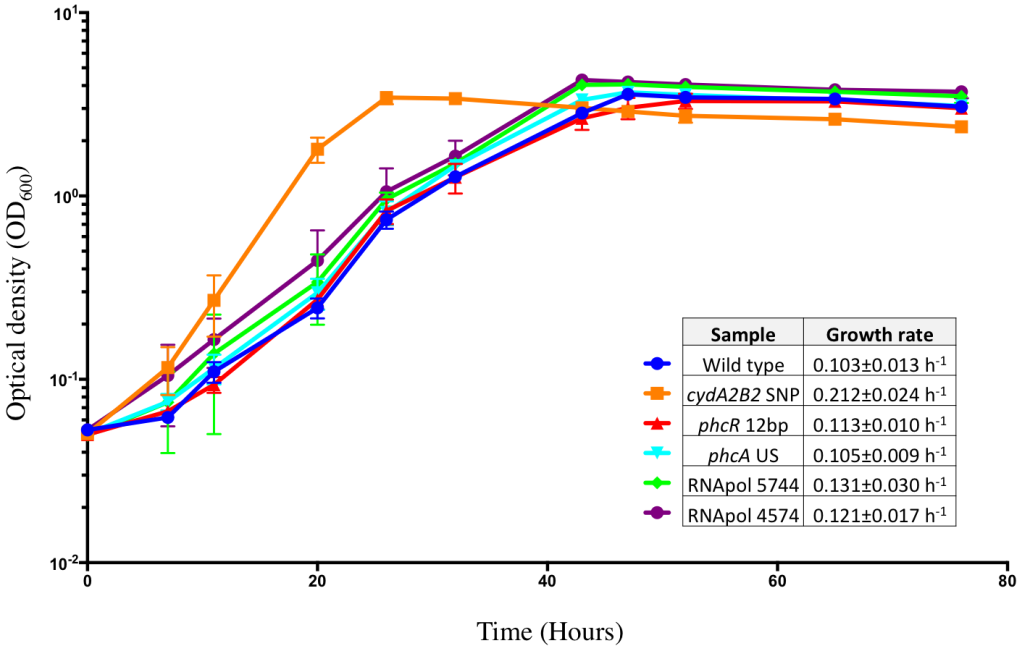


Figure 3.17. Single knock-in mutant growth compared to wild type in the presence of 50% CO/ 50% air.

C. necator H16 and its mutant derivatives were grown at 30 °C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air (2 bar total pressure). Blue circles, wild type; orange squares, *C. necator cydA2B2* SNP; red triangles, *C. necator phcR* 12bp; cyan triangles, *C. necator phcA* US; green diamonds, *C. necator* RNApol 5744; purple circles, *C. necator* RNApol 4574. The inset shows growth using a logarithmic scale. Error bars represent the standard deviation of the mean for three independent replicates.

The growth curves with CO followed a time course of 76 hours, with the wild type reaching its peak OD_{600} after 48 hours. In contrast *C. necator cydA2B2* SNP reached its peak OD_{600} after 30 hours and was the fastest growing mutant, displaying a significantly

higher growth rate than the wild type (0.212±0.024 h^{-1} versus 0.103±0.01 h^{-1}, p=0.0042) (Fig. 3.17). *C. necator phcR* 12bp was the only mutant not displaying improved growth compared to the wild type (0.113±0.013 h^{-1}), reaching its highest OD_{600} at 52 hours. *C. necator phcA* US reached its peak OD_{600} at the same time as the wild type but *C. necator* RNApol *5744* and RNApol *4574* both appeared to display a slight growth advantage due to a slightly shorter lag phase, reaching peak OD_{600} at 43 hours; however, maximum growth rates for all three aforementioned strains were not significantly different to that of the wild type (0.104±0.009 h^{-1}, 0.131±0.030 h^{-1} and 0.1209±0.017 h^{-1}, respectively; $p < 0.01$). The slightly improved growth of these mutants as seen in Fig. 3.17 was due to achieving a slightly higher final OD_{600} and/or somewhat faster growth compared to the wild type in the later stages of growth.

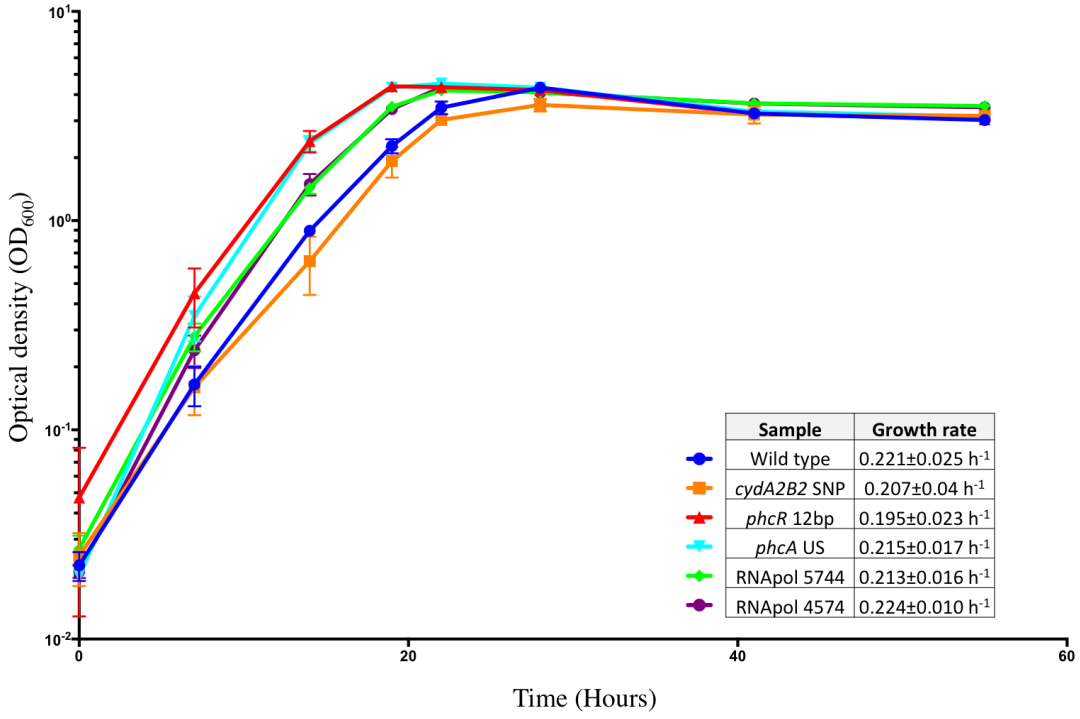


Figure 3.18. Figure 3.18. Single knock-in mutant growth compared to wild type in the presence of 50% N_2/ 50% air.

C. necator H16 and its mutant derivatives were grown at 30 °C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% N_2 /50% air (2 bar total pressure). Blue circles, wild type; orange squares, *C. necator cydA2B2* SNP; red triangles, *C. necator phcR* 12bp; cyan triangles, *C. necator phcA US*; green diamonds, *C. necator* RNApol 5744; purple circles, *C. necator* RNApol 4574. The inset shows growth using a logarithmic scale. Error bars represent the standard deviation of the mean for three independent replicates.

The control experiment with N_2 was started with an approximate OD_{600} of 0.01 and followed a time course of 55 hours, with the wild type reaching peak OD_{600} after 28 hours (Fig. 3.18). Interestingly *C. necator cydA2B2* SNP reached its highest OD_{600} at the same time as the wild type, indicating there is no growth advantage for this strain in the absence of CO. The obtained growth rates for wild type and *cydA2B2* mutant were 0.222±0.025 h^{-1} and 0.207±0.04 h^{-1}, respectively. In contrast, the remaining four strains appeared to grow slightly faster compared to the wild type. *C. necator* Δ*phcR* 12bp and *C. necator* Δ*phcA* US reached peak OD_{600} after 19 hours, and *C. necator* RNApol 5744 and RNApol 4574 after 22 hours; however the maximum growth rate of these strains was not significantly higher than that of the wild type ($p \geq 0.1$) with growth rates of 0.196±0.023 h^{-1} and 0.215±0.017 h^{-1} for the former two and 0.2125±0.016 h^{-1} and 0.224±0.010 h^{-1} for the latter, respectively ($p \geq 0.3$) Similar to what was observed in the presence of CO, the improved growth of these mutants as seen in Fig. 3.17 was due to somewhat faster growth in the later stages of the growth cycle and, in the case of *C. necator phcR* 12bp, differences in the starting inoculum.

3.2.4.4. Growth comparison of generated double and triple knock-in mutants with wild type *C. necator*

Having profiled the growth of the generated single knock-in mutant strains, revealing that *C. necator cydA2B2* SNP exhibited by far the greatest resistance to CO, it was necessary to assess whether additional mutations in this strain further increased this resistance. The 12 bp deletion within *phcR*, the SNP upstream of *phcA* and the SNPs present in the RNA polymerase sigma-54 factor gene (positions 4574 and 5744) were chosen for this purpose. The mutations *phcA* US and RNApol 4574 were chosen to generate a triple knock in strain since this reflected the mutational profile of the evolved isolate with highest growth rate (Evo 6) (3.2.3.4, Table 3.3), possibly indicating that this combination of mutations gave higher CO resistance. For each double and triple knock-in mutant, three colonies were grown in 5 ml F-MM alongside *C. necator cydA2B2* SNP. These were then subcultured in 20 ml F-MM in serum bottles with 50% CO and 50% air or 50% N_2 and 50% air. The experiment was initiated with a starting OD_{600} of 0.05. The growth curves with CO followed a time course of 76 hours, with *C. necator cydA2B2* SNP reaching its peak OD_{600} after 30 hours (Fig. 3.19), along with derivative strains *C. necator cyd* + phcR 12bp, *C. necator cyd* + phcA US and *C. necator cyd* + RNApol 5744 and *C. necator cyd* + RNApol 4574 (Fig. 3.19).

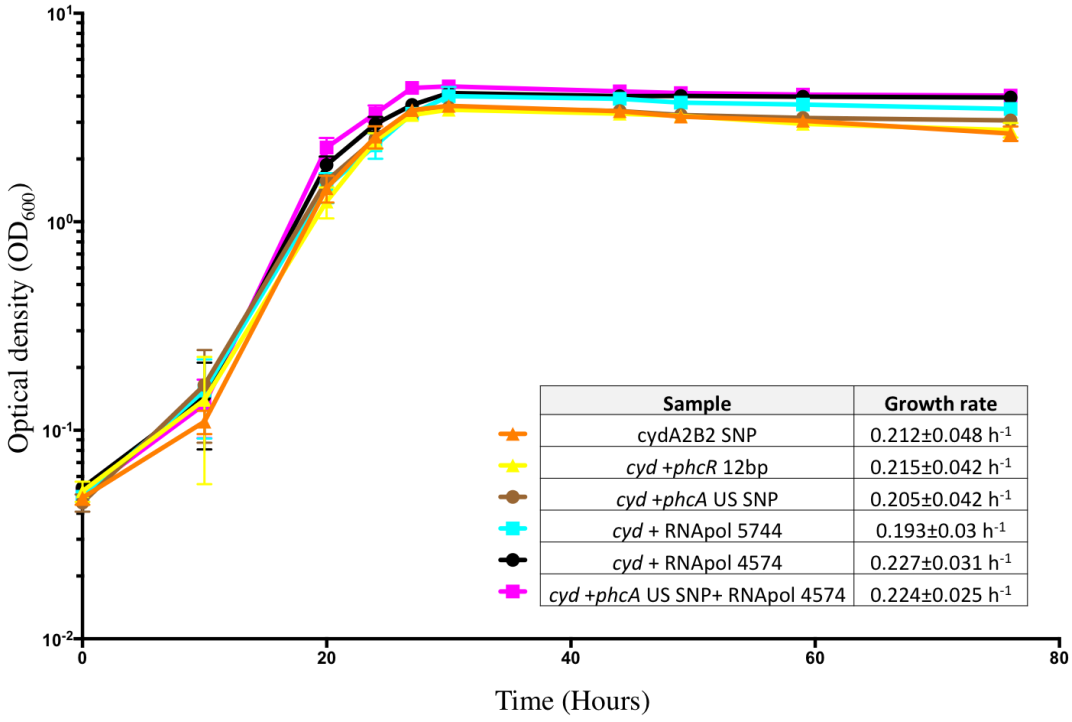


Figure 3.19. Double and triple knock-in mutant growth compared to the *cydA2B2* SNP strain in the presence of 50% CO/50% air.

C. necator mutant derivatives were grown at 30 °C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air (2 bar total pressure). Orange triangles, *cydA2B2* SNP; yellow triangles *cyd* + *phcR* 12bp; brown diamonds, *cyd* + *phcA* US; turquoise diamonds, *cyd* + RNApol 5744; black circles, *cydA* + RNApol 4574; pink squares *cyd* + *phcA* US SNP + RNApol 4574. The inset shows growth using a logarithmic scale Error bars represent the standard deviation of the mean for three independent replicates.

The obtained growth rates under these conditions were 0.212±0.048 h-1, 0.215±0.042 h^{-1}, 0.205±0.042 h^{-1}, 0.193±0.03 h^{-1}, 0.227±0.031 h^{-1} and 0.224±0.025 h^{-1} for *C. necator cydA2B2* SNP, *C. necator cyd* + *phcR* 12bp, *C. necator cyd* + *phcA* US, *C. necator cyd* + RNApol 5744, *C. necator cyd* + RNApol 4574, *C. necator cyd* + *phcA* US + RNApol4574 respectively. The difference in growth rate between the derivate strains and *C. necator cydA2B2* SNP was not statistically significant ($p > 0.05$). However, the *C. necator cyd* + *phcA* US + RNApol 4574 strain reached its highest OD_{600} after 27 hours and along with *C. necator cyd* + RNApol 5744 and *C. necator cyd* US + RNApol 4574 also reached a higher peak OD_{600} than *C. necator cyd* SNP.

The control experiment with N_2 was started with an OD_{600} of 0.01 and followed a time course of 55 hours, with the *cydA2B2* strain showing a growth rate of 0.191±0.025 h^{-1}

and reaching peak OD_{600}, the lowest observed for all tested strains, after 28 hours (Fig. 3.20), similar to what was observed before (Fig. 3.18). The remaining strains, *C. necator cyd* + *phcR* 12bp, *C. necator cyd* + *phcA US*, *C. necator cyd* + *RNApol 5744*, *C. necator cyd* + RNApol 4574 and *C. necator cyd* + *phcA* + RNApol 4574 reached maximum growth rates of 0.202±0.031 h^{-1}, 0.216±0.05 h^{-1}, 0.192±0.03 h^{-1}, 0.195±0.023 h^{-1} and 0.192±0.012 h^{-1}, respectively.

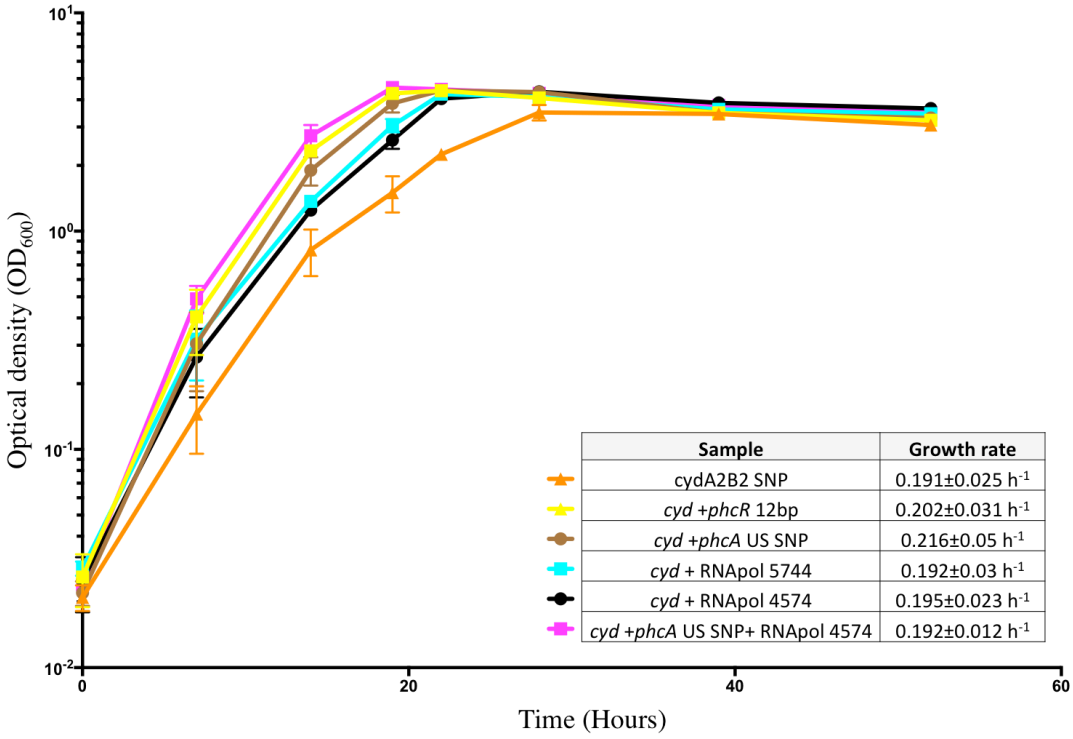


Figure 3.20. Double and triple knock-in mutant growth compared to the *cydA2B2* SNP strain in the presence of 50% N_2/50% air.

C. necator mutant derivatives were grown at 30 °C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% N_2/50% air (2 bar total pressure). Blue circles, wild type; orange triangles, *cyd* SNP; yellow triangles *cyd* SNP + *phcR* 12bp; brown diamonds, *cyd* + *phcA* US; turquoise diamonds, *cyd* + RNApol 5744; black circles, *cyd* + RNApol 4574; pink squares *cyd* + *phcA* US + RNApol 4574. The inset shows growth using a logarithmic scale. Error bars represent the standard deviation of the mean for three independent replicates.

The observed differences were not significantly different from the *C. necator cydA2B2* SNP parent strain ($p > 0.5$). However, peak OD_{600} were higher or reached earlier, as these strains maintained faster growth during the later stages before entering stationary phase. Based on this analysis, *C. necator cyd* + *phcR 12bp*, *C. necator cyd* + *phcA* US and *C. necator cyd* + *phcA* + RNApol *4574* were the best performing strains under the

chosen condition. The results suggested that at least some of the analysed mutations, or combinations thereof, had no direct effect on CO resistance but generally improved *in vitro* growth performance.

3.2.5. Analysis of cytochrome *bd* ubiquinol oxidase in *C. necator*

The single point mutation (G→A) upstream of the cytochrome *bd* ubiquinol oxidase operon *cydA2B2* appears to be the only mutation that gives *C. necator* significantly increased resistance to CO. Since the mutation is not in a protein-encoding region, it may be altering operon expression; therefore, further investigation into this operon was carried out, using three approaches. First, the upstream intergenic region of the operon was tested for promoter activity using an RFP-based reporter plasmid to elucidate whether introduction of the point mutation increased the level of expression as previously hypothesised (3.2.3.4.). Second, the operon was overexpressed in an otherwise wild type background, again to determine the resulting effect on CO tolerance. Additionally, the two subunit-encoding genes of the operon were inactivated to determine how this impacted on CO tolerance.

3.2.5.1. RFP expression analysis of *cydA2B2* intergenic region in *C. necator*

To ascertain whether the G→A mutation influenced gene expression of *cydA2B2*, two RFP-based reporter plasmids were constructed containing the 616 bp upstream region of *cydA2B2* (Fig. 3.21). One plasmid harboured the G→A mutation in this region (p006_*cyd* SNP) whilst the other contained the wild type sequence (p006_*cyd* US). Both plasmids were cloned in *E. coli* DH5 alpha to give *E. coli* p006_*cyd* US and *E. coli* p006_*cyd* SNP, respectively. The plasmids were transferred into *E. coli* S17-1, which was conjugated with *C. necator* H16 to create *C. necator* p006_*cyd* US and *C. necator* p006_*cyd* US. All transformed cells were confirmed by colony PCR.

Following confirmation by colony PCR, three positive colonies for each plasmid were grown overnight in 5 ml F-MM + chloramphenicol (50mg/ml) alongside *C. necator* wild type and *C. necator* containing an empty vector (pEH00E). These cultures were diluted back to an OD_{600} of 0.05 in fresh serum bottles containing 20 ml F-MM, chloramphenicol to maintain the respective plasmids (50 μg/ml) and 50% CO/50% air or 50% N_2/50% air in the headspace.

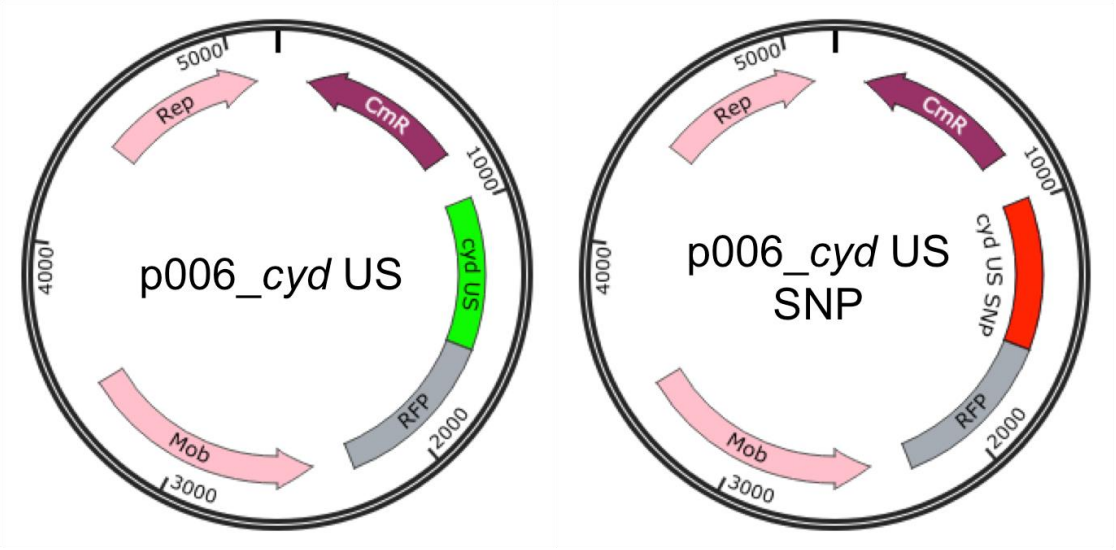


Figure 3.21. *cydA2B2* transcriptional reporter plasmid.

The native upstream region of *cydA2B*2 was amplified from *C. necator* wild type chromosomal DNA with Phusion polymerase, using CytoD inter US F and CytoD inter US R, which contain AatII (Forward) and NdeI (Reverse) restriction sites respectively. This PCR fragment and pEH006 were digested with these enzymes and ligated together using T4 DNA ligase to generate p006_*cyd* US (left). In this construct, expression of the *rfp* reporter gene is under control of a potential wild type *cydA2B*2 promoter located in the amplified intergenic region. The G→A SNP carrying upstream region of *cydA2B2* was amplified from *C. necator cydA2B2* SNP chromosomal DNA with Phusion using the same primer pair. This PCR fragment was ligated into pEH006 to generate p006_*cyd* SNP. In this construct, expression of the *rfp* reporter gene is under control of a potentially modified *cydA2B*2 promoter (or a newly generated one) located in the amplified intergenic region. CmR, chloramphenicol resistance gene; *cyd* US, *cydA2B*2 upstream region (WT: green, SNP containing: red); RFP, ref fluorescent protein gene; Mob, functions for plasmid mobilisation; Rep, derived from pBBR1MCS, functions for plasmid replication (plasmid maps designed using SnapGene).

Samples were taken at 1, 4 and 24 hours to measure fluorescence and optical density for each sample using a Tecan plate reader (Fig. 3.22). Fluorescence was divided by the optical density to give normalised fluorescence. At all three time points, in the presence of both CO and N_2, *C. necator* p006_*cyd* SNP expressed 400-fold more fluorescence than *C. necator* p006_*cyd* US ($p < 0.05$), which produced no higher fluorescence than the empty plasmid-containing *C. necator* p006E control ($p > 0.05$) (Fig. 3.22). Similarly, *C. necator* p006_*cyd* SNP did not produce significantly more fluorescence when cultured in the presence of CO than N_2 and levels were comparable to that of the empty vector. These results suggests that the G→A mutation increases gene expression in *C. necator* regardless of whether CO is present or absent, possibly

by increasing the promoter strength or by enhancing or inhibiting transcriptions factor binding in this intergenic region.

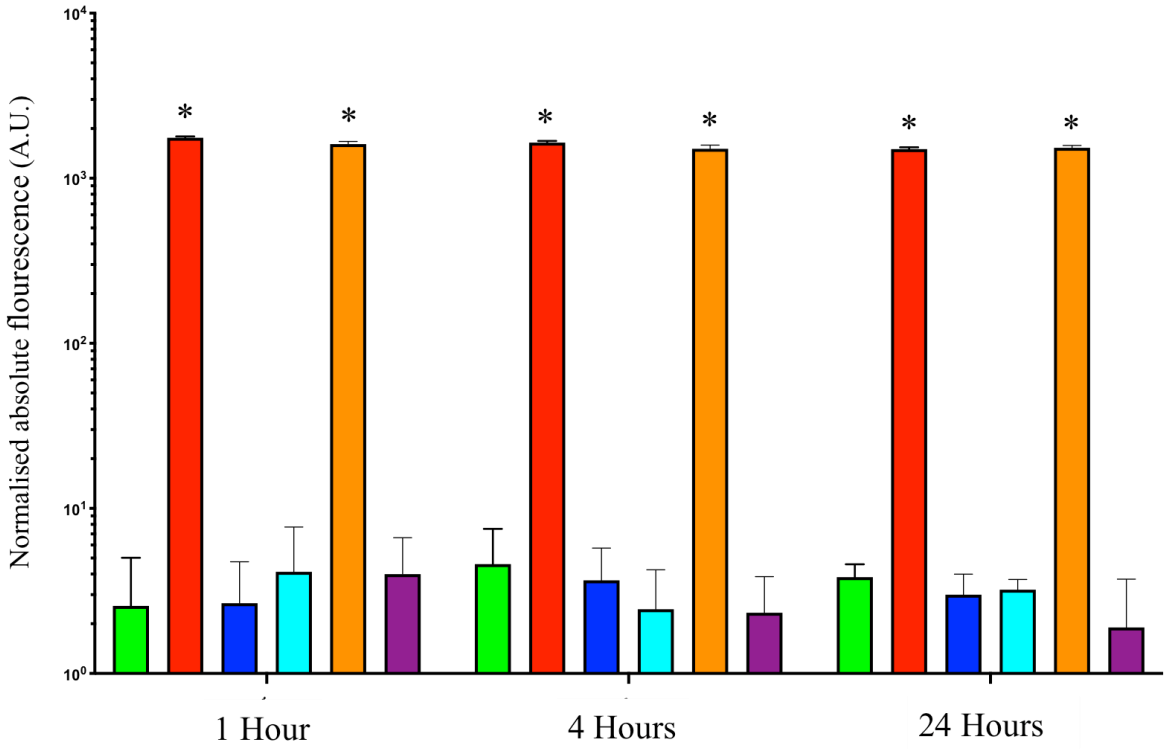


Figure 3.22. RFP expression of *C. necator* p006_*cyd* US, *C. necator* p006_*cyd* SNP and *C. necator* p006E in 50% CO / 50% air and 50% N_2/50% air.

Strains were grown at 30 °C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air (2 bar total pressure). Expression of the *rfp gene* was monitored after 1 h, 4 h and 24 h by determining RFP fluorescence for *C. necator* p006_*cyd* US (CO, green; N_2, cyan), *C. necator* p006-_*cyd* SNP (CO, red; N_2, orange) and *C. necator* p006E (CO, blue; N_2, purple). Average normalised absolute fluorescence are given for each time point. Error bars represent standard deviations from the mean for three biological replicates. There was no significant difference between samples taken from CO and N_2 grown cultures for *C. necator* p006_*cyd* US (1 hour, p = 0.568; 4 hours, p = 0.340; 24 hours, p = 0.302) and *C. necator* p006_*cyd* SNP (1 hour, p = 0.022; 4 hours, p= 0.060084; 24 hours, p = 0.451338). There was no significant difference between the negative control (C. *necator* p006E) and *C. necator* p006_*cyd* US in the presence of CO (1 hour, p = 0.480; 4 Hours, p = 0.272; 24 hours, p = 0.156) or N_2 (1 hour, p = 0.482; 4 hours, p = 0.469; 24 hours, p = 0.146). There was a significant difference (indicated by the asterisk) between *C. necator* p006_*cyd* SNP and *C. necator* p006_*cyd* US in the presence of CO ($p < 0.00001$ for all time points) and N_2 ($p < 0.00001$ for all time points).

3.2.5.2. RFP expression analysis of the *cydA2B2* intergenic region in *E. coli*

To test whether introduction of the G→A mutation upstream of the *C. necator cydA2B2* operon would also change gene expression in *E. coli* DH5α, activity of the two generated *rfp* reporters was also analysed in *E. coli* p006_*cyd* US and *E. coli* p006_*cyd*

SNP, with *E. coli* p006E as an empty plasmid control. Three positive colonies for each strain were grown overnight in 5 ml F-MM + chloramphenicol (25 mg/ml). These cultures were diluted back to an OD_{600} of 0.05 in fresh serum bottles containing 20 ml F-MM, chloramphenicol (50 μg/ml) and 50% CO/50% air or 50% N_2/50% air in the headspace.

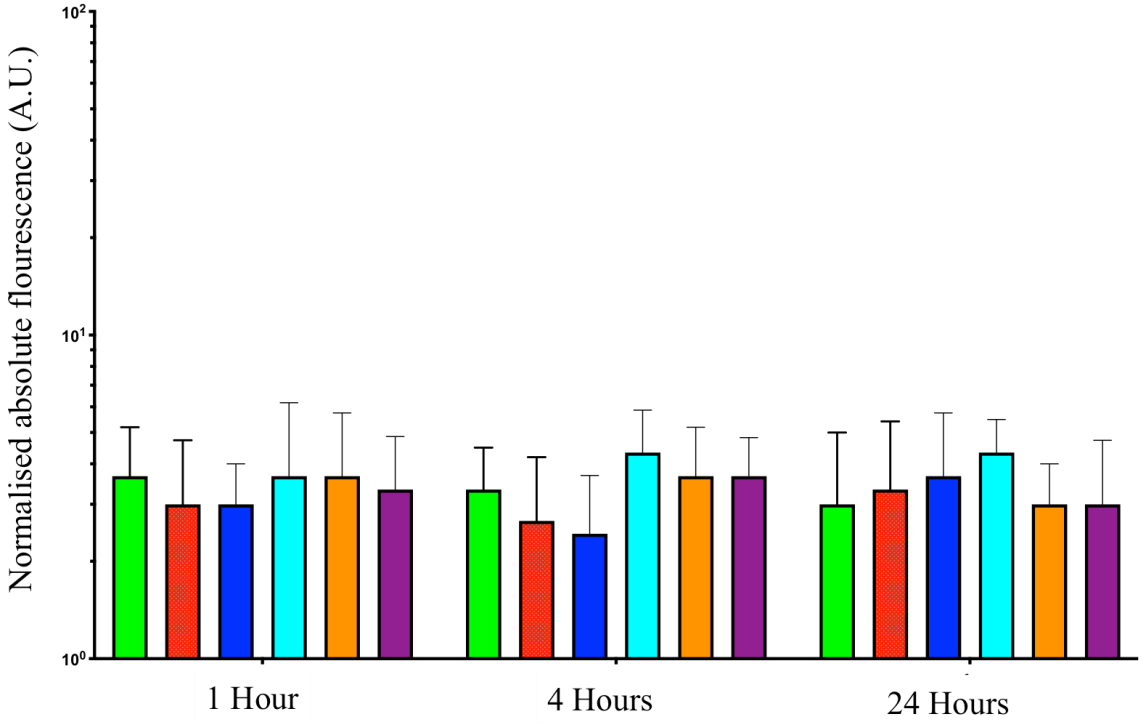


Figure 3.23. RFP expression of *E. coli* p006_*cyd* US, *E. coli* p006_*cyd* SNP and *E. coli* p006E in 50% CO / 50% air and 50% N_2 / 50% air.

E. coli DH5α strains were grown at 37 °C in 20 ml G-MM contained in 150 ml serum bottles under an atmosphere of 50% CO / 50% air (2 bar total pressure). Expression of the *rfp gene* was monitored after 1 h, 4 h and 24 h by determining RFP fluorescence in for *E. coli* p006_*cyd* US (CO, green; N_2, cyan), *E. coli* p006_*cyd* SNP (CO, red; N_2, orange) and *E. coli* p006E (CO, blue; N_2, purple). Average normalised absolute fluorescence are given for each time point. Error bars represent standard deviations from three biological replicates. There was no significant difference between any of the strains and conditions ($0.83 \geq p \geq 0.14$).

Time points were taken at 1, 4 and 24 hours using a Tecan plate reader to measure fluorescence and absorbance for each sample (Figs. 3.23). Fluorescence was divided by the optical density to give normalised fluorescence. For all three time points in the presence of CO and N_2, the differences in fluorescence were not significant between

and *E. coli* p006_*cyd* US and *E. coli* p006_*cyd* SNP, indicating that the G→A mutation does not increase gene expression in *E. coli* DH5α.

3.2.5.3. Overexpression of cytochrome *bd*

To assess whether overexpression of *cydA2B2* would positively influence CO tolerance, an overexpression plasmid containing the *cydA2B2* operon was constructed and introduced into *C. necator* by conjugation via *E. coli* S-17-1. The plasmid pEH006 (Hanko *et al.* 2017) was used as the vector backbone for this work as it contains an arabinose-inducible system which can be used to overexpress genes cloned downstream of an inducible pBAD promoter as well as a ribosome binding site (RBS) (Fig. 3.24).

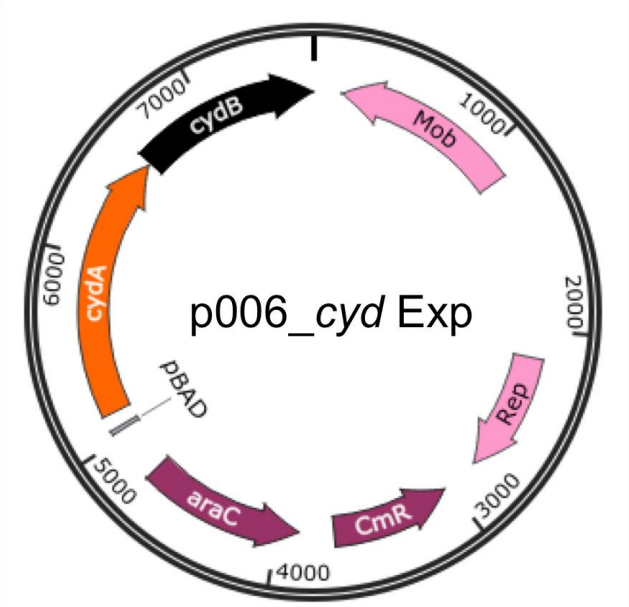


Figure 3.24. Construction of the cytochrome *bd* ubiquinol oxidase operon overexpression vector pEH006_*cydA2B2*.

The *cydA2B2* operon was amplified from the *C. necator* genome with Phusion polymerase, using primers containing overlapping regions with pEH006. Restriction digest was performed on pEH006 with NdeI and BamHI to remove the RFP gene. The digested backbone and the PCR amplified operon were used in a HiFi-assembly reaction to generate the overexpression plasmid pEH006_CydAB Exp. CmR, chloramphenicol resistance gene; cydA gene (orange) and cydB gene (black): *cydA2B*2 operon; Mob, functions for plasmid mobilisation; Rep, derived from pBBR1MCS, functions for plasmid replication; araC, encoding arabinose-responsive AraC transcriptional regulator; pBAD promoter (Plasmid map designed using SnapGene).

A growth comparison was performed between *C. necator* pEH006_*cydA2B2* Exp (*C. necator* p006_*cyd* Exp) and *C. necator* containing pEH006 (*C. necator* p006*)*. For each plasmid containing strain, three colonies were individually grown in 5 ml F-MM

containing 50 µg/ml chloramphenicol (Cl). These were then subcultured in 20 ml of the same medium in serum bottles under an atmosphere of 50% CO/50% air. The experiment was started with an OD_{600} of 0.05.

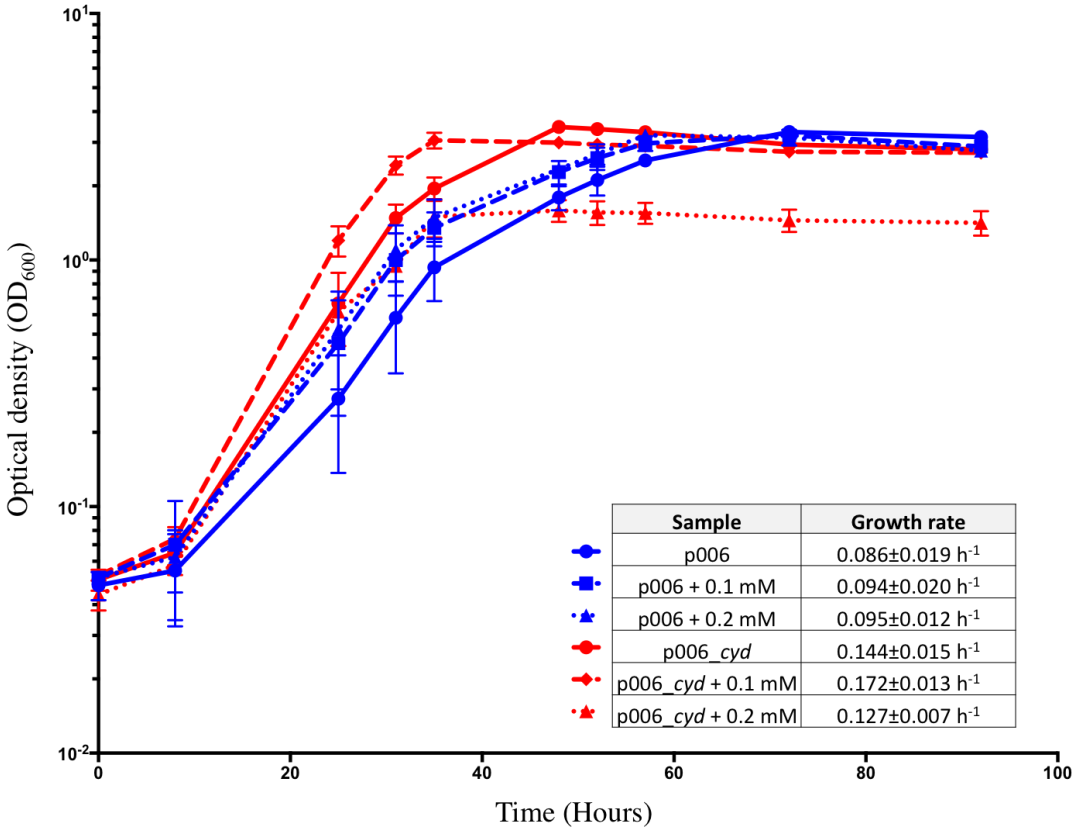


Figure 3.25. Growth of *C. necator* p006_*cyd* Exp and *C. necator* p006 with arabinose induction in the presence of 50% CO/50% air.

C. necator p006_*cyd* and control strain *C. necator* p006 were grown at 30°C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air (2 bar total pressure). Expression of *cydA2B2* from the pBAD promoter was induced with the indicated arabinose concentrations; 0 mM, red line, 0.1 mM, red dashed line, 0.2 mM, red dotted line. *C. necator* p006 was induced with the same arabinose concentrations: 0 mM, blue line, 0.1 mM: blue dashed line, 0.2 mM: blue dotted line. The inset shows growth using a logarithmic scale. Error bars represent the standard deviation of the mean for three independent experiments.

The growth curves followed a time course of 90 hours, with *C. necator* p006 reaching stationary phase after 72 hours and *C. necator* p006_*cyd* after 48 hours (Fig. 3.25). In the absence of arabinose *C. necator* p006_*cyd* displayed significantly faster growth than the control (0.143±0.015 h^{-1} compared to 0.086±0.019 h^{-1}, p = 0.038). A possible explanation for this could be the presence of low-level background expression of *cydA2B2*in the absence of inducer.

Addition of 0.1 mM arabinose slightly improved growth (0.094±0.020 h^{-1}) of the *C. necator* p006 control, but this was not statistically significant when compared to rate obtained in the absence of arabinose ($p > 0.05$)(Fig. 3.25). The same concentration significantly increased growth rate of *C. necator* p006_*cyd* Exp (0.1719±0.014 h^{-1}; $p < 0.05$) presumably due to increased expression of *cydA2B2.* Again this growth rate was significantly higher than that of the p006 control at the same arabinose concentration ($p = 0.020$). In the presence of 0.2 mM arabinose, *C. necator* p006 displayed no further improvement in growth (0.094±0.011 h^{-1}), however *C. necator* p006_*cyd* Exp exhibited growth inhibition, reaching a lower peak OD_{600} than with 0 or 0.1 mM arabinose and a reduced growth rate (0.127±0.007 h^{-1}). Taken together the obtained results suggested that induced expression of the cytochrome *bd* ubiquinol oxidase genes increased CO tolerance. However, it seems that higher expression of *cydA2B2* at 0.2 mM arabinose had detrimental effects on *C. necator.* The addition of arabinose likely contributes to the growth of *C. necator,* which may explain the faster growth of *C. necator* p006 at both arabinose concentrations.

3.2.5.4. Knock-out mutant generation

To assess whether the lack of cytochrome *bd* impacts on the capacity of *C. necator* to tolerate CO, a knock-out plasmid was constructed using the same method described in 3.2.4.1, to generate an knock-out of the *cydA2B2* operon using the protocol described in 3.2.4.2. Patched colonies unable to grow with tetracycline were screened by colony PCR and confirmed by Sanger sequencing of PCR fragments. The deletion generated in this manner encompassed the entire operon apart from the first 7 bp of *cydA2* and the last 7 bp of the *cydB2* gene.

3.2.5.5. Growth comparison of the cytochrome *bd* knock-out with wild-type *C. necator* and SNP knock-in strain *cydA2B2* SNP

Growth of the generated *cydA2B2* mutant (*C. necator cydA2B2* KO) was compared to that of the H16 wild type and the generated, defined SNP knock-in strain *cydA2B2* SNP (3.2.4.2.). For each of these strains, three colonies were individually grown in 5 ml F-MM. These were then subcultured in 20 ml F-MM in serum bottles with 50% CO/50% air or 50% N_2/50% air. The CO the experiment was initiated with a starting OD_{600} of 0.05.

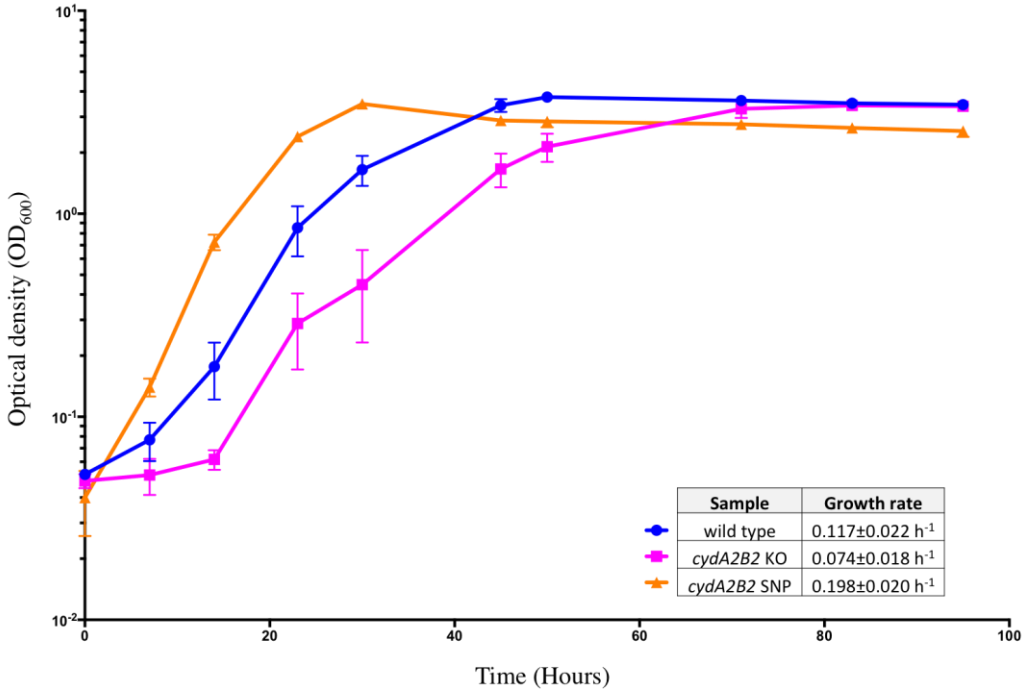


Figure 3.26. Growth of the *cydA2B2* knock-out mutant compared to the H16 wild type and *cydA2B2* SNP strains in the presence of 50%CO/ 50% air.

C. necator H16 and its mutant derivatives were grown at 30°C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air (2 bar total pressure). Blue circles, wild type; orange triangles, *C. necator cydA2B2* SNP, pink squares, *C. necator* Δ*cydA2B2*. Inset displays the natural log plot for each sample. Error bars represent the standard deviation of the mean for three independent replicates.

The growth curves followed a time course of 95 hours, with the wild type strain reaching stationary phase after 50 hours and the *cydA2B2* SNP strain after 30 hours (Fig. 3.26). In contrast, the *cydA2B2* KO strain reached stationary phase after 83 hours and displayed an increased lag phase in comparison to the wild type strain. The growth of Δ*cydA2B2* indicated that lack of cytochrome *bd* ubiquinol oxidase activity was detrimental to growth of *C. necator* when in the presence of CO. This was reflected by the obtained growth rates for the *cydA2B2* SNP strain, wild type and Δ*cydA2B2* mutant, which achieved 0.1983±0.020 h^{-1}, 0.117±0.022 h^{-1} and 0.073±0.018 h^{-1}, respectively. For both mutants these differences were statistically significant when compared to the wild type ($p = 0.042$ and $p = 0.026$, respectively).

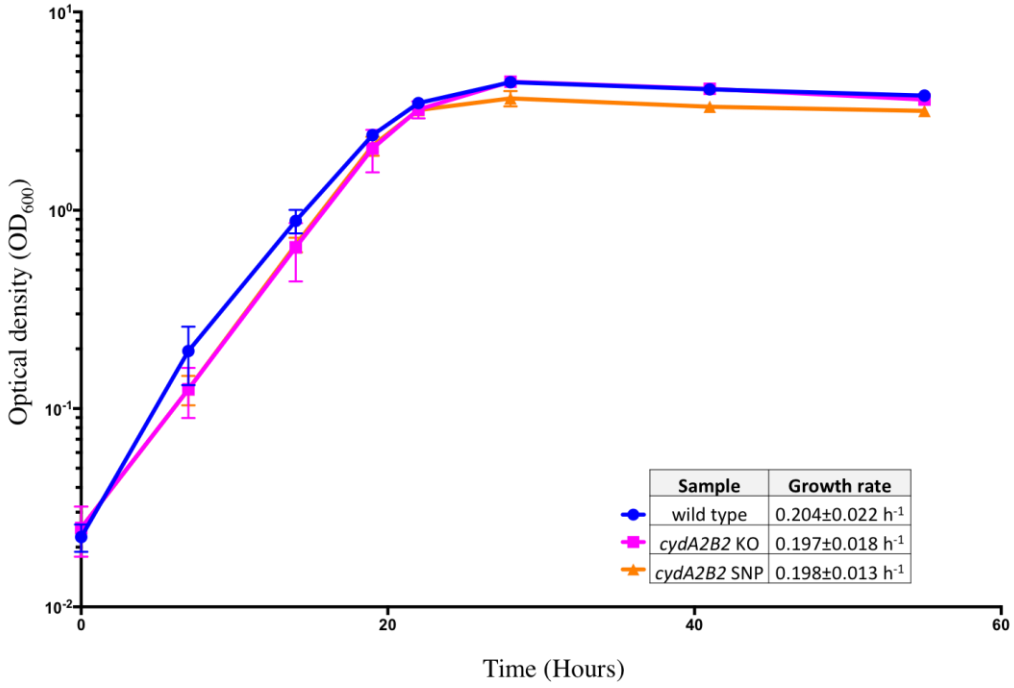


Figure 3.27. Growth of the *cydA2B2* knock-out mutant compared to the H16 wild type and *cydA2B2* SNP strains in the presence of 50% N_2/50% air.

C. necator H16 and its mutant derivatives were grown at 30°C in 20 ml F-MM contained in 150 ml serum bottles under an atmosphere of 50% N_2/50% air (2 bar total pressure). Blue circles, wild type; orange triangles, *C. necator cydA2B2* SNP; pink squares, *C. necator* Δ*cydA2B2*. Inset displays the natural log plot for each sample. Error bars represent the standard deviation of the mean for three independent replicates.

The N_2 experiment was started with an OD_{600} of 0.01, with the growth curves following a time course of 55 hours (Fig. 3.27). The wild type, *cydA2B2* SNP and *cydA2B2* KO strains reached their peak OD_{600} after 28 hours, with growth rates of 0.205±0.022 h^{-1}, 0.198±0.013 h^{-1} and 0.197±0.014 h^{-1}, respectively. There was no significant difference between the knock out strain and the wild type strain ($p > 0.580$), indicating that a lack of cytochrome *bd* ubiquinol oxidase does not noticeably impact on the growth of *C. necator* in the absence of CO.

3.3. Discussion

The overall aim of this study was to isolate and characterise mutant strains of *C. necator* that have adapted to growth with CO, to gain insights into potential molecular and physiological resistance mechanisms as well as strategies to engineer better performing strains. Initial experiments with increasing CO concentrations had revealed a remarkable tolerance of the wild type, but also that growth at industrially relevant concentrations was suboptimal. This finding was expected given that growth of *C. necator* is significantly inhibited by CO concentrations exceeding industry levels (Cypionka and Meyer 1982). Following successive subculturing in the presence of an inhibitory CO concentration, mutant *C. necator* strains exhibiting enhanced growth on CO were isolated. Three strains were initially isolated following eight months of subculturing, and their whole genome analysis revealed an array of mutations found throughout the genome (Table 3.1). Many of the mutations were found in or upstream of genes not suspected to be involved with CO resistance; however, several sites were mutated that appeared to be good candidates for further investigation. An eight months regime was chosen to establish a bench mark, i.e. a time limit for developing maximum CO tolerance (within the time constraints of this study), after which only small incremental increases in tolerance might be expected.

To observe whether mutations in the same gene regions would reappear, the initial subculturing experiment (3.2.2) was repeated, with ten strains being isolated following 15 rounds of re-inoculation. This was also done to establish the overall time course of CO adaptation. Following whole genome analysis, various mutations were found across the ten isolates, many of which did not appear in the original experiment (3.2.2, Table 3.3). However, two of these mutations precisely matched those found previously, including a single point mutation upstream of the *cydA2B2* operon, encoding cytochrome *bd* ubiquinol oxidase (G→A) and an in-frame 12 bp deletion within the *phcR* gene. Surprisingly, the former was common to all isolates in the second evolution experiment (3.2.3) and was found in all three strains from 3.2.2 (Table 3.1 & 3.3) whereas the latter was found in five isolates from 3.2.3 and in two isolates from 3.2.2. There are two possible explanations for the ubiquitous presence of the G→A mutation. Since the starter culture arose from a single colony, it was presumed to be homogenous leading to the assumption that the mutation may have arisen independently across all

evolved isolates. However, the mutation could have arisen very early following pre-culturing of the single colony or was present in the stock culture, meaning that the single colony chosen arose from heterogeneous cell aggregates rather than a single cell. The latter explanation is unlikely it raises the question of why enrichment of this mutation did not happen sooner given its demonstrated increase in growth rate/fitness (Fig. 3.17).

Cytochrome *bd* is a prokaryotic respiratory quinol:O_2 oxidoreductase, which catalyses the reduction of O_2 to 2 H_2O using quinols as reducing substrates (Jünemann 1997, Borisov *et al.* 2011). The enzyme contains three hemes as cofactors, including two *b*-type hemes (hemes b_{558} and b_{595}) and heme *d*, where O_2, NO or CO are able to bind (Giuffrè *et al.* 2014). In various organisms, reduction of O_2 is performed by cytochrome *bd* in stressful conditions such as oxygen-limiting circumstances or iron deficiency (Ingledew & Poole 1984, Poole & Cook 2000, Boristov *et al.* 2009, Trutko *et al.* 2003). The enzyme serves to inhibit the degradation of O_2 sensitive enzymes such as nitrogenases in O_2 poor environments and enhances tolerance to nitrosative stress (Kelly *et al.* 1990, Hill *et al.* 1990, Borisov *et al.* 2004). Increased expression of cytochrome *bd* in *E. coli* has been observed at low O_2 concentrations and other stressful conditions such as media alkalisation, high temperatures and high hydrostatic pressure (Cotter *et al.* 1990, Avetisyan *et al.* 1992, Wall *et al.* 1992, Tamegai *et al.* 1998). CO has been shown to prevent the O_2 reactivity of *E. coli* cytochrome *bd* subunit I but is unable to inhibit the catalase activity of the enzyme (Borisov *et al.* 2013).

C. necator H16 contains two different cytochrome *bd* ubiquinol oxidases, encoded by *cydA1B1* and *cydA2B2* (Pohlmann *et al.* 2006). Since the G→A mutation upstream of *cydA2B2* is the only mutation common to all evolved isolated strains of *C. necator*, it appeared to be the mostly likely candidate to enable increased CO resistance. The mutation potentially affects the expression of the two subunits (I & II) comprising the complex, giving cells with this mutation a growth advantage when in the presence of CO. To support this claim, the mutation was artificially recreated in the H16 wild type background to generate *C. necator cydA2B2* SNP, which exhibited the fast growing phenotype of the evolved isolates, when grown with CO (Fig. 3.17). *C. necator cydA2B2* SNP did not show a growth advantage when cultured in the absence of CO (Fig. 3.18). This finding crucially shows that the enhanced growth phenotype of *C.*

necator cydA2B2 SNP is dependent on the presence of CO and is not a result of the lack of oxygen in the headspace of the used serum flask bottles (containing only 50% air) or an adaptation to other conditions specific to the chosen selection regime, such as use of fructose in minimal medium. If this experiment was to be repeated, serum bottles containing F-MM with 50% N_2 instead of CO could be used to culture *C. necator* in tandem with CO containing cultures. Sequencing of isolates growing in the presence of N_2 would reveal which SNPs are merely the result of the chosen media.

Given that the identified SNP was located 209 bp upstream of *cydA2B2*, it was hypothesised that either it had created a previously unrecognised promoter sequence, or enhances expression from an existing promoter due to higher affinity of RNA polymerase or increased or reduced binding of transcriptional activators or repressors, respectively (Dormeyer *et al.* 2017, Commichau *et al.* 2007, Belitsky *et al.* 2000). The latter hypothesis appears more likely, given the very strong effect this mutation has on operon expression as measured by RFP reporter expression and absence of expression in *E. coli* (Fig. 3.22 & 3.23). Additionally, a promoter site is predicted in the upstream region by BPROM software when this SNP is present (Fig. 3.5). Therefore, increased transcription of *cydA2B2* may lead to higher membrane concentrations of cytochrome *bd* ubiquinol oxidase subunits. This hypothesis is supported by the enhanced growth phenotype of *C. necator* in the presence of CO when *cydA2B2* was overexpressed (Fig. 3.25). Additionally, when *cydA2B2* was inactivated, *C. necator* only displayed increased sensitivity in the presence of CO (Fig. 3.26 & Fig. 3.27). Cytochrome *bd* has been shown to exhibit CO binding properties in *C. necator* and the cytochrome patterns displayed by the organism vary according to different growth phases and conditions (Kömen et al, 1991).

Analysis of *C. necator* protein concentration indicates that cytochrome *bd* is expressed at extremely low levels in early growth phase cells, but substantially increases in concentration in late growth phase cells (Kömen et al, 1991). This expression pattern of presumably both *cyd* homologues (*cydA1B1* and *cydA2B2*), potentially causes wild type *C. necator* to become inhibited by CO during lag phase but the increased expression of *cydA2B2* may enable *C. necator cydA2B2* SNP and the evolved isolates to resist CO and overcome the repression of growth. However, both the evolved *C.*

necator isolates and *C. necator cydA2B2* SNP displayed a reduction in the maximum OD_{600} in comparison to the wild type, a characteristic which has also been observed when *cydAB* is overexpressed in *Corynebacterium glutamicum* (Kabus *et al.* 2007). This phenotype is potentially caused by a bioenergetic inefficiency between cytochrome *bd* oxidase and the cytochrome bc_1-aa_3 supercomplex, which has been documented in *E. coli* (Calhoun *et al.* 1993).

The pattern of cytochrome expression in CODH-engineered *C. necator* possibly hinders the utilisation of CO (Heinrich *et al.* 2018). *C. necator* expressing CODH is unable to oxidise CO following heterotrophic growth, however after growth on syngas containing CO_2, CO oxidation is possible. *C. necator* cytochromes show a moderate affinity for CO when cultivated with CO_2/H_2 (Probst & Schlegel 1976), which suggests that the cytochrome pattern requires time to adapt to CO. The potentially increased concentration of cytochrome *bd* in the artificially mutated and evolved *C. necator* strains may explain this phenomenon. Future characterisation work to be carried out on *cydA2B2* would involve complementing the *cydA2B2* knockout strain to restore the wild type growth phenotype in the presence of CO. The *cydA2B2* overexpression growth comparison would also be repeated with a strain containing a control plasmid with no RFP gene so that the effect of *cydA2B2* upregulation can be more accurately monitored. Characterisation of *cydA1B1* could also be performed to elucidate whether the encoded terminal oxidase is involved with CO tolerance.

The only other mutation found in relation to the respiratory chain was a point mutation (A→G) changing a serine to a proline within cytochrome *c* biogenesis protein ResB in two isolates from 3.2.2 (Table 3.1). Bacterial cytochrome *c* oxidase is a key component of the respiratory chain, which couples proton translocation to oxygen reduction (Iwata 1998), in contrast to cytochrome *bd* oxidases, which have been shown not pump protons (Kolonay *et al.* 1997, Bertsova *et al.* 1997). ResB, along with ResA and ResC form the biogenesis system II, where ResB and ResC form the cytochrome *c* synthase (Simon & Hederstedt 2011). The mutation found in the ResB gene may affect the formation and function of *C. necator* cytochrome *c* in the presence of CO. However, since this mutation was not found following sequencing of the ten isolates in 3.2.3 and was only

found in two of the original isolates, it was not believed to affect CO tolerance so was not characterised in the wild type background or in the absence of CO.

From several other mutations found across the ten isolates in 3.2.3, a pattern emerged suggesting a link between CO resistance and the Phc cell density sensing system of *C. necator*. This system consist of PhcR/PhcS, which regulates PhcA and senses the endogenous signal molecule 3-hydroxypalmitic acid methyl ester (3-OH PAME), produced by PhcB (Clough *et al.* 1997, Flavier *et al.* 1997). The gene for the hybrid sensor histidine kinase/response regulator of the system, PhcR, contained a 12 bp deletion in five isolates and the gene for the LysR- type transcriptional activator, PhcA, contained an in frame point mutation in one isolate and a point mutation upstream (A→C) in another (Table 3.3). PhcR and PhcS together form an atypical two-component system (Clough *et al.* 1997). The system negatively controls levels of active PhcA, whereby the sensor kinase, PhcS phosphorylates PhcR. The enzyme PhcB, synthesises 3-OH PAME, which is sensed by PhcS, regulating the amount of PhcA in a density-dependent manner (Flavier *et al.* 1997). PhcA has been shown to activate motility in *C. necator* (Garg *et al.* 2000). It is not clear whether the 12 bp deletion (four amino acids), corresponds to the receiver or transmitter domain of PhcR, when the *C. necator* protein sequence is compared with the analysed sequence of the closely related *R. solanacearum* (Clough *et al.* 1997). The deletion may either disrupt the phosphorylation of PhcR or prevent it transmitting a signal to PhcA, therefore stopping repression of PhcA. The mutation upstream of *phcA* may cause the gene to become downregulated, allowing decreased activation of downstream genes due to decreased levels of PhcA. When these two mutations were artificially created in the wild type strain, neither resulting mutant displayed a faster growing phenotype compared to the wild type when exposed to CO (Fig 3.17). However, both mutant strains exhibited improved growth than the wild type in the absence of CO (Fig. 3.18), indicating that these mutations must allow *C. necator* to overcome metabolic restrictions, possibly through reduced activity of PhcA which is performing functions not needed under the given growth conditions.

Several mutations found in 3.2.3 were located in or upstream of genes not suspected to be involved with CO resistance (Table 3.3), however the mutations in the RNA

polymerase sigma-54 factor gene (RpoN) improved the growth of *C. necator* with CO and in the absence of CO (Fig 3.17 & 3.18). Again, this indicated that the observed growth improvement was not linked to increased CO resistance. Bacterial sigma-54 factors promote the binding of RNA polymerase to specific promoters, to express a wide variety of genes and has been linked to several functions in *C. necator* including hydrogen oxidation and denitrification (Merrick 1993). RpoN recognises the soluble [Ni–Fe] uptake hydrogenase (SH) promoter of *C. necator,* P_{SH} (Schwartz *et al.* 1998). The mutations found in this gene could negatively influence the binding of RNA polymerase to this hydrogenase and multiple other genes, enhancing the growth of *C. necator,* regardless of whether CO is present or absent, due to a reduced metabolic burden.

In this study, *C. necator* has been evolved to tolerate high concentrations of CO by evolution through successive subculturing. From the mutations identified in the evolved strains, only one point mutation upstream of one of the two cytochrome *bd* ubiquinol oxidase operons, *cydA2B2,* gave *C. necator* increased tolerance to CO. The *C. necator cydA2B2* SNP strain demonstrably shows enhanced CO resistance as it can overcome the inhibitory growth effects CO has on the wild type strain. This strain therefore provides a potential solution to the problem of toxicity posed by fermentation with gas mixtures containing CO such as syngas in *C. necator.* This strain will serve as a chassis for the inclusion of *cox* genes via conjugation of the megaplasmid pHCG3 from *Oligotropha carboxidovorans,* which will be described in the next chapter.

Chapter 4: Conjugative transfer of the megaplasmid pHCG3

4.1. Introduction

Numerous proteobacteria known as 'carboxydovores' possess *cox* genes encoding functional carbon monoxide dehydrogenase (CODH), giving them the ability to oxidise CO to CO_2 (King & Weber 2007). These organisms are unable to assimilate the resulting CO_2 for energy and growth, as they lack a relevant metabolic pathway such as the Calvin-Benson-Bassham (CBB) cycle. Conversely, *C. necator* possess *cbb* genes so can grow efficiently with CO_2 + H_2 as a sole source of carbon and energy (Friedrich & Schwartz 1993) but lacks the ability to oxidise CO, possibly due to the absence of a functional CODH. The carboxydotroph *O. carboxidovorans* has the capacity to utilise CO as a sole source of carbon and energy due to the presence of complete *cox* and *cbb* operons on its megaplasmid pHCG3 (Schübel *et al.* 1995, Fuhrmann *et al.* 2003, Paul *et al.* 2008, Volland *et al.* 2011). The main objective of this project was to give *C. necator* the ability to utilise CO through the transfer of pHCG3. In addition to increased CO tolerance, this was seen as an important property for the organism to acquire since high concentrations of CO are present in several industrial waste gases such as syngas, generated from bio and municipal waste (National Energy Technology Laboratory 2018).

Concurrently with this project, efforts were made to introduce the *cox* gene cluster from pHCG3 into *C. necator* on a shuttle vector and through genomic integration, however neither of these approaches provided *C. necator* with the ability to utilise CO (Alagesan & Malys, unpublished findings). During the course of this project, a study was published documenting a similar approach whereby *C. necator* was able to express plasmid-borne *cox* genes from pHCG3 and oxidise CO (Heinrich *et al.* 2018). However, *C. necator* containing these plasmids could still not use CO as a sole source of carbon and energy. Utilisation of CO for growth was only achieved by supplying *C. necator* with an additional source of hydrogen as a reducing equivalent (Heinrich *et al.* 2018). There are several potential reasons why these engineered strains were unable to utilise the reducing power of CO. The native hydrogenases of *C. necator* may be insufficient to reduce the resulting hydrogen from the oxidation of CO and

might require *hox* genes such as those harboured on pHCG3. This cluster is believed to sense and utilise H_2 hydrogenotrophically, employing a [NiFeS]-hydrogenase encoded by the *hoxSLZ* subcluster (Santiago & Meyer 1997). The previous study by Heinrich *et al.* (2018) used a subset of pHCG3-encoded *cox* genes, including structural genes of the CODH, *coxL* (large subunit), *coxM* (medium subunit) and *coxS* (small subunit) as well as the posttranslational maturation genes of the enzyme's binuclear active site (*coxDEFG*). These constructs excluded several pHCG3-encoded *cox* genes, including *coxC,* which may be involved with transcriptional regulation of the *cox* cluster. Additionally, the genes encoding putative transmembrane proteins CoxB, CoxH and CoxK were not included (Volland *et al.* 2011). None of these proteins have been characterised in *O. carboxidovorans,* but could be involved with anchoring CODH to the cytoplasmic membrane (Fuhrmann *et al.* 2003). Recently, *coxMSLDEFG* were expressed in *E. coli,* with *coxI* and *coxH,* enabling heterologous expression of CODH (Kaufmann *et al.* 2018). This study revealed that CoxI, along with CoxDEFG is essential for the insertion of the molybdopterin cytosine dinucleotide cofactor (MCD) into the bimetallic [MoO2S-Cu] cluster of CoxL. Although *coxBC* and *coxHIK* are not essential for CODH expression in *C. necator* (Heinrich *et al.* 2018), their inclusion may improve the reported sub-optimal efficiency of the enzyme. The transfer of pHCG3 into *C. necator* via conjugation may provide additional as yet unidentified functions potentially required for efficient CO utilisation.

pHCG3 contains a *tra/trb* operon, which encodes putative functions required for plasmid mobilisation and transfer via conjugation (Paul *et al.* 2011, Fuhrmann *et al.* 2003). Contained within this cluster and elsewhere on pHCG3 are genes for LuxRI-family quorum sensing (QS) systems, suggesting that conjugal transfer of this plasmid is controlled by cell-cell communication (Fuhrmann *et al.* 2003, Paul *et al.* 2008, Volland *et al.* 2011). This is similar to what has been observed for the *Agrobacterium tumefaciens* Ti plasmid, to which the *repABC* replicon of pHCG3 is most similar (Fuhrmann et al. 2003). Conjugal transfer of the Ti plasmid requires the transcriptional activator TraR and a population-dependent QS signal, an *N*-acyl-homoserine lactone (AHL) quormone. (Piper *et al.* 1993, Zhang *et al.* 1993). *A. tumefaciens* TraR binds AHLs, produced by the AHL synthase TraI. This dimerises

and stabilises the activator, enabling initiation of transcription of the *tra/trb* operon (Zhu & Winans 1999, 2001, Qin *et al.* 2000, Fuqua & Winans 1996, He *et al.* 2003). QS systems such as these are in some instances also controlled by an antiactivator, TraM, an 11-kDa protein which binds to the DNA-binding domain of TraR (Hwang et al., 1999; Luo et al., 2000; Swiderska et al., 2001). TraM inhibits the activity of TraR, thus preventing transcription of *tra/trb* operon (Fuqua & Winans 1994, Fuqua *et al.* 1995, Hwang *et al.* 1995, Danino *et al.* 2003, Luo *et al.* 2000). When *traM* is mutated, *A. tumefaciens* produces high levels of AHL, which gives a hyperconjugal phenotype for the Ti plasmid (Hwang *et al.* 1995, Piper & Farrand 2000). For efficient transfer of pHCG3 into *C. necator*, deactivation of *traM* was therefore to be attempted in *O. carboxidovorans,* to boost the conjugative capacity of the plasmid.

4.1.1. Aims of this study

The purpose of this study was to examine the conjugative functions of pHCG3 and to ultimately transfer the plasmid into *C. necator.* It was hypothesised that possession of this plasmid might equip *C. necator* with the ability to utilise CO as a sole source of carbon and energy. Specific objectives were:

- Enhancing the conjugative functions of pHCG3, by knocking out the gene for the putative repressor of the quorum sensing system, TraM.
- Characterisation of the pHCG3 AHL synthase genes to elucidate the types of AHL they produce. If required, synthetic versions of these signal molecules might then be added during conjugation experiments to further boost the efficiency of plasmid transfer.
- Equipping the megaplasmid with an antibiotic marker to select for its presence following conjugation.
- To first attempt transfer of modified pHCG3 into *A. tumefaciens*, a relative of *O. carboxidovorans*, to demonstrate that the plasmid can be conjugated within the same proteobacterial subclass before initiating further experiments to achieve transfer and maintenance into *C. necator.*
- Characterisation of putative CO transcription factor binding sites upstream of *coxM* in *O. carboxidovorans* to elucidate whether they are functional in *C. necator* and therefore able to drive expression of the *cox* operon following pHCG3 transfer.

4.2. Results

4.2.1. Analysis of *cox* genes in *C. necator* H16

Before attempting the transfer of the *O. carboxidovorans* pHCG3 megaplasmid, it was of interest to establish whether *C. necator* contained potentially functional *cox* gene homologues, encoding the CODH subunits and associated functions. Identification of these genes and their analysis are described below.

4.2.1.1. Identification of putative *C. necator cox* gene clusters

The complete *cox* gene cluster of *O. carboxidovorans* OM5, located on its megaplasmid pHCG3 contains *coxBCMSLDEFGHIK*, which encodes carbon monoxide dehydrogenase (CODH), functioning to utilise CO (*coxMSL*), in addition to genes required for posttranslational modification of CODH and for possible membrane anchoring of the enzyme (*coxDEFG*) (Fuhrmann *et al.* 2003, Volland *et al.* 2011, Paul *et al.* 2008, Pelzmann *et al.* 2009). Use of the NCBI protein blast server revealed that *C. necator* H16 did not contain a complete *cox* gene cluster analogous to

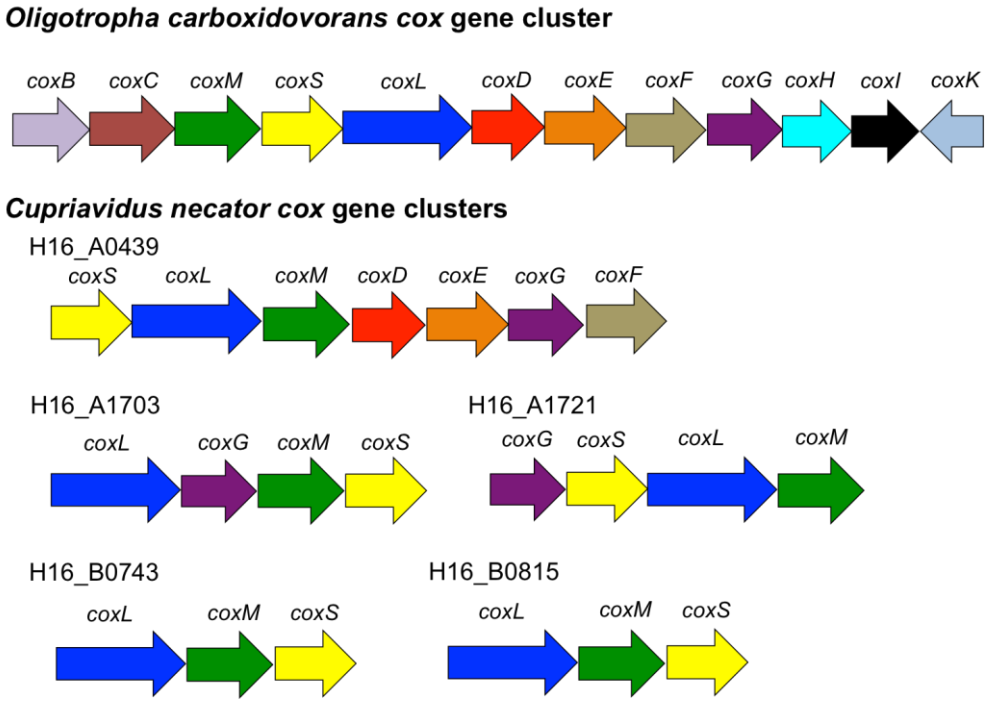


Figure 4.1. Comparison of cox gene clusters identified in *O. carboxidovorans* OM5 and *C. necator* H16.

All *C. necator cox* genes across the five clusters showed at least 50% coverage and 24% identity to *O. carboxidovorans* genes. The *C. necator* cluster beginning with *coxS* (H16_A0439) is the most complete in comparison to *O. carboxidovorans*.

O. carboxidovorans; however, five incomplete clusters were identified which contained putative *cox* genes (Fig. 4.1). Several *coxL* homologues have been found in *C. necator* previously (Cramm *et al.* 2009), however only these five loci contained genes for the medium and small subunits (*coxMS*). All clusters contained genes for the structural components of CO dehydrogenase, CoxS, CoxM and CoxL, however these were all arranged in a different order to *O. carboxidovorans* OM5 (Fig. 4.1). The largest cluster beginning with *coxS* (H16_A0439) contained orthologs of genes encoding the proteins required for posttranslational modification of CoxL*; coxF,* which binds copper forming a complex with *coxE,* a MoxR-like AAA+ ATPase chaperone (Pelzmann *et al.* 2014). Additionally, *coxD* is present, which is responsible for unfolding CoxL to allow its [MoO_3] centre to be modified through the addition of sulphur and copper (Pelzmann *et al*. 2009). These findings are consistent with previous analyses (Cramm *et al.* 2009).

The gene for the membrane-anchoring protein CoxG was also located in the largest cluster and was also found in two shorter clusters, adjacent to one another and with different orientations on chromosome 1 of *C. necator* (H16_A1703 & H16_A1721; Fig. 4.1). Both of these shorter clusters harboured *coxS, coxM* and *coxL* but arranged in a different order to the H16_A0439 cluster and to each other. The three structural CO dehydrogenase genes were also present in two additional *coxLMS* clusters found 81 kb apart on chromosome 2 of *C. necator* (H16_B0743 & H16_B0815) (Fig. 4.1).

4.2.1.2. Sequence alignments

Using Clustal Omega (Sievers & Higgins 2013) and edited using Jalview (Waterhouse *et al.* 2009), the protein sequences for the structural subunits, CoxS, CoxM and CoxL from all five *cox* clusters found in *C. necator* were aligned with the sequences from *O. carboxidovorans* OM5 to determine whether key amino acids in metal ion-binding and FAD-binding sites were present in these proteins. Cox sequences that represented the closest match to *O. carboxidovorans* from *Rhizobium sp.* APP43 and *Labrenzia aggregata* were also included in the alignment for comparison. The bimetallic centre of the active site in the [$CuSMoO_2$]-protein, CoxL, contains copper and molybdenum ion-binding sites (Rovaletti *et al.* 2019, Hille *et al.* 2015, Schübel *et al.* 1995, Dobbek *et al.* 2002, Gnida et al. 2003). In *O. carboxidovorans*, *Rhizobium sp.* APP43 and *L. aggregata* the copper binding site is a

cysteine residue, however in all *C. necator* protein sequences, the amino acid at this position is a glycine (Fig. 4.2), indicating that it has not been conserved, in agreement with previous findings (Cramm *et al.* 2009). However, the molybdenum ion-binding site, which forms part of the molybdopterin cytosine dinucleotide cofactor (MCD) (Dobbek *et al.* 2002), is conserved in all *C. necator* sequences.

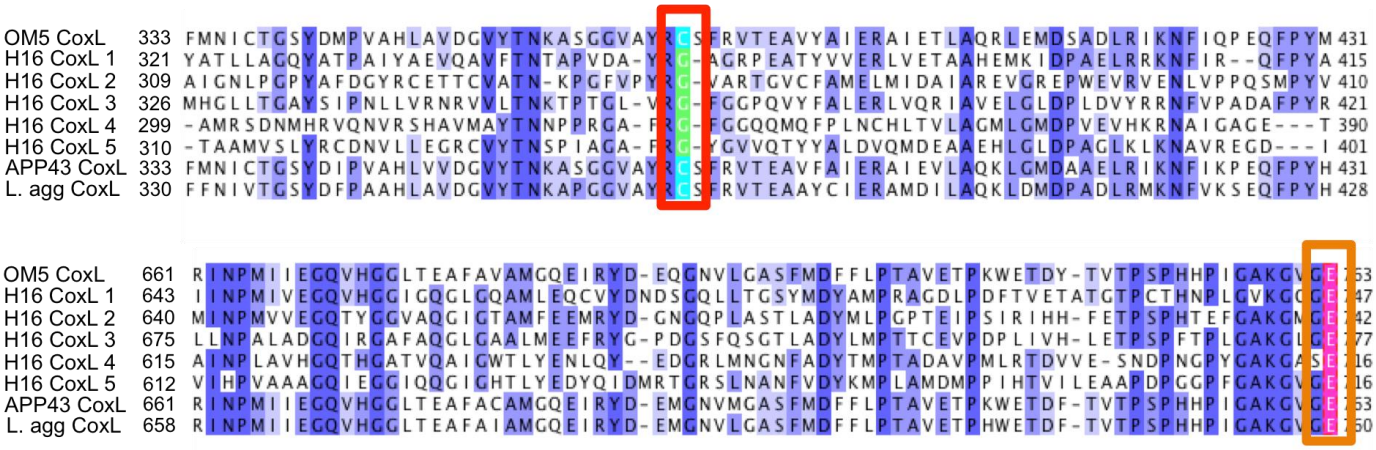


Figure 4.2. CoxL sequence alignment.

Two sections of the protein sequence, which contain ion binding sites in *O. carboxidovorans* CoxL aligned with homologues from each *cox* cluster found in *C. necator* H16 CoxL sequences from clusters, H16_A0439 (CoxL 1), H16_A1703 (CoxL 2), H16_A1721 (CoxL 3), H16_B0743 (CoxL 4) and H16_B0815 (CoxL 5) and from *Rhizobium sp.* APP43 and *L. aggregata*. Copper ion-binding sites are highlighted in red and molybdenum-ion binding site highlighted in orange. Alignment generated using Clustal Omega and Jalview. Residues highlighted in blue indicate positions matching *O. carboxidovorans* consensus sequence, with darker shades reflecting the greater the number of similar residues across all sequences.

The flavoprotein CoxM contains two FAD-binding sites in *O. carboxidovorans;* both consist of five amino acid residues (Dobbek *et al.* 1999). The first site contains alanine, two glycines, histidine and serine in *O. carboxidovorans* and *Rhizobium sp*. APP43. In all five *C. necator* CoxM sequences, the two glycine residues are conserved and the alanine residue is conserved in all sequences but H16_A0439 (Fig. 4.3). Histidine is conserved in only H16_A0439 and serine only in H16_1721. The second FAD-binding site in *O. carboxidovorans* and *Rhizobium sp.* APP43 contains threonine, isoleucine, two glycines and asparagine. Threonine and one glycine residue are conserved in all the sequences, however the other residues are only partially conserved with only one sequence comprising all amino acids in the consensus sequence (H16_A1703) (Fig 4.3). The small subunit, CoxS contains iron-sulphur ion binding sites [2Fe-2S] in *O. carboxidovorans,* the ligands of which consist of eight

cysteine residues (Schübel *et al.* 1995), all of which are conserved in CoxS sequences of *C. necator* (Fig. 4.4).

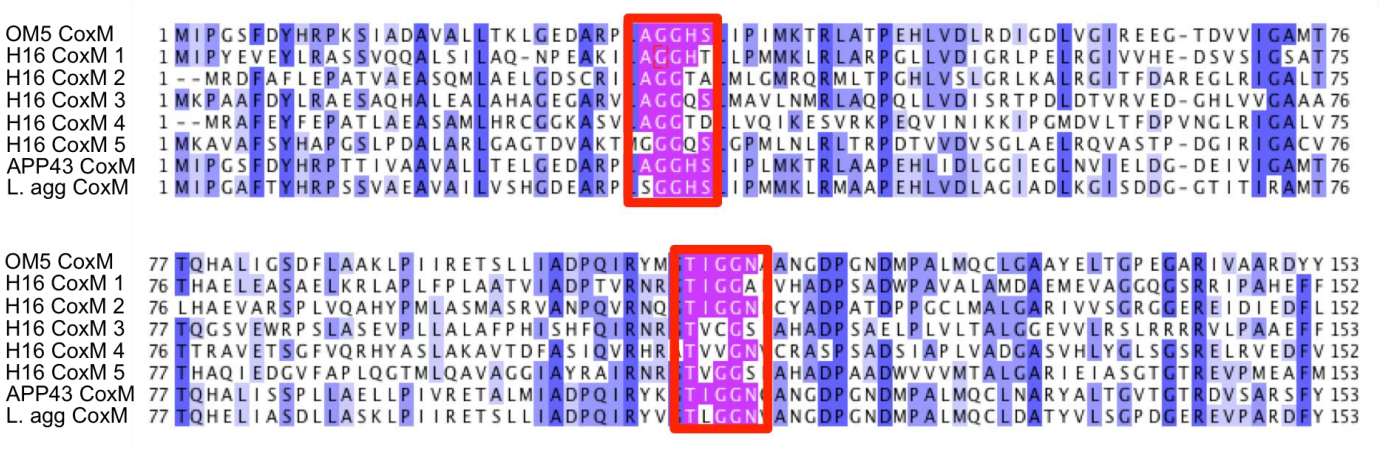


Figure 4.3. CoxM sequence alignment.

The protein sequence of CoxM found in *C. necator* H16, cluster clusters, H16_A0439 (CoxM 1), H16_A1703 (CoxM 2), H16_A1721 (CoxM 3), H16_B0743 (CoxM 4) and H16_B0815 (CoxM 5) and from *Rhizobium sp.* APP43 and *L. aggregata* aligned with CoxM from *O. carboxidovorans* OM5. FAD binding sites are highlighted in red. Alignment generated using Clustal Omega and Jalview. Residues highlighted in blue indicate positions matching *O. carboxidovorans* consensus sequence, with darker shades reflecting the greater the number of similar residues across all sequences.

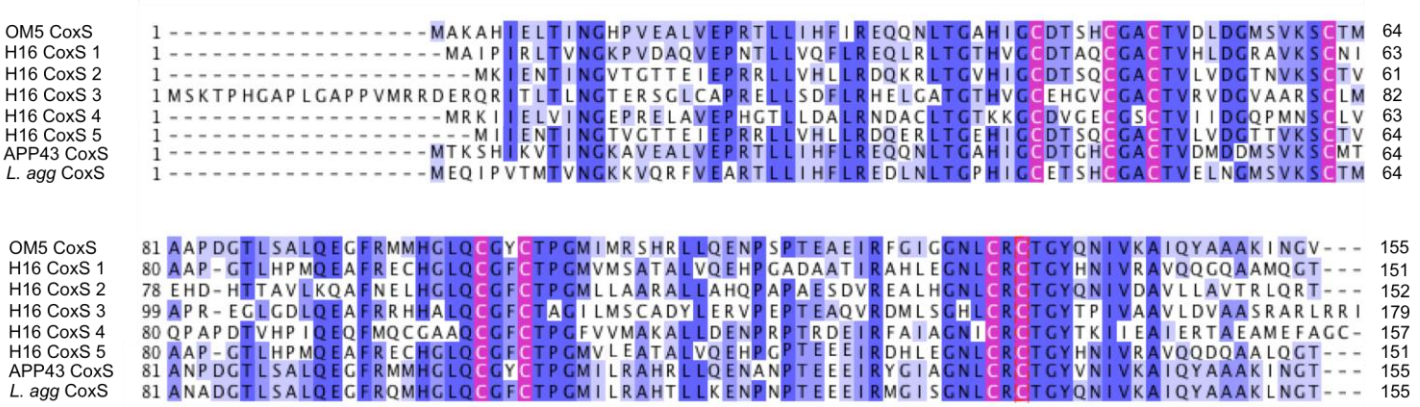


Figure 4.4. CoxS sequence alignment.

Two sections of the CoxS protein found in *C. necator* H16 clusters, H16_A0439 (CoxS 1), H16_A1703 (CoxS 2), H16_A1721 (CoxS 3), H16_B0743 (CoxS 4) and H16_B0815 (CoxS 5) and from *Rhizobium sp.* APP43 and *L. aggregata* aligned with CoxS from *O. carboxidovorans* OM5. Conserved ligands of [2Fe-2S] centres highlighted in purple. Alignment generated using Clustal Omega and Jalview. Residues highlighted in blue indicate positions matching *O. carboxidovorans* consensus sequence, with darker shades reflecting the greater the number of similar residues across all sequences.

The lack of a complete *cox* gene cluster and the fact that the copper ion-binding sites in CoxL are not conserved in *C. necator* may explain why the organism cannot use

CO as a sole carbon source and displays no CODH activity (Heinrich *et al.* 2018). However, since the FAD binding sites in *coxM* are partially conserved, the iron-sulphur ion binding sites in *coxS* are fully conserved and the posttranslational modification genes are present, this suggests that *C. necator* may have been a CO-oxidiser in the past.

4.2.2. Modifying pHCG3 to allow enhanced transfer and efficient detection

Having established that *C. necator* is lacking a number of functional cox genes, there were sufficient grounds to attempt conjugal transfer of the *O. carboxidovorans* pHCG3 megaplasmid into the organism to potentially allow production of active CODH. However, this was seen as a major undertaking requiring careful preparation, including slight modification of the plasmid to increase the chances of success. These included the introduction of an antibiotic resistance marker to allow easy monitoring of successful transfer as well as changes to the postulated conjugal regulatory system to boost transfer ability as described below.

pHCG3 contains two putative AHL synthase genes, *traI* and *solI* (Volland *et al.* 2011). *solI* is located next to a *luxR* homologous gene (*solR*), which presumably encodes the receptor for SolI-produced AHL(s), whereas *traI* is the first gene in the *trb* cluster, suggesting this gene also plays a role in plasmid transfer (Fuhrmann *et al.* 2003). The gene for the transcriptional activator and receptor for TraI-produced AHLs, *traR*, is located between the *tra* and *trb* gene clusters next to the *traM* gene on pHCG3, resembling the system present on the *A. tumefaciens* Ti plasmid, although with a different cluster organisation (Wetzel *et al.* 2015). Expression of *traI* is presumably regulated by the TraR binding activity of TraM, which has been described in *A. tumefaciens* (Hwang et al., 1999; Luo et al., 2000; Swiderska et al., 2001). The type of AHLs both TraI and SolI produce in *O. carboxidovorans* and whether they play a role in the induction of conjugation, via activation of the *tra/trb* operons, are currently unknown. To identify these signal molecules, *traI* and *solI* were cloned into a derivative of pBBR1MCS-2-PphaC-eyfp-c1 (Pfeiffer & Jendrossek 2011) and expressed in *E. coli*. The heterologously produced AHLs were then analysed by bioassay and mass spectrometry.

4.2.2.1. Plasmid construction for overexpression of AHL synthase genes

The AHL synthase genes *traI* and *solI* were amplified by PCR with Phusion polymerase from pHCG3 contained in total DNA preparations of *O. carboxidovorans*, using a forward primer with an NdeI restriction site and a reverse primer with an NheI restriction site (Fig. 4.5). These fragments were ligated separately into the vector pBBR1MCS-2-PphaC-eyfp-c1 (Pfeiffer & Jendrossek 2011), downstream of a strong *E. coli* promoter to create pBBR1_*solI* and pBBR1_*traI*, which were introduced into *E. coli* DH5α (Fig. 4.5). Kanamycin (kan) resistant colonies were selected for after streaking on agar plates and confirmed by colony PCR. Both plasmids were miniprepped and confirmed by Sanger sequencing.

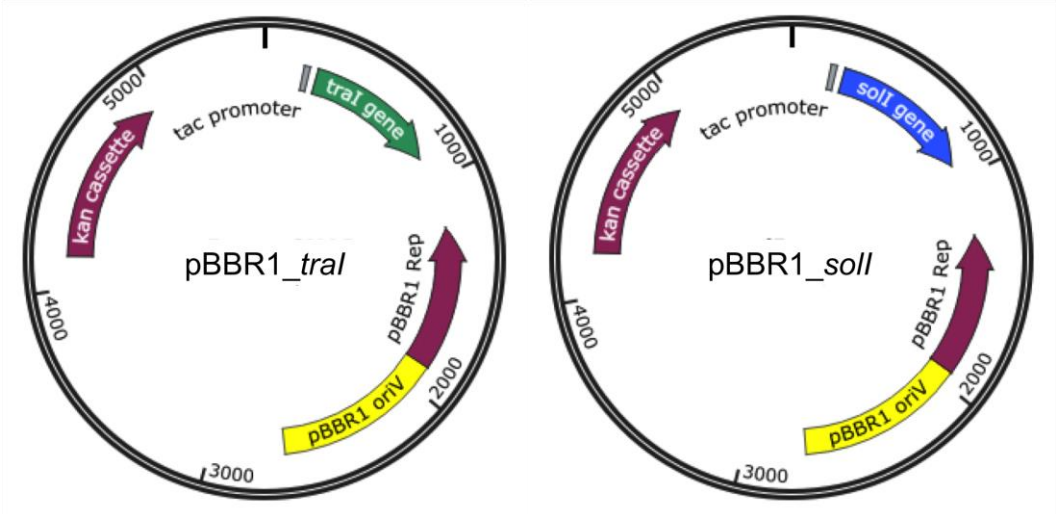


Figure 4.5. Plasmid maps of AHL synthase overexpression vectors pBBR1_traI (right) and pBBR1_solII (left).

solI amplified by PCR using primers SolIF and SolIR and *traI* using TraIF and TraIR, both forward primers containing an NdeI site and both reverse primer containing an NheI site. Both fragments were ligated separately into NdeI and NheI sites on pBBR1_rbs1 using T4 ligase, upstream of the strong tac promoter.

4.2.2.2. Detection of AHLs produced by *E. coli* pBBR1_*traI* and *E. coli* pBBR1_*solI* using a biosensor

To observe whether *E. coli* pBBR1_*traI* and *E. coli* pBBR1_*solI* could produce any detectable levels of AHL, the strains were T-streaked on LB agar plates against a biosensor, *E. coli* DH5α plasmid pSB403 (Fig. 4.6, Winson *et al.* 1998), and incubated at 37^{o}C for 18 hours (Fig. 4.7). This plasmid contains a *luxRI'*::*luxCDABE* fusion, constituting an AHL-inducible reporter system that will expresses luciferase, displaying a bioluminescent phenotype when exposed to AHLs produced by other

species. It is particularly suited for the detection of AHLs C_4, C_6, C_8, C_{10} and C_{12} (Winson *et al.* 1998).

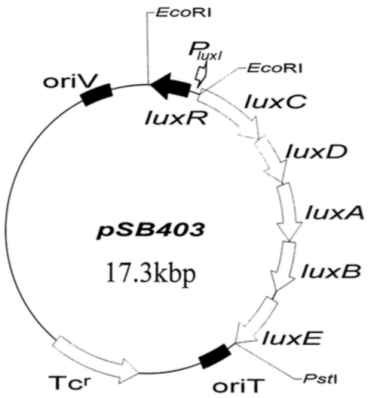


Figure 4.6. Map of the *N*-acyl homoserine lactone sensor plasmid pSB403.

The plasmid contains a *luxRI'*::*luxCDABE* fusion, a broad host range replication origin and a tetracycline resistance marker. The *luxCDABE* genes encode luciferase and functions to generate long-chain fatty aldehyde, the luciferase substrate. The plasmid confers a bioluminescent phenotype in the presence of AHLs activating the encoded LuxR homologue (Figure taken from Winson *et al.* 1998).

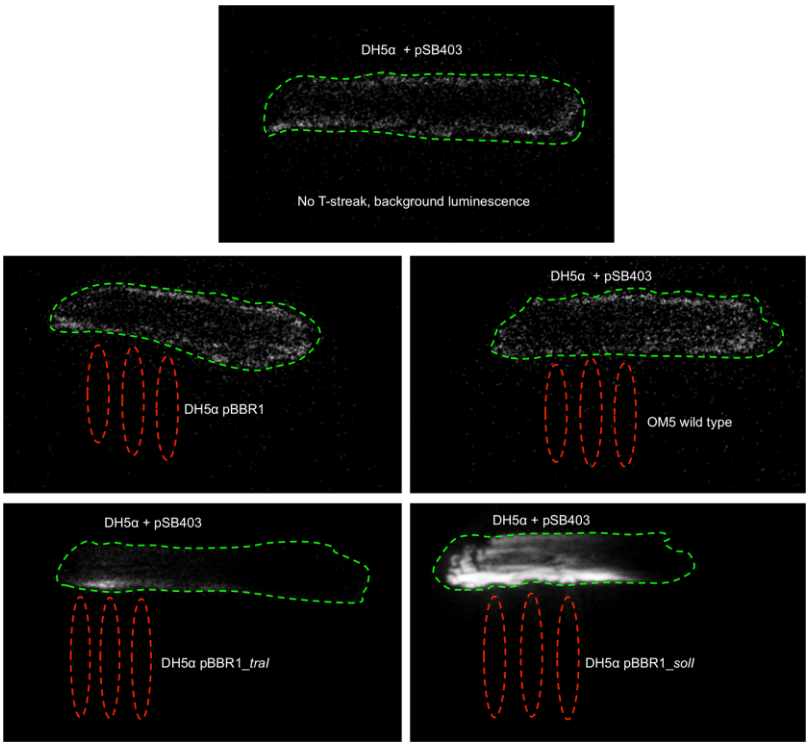


Figure 4.7. Bioluminescent response of *E. coli* DH5α pSB403 when T-streaked against strains overexpressing putative AHLs synthase genes *traI* and *solI*.

E. coli DH5α pSB403 without T-streak (Top) and streaked against DH5α pBBR1 (empty vector) (middle left), *O. carboxidovorans* OM5 wild type (middle right), DH5α pBBR1_*traI* (bottom left), DH5α pBBR1_*solI* (bottom right), the latter two showing increased luminescence at position of T-streak. Red circles indicate position of streaked cells against DH5α pSB403. Images captured by a Hamamatsu imager using Wasabi.

E. coli DH5α pSB403 displays some background bioluminescence in the absence of AHLs but is induced by both *E. coli* pBBR1_*traI* and *E. coli* pBBR1_*solI* cross-streaks (Fig. 4.7). Wild type *O. carboxidovorans* failed to induce any bioluminescence, implying that *traI* and *solI* are not expressed under the tested conditions. *E. coli* pBBR1_*solI* induced stronger luciferase expression than *E. coli* pBBR1_*traI*, which suggest that the former may produce AHLs at a higher concentration or form more active AHL variants.

4.2.2.3. Mass spectrometry analysis of *traI* and *solI*

To elucidate the types of AHLs TraI and SolI produce and at what concentration, the supernatants of *E. coli* pBBR1_*traI* and *E. coli* pBBR1_*solI* cultures were analysed by mass spectrometry. These strains were grown with DH5α pBBR1 (empty vector control) in M9 minimal medium to an OD_{600} of 3 and wild type OM5 in P-MM to an OD_{600} of 0.7, then centrifuged at 8000 rpm.

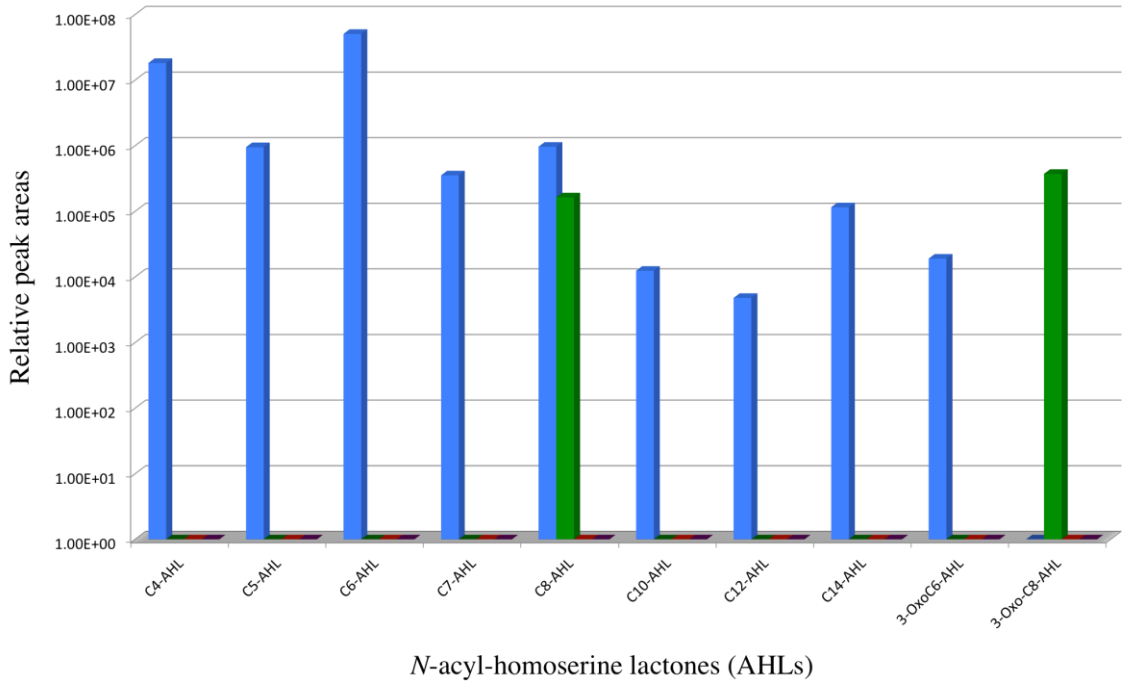


Figure 4.8. Relative amounts of AHLs produced by E. *coli* expressing *traI* and *solI*.
Relative peak areas for specific AHLs present in culture supernatants were determined by LC-MS/MS for *E. coli* producing *O. carboxidovorans* TraI and SolI synthases, respectively. For TraI (green bars), C_8 and 3-oxo-C_8 AHLs were detected at relative peak area sizes of 1.63E+5 and 3.71E+5 respectively. For SolI (blue bars), C_6 and C_4 detected the most at relative peak areas of 5.05E+7 and 1.82E+7. C_5, C_7, C_8, C_{10}, C_{12}, C_{14} and 3-oxo-C_6 were present at relative peak area sizes of 9.47E+5, 3.52E+5, 9.65E+5, 1.24E+4, 4.78E+03, 1.15E+05 and 1.90E+4, respectively. No AHLs were detected in the empty vector control (red bars) or *O. carboxidovorans* supernatants (purple bars).

The supernatants were filter sterilised (0.2 μm) and then analysed by liquid chromatography (LC) mass spectrometry (MS) /MS as described in the Methods (2.4.3) to detect any AHLs synthesised by the plasmid encoded SolI and TraI synthases. LC MS/MS was carried out by Dr Nigel Halliday (University of Nottingham). By screening the LC eluent, AHL-specific peaks were detected in the supernatants for the strains harbouring *traI* and *solI* but none in the empty vector control or wild type OM5 (Fig. 4.8). This indicated the range of AHLs TraI and SolI may synthesise in *O. carboxidovorans*, though not necessarily in the concentrations detected. SolI produced predominantly C_4 and C_6 AHLs but signals were also detected for C_5, C_7, C_8, C_{10}, C_{12}, C_{14} and 3-oxo-C_6 AHLs whereas TraI was very specific and produced only C_8 and 3-oxo-C_8 AHLs (Fig. 4.8).

4.2.2.4. Plasmid construction for a *traM* knockout

Having established the roles of *traI* and *solI* as functional AHL synthases but with no evidence for their expression in their native organism under the employed conditions, the *traM* gene was inactivated to simultaneously boost the expression of AHL synthase genes in *O. carboxidovorans* and to tag pHCG3 with an antibiotic resistance marker. The suicide plasmid pLO3 (Lenz & Friedrich 1998) was chosen as the vector backbone to construct a knock-out plasmid as it contains the pBR322 origin, which is

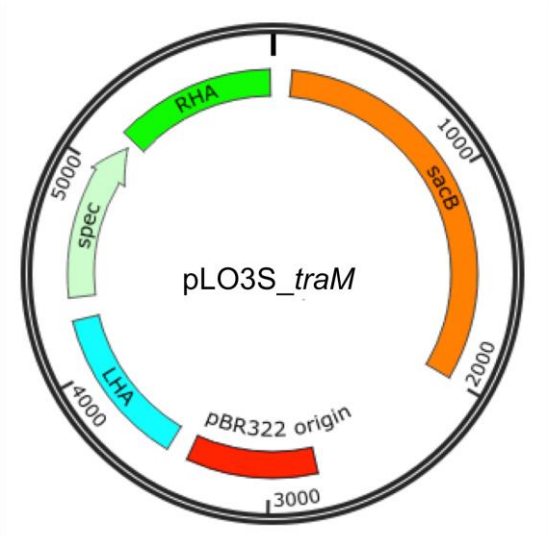


Figure 4.9. Map of pLO3S_*traM*: a suicide plasmid for the replacement of *traM* with a spectinomycin cassette.

The left hand homology arm (LHA: blue) was ligated into SacI and SalI sites upstream of spec cassette and the right hand arm (RHA: green) was ligated into SpeI and AatII sites downstream of the cassette (light green). The plasmid contains a *sacB* counter selection marker and the pMB1 origin from pBR322.

non-functional in *O. carboxidovorans*. This plasmid would replace *traM* with a spectinomycin (spec) resistance marker. The tet region of pLO3 was replaced with spec resistance cassette with additional restriction sites at the 3' and 5' end to insert homology arms (pLO3S). Spectinomycin was chosen as selective agent due to *O. carboxidovorans* showing greater sensitivity towards this antibiotic. Homology arms (750 bp in length) upstream and downstream of the deletion site were amplified from pHCG3 by PCR and ligated into pLO3S with the chosen restriction sites to generate pLO3S_*traM* (Fig. 4.9). The plasmid was introduced into *E. coli* DH5α, miniprepped and used to transform the conjugative *E. coli* strain S17-1.

4.2.2.5. Attempted knock-out of *traM* in *O. carboxidovorans*

Conjugation was performed between *E. coli* S17-1 pLO3S_*traM* and wild type *O. carboxidovorans* OM5 on agar plates containing 75% pyruvate (0.2%) minimal medium (P-MM) and 25% LB (P-MM/LB), a medium on which both organisms could grow sufficiently. Potential transconjugants were selected for on P-MM agar plates containing spectinomycin (150 μg/ml) for presence of the suicide plasmid. After 4-6 days, single colonies were picked and re-streaked twice on selective plates and single crossover was confirmed by colony PCR (Fig. 4.10). For this, two primer pairs were used to verify whether the LHA or RHA had initiated the single crossover integration of the suicide plasmid into the genome (see Materials and Methods for diagrams).

Single crossover mutants were grown in 5 ml P-MM medium supplemented with spectinomycin and allowed to grow to an OD_{600} of about 0.8 μl. Five μl of this culture was used to inoculate 5 ml P-MM with 15% sucrose, which was allowed to grow to an OD_{600} of about 0.8. Serial dilutions of this culture were made with 10^{-5} and 10^{-6} dilutions being plated onto P-MM 15% sucrose agar plates. These steps allowed for the selection of cells in which the second crossover event of the integrated construct had occurred due to *sacB* encoding levansucrase, catalysing the formation of toxic levans in presence of sucrose, as previously described (3.2.4.2). This will result in cells which either have reverted to the wild type form through simple reversal of the single crossover, or have generated the desired mutation.

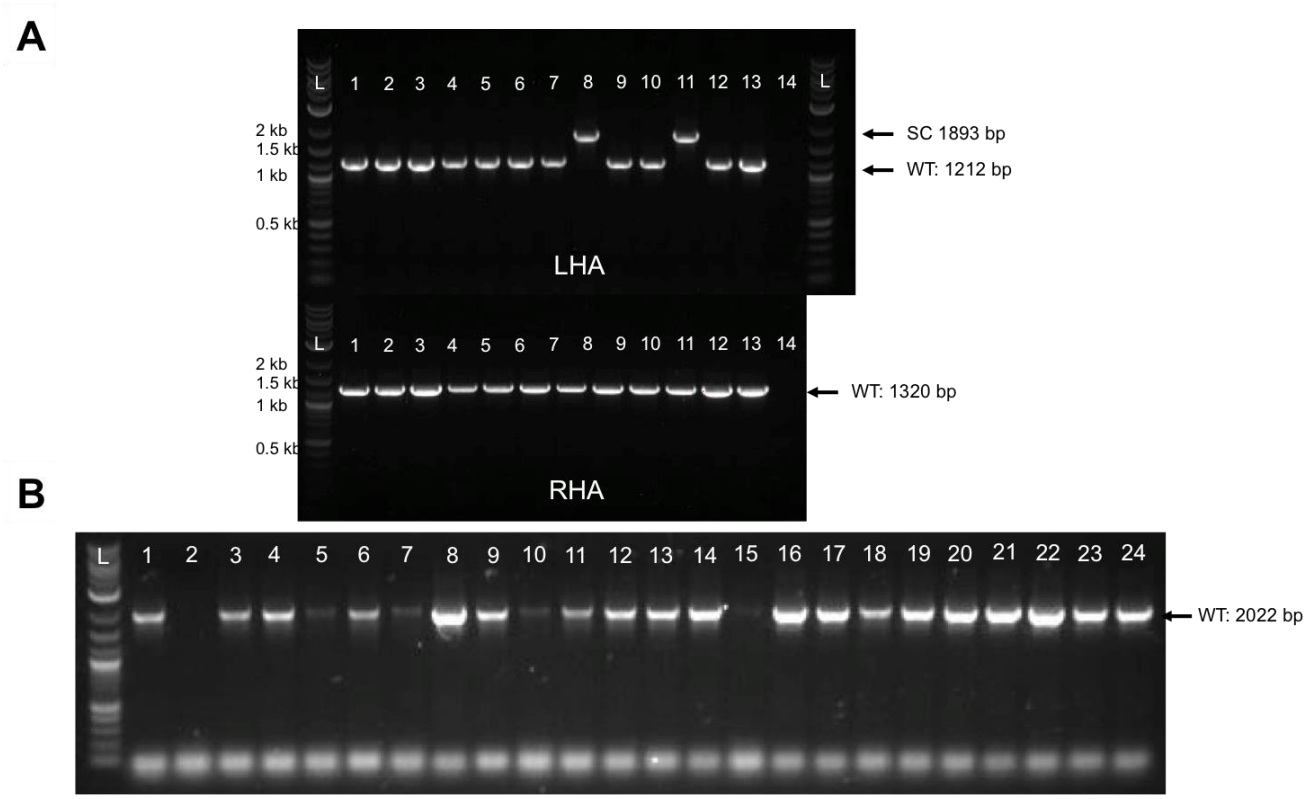


Figure 4.10. PCR confirmation of single crossover (A) and failure of second crossover (B) of pLO3S_*traM* in *O. carboxidovorans.*

Two primer pairs, TraM HA seqF / TraM seqR and TraM seqF / TraM HA seqR, were used to amplify the single crossover region of LHA (A: top) and RHA (A: bottom) regions, respectively. Lanes 1 to 13 contain screened spectinomycin-resistant *O. carboxidovorans* colonies of which 8 and 11 display a mutant phenotype for LHA crossover. Lane 13, negative control with wild type *O. carboxidovorans*; lane 14, blank PCR reaction. Primer pair TraM HA seqF / TraM HA seqR was used to amplify the putative knockout region (B). A subset of screened colonies for second crossover event. Lane 1 contains wild type OM5 and lane 2 contains a blank PCR reaction and lanes 2 to 24 contain possible knockout candidates. The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

After 2 days, single colonies were transferred onto P-MM + spec plates and screened by PCR for the replacement of *traM* with the spectinomycin cassette. All colonies screened displayed the wild type band size (Fig. 4.10) indicating that no double-crossover had taken place. After multiple unsuccessful repetitions with different single crossover mutants, this method was no longer used. A different plasmid, which was concurrently constructed, was used to inactivate *traM* via a single crossover.

4.2.2.6. Plasmid construction for *traM* truncation

A 5' and 3' truncated version of *traM* (*traMt*) was amplified from pHCG3 by PCR using Phusion polymerase and primers containing SacI and XbaI sites (Fig. 4.11). These primers were designed to bind downstream of the *traM* ATG start codon and upstream of the stop codon, creating a truncated gene product which following single crossover integration will generate two truncated non-functional copies of the *traM* protein (Fig. 4.11).

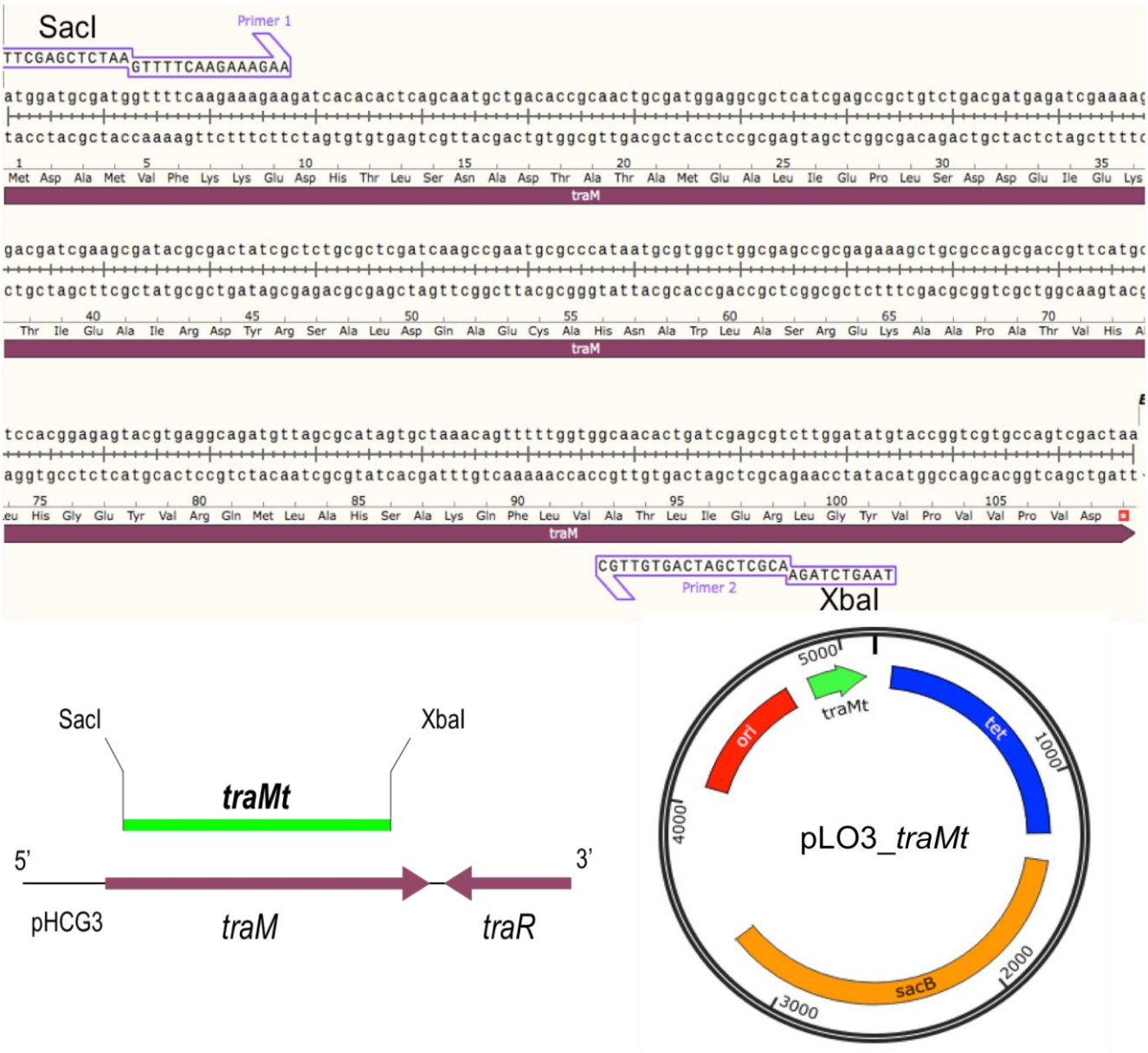


Figure 4.11. Cloning of *traMt* into pLO3.

Schematic of the *traM* sequence depicting the primer binding of TraMt_F and TraMt_R, which were used to amplify *traMt* from pHCG3 (Top). PCR product *traMt* cloned into pLO3 using SacI and XbaI sites with T4 ligase to create pLO3_*traMt* (bottom).

One truncated gene will lack the first four codons of *traM* thus preventing its translation. The other will lack the last 11 codons, eliminating essential residues of the C-terminus from the resulting protein. This fragment was cloned into pLO3 using the aforementioned restriction sites to create pLO3_*traMt* (Fig. 4.11). pLO3_*traMt* was introduced into *E. coli* S-17.

4.2.2.7. Truncation of *traM* in *O. carboxidovorans*

E. coli S-17 pLO3_*traMt* was conjugated with *O. carboxidovorans* on P-MM/LB agar plates at a ratio of 1:1. The mating spot was transferred to P-MM agar plates with tetracycline (150 μg/ml). Resulting colonies were restreaked and screened for single crossover of pLO3_*traMt* at the *traM* locus on pHCG3. Colony PCR was performed using a forward primer annealing upstream of *traM* on pHCG3 and a reverse primer for the tet-marker on pLO3. This produced a 1.8 kb product, which matched the predicted fragment size and was confirmed by Sanger sequencing (Fig. 4.12), whereas no product was obtained for the wild type. The megaplasmid DNA of this strain (OM5_*traM*) was isolated with a commercial genome isolation kit (Sigma Aldrich).

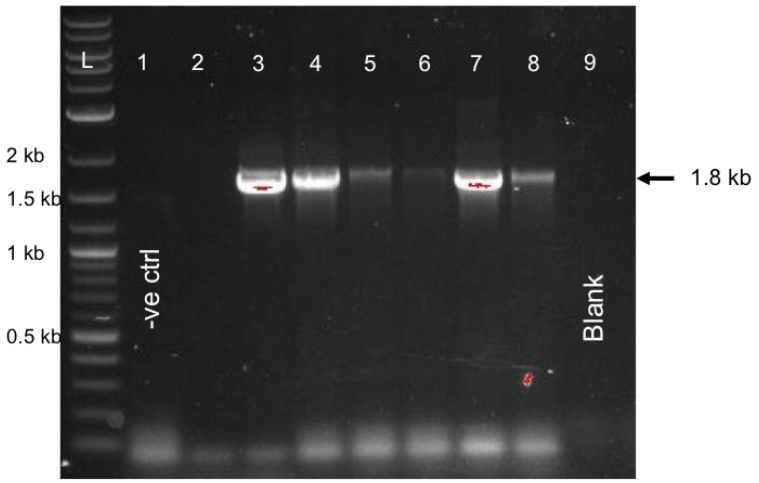


Figure 4.12. PCR confirmation of pLO3_*traMt* single crossover.

Primer TraM seq F binding to the *traM* upstream region and plasmid-binding primer Tet R were used to amplify single crossover regions. Lane 1, wild type *O. carboxidovorans*; lanes 2 to 8, tet resistant single crossover candidates; lane 9, blank PCR reaction. The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

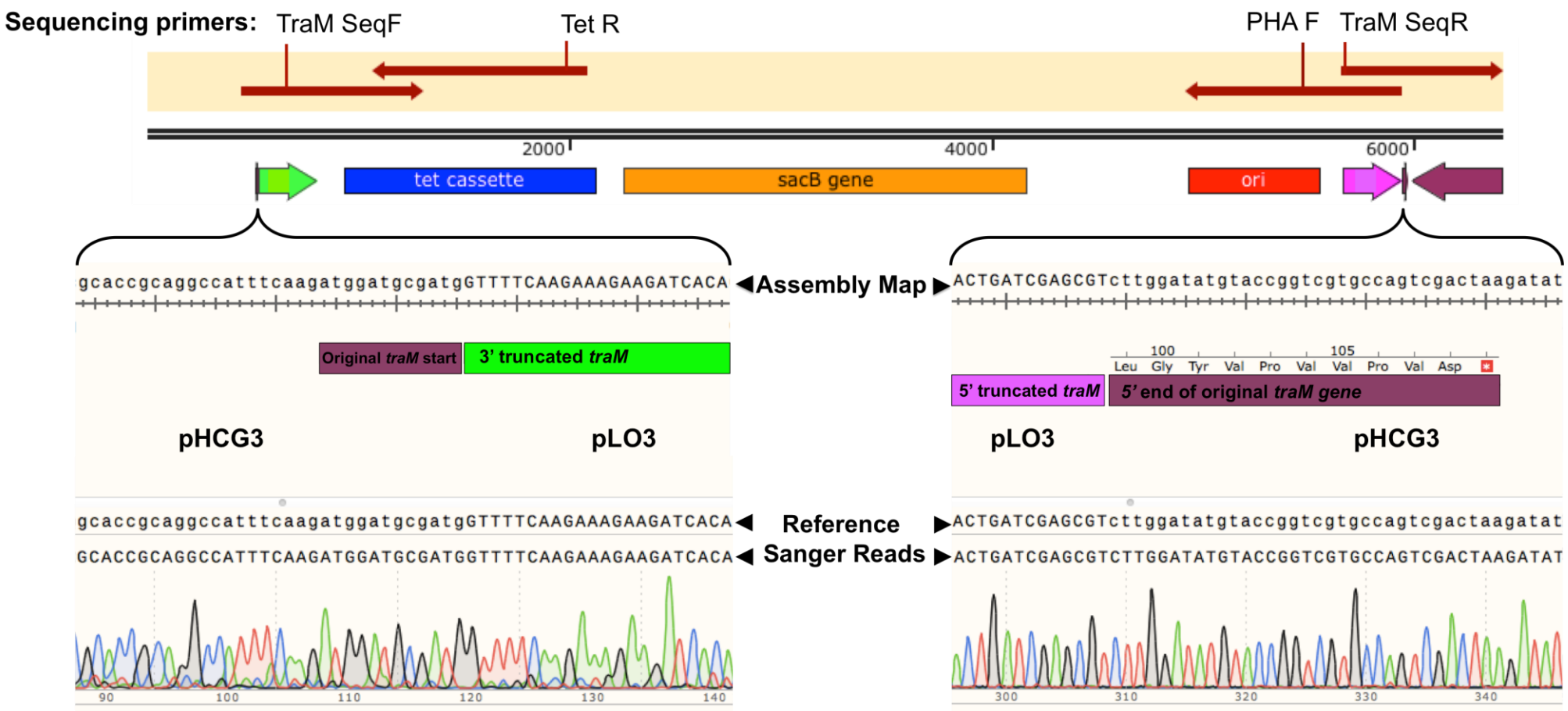


Figure 4.13. Sequencing of pLO3_*traMt* insertion region.

SnapGene reconstruction of single crossover of pLO3_*traMt* into pHCG3 with aligned sequencing reads (red arrows) (top). The two truncated copies of *traM* are indicated in green (A) and pink (B). The original sequence for the 3' *traM* start sequence and 5' end sequence are indicated in purple. Chromatograms of insertion regions at the 3' end of *traM* from TraM seqF primer (bottom left) and the 5' end of *traM* from PHA F primer (bottom right). The intact start ('traM start', left) and end (traM terminator, right) of the original *traM* gene are shown in dark purple.

The entire single crossover region (about 7kb) was amplified by PCR from the genome-prepped DNA using Phusion polymerase and sent for Sanger sequencing to further confirm integration of the pLO3 *traMt* and the resulting generation of two truncated *traM genes.* The predicted sequence of the plasmid knock-in region was created in SnapGene and the sequencing reads obtained were assembled and aligned to this map (Fig. 4.13). The assembled sequence of the mutated region as obtained by Sanger sequencing matched the predicted sequence, confirming that two non-functional copies of *traM* were now present in the mutant, one with a missing start codon and the other with a missing terminal sequence. The generated megaplasmid variant was designated pHCG3_*traM* and the strain carrying it OM5_*traM*. Whilst mutation via single crossover may not be viewed as ideal, due to an expected lack of stability, it was considered an advantage in the current study, as its potential reversibility would allow simple removal of the tet resistance marker and restoration of the native *traM* in the envisaged target organism, *C. necator*.

4.2.2.8. Mass spectrometry analysis of AHLs produced by OM5_*traM*

If both *traM* copies had been truncated as described above and were therefore non-functional in *O. carboxidovorans* OM5_*traM*, AHL production should occur provided native TraM acts as anti-activator of TraR and the *tra*/*trb* operon in this species, too, analogous to its described role in *A. tumefaciens* (Fuqua & Winans 1994, Fuqua et al. 1995, Hwang et al. 1995, Danino et al. 2003, Luo et al. 2000). To test this hypothesis, OM5_*traM* and wild type *O. carboxidovorans* OM5 were grown in P-MM to an OD_{600} of 0.8 and centrifuged at 8000 rpm for 5 minutes. The supernatants were filter-sterilised and analysed by LC-MS/MS. Whilst no AHL-specific signals were detected for the wild type, OM5_*traM* supernatants were found to contain C_8, 3-OH-C_8, 3-oxo-C_8, C6, 3-OH-C_6, 3-oxo-C_6, C_{10}, 3-OH-C_{10} and 3-oxo-C_{10} (Fig. 4.14).

The largest peak areas detected were for C_8 and 3-oxo-C_8, which matched the AHL profile observed in *E. coli* pBBR1_*traI*, though at different concentrations. These results suggested that in agreement with the above hypothesis truncated TraM was unable to bind to and inhibit TraR thus enhancing the expression of *traI* and boosting the production of AHLs. In analogy to what has been observed for *A. tumefaciens* Ti plasmids (Hwang *et al.* 1995, Piper & Farrand 2000), the increased production of AHLs in OM5_*traM* together with the lack of TraR inhibition was hypothesised to

enhance the conjugative capacity of pHCG3. Therefore, OM5_*traM* was used in mating experiments with *C. necator* and *A. tumefaciens* to determine whether pHCG3_*traM* is transferrable by conjugation between species. Strains containing pHCG3_*traM* can be selected for using tetracycline and tested for the presence of the plasmid by PCR screening of several regions. *A. tumefaciens* C58 was chosen for this purpose, as avirulent and Ti plasmid-free variant strains were available (A676 and A677) which in previous work had been shown to accept rhizobial megaplasmids following interspecies conjugation (Danino *et al.* 2003).

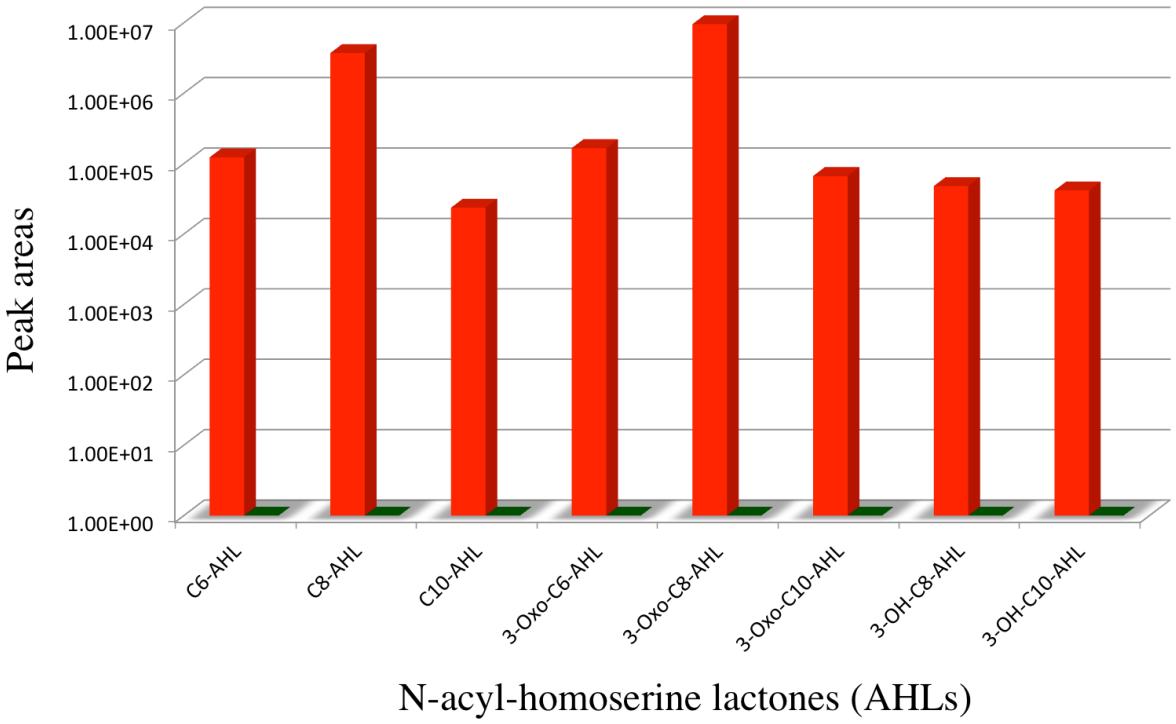


Figure 4.14. Mass spectrometry identification of AHLs produced by OM5_*traM*.
Relative peak areas for specific AHLs present in culture supernatants were determined by LC-MS/MS. OM5_*traM* displayed signals specific for C_8, 3-oxo-C_8 and 3-OH-C_8 AHL with relative peak areas of 3.72E+6, 9.56E+6 and 4.80E+4, respectively; C_6 and 3-oxo-C_6 AHL signals showed relative peak areas of 1.23E+5 and 1.66E+5, respectively; C_{10}, 3-oxo-C_{10} and 3-OH-C_{10} AHL signals showed relative peak areas of 2.37E+4, 1.66E+5 and 4.16E+4, respectively. No peak areas were detected for wild type OM5 (green bars).

4.2.3. Investigation of pHCG3 transfer into *A. tumefaciens*

The purpose of transferring pHCG3_*traM* into *A. tumefaciens* was twofold; firstly to prove that the plasmid was transferrable between proteobacterial species of the same subclass and secondly to house the plasmid in a suitable intermediate species for further genetic manipulation such as integration of a *C. necator*-compatible origin of replication. For instance, integration of *rep28* via pLO3 into pHCG3_*traM* could then be attempted using single and double crossover with homology arms corresponding to a suitable locus on pHCG3, via *sacB*-mediated counter selection. The latter method was unsuccessful in *O. carboxidovorans* in this study (4.2.2.5), however it has been proven to work in *A. tumefaciens* (Morton & Fuqau 2012). A676 and A677 were ideal candidate strains for this work because they carried additional resistances towards rifampicin and streptomycin, respectively, which could be used to eliminate the conjugal OM5 pHCG3_*traM* donor. AHL production, if successful transfer and expression of pHCG3_*traM* genes were achieved, also could be monitored since these strains are presumably unable to produce their own AHLs due to the lack of the native Ti plasmid. Additionally CO and $CO_2 + H_2$ utilisation could be monitored, since they likely cannot utilise these substrates for growth as the genome of *A. tumefaciens* C58 does not harbour any *cox* or *cbb* genes (Goodner *et al.* 2001).

4.2.3.1. Conjugation with *A. tumefaciens*

OM5_*traM* was conjugated with plasmid-less *A. tumefaciens* strains A676 and A677 (Danino *et al.* 2003) using the conjugation method previously described (2.2.5.3). Following incubation for 18 hours, the mating spot was resuspended from the filter in 0.9 ml PBS and transferred to glucose minimal medium (G-MM; 0.2% glucose) + tet agar plates to select for A676 (+ rifampicin at 20 µg/ml) or A677 (+ streptomycin at 400 µg/ml) transconjugants, since *O. carboxidovorans* is unable to grow using glucose as a carbon source. The resulting colonies growing on the selective media were restreaked twice onto fresh plates. These colonies were screened by PCR using primers 40seq F and 40seq R, which anneal to a pHCG3 encoded integrase gene (locus tag: OCA5_pHCG300040) for a first preliminary confirmation of successful transfer (Fig. 4.15). The conjugation frequency was taken as the number of transconjugants per 10^8 recipients (Danino *et al.* 2003) (Table 4.1). The frequency of transfer was 9.41×10^{-8} and 4.76×10^{-8} per recipient for A676 and A677 respectively.

4.2.3.2. PCR confirmation of plasmid transfer into *A. tumefaciens* and absence of the conjugal *O. carboxidovorans* donor.

Table 4.1. Conjugation frequency of pHCG3_*traM* for *A. tumefaciens* A676 and A677

Strain	No. of recipients per conjugation	Average No. of positive colonies per conjugation	Conjugation frequency	Standard deviation	No. of colonies for negative control
A676	8.50×10^{7}	8	9.41×10^{-8}	$\pm 1.66 \times 10^{-8}$	0
A677	1.05×10^{8}	5	4.76×10^{-8}	$\pm 2.69 \times 10^{-8}$	0

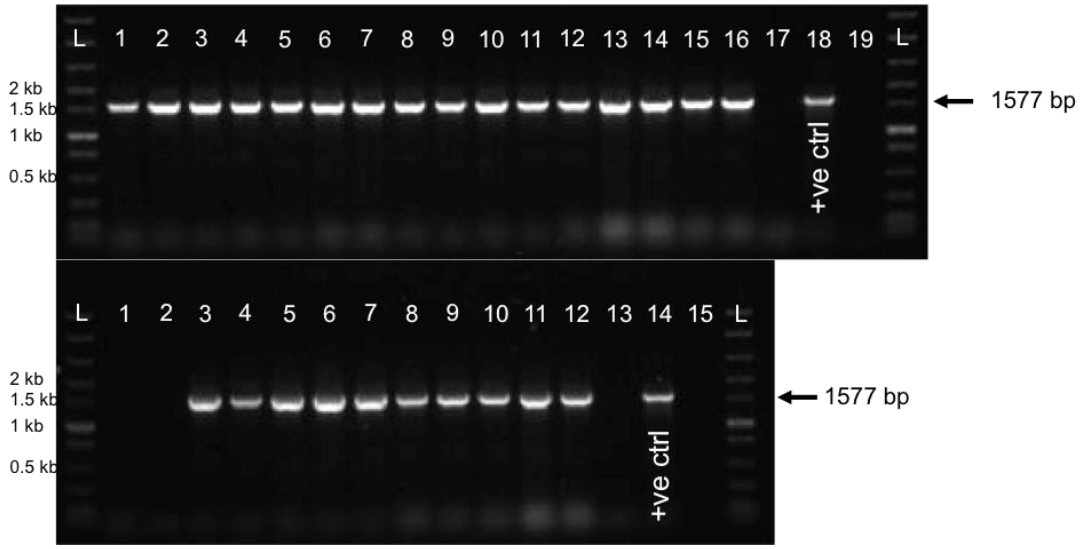


Figure 4.15. PCR confirmation of pHCG3_*traM* presence in A676/A677.
The integrase gene OCA5_pHCG300040 (1577 bp) was amplified from pHCG3 using primer pair 40 seqF / 40 seqR. Samples screened in lanes 1 to 16 for A676 (top), with wild type A676 control (lane 17), OM5_*traM* (lane 18) and a blank PCR reaction (lane 19). Samples screened in lane 1 to 12 for A677 (bottom), with wild type A677 control (lane 13), OM5_*traM* (lane 14) and a blank PCR reaction (lane 15). The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

A subset of selected positive colonies were screened using two primer pairs (FabG F / FabG R and FtsA F / FtsA R) annealing to genes on the circular chromosome of *A. tumefaciens* and two primer pairs (RecF F / RecF R and Rbfa F / Rbfa R) annealing to genes on the circular chromosome of *O. carboxidovorans* to confirm that the obtained isolates were indeed *A. tumefaciens* and that there was no conjugal donor contamination in the samples screened (Fig 4.16 & 4.17). The former genes were chosen, as there were no matching homologues in *O. carboxidovorans.* As expected, DNA fragments of identical size could be amplified from the obtained transconjugant

DNA with the *A. tumefaciens*-specific primer sets, whereas no PCR fragments or unspecific fragments of much reduced size were obtained using the *O. carboxidovorans*-specific primer pairs.

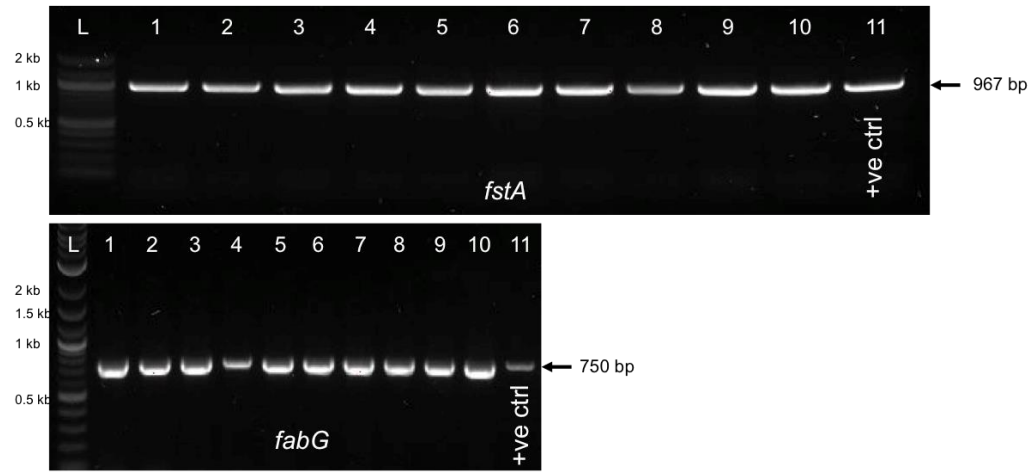


Figure 4.16. PCR confirmation of A676/A677 transconjugant chromosomal DNA.
PCR reactions using primers FabG F & Fab R (Atu3026, top panel) and FtsA F & FtsA R (Atu2087, bottom panel), which both anneal to the *A. tumefaciens* circular chromosome. Samples in lanes 1-5 contain A676 transconjugants and lanes 6-10 contain A677 transconjugants and display bands corresponding to the positive control (lane 11, parental A676). The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

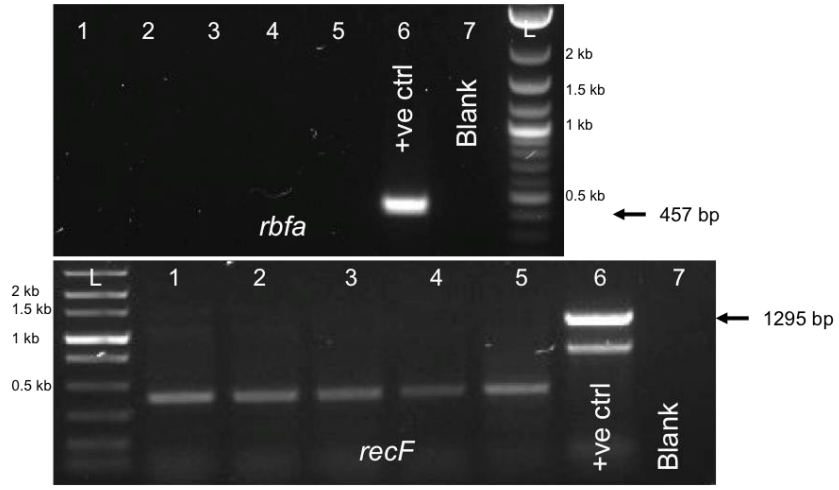


Figure 4.17. PCR analysis to confirm absence of contaminating *O. carboxidovorans* DNA in isolated transconjugant A676/A676 chromosomal DNA.
Two PCR reactions using primers RecF F & R (specific for *recF*, locus tag OCA5_c00040, top panel) and RfbA F & R (specific for *rfbA*, locus tag OCA5_c00280, bottom panel) which both anneal to the *O. carboxidovorans* chromosome. Samples in lanes 1-5 do not display bands corresponding to the positive control (lane 6, wild type *O. carboxidovorans*); lane 7 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

Selected positive colonies confirmed as described above were further screened by PCR and the resulting products subjected to Sanger sequencing to confirm the presence of four distinctly separate regions of pHCG3_*traM*. The primer pairs 40_seq F / 40_seq R, 220_seqF / 220_seq R, 880seqF /880_seqR and 1000_seq F / 1000_seq R were directed against pHCG3 genes for an integrase, hypothetical protein A, hypothetical protein B and hypothetical protein C (Fig. 4.18).

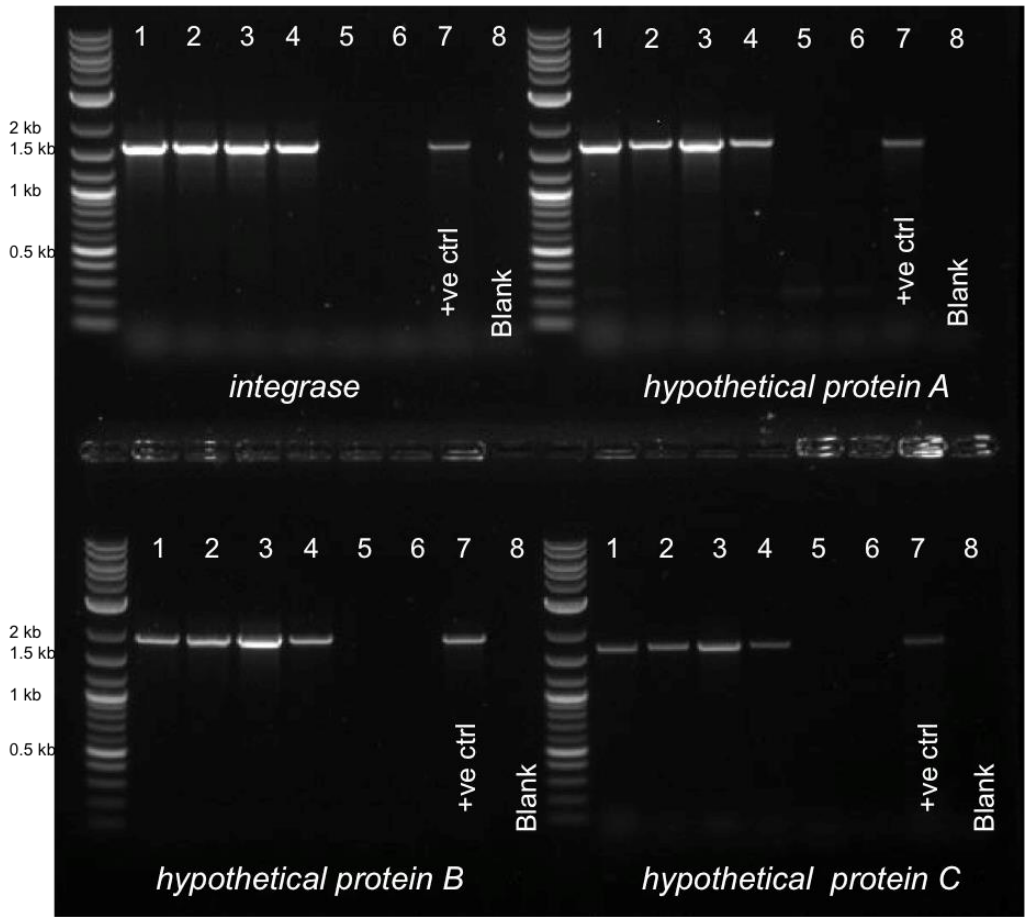


Figure 4.18. PCR confirmation of multiple regions of pHCG3_traM in A676/A677.
Integrase gene fragment, 1577 bp (OCA5_pHCG300040) (top left); hypothetical protein A gene, 1582 bp (OCA5_pHCG300220) (top right); hypothetical protein B gene, 1839 bp (OCA5_pHCG880) (bottom left); hypothetical protein C gene, 1657 bp (OCA5_pHCG301000) (bottom right). For each region of pHCG3_*traM*, lanes 1 and 2 contain A676 pHCG3_ *traM* and lanes 3 and 4 contain A677 pHCG3_ *traM*, all of which show a positive band corresponding in length to the positive control for all regions screened (wild-type OM5, lane 7). Lanes 5 and 6 contain the negative controls wild type A676 and A677 respectively and lane 7 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L). Sanger sequencing of the obtained PCR products confirmed that the above genes had been correctly amplified in all cases.

Furthermore, the 16S ribosomal RNA gene was amplified from each positive transconjugant using primer pair U113F and U191R and sent for Sanger sequencing, which confirmed that the samples were *A. tumefaciens.* The total DNA of two confirmed isolates (representing an A676 and an A677 transconjugant, respectively) and the parental A676 strain was isolated as described in the Methods (2.3.1.2).

4.2.3.3. Whole genome sequencing

Illumina next generation sequencing was performed on the selected A676 pHCG3_*traM* and A677 pHCG3_*traM* isolates and the parental A676 strain. This was carried out to compare the genomes of the transformed strains with the original wild type in order to further confirm presence and integrity of the megaplasmid in *A. tumefaciens*. The sequences were processed using the software CLC Genomics Workbench 11 and mapped onto the most recently published *O. carboxidovorans* genome (Volland *et al.* 2011) as described in the Methods (2.6.3).

Table 4.2. Illumina sequencing reads of A676/677 pHCG3_*traM* mapped to the wild type *O. carboxidovorans* OM5 and *A. tumefaciens* A676 genomes.

Reference covered given as a percentage of the reference covered by reads. Average coverage given as the average number of reads at any position on the reference.

Reference	A676 pHCG3_*traM*	A677 pHCG3_*traM*
OM5 pHCG3		
Reference covered (%)	100	100
Average coverage	478.75	248.11
SNPs	0	2
OM5 circular chromosome		
Reference covered (%)	11	10
Average coverage	4.24	2.1
SNPs	/	/
A676 circular chromosome		
Reference covered (%)	100	100
Average coverage	40.93	81.76
SNPs	79	73

This mapping indicated that pHCG3_*traM* was present whereas OM5 chromosomes 1 and 2 were absent from the gDNA samples of isolates A676 pHCG3_*traM* and A677 pHCG3_*traM* (Table 4.2, see Appendix for read coverage graphs). Although there was some suggested coverage for these chromosomes, further analysis revealed that

the responsible reads were in fact mapped to matching stretches of DNA within homologous genes of the *A. tumefaciens* chromosome, which was expected given how closely related the two bacteria are. This demonstrated that there was no bacterial contamination from *O. carboxidovorans* leading to a false positive result for pHCG3_*traM* presence. When mapped onto the *A. tumefaciens* C58 genome, reads for A676 pHCG3_*traM* and A677 pHCG3_*traM* only aligned with the circular chromosome (Table 4.2, see appendix). A676/677 pHCG3_*traM* did not map to the At or Ti plasmid of *A. tumefaciens*, since the latter have been cured of both of these plasmids. Following mapping onto the *A. tumefaciens* C58 genome and analysis of unique mutations, A676 pHCG3_*traM* and A677 pHCG3_*traM* showed several unique SNPs in comparison to parental A676 which might have occurred following conjugation (Table. 4.2). These SNPs were not identical in both strains but were located in similar regions across the circular genome. Parental A676 also displayed a number of unique mutations in comparison to the published C58 genome, which were also present in A676 pHCG3_*traM* and A677 pHCG3_*traM* (data not shown).

4.2.3.4. Monitoring of autotrophic growth with CO and CO_2

To observe whether presence of pHCG3_*traM* enabled *A. tumefaciens* to grow using CO or CO_2 as a sole or additional carbon sauce, growth comparisons were performed between wild type A676 and A676 pHCG3_*traM* (Fig. 4.19). Three colonies of A676 pHCG3_*traM* and three colonies of parental A676 were cultured in 5 ml G-MM + tet for 18 hours and spun down at 5000 rpm for 4 minutes. The resulting cell pellets were washed with PBS twice to remove any residual media containing glucose. The obtained cell pellets were resuspended and split into serum bottles containing a growth limiting concentration of glucose (0.02%, approximately 1.1 mM), allowing growth to an OD_{600} of about 0.2. The bottles contained MM + tet (15 μg/ml) for A676 pHCG3, provided in 20 μl of ethanol to select for megaplasmid maintenance, and an equivalent amount of ethanol for the parental A676 control culture with either 50% CO/50% air or 50% N_2/50% air or 80% H_2/10% CO_2/10% air. The growth comparison was performed with a starting OD_{600} of 0.01.

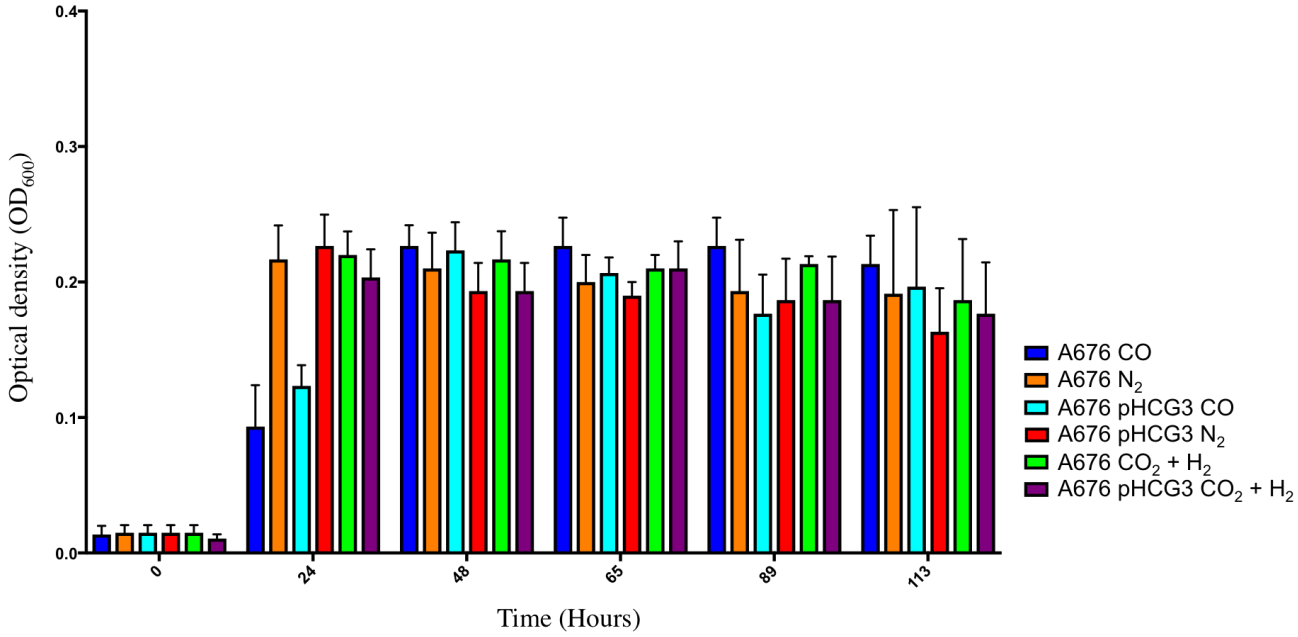


Figure 4.19. Growth of A676 pHCG3_*traM* in comparison to the parental strain A676 with either CO or CO_2 + H_2.

A676 was grown at 30 °C in 20 ml MM containing 0.02% glucose and contained in 150 ml serum bottles under an atmosphere of 50% N_2/50% air ("N_2") or 50% CO/50% air ("CO") and 80% H_2 /10% CO_2/10% air ("CO_2+H_2") (2 bar total pressure). Blue bars, parental A676 (CO); green bars, parental A676 (CO_2+ H_2); orange bars, parental A676 (N_2); cyan bars A676 pHCG3_*traM* (CO); purple bars, A676 pHCG3_*traM* (CO_2+H_2); red bars, A676 pHCG3_*traM* (N_2). Error bars represent the standard deviation of the mean for three independent replicates.

The growth curves followed a time course of 113 hours, with wild type A676 and A676 pHCG3_*traM* showing some growth in all conditions. However, growth for both strains was similar under all conditions and neither the presence of CO or CO_2 allowed cultures to grow to an OD_{600} greater than when just N_2 was present (Fig. 4.19). This indicated that the presence of pHCG3 did not give *A. tumefaciens* the ability to utilise CO or CO_2 as a carbon source for growth or at least not to the extent that it visibly improved growth.

4.2.3.5. Mass spectrometry analysis of AHLs produced by A676

To observe whether presence of pHCG3_*traM* enabled *A. tumefaciens* to produce AHLs (as described in section 4.2.2.8. for *O. carboxidovorans*), A676 pHCG3_*traM* and A677 pHCG3_*traM* and parental A676 were grown in G-MM (0.2%) to an OD_{600} of 1.8 and centrifuged at 8000 rpm for 5 minutes.

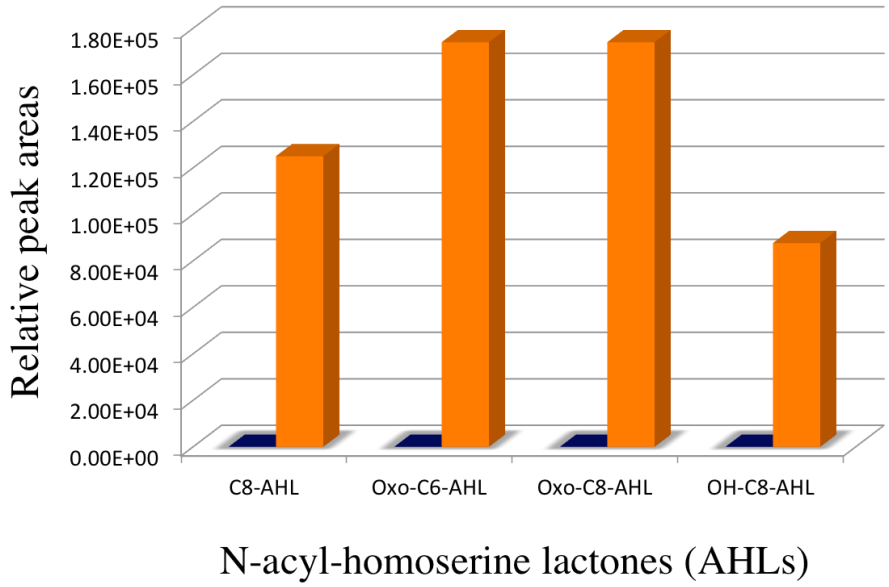


Figure 4.20. Mass spectrometry detection of AHLs produced by A676 pHCG3_*traM*.
Relative peak areas for specific AHLs present in culture supernatants were determined by LC-MS/MS. No peak areas were detected for parental A676 (blue bars) but A676_*traM* (orange bars) displayed specific signals for C_8, 3-oxo-C_8, 3-OH-C_8 and 3- OH-C_6 with relative peak of 1.25 x 10^5, 1.74 x 10^5, 8.76 x 10^4 and 1.74 x 10^5, respectively.

The supernatants were filter-sterilised and analysed by LC-MS/MS (Fig. 4.20) as previously described by Dr Nigel Halliday (University of Nottingham). No AHL-specific signals were detected for parental A676, which was expected given that it is cured of its AHL-producing plasmids (Danino *et al.* 2003). A676 pHCG3_*traM* supernatants were found to contain C_8, 3-OH-C_8, 3-oxo-C_8 and 3-oxo-C6. This indicated that the QS genes are being actively transcribed in *A. tumefaciens* and the *traM* truncation is still intact as demonstrated by whole genome sequencing.

4.2.4. Investigation of pHCG3_*traM* transfer into a *Pseudomonas* species

During the conjugation experiments described in 4.2.3.1., the first attempted conjugation yielded numerous colonies on the selective agar plates (Fig. 4. 21). For this experiment LB + tet (15 µg/ml) was used as a selective medium. These colonies were screened positively for the presence of pHCG3 using the same approach as described under 4.2.3.2. and it was presumed that the plasmid had been transferred into *A. tumefaciens.* However, following PCR screening for regions corresponding to the *A. tumefaciens* genome, it was clear that this was a different microorganism.

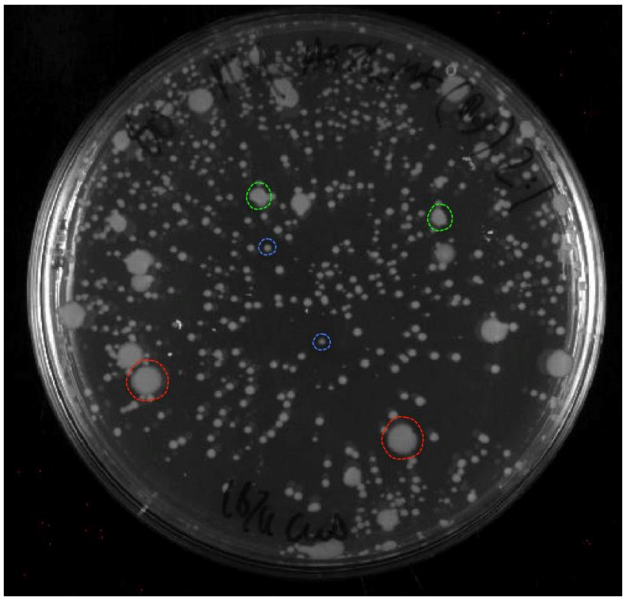


Figure 4.21. Contamination observed on selective plates for initial conjugations between OM5 pHCG3_*traM* and *A. tumefaciens* A676.

Larger colonies show pHCG3_*traM* containing *P.* cws transconjugants (red) and smaller colonies indicate the slower growing parental *P.* cws strain (green). Smallest colonies indicate A676 (blue). LB + tet (15 µg/ml) was used as selective medium.

Additionally PCR results tested negative for the *O. carboxidovorans* genome (data not shown). Following PCR and Sanger sequencing of the 16S ribosomal RNA gene, the most likely candidate appeared to be a species of *Pseudomonas*, here referred to as *Pseudomonas* strain "cws" (*P.* cws). The megaplasmid-containing strain was named *P.* cws pHCG3_*traM*_A. From the colonies screened for this conjugation, several appeared to grow more slowly on the selective tetracycline containing medium, forming smaller colonies. These tested negative for pHCG3_*traM* but appeared to be *P.* cws following 16S rRNA gene sequencing. This accidental isolate was chosen as to

act as the wild type for further conjugation experiments to determine whether pHCG3_*traM* was transferrable, seemingly into a different subclass (γ) of proteobacteria. Additionally, three other species, which were the closest match to the obtained 16S sequence of *P.* cws were chosen as recipients for pHCG3_*traM*. *E. coli* DH5α and *A. tumefaciens* A676 were also included as additional negative and positive conjugation controls respectively.

4.2.4.1. Conjugation with *P.* cws and other species

O. carboxidovorans OM5 pHCG3_*traM* was conjugated with the *P.* cws, *Pseudomonas azotoformans* (DSM 18862), *Pseudomonas lactis* (DSM 29167), *Pseudomonas paralactis* (DSM 29164), *E. coli* DH5α and *A. tumefaciens* A676, respectively, as previously described (2.2.5.3). Following incubation for 18 hours, the mating spots were resuspended from the filters in 0.9 ml PBS and transferred onto G-MM + tet agar plates to select for transconjugants.

Table 4.3. Conjugation frequency of pHCG3_*traM* for *P.* cws and other selected proteobacteria

Strain	No. of recipients per conjugation	Average No. of positive colonies per conjugation	Conjugation frequency	Standard deviation	No. of colonies for negative control
P. cws	8.3×10^{7}	4	4.82×10^{8}	$\pm 1.7 \times 10^{-8}$	0
P. lactis DSM 29167	1.10×10^{8}	0	0	0	0
P. paralactis DSM 29164	8×10^{7}	0	0	0	0
P. azotoformans DSM 18862	8.9×10^{7}	0	0	0	0
E. coli DH5α	1.04×10^{8}	0	0	0	0
C. necator H16	7.8×10^{7}	0	0	0	0
A. tumefaciens A676	9×10^{7}	6	6.67×10^{-8}	$\pm 3.14 \times 10^{-8}$	0

Apart from the A676 positive control, only *P.* cws conjugations resulted in tetracycline resistant colonies, indicating that this strain is uniquely capable of both receiving and replicating the plasmid. The resulting colonies growing on the selective glucose minimal media (G-MM) were restreaked twice onto fresh plates. The obtained purified colonies were screened by PCR for megaplasmid presence using primers 40_seqF and 40_seqR (OCA5_pHCG300040) (Fig. 4.22). The conjugation

frequency was taken as the number of transconjugants per 10^8 recipients (Table 4.3). The frequency of transfer was 4 x 10^{-8} per recipient for *P.* cws and 6 x 10^{-8} for A676. No colonies grew on selective plates for *Pseudomonas azotoformans* (DSM 18862), *Pseudomonas lactis* (DSM 29167) and *Pseudomonas paralactis* (DSM 29164) *or E. coli* DH5α conjugated with OM5 pHCG3_*traM* suggesting the megaplasmid could not be transferred or was destroyed by native restriction systems.

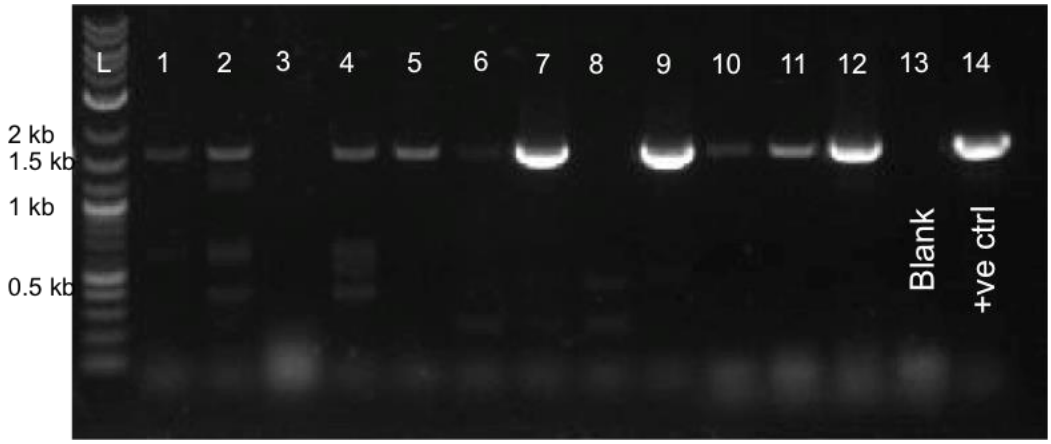


Figure 4.22. PCR confirmation of pHCG3_*traM* presence in *P.* cws transconjugants.
Primers 40seq F and 40seq R were used to amplify a pHCG3-specific integrase gene fragment of 1577 bp (locus tag OCA5_pHCG300040). Lanes 1 to 12 contain tetracycline resistant *P.* cws candidates, lane 14 contains OM5_*traM*, and lane 13 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

4.2.4.2. PCR screening

Four selected positive *P.* cws colonies were screened by PCR for the presence of other regions of pHCG3 using the primers 40_seq F & 40_seq R, 220_seqF & 220_seq R, HoxA F & HoxA R and TraG seq F & TraG seq R and the resulting products were further confirmed by Sanger sequencing (Fig. 4.23). The 16S ribosomal RNA gene was also amplified from each sample and sent for Sanger sequencing, which confirmed that the samples were *P.* cws. The obtained transconjugate strain was named *P.* cws pHCG3_*traM*_B and was cultured in LB + tet liquid media from which the genomic DNA was isolated (Sigma Aldrich kit).

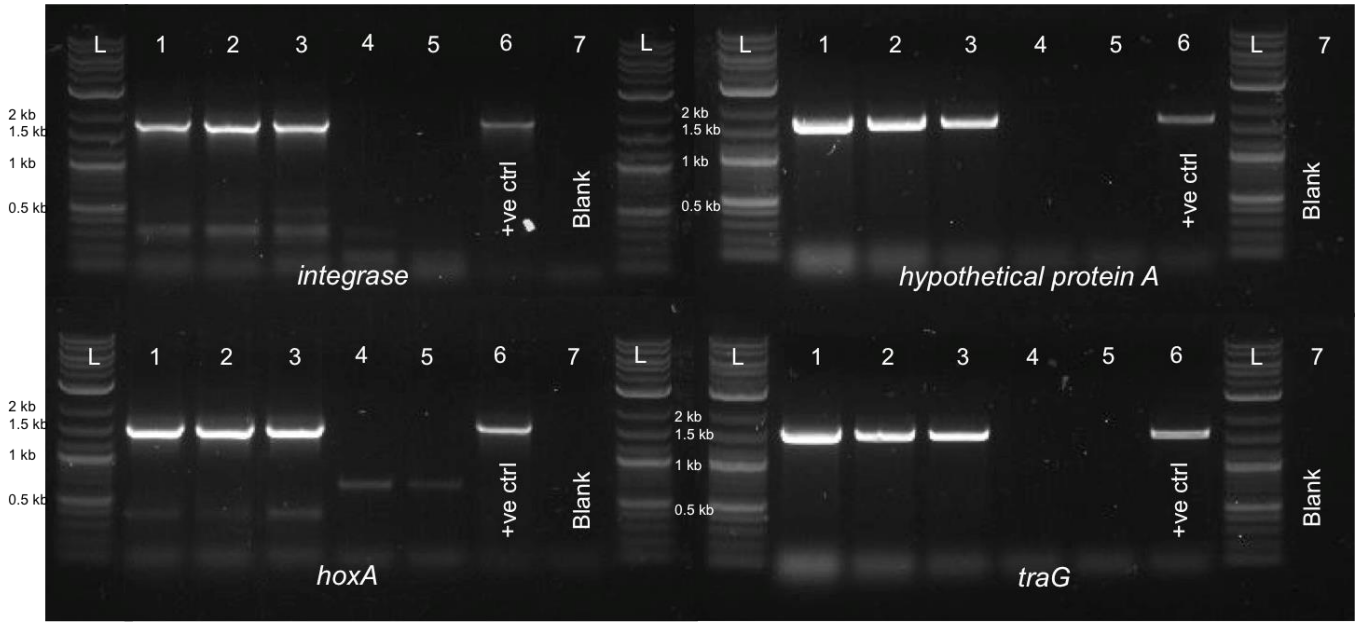


Figure 4.23. PCR amplification of multiple regions of pHCG3_*traM* confirming presence in *P.* cws.

Amplification of the integrase gene, 1577 bp (OCA5_pHCG300040) (top left); hypothetical protein A gene, 1582 bp (OCA5_pHCG300220) (top right); *hox A*, 1500 bp OCA5_pHCG300530) (bottom left); *traG*, 1588bp (OCA5_pHCG301040) (bottom right). For each region of pHCG3_*traM*, lanes 1-3 contain putative *P.* cws pHCG3_*traM* transconjugants, lanes 4 and 5 contain wild type *P.* cws, lane 6 contains OM5_*traM* and lane 7 contains a blank PCR reaction. The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

4.2.4.3. Whole genome sequencing

Since there was no reference genome for *P.*cws, a combination of Illumina and Oxford Nanopore next generation sequencing was performed on this strain and the original pHCG3_*traM* containing isolate (*P.* cws pHCG3_A). This method produces hybrid assemblies, bridging over repeated DNA region, enabling the scaffolding of contigs with high per-base quality to enable accurate base calling across the genome (www.microbesNG.com). Therefore, the resulting quality of the wild type *P.* cws assembly permitted its use as a reference genome, which *P.* cws pHCG3_*traM*_A and *P.* cws pHCG3_*traM*_B were mapped onto. Standard Illumina next generation sequencing was performed on *P.* cws pHCG3_*traM*_B. Reads of both pHCG3 containing strains were also mapped against the most recently published *O. carboxidovorans* genome (Volland *et al.* 2011). This mapping clearly indicated that pHCG3_*traM* is present with 100% of the pHCG3 reference genome covered. The mapping also suggested that OM5 chromosome 1 and 2 were absent from both

samples (Table 4.4). This indicated that there was no bacterial contamination from *O. carboxidovorans* leading to a false positive result for pHCG3_*traM* presence. Although very low coverage for these chromosomes was detected, further analysis revealed that the responsible reads were mapped to matching stretches of DNA within homologous genes of the *P.* cws chromosome. When mapped against the assembled *P.* cws genome, both pHCG3_*traM*-carrying strains displayed 100% reference coverage and a number of unique mutations (Table 4.4).

Table 4.4. Illumina sequencing reads of *P.* cws pHCG3_*traM* mapped to wild type *O. carboxidovorans* and *A. tumefaciens* genomes.

Reference covered given as a percentage of the reference covered by reads. Average coverage given as the average number of reads at any position on the reference.

Reference	*P.* cws pHCG3_*traM*_A	*P.* cws pHCG3_*traM*_B
OM5 pHCG3		
Reference covered (%)	100	100
Average coverage	117.58	41.94
SNPs	0	0
OM5 circular		
Reference covered (%)	8	4
Average coverage	1.97	0.7
SNPs	/	/
P. cws		
Reference covered (%)	100	100
Average coverage	117.25	44.18
SNPs	60	70

4.2.4.4. Phylogeny of *Pseudomonas* cws

To analyse the phylogeny of *P.* cws, the RNA polymerase sigma factor gene, *rpoD* was chosen for comparison to other species of *Pseudonomas* from three different phylogenetic groups. The *rpoD* sequence obtained from the assembled *P.* cws genome was used to search the GenBank database using the BLASTN algorithm to find the closest matches to *P.* cws (Benson *et al.* 2013, Altschul *et al.* 1990). An alignment was generated using Clustal Omega (Sievers & Higgins 2013), the algorithm of which was used to construct a phylogenetic tree using the Interactive tree of life (iTOL) software (Letunic & Bork 2006, 2019). From this analysis, *P.* cws appears to be most closely related to *P. fluorescens* strain PF 36, falling into the *Pseudomonas*

fluorescens group and most distantly related to species of the *Pseudomonas putida* group (Fig. 4.24). The *P. fluorescens* group comprises species which have been isolated from a range of environments including soil, water, plant tissue (Andersen *et al.* 2000, Mirand & Zemelman 2002, Brown *et al.* 2012), many of which have been shown to exhibit plant growth promotion activity (Kang *et al.* 2006, Raaijmakers *et al.* 2009, Shinde *et al.* 2017).

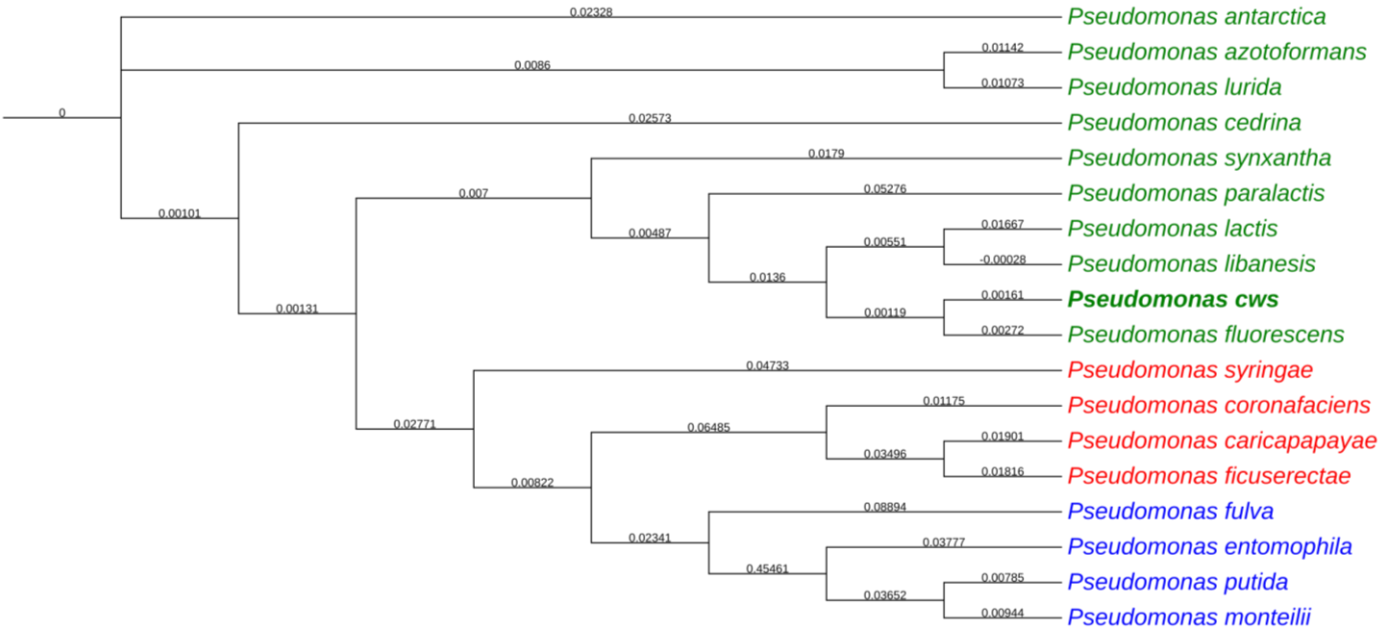


Figure 4.24. Neighbour-joining tree based on *rpoD* sequences from different *Pseudomonas* species
The tree shows the phylogenetic position of the novel *P.* cws strain and species chosen as reference strains, which include strains from the *P. fluorescens* (green), *P. syringae* (red) and *P.* putida groups (blue). Branch lengths are not to scale. Lengths of each branch are displayed on each node of the tree. Tree generated using iTol (Interactive tree of life).

Table 4.5. Similarities of the 16S *rRNA* and *rpoB* sequences of *Pseudomonas* strains to *P.* cws

Strain	**16S *rRNA* (%)**	***rpoB* (%)**	**Reference**
Pseudomonas sp. J380	100	99.8	Santander *et al.* unpublished
Pseudomonas sp. MYb193	100	99.19	Zimmermann *et al.* unpublished
Pseudomonas sp. FDAARGOS_380	99.8	98.72	Campos *et al.* unpublished
Pseudomonas fluorescens A506	99.8	98.6	Loper *et al.* 2012
Pseudomonas fluorescens strain Pt14	99.74	98.6	Rani & Lal unpublished
Pseudomonas libanensis strain DMSP-1	99.67	98.8	Zeng unpublished
Pseudomonas azotoformans strain S4	99.8	97.5	Fang *et al.* unpublished

To ascertain the identity of *P.* cws, the 16S *rRNA* and *rpoB* sequences obtained from the assembled *P.* cws genome were used to search the GenBank database to find the *Pseudomonas* strain displaying the greatest similarity to *P.* cws, as demonstrated by Bondoso *et al.* (2013). From this analysis, the strain showing the closest match was an environmental isolate; *Pseudomonas* sp. J380 (Table 4.5). This GenBank search was not suitable to identify the closest type strain match to *P.* cws, therefore a search was performed with the entire *P.* cws genome using the EZBioCloud database. Following whole genome comparison based on average nucleotide identity (ANI) using TrueBac ID (Ha *et al.* 2019), the closest type strain match to *P.* cws was *Pseudomonas carnis* B4-1^T (Table 4.6). According to Richter and Ramon Rosselló-Móra (2009), the ANI threshold boundary for species circumscription is 95–96%, therefore *P.* cws can be classified as a strain of *P. carnis* with an ANI of 98.44%.

Table 4.6. Average nucleotide identity and 16S *rRNA* sequence identity of *P.* cws to other *Pseudomonas* species

Strain	ANI (%)	Coverage (%)	16S *rRNA* (%)	Reference
Pseudomonas carnis B4-1^T	98.44	82.5	99.86	Lick *et al.* 2020
Pseudomonas paralactis WS4992^T	91.66	70.6	99.86	von Neubeck *et al.* 2017
Pseudomonas azotoformans DSM 18862^T	89.19	56.2	99.86	Skerman *et al.* 1980
Pseudomonas synxantha DSM 18928^T	90.94	66.8	99.59	Skerman *et al.* 1980
Pseudomonas libanensis CIP 105460^T	90.94	68	99.52	Dabboussi *et al.* 1999
Pseudomonas gessardii CIP 105469^T	88.33	45.2	99.45	Verhille *et al.* 1999
Pseudomonas poae DSM 14936^T	88.76	51.1	98.97	Behrendt *et al.* 2003
Pseudomonas salomonii CFBP 2022^T	89.14	57.3	98.77	Gardan *et al.* 2002
Pseudomonas palleroniana CFBP 4389^T	88.72	56.5	98.72	Gardan *et al.* 2002

4.2.4.5. Monitoring of autotrophic growth

To observe whether presence of pHCG3_*traM* enabled *P.* cws to grow using CO or CO_2 as a sole or additional carbon source, growth comparisons were performed between wild type *P.* cws and *P.* cws + pHCG3_*traM*_A. Pre-cultures were prepared in the same way as described for 4.2.3.4. Cells were harvested by centrifugation, resuspended and split into serum bottles containing a growth limiting concentration of glucose (0.02%, approximately 1.1 mM), allowing growth to an OD_{600} of about 0.2. The bottles contained G-MM (+15 μg/ml tet for *P.* cws pHCG3_*traM*_B) with either

50% CO/50% air or 50% N_2/50% air or 80% H_2/10% CO_2/10% air. The growth comparison was performed with a starting OD_{600} of 0.01.

The growth curves followed a time course of 124 hours, with wild type *P.* cws and *P.* cws pHCG3_*traM*_A showing similar growth in all conditions. With 0.02% glucose alone, a maximum OD_{600} of about 2 was achieved; however, neither the presence of CO or CO_2 + H_2 allowed for further growth (Fig. 4.25). This indicated that the presence of pHCG3_*traM* did not give *P.* cws the ability to utilise CO or CO_2 as a carbon source for growth.

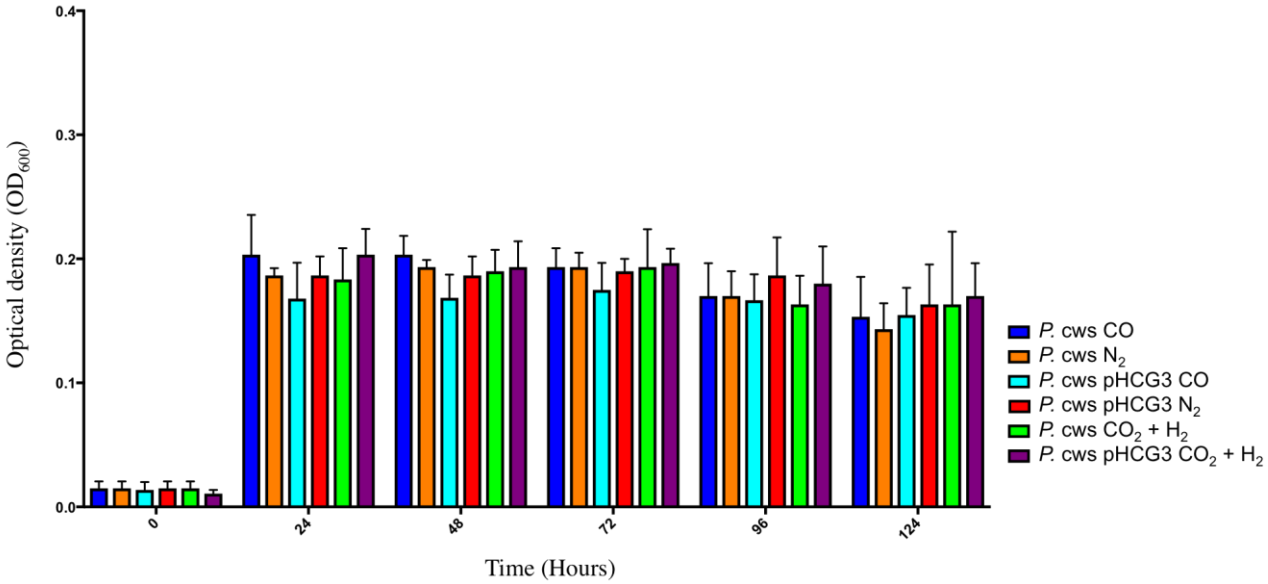


Figure 4.25. Growth of *Pseudomonas* cws pHCG3_*traM*_A growth in comparison to the parental strain with either CO or CO_2 + H_2.

Strains were grown at 30°C in 20 ml MM containing 0.02% glucose and contained in 150 ml serum bottles under an atmosphere of 50% N_2/50% air ("N_2"), 50% CO/50% air ("CO") and 80% H_2/10% CO_2/10% air ("CO_2+H_2"), respectively (2 bar total pressure). Blue bars, wild type *P.* cws (CO); green bars, wild type *P.* cws (CO_2 + H_2); orange bars, wild type *P.* cws (N_2); cyan bars *P.* cws pHCG3_*traM*_A (CO); purple bars, *P.* cws pHCG3_*traM*_A (CO_2+ H_2); red bars, *P.* cws pHCG3_*traM*_A (N_2). Error bars represent the standard deviation of the mean for three independent replicates.

4.2.4.6. Mass spectrometry analysis of *P.* cws supernatants

To observe whether presence of pHCG3_*traM* enabled *P.* cws to produce AHLs (as described in section 4.2.2.8. for *O. carboxidovorans*), *P.* cws pHCG3_*traM*_A and *P.* cws pHCG3_*traM*_B and parental *P.* cws were grown in G-MM (0.2%) to an OD_{600} of 2.0 and centrifuged at 8000 rpm for 5 minutes. The supernatants were filter-

sterilised and analysed by LC-MS/MS. No AHLs were detected in any of the culture supernatants indicating that *P.* cws was not sufficiently able to either transcribe and/or translate the QS genes on pHCG3 (data not shown).

4.2.5. Investigation of pHCG3 transfer in *C. necator*

Since transfer of pHCG3 from its host into a distantly related organism was established with *P.* cws, conjugation into the genetically more similar *C. necator* was thought to be possible.

4.2.5.1. Conjugation with *C. necator*

OM5_pHCG3_*traM* was conjugated with *C. necator* H16 by mating equal volumes of agar plate-grown cells resuspended in P-MM / LB and pipetted onto a 0.2 µm nitrocellulose filter on P-MM / LB agar plates. Following this, the mating spot was transferred to an F-MM + tet agar plate to select for *C. necator* transconjugants, since *O. carboxidovorans* is unable to grow using this medium (data not shown). No growth was seen, indicating that the plasmid could not be transferred to *C. necator* under the chosen conditions. The tet cassette used for selection of pHCG3_*traM* is functional in *C. necator*, as shown by its use in Chapter 3 and other published studies (Juengert *et al.* 2018, Bresan & Jendrossek 2017, Kleihues *et al.* 2000 Lenz & Friedrich 1998). Therefore, the absence of tet resistant colonies could not be explained by a lack of gene expression.

4.2.5.2. Analysis of *repABC*

A possible explanation for the failure of pHCG3_*traM* to be transferred is that its origin of replication may not be functional in *C. necator*. pHCG3 harbours a conserved *repABC* operon, within which its replication origin is contained. The *repABC* operon is responsible for the replication and segregation of pHCG3 and encodes the proteins RepA, RepB and RepC. Most *repABC* plasmids employ a Type I segregation system with Walker-type ATPase (ParA superfamily) (Cevallos *et al.* 2008). *C. necator* must be able to replicate pHCG3 in order to utilise its *cox* genes, therefore the functionality of *repABC* in *C. necator* was tested.

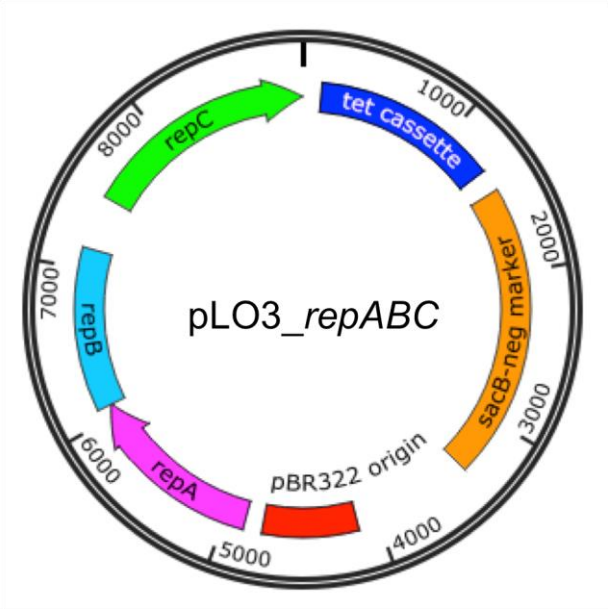


Figure 4.26. The pHCG3 *repABC* operon cloned into pLO3.

repABC was amplified with primers RepF and RepR, containing AscI and XbaI restriction sites. The resulting 4115bp fragment containing the origin of replication, the *repABC* promotor and ParS binding sites, was cloned into AscI and XbaI sites on pLO3 to generate pLO3_*repABC.*

To establish whether *C. necator* can utilise *repABC,* the operon was cloned into the suicide vector pLO3 using primers containing AscI and XbaI sites, to generate a *repABC*-type replicon (pLO3_*repABC*) (Fig. 4.26). This was introduced into *E. coli* S17-1, which was conjugated with *C. necator* as described in section 2.2.5.1. pLO3 contains a tet resistance marker and the pBR322 origin, which is unable to function in *C. necator* (Fig. 4.25). Therefore, if the introduced *repABC* operon and replication origin is functional in *C. necator*, pLO3_*repABC* should be able to replicate, giving tet resistance colonies on selective agar media containing the antibiotic. In control experiments, wild type *C. necator* was also plated on selective media to screen for background resistance to tetracycline. No colonies grew on the selective agar plates following conjugation with *E. coli* S17-1 pLO3_*repABC*, suggesting that the encoded RepABC replication machinery did not function in *C. necator.* In order for pHCG3 derivatives to replicate in *C. necator*, a functional origin must therefore be introduced.

4.2.5.3. Analysis of pMOL28 replication unit

The replicon from the *C. metallidurans* megaplasmid pMOL28 encodes RepA28, ParA and ParB (Juhnke *et al.* 2002) and has been previously characterised in the host organism (Taghavi *et al.* 1996). RepA28 displays 31.6% identity to *C. necator* partitioning proteins and ParA displays 40.5% identity to *C. necator* ParA family protein (Schwartz *et al.* 2003). This replicon has previously been cloned into a modular vector containing a *colE1* origin (pMTL72111_E1), which is non-functional in *C. necator* (Ehsaan unpublished) (Fig. 4.27).

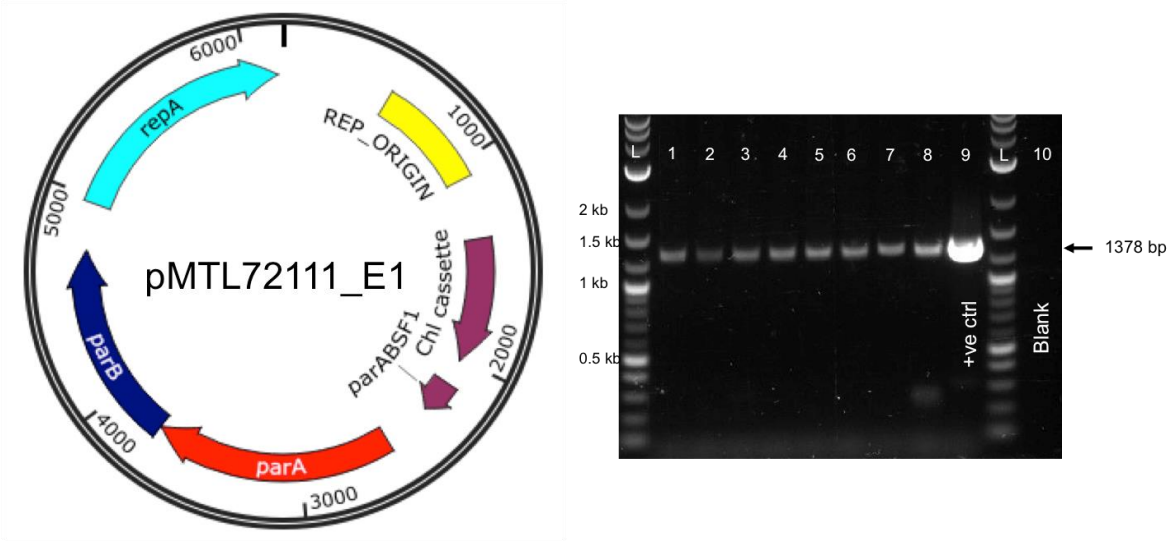


Figure 4.27. pMTL72111_E1 plasmid and PCR confirmation of its presence in *C. necator*.

pMTL72111_E1 contains *colE1* origin, a *catP* chloramphenicol resistance cassette and the *C. metallidurans* pMOL28 replicon consisting of *repA28* (turquoise), *parA* (red) and *parB* (blue) (left). Primers annealing to the pMOL28 replicon, rep28_intF and rep28_intR used to amplify a fragment from pMTL72111_E1 (right). Lanes 1 to 8 contain candidate chloramphenicol resistant *C. necator* colonies, lane 9 contains positive control with pMTL72111_E1 and lane 10 contains a blank PRC reaction. The size of PCR products was estimated by 1% agarose gel electrophoresis using a 2-log ladder (L).

This plasmid was introduced into *E. coli* S17-1 and conjugated with *C. necator* to test its functionality. After 6 hours, the mating spot was transferred on selective F-MM + chl plates and colonies grew within 48 hours. Colonies screened by PCR tested positive for the presence of the plasmid (Fig. 4.27), indicating that *repA28* is functional in *C. necator*. A suicide plasmid containing *repA28* with flanking homology arms was to be constructed and integrated into pHCG3 however time constraints towards the end of this project meant that this work was not completed.

4.2.6. Analysis of pLO3 functionality

Replication origin of pLO3 is not functional in *C. necator* as demonstrated (4.2.5.2), however it was necessary to test its functionality in the *Pseudomonas* and *A. tumefaciens* strains used preciously. Due to the method by which *traM* was knocked out (4.2.2.7), the backbone of pLO3 remained intact, and therefore its origin (derived from pBR322) could still be functional and potentially contribute to replication of pHCG3. The functionality of pLO3 was tested in the strains able to harbour pHCG3_*traM*, A676, A677 and *P.* cws by conjugation with S17-1 pLO3 (as described in the Methods 2.2.5) and subsequent screening for tet resistance on selective MM agar plates. All strains failed to grow on selective media, thus it was assumed that the pBR322 origin was non-functional.

4.2.7. Characterisation of putative CO transcription factors

The findings published during the completion of this work demonstrated that *C. necator* is able to express a set of *cox* genes from pHCG3 under the control of an inducible promoter to oxidise CO to CO_2 (Heinrich *et al.* 2018). *cox* gene expression was previously attempted by introducing a plasmid containing the full *cox* cluster from pHCG3 into *C. necator* and by integrating this cluster together with the pHCG3 *hox* gene cluster into its genome (Alagesan & Malys, University of Nottingham, unpublished findings). These strains failed to grow with CO as a sole carbon source, which, following publication of the above study, prompted speculation as to whether the *cox* operon contained transcription factor binding sites (TFBSs), which are responsive to CO and whether these were not expressed following introduction into *C. necator*. Four RFP-expressing reporter plasmids were constructed containing different regions upstream of *coxM* on pHCG3, which included the gene for a proposed CO-responsive transcription factor (COTF) encoded by *coxC* (Fuhrmann *et al.* 2003, Paul *et al.* 2008, Volland *et al.* 2011). These regions were chosen due to presence of putative promoter sequences located intergenically as well as within *coxB* and *coxC*, which were predicted by BPROM (www.softberry.com). The generated reporter constructs were transferred into *E. coli*, *C. necator* and *O. carboxidovorans*, which were cultured in minimal media in the presence and absence of CO and further analysed as described below.

4.2.7.1. Construction of CO responsive RFP reporters

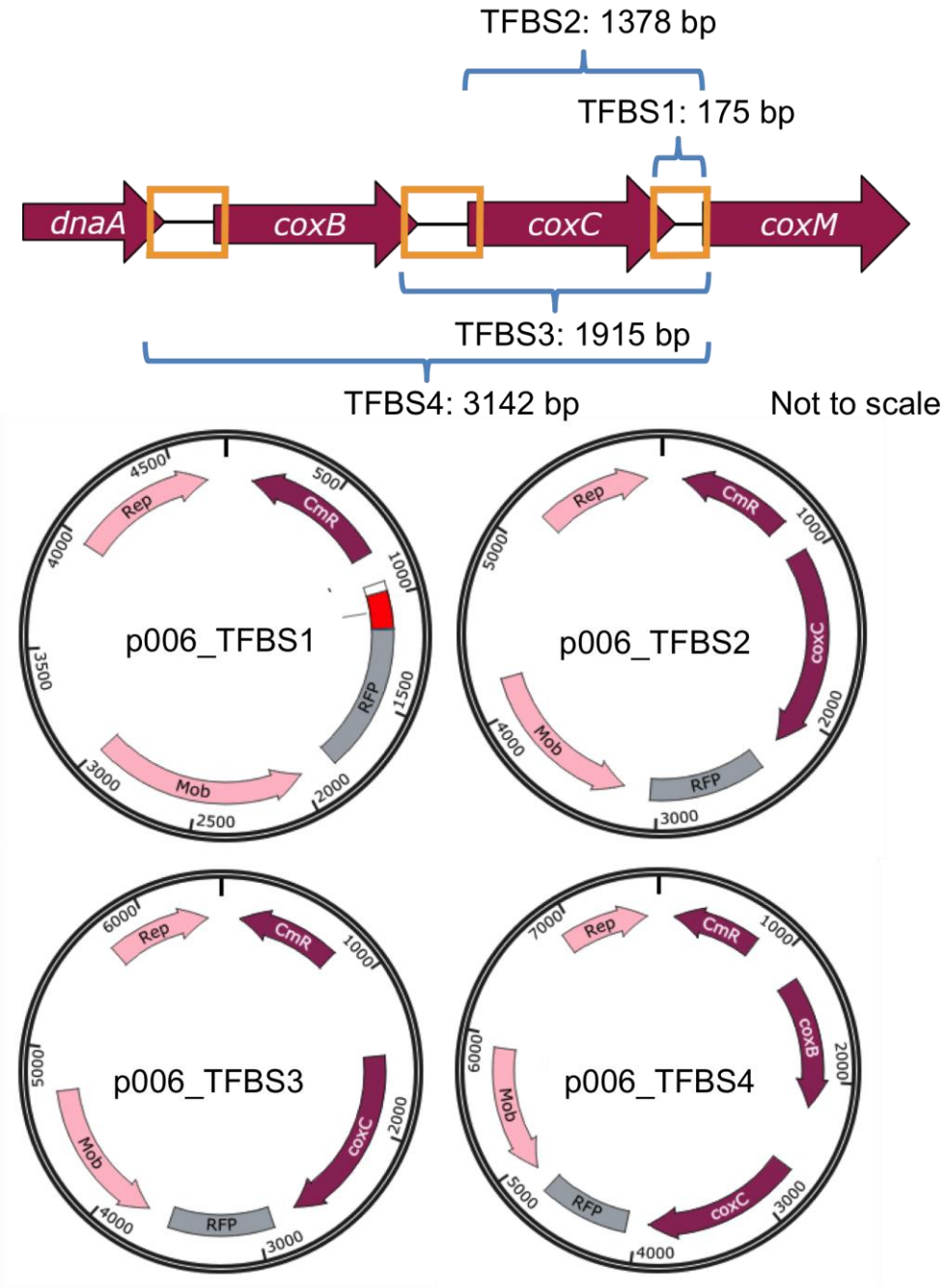


Figure 4.28. Reporter plasmids to test CO-responsiveness of *coxM expression.*

Top panel: four fragments hypothesised to contain COTF binding sites (orange squares) were amplified by PCR using Phusion polymerase with a forward primer containing an AatII and a reverse primer containing an NdeI 5' restriction site. The regions tested were the intergenic region directly downstream of *coxC* to the start of *coxM* (TFBS 1), the region from the start of *coxC* to the start of *coxM* (TFBS2), the intergenic region downstream of *coxB* to the start of *coxM* (TFBS3) and the region upstream of *coxB* to the start of *coxM* (TFBS4). Middle and bottom panels: maps of the four plasmids carrying the indicated *coxM* upstream regions, p006_TFBS1, p006_TFBS2, p006_TFBS3 and p006_TFBS4.

Four *cox* cluster regions hypothesised to contain COTF (CoxC) binding sites were amplified by PCR as shown in Fig. 4.26 and fragments ligated into AatII and NdeI sites of the RFP reporter pEH006 (Hanko *et al.* 2017), thus replacing the arabinose inducible system on this plasmid. This generated p006_TFBS1, p006_TFBS2, p006_TFBS3 and p006_TFBS4 (Fig. 4. 28), which were all confirmed by PCR and Sanger sequencing. These plasmids were cloned in *E. coli* DH5α, then introduced along with a promoterless control plasmid, pEH006E (p006E) into *E. coli* S17-1. pEH006E cannot induce expression of RFP, therefore any fluorescence measured from strains containing this plasmid, is background autofluorescence (Hanko *et al.* 2017). These strains were conjugated with wild type *C. necator* H16 and *O. carboxidovorans* OM5.

4.2.7.2. *coxM*-RFP reporter expression in *E. coli* DH5α

E. coli p006_TFBS1, *E. coli* p006_TFBS2, *E. coli* p006_TFBS3, *E. coli* p006_TFBS4, the negative control *E. coli* p006E and plasmidless *E. coli* DH5α were all cultured in M9 G-MM + chloramphenicol (25 μg/ml) and diluted to a starting OD_{600} of 0.05 in serum bottles containing the same growth medium with 50% CO/50% air or 50% N_2/50% air in the headspace. Measurements were taken after 24 hours using an infinite pro M1000 plate reader (Tecan) to determine fluorescence and absorbance for each sample (Fig. 4.29).

There was no significant difference between the fluorescence observed from the promoter-less control *E. coli* p006E and any of the samples ($p > 0.05$) (Fig. 4.29). This indicated that none of the postulated COTF-dependent regions were active in *E. coli* in the presence and absence of CO.

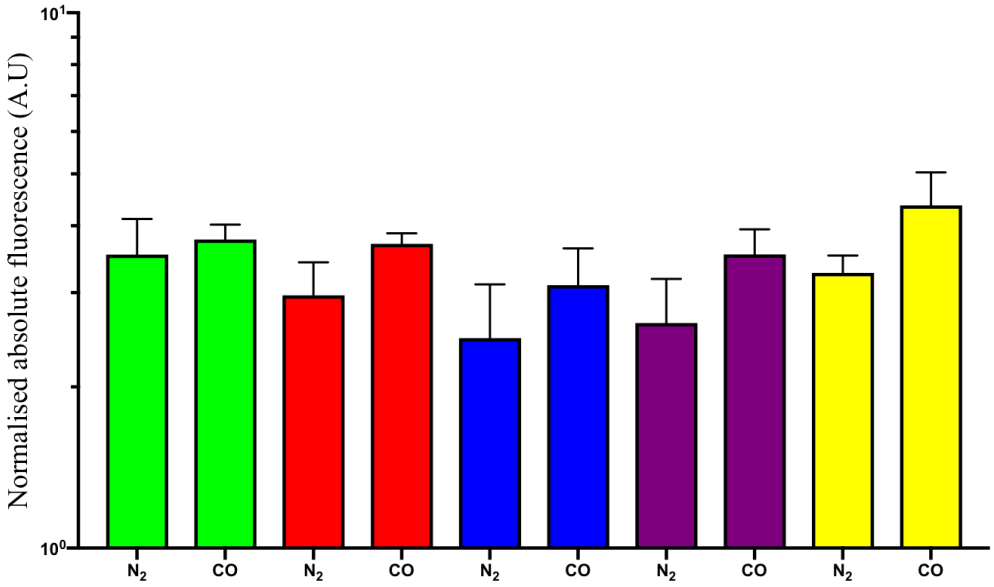


Figure 4.29. Expression of *cox*-rfp *reporter* plasmids in *E. coli* DH5α.
Strains were grown at 30 °C in 20 ml G-MM + chloramphenicol contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air ("CO") or 50% N_2/50% air ("N_2") (2 bar total pressure). Expression of the *rfp* gene was monitored after 24 h by determining RFP fluorescence in *E. coli* p006_TFBS1 (green), *E. coli* p006_TFBS2 (red), *E. coli* p006_TFBS3 (blue), *E. coli* p006_TFBS4 (purple) and *E. coli* p006E (yellow). Error bars represent standard deviations from three biological replicates. There was no significant difference in the absence or presence of CO between the *E. coli* p006E control and *E. coli* p006_TFBS1 (N_2, p = 0.51; CO, p = 0.22), *E. coli coli* p006_TFBS2 (N_2, p = 0.37; CO, p = 0.17), *E. coli coli* p006_TFBS3 (N_2, p = 0.12; CO, p = 0.061) or *E. coli* p006_TFBS4 (N_2, p = 0.14; CO, p = 0.14).

4.2.7.3. *coxM*-RFP reporter expression in in *C. necator* H16

C. necator p006_TFBS1, *C. necator* p006_ TFBS2, *C. necator* p006_ TFBS3, *C. necator* p006_ TFBS4, *C. necator* p006E and wild type *C. necator* H16 were all cultured in F-MM + chloramphenicol (50 μg/ml) and diluted to a starting OD_{600} of 0.05 in serum bottles containing the same growth media with 50% CO/50% air or 50% N_2/50% air in the headspace. The samples were measured as in 4.2.5.1.

There was no significant difference between the fluorescence observed from the promoter-less control *C. necator* p006E and any of the samples measured ($p > 0.05$) (Fig. 4.30). This indicated that none of the postulated TFBS-dependent regions were active in *C. necator* in the presence and absence of CO.

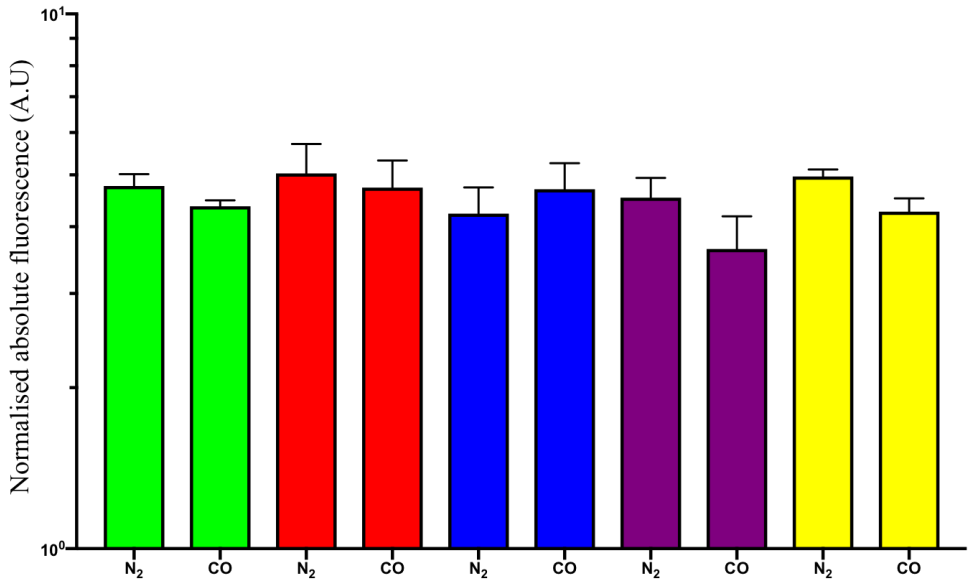


Figure 4.30. Expression of *cox*-rfp reporter plasmids in *C. necator* H16.
Strains were grown at 30 °C in 20 ml F-MM + chl contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air ("CO") or 50% N_2/50% air ("N_2") (2 bar total pressure). Expression of the *rfp* gene was monitored after 24 h by determining RFP fluorescence in *C. necator* H16 p006_TFBS1 (green), *C. necator* p006_TFBS2 (red), *C. necator* p006_TFBS3 (blue), *C. necator* p006_TFBS4 (purple) and *C. necator* p006E (yellow). Error bars represent standard deviations from three biological replicates. There was no significant difference in the absence or presence of CO between the *C. necator* p006E control and *C. necator* p006_TFBS1 (N_2, p = 0.3046; CO, p = 0.5655), *C. necator coli* p006_TFBS2 (N_2, p = 0.8766; CO, p = 0.2737), *C. necator coli* p006_TFBS3 (N_2, p = 0.0731; CO, p = 0.2866) or *C. necator coli* p006_TFBS4 (N_2, p = 0.1574; CO, p = 0.1443).

4.2.7.4. *coxM*-RFP reporter expression in in *O. carboxidovorans*

O. carboxidovorans strains OM5 p006_TFBS1, OM5 p006_TFBS2, OM5 p006_TFBS3, OM5 p006_TFBS1, OM5 p006E and wild type OM5 were all cultured in P-MM + chloramphenicol (25 μg/ml) and diluted to a starting OD_{600} of 0.05 in serum bottles containing the same growth medium with 50% CO/50% air or 50% N_2/50% air in the headspace. The samples were measured as in 4.2.5.1.

There was no significant difference between the fluorescence observed from the negative control strain OM5 p006E and any of the other samples measured (p > 0 .05) (Fig. 4.31). This suggested that none of the postulated COTF-dependent regions were active or that the reporter system was non-functional in *O. carboxidovorans* in the presence and absence of CO. Therefore, the original pEH006 plasmid was conjugated into *O. carboxidovorans* and *C. necator* as positive control via *E. coli* S17-1 and the

resulting strains *C. necator* p006 and OM5 p006, were induced with arabinose alongside *E. coli* 17-1 p006 (1 mM). In this plasmid, *rfp* expression is controlled by AraC and the *araBAD* promoter and hence *rfp* expression can be induced through arabinose addition. p006 was additionally transformed into *E. coli* DH5α, which along with *C. necator* were able to produce RFP; however, no florescence was observed in *O. carboxidovorans* demonstrating that the system was for some reason not functional in this background (Data not shown).

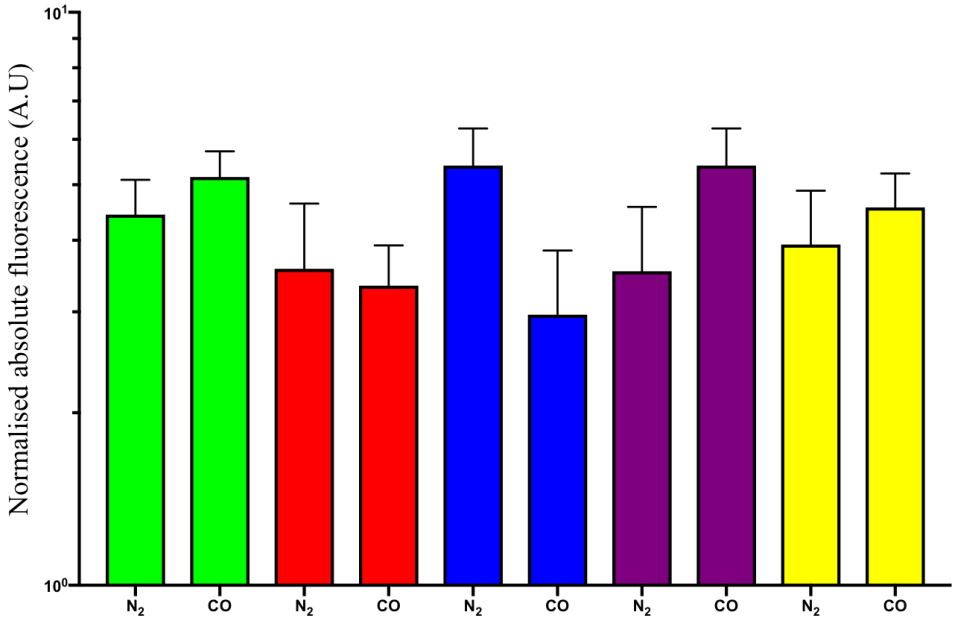


Figure 4.31. Expression of *cox*-rfp reporter plasmids in *O. carboxidovorans* OM5.
Strains were grown at 30 °C in 20 ml P-MM + cl contained in 150 ml serum bottles under an atmosphere of 50% CO/50% air ("CO") or 50% N_2/50% air ("N_2") (2 bar total pressure). Expression of the *rfp* gene was monitored after 24 h by determining RFP fluorescence in OM5 p006_TFBS1 (green), OM5 p006_TFBS2 (red), OM5 p006_TFBS3 (blue), OM5 p006_TFBS4 (purple) and OM5 p006E (yellow). Error bars represent standard deviations from three biological replicates. There was no significant difference in the absence or presence of CO between the OM5 p006E control and OM5 p006_TFBS1 (N_2, p = 0.4970; CO, p = 0.3044), OM5 p006_TFBS2 (N_2, p = 0.6801; CO, p = 0.0737), OM5 p006_TFBS3 (N_2, p = 0.1202; CO, p = 0.0652) or OM5 p006_TFBS4 (N_2, p = 0.6488; CO, p = 0.2586).

4.3. Discussion

The ultimate aim of this study was to investigate the transfer capabilities of the *O. carboxidovorans* megaplasmid pHCG3 with the intention of introducing the plasmid into *C. necator* via conjugation. The purpose of this was to equip *C. necator* with a functioning operon of *cox* genes, which might enable utilisation of CO. This would enable greater exploitation of the available carbon found in syngas and other industrial waste gases, which are highly desirable feedstocks for *C. necator* for the production of bioproducts. As mentioned in the introduction of this chapter, a set of *cox* genes from pHCG3 have been expressed in *C. necator* allowing oxidation of CO to CO_2 (Heinrich *et al.* 2018). This was published during the completion of this project, but efforts were still made to explore the transfer of pHCG3, since *C. necator* was still not able to use CO as a sole source of carbon and energy in the aforementioned study.

It was hypothesised that by knocking out the function of TraM, the conjugative capacity of pHCG3 could be boosted through AHL production. Indeed, it was discovered that the AHL synthase gene *traI* led to the production of C_8 and 3-oxo-C_8-AHL when expressed in *E. coli* (Fig 4.8). No AHLs were detected in wild type *O. carboxidovorans*; however, the *traM* truncation mutant (OM5 *traM*), produced eight different AHLs with C8 and 3-oxo-C8-AHL displaying the largest peak areas (i.e. relative amounts), the latter two matching the profile of *E. coli*-expressed *traI,* albeit with lower amounts (Fig 4.14). This indicates that TraM function directly affects *traI* expression in this organism, presumably through the binding and inactivation of TraR, as observed for *A. tumefaciens* (Luo *et al.* 2000).

Indeed, the identified *O. carboxidovorans* system as a whole appears to be analogous to the quorum sensing system of *A. tumefaciens*, where 3-oxo-C8 AHL binds to TraR, activating *traI* expression (Piper & Farrand 2000, Zhang *et al.* 1993). Thereby a positive feedback mechanism is created, amplifying production of 3-oxo-C8 AHL, which then activates TraR. *A. tumefaciens* TraR also activates transcription of the *traM* gene, then TraM negatively regulates TraR by forming a TraM-TraR antiactivation complex, preventing binding to *traI* and potentially inhibiting *A. tumefaciens* from engaging in conjugal transfer at low cell densities (Fuqua *et al.*

1995, Hwang *et al.* 1995, Luo *et al.* 2000, Churchill & Chen 2011). Consistent with the findings presented, when *traM* is mutated, *A. tumefaciens* also produces easily detectable levels of AHL, which gives a hyperconjugal phenotype for the Ti plasmid (Hwang *et al.* 1995, Piper & Farrand 2000). Due to the production of AHLs in OM5 pHCG3_*traM*, it was assumed that the truncated form of *traM* was non-functional and unable to bind TraR. However to confirm this, western blot analysis might be performed as demonstrated by Luo *et al.* (2000), where the binding capacity of TraM to TraR was monitored under different temperatures or by the multitude of alternative methods that exist to demonstrate and quantify protein-protein interaction such as pull-down assays, surface plasmon resonance, isothermal titration calorimetry, or two-hybrid screens (Miura 2018).

The other *O. carboxidovorans* AHL synthase gene, *solI*, produced nine AHLs when expressed in *E. coli* (Fig. 4.8) and caused strong induction of the LuxRI system encoded by the pSB403 reporter plasmid (Fig. 4.7). This gene is likely regulated by the LuxR family transcriptional regulator (*solR*), which overlaps with its 5' end (Fuhrmann *et al.* 2003, Volland *et al.* 2011). Of the nine AHLs produced by SolI in *E. coli*, only C_6, 3-oxo-C_6, C_8 and C_{10} match those seen in the *O. carboxidovorans* OM5_*traM* mutant (Fig. 4.14), one of which is also produced by TraI, therefore it may seem unlikely that TraM affects the regulation of SolR. However, the spectrum of AHLs produced by LuxI homologues can be determined by their substrate specificity, i.e. the range of acyl-acyl-carrier-proteins (acyl-ACPs) they are able to use, or by the availability of these substrates (Li & Nair 2012, Hanzelka & Greenberg 1996). Hence the AHL spectrum and ratios observed after expression of *luxI*-type genes in *E. coli* may differ from the expression seen in the native background, so TraM may regulate both TraR and SolR. However, if the TraM only regulates TraR there are six additional AHLs TraM must regulate, directly or indirectly, which may be produced by other AHL synthases encoded elsewhere on the genome of *O. carboxidovorans.* One possible candidate is a putative autoinducer synthase gene (OCA5_pOC16701320) located on the pOC167 megaplasmid (Volland *et al.* 2011). Only C_8, 3-oxo-C_8, 3-OH-C_8 and 3-oxo-C_6 were detected in A676 pHCG3_*traM* (Fig. 4.21), which suggests that this enzyme is responsible for the additional AHLs present in OM5 pHCG3_*traM* and also *A. tumefaciens* A676_*traM*. Further experimentation

would be required to establish the precise functioning and potential QS hierarchy in OM5, such as the characterisation of OCA5_pOC16701320 in *E. coli* and the inactivation of *traR, traI, solR* and *solI* in *O. carboxidovorans.*

The modified version of pHCG3, pHCG3_*traM*, could not be successfully conjugated into wild type *C. necator* from OM5 *traM.* A possible explanation for this was that the megaplasmid was rapidly degraded in the recipient due to native restriction systems or CRISPR/Cas9 immunity, which has been shown to be functional in *C. necator* (Xiong *et al.* 2018). Additionally, efficient transcription and translation were not possible due to a lack of promoter recognition or activation and differences in codon usage. Another possibility was a lack of suitable selection due to the tet cassette not being expressed, prompting questioning of the functionality of the replication origin contained within *repABC.* The *repABC* operon is responsible for the replication and segregation of a particular group of low copy-number plasmids and encodes the proteins RepA, RepB and RepC (Pinto *et al.* 2012). Most of the plasmids carrying this operon employ a Type I segregation system employing a Walker-type ATPase (ParA superfamily) and *parS* binding protein (ParB superfamily)(Cevallos *et al.* 2008). ParA polymerises into filaments, which segregate the plasmids, and ParB binds to ParA and centromere-like *parS* sites, to mediate between the plasmids and filaments (Chai & Winans 2005, Ebersbach *et al*. 2005). The pHCG3-associated *repABC* operon was found in this study to be non-functional in the β-proteobacterium *C. necator*, supporting the consensus that the system only functions in the α-proteobacteria as it is conserved in the majority of large plasmids in this taxonomic group (Pinto *et al.* 2012, Döhlemann *et al.* 2017). Many RepABC family plasmids initiate replication from an origin situated within *repC* (Cervantes-Rivera *et al.* 2011). However, in pHCG3 the origin of replication is located within the *repB-repC* intergenic region (Fuhrmann *et al.* 2003). RepC-type proteins are reportedly not homologous to any other replication initiators and were only thought to be found in α-proteobacteria (Pinto *et al.* 2012).

A RepABC-like plasmid has reportedly been found in β-proteobacteria in the order *Burkholderiales* (Shintani *et al.* 2015). Putative plasmid-like replication origins have also been found in the *C. necator* genome in addition its chromosomal origins,

including a putative RepA binding site in the 5'-flanking region of *repA*, with adjacent *parA* and *parB* orthologs (Pohlmann 2006). The *C. metallidurans* megaplasmid pMOL28 contains a 4112 bp replication operon (Juhnke *et al.* 2002), which was shown to function in *C. necator* (Fig. 4.27). This result appeared likely since the proteins encoded in this replicon display significant identity to partitioning proteins found in *C. necator* (Schwartz *et al.* 2003). The next step of this work was to integrate this operon into pHCG3_*traM* and conduct future conjugation experiments with *C. necator*. To conduct this modification, transfer of pHCG3_*traM* into the well-characterised and easily modifiable *A. tumefaciens* was carried out.

It was proven by PCR and whole genome sequencing that pHCG3_*traM* was successfully transferred into plasmid-less strains of *A. tumefaciens* (A676 and A677) from OM5_*traM*, demonstrating that this megaplasmid can indeed be conjugated (Table 4.2). This result was expected given that transfer of the RepABC-type pRL1JI from *Rhizobium leguminosarum* bv. *Viciae*, has been demonstrated in A677. (Danino *et al.* 2003). Transfer of pRL1JI has been monitored between A894 (an A677 derivative) and A676, which is transferred at a lower frequency than pHCG3_*traM* (Danino *et al.* 2003). However when the recipient expressed AHLs via the CinR/CinI quorum sensing system, the conjugation frequency could be dramatically increased. Transfer of the pSym megaplasmid into *A. tumefaciens* from the α-proteobacterium *Rhizobium trifolii* has been previously shown (Hooykaas *et al.*, 1981, Hooykaas *et al.*, 1982). Wild type *A. tumefaciens* C58, of which A676/A677 are derivatives, possesses the Ti and At plasmids which are both RepABC-type and possess the same conjugation system as pHCG3 (Class I) (Danino *et al.* 2003, Wood *et al.* 2001, Wetzel *et al.* 2015). The intergenic *repB-repC* region of these plasmids exerts incompatibility (Pappas 2008). However, since they have been cured from A676 or A677, this may have facilitated the transfer of pHCG3_*traM*.

The pHCG3_*traM* megaplasmid was also successfully transferred into *P.* cws, demonstrating for the first time that a RepABC family plasmid can be maintained in a γ-proteobacterium. *P.* cws was found to be a member of the *Pseudomonas fluorescens* group (Fig. 2.24) and was identified as a species of *Pseudomonas carnis* (Lick *et al.* 2020), following a whole genome comparison (Table 4.6). pHCG3_*traM* could not be

transferred into other closely related species from the *P. fluorescens* group (Table 4.3), potentially due to issues with plasmid stability, replication, plasmid compatibility or due to native restriction systems. To investigate this further, functionality of the replication origin of pHCG3 could be tested in these strains using pLO3_*repABC.* Stable maintenance of pHCG3_*traM* in *P.* cws was unexpected given RepABC plasmids are reportedly found exclusively in α-proteobacteria (Pinto *et al.* 2012) with the possible exception of certain β-proteobacteria (Shintani *et al.* 2015). Megaplasmid transfer has been demonstrated between species of *Pseudomonas* (Romanchuk *et al.* 2014, Dougherty *et al.* 2014), but to date there has been no documentation of γ-proteobacteria containing a *repABC* plasmid. However, the megaplasmid p12969-DIM from *Pseudomonas putida* carries an additional potential *repABC*/*mazFE* replication system, possibly to expand its host range (Sun *et al.* 2016). Additionally several other γ-proteobacteria possess homologues of Rep proteins, including *Streptomyces sp.* whose homologues displays around 76% similarity to *O. carboxidovorans* RepC (Tarkka *et al.* 2015) and *Pseudomonas bauzanensis* strain W13Z2 V14_32 whose RepC displays 50 % similarity (Wang *et al.* 2014). *Blastopirellula* sp. isolate NORP48 contains a homologue of RepA, which displays 39.18% similarity to *O. carboxidovorans* RepA (Tully *et al.* 2018). One could speculate that some RepC homologues have evolved to function in certain γ-proteobacteria and are consequently able to house RepABC family plasmids, since *repC* is solely capable of initiating their replication (Cervantes-Rivera *et al.* 2011, Pinto *et al.* 2012).

Whilst pHCG3_*traM* could be maintained in *A. tumefaciens* and *P.* cws, its presence did not enable either of these species to utilise CO or CO_2 as a carbon source, at least not at levels detectable through simple growth experiments (Fig. 4.19 & 4.24). Both the *cox* or *cbb* gene clusters displayed 100% coverage and contained no SNPs in either strain (Table 4.2 & 4.4), therefore the genetic information for CODH and CO_2 fixation via the Calvin-Benson-Bassham cycle was complete. Presumably, therefore, both species may have been unable to transcribe the relevant gene clusters or produce the encoded proteins. To elucidate whether *cox* and *cbb* genes are actively transcribed and at what level, RNA-seq or qPCR could be performed on these strains. The head space of cultures of the pHCG3_*traM* containing strains could also be analysed by gas

chromatography to observe changes in the volumes of CO or CO_2 to elucidate whether they are being metabolised. Additionally, experiments employing $^{13}CO_{(2)}$ may enable a more sensitive approach to analysing functionality of the introduced enzymes and pathways, as demonstrated by Heinrich *et al.* (2018). Furthermore, CODH enzyme assays could be performed on cell lysates to determine levels of CO DH activity (Meyer & Schlegel 1978). Specific issues such as absence or insufficient expression of pathway components could then be addressed through further genetic engineering, in combination with strain evolution. For instance, once pathways are functional and allow incorporation of CO-derived carbon at a low level, strains may be further evolved through growth in continuous culture, in the presence of excess CO, but limited with respect to an organic carbon source.

None of the selected putative COTF binding regions upstream of selected *cox* genes on pHCG3 were able to induce RFP expression when cloned into reporter constructs in the presence or absence of CO in *E. coli* or *C. necator*, suggesting that these regions do not function in these organisms (Fig. 4.28 & 4.29). Lack of *cox* gene expression may be one of reasons why no growth on CO and CO_2 was observed, as discussed above. The same RFP reporter failed to be expressed in the native organism, *O. carboxidovorans* (Fig. 4.30). However, it was unclear whether these region were inactive in this organism since the positive control plasmid containing the RFP-inducible system did not appear to function either. The reasons for this are unclear. The transcription factor binding sites may lie outside of these regions on pHCG3, which could be tested using the same system. Alternatively, *rfp* expression might have been poor due to other reasons, such as unfavourable codon usage. However, in principle there is no need to employ the native regulatory system to drive *cox* gene expression. An inducible promoter system could be integrated into pHCG3 to facilitate *cox* gene expression, in a similar fashion to that described by Heinrich *et al.* (2018).

To summarise, in this study the *O. carboxidovorans* megaplasmid pHCG3 was modified to constitutively produce AHLs to boost its conjugative functions. The AHLs produced by pHCG3 and the AHLs under regulation of TraM were identified. This demonstrated that that TraM governs the expression of *traI* in *O.*

carboxidovorans. Putative COTFs at the start of the *cox* operon were found to be non-functional in *C. necator*, indicating why previous attempts to integrate and express *cox* genes in this organism were unsuccessful. pHCG3 was transferred successfully into plasmid-less strains of *A. tumefaciens* (A676 and A677) and a previously unknown species of *Pseudomonas* (*P.* cws), which was confirmed by next generation Illumina sequencing. The megaplasmid could not be transferred into *C. necator*, which is likely due to the non-functionality of the pHCG3 replicon *repABC,* which was demonstrated. Future work to integrate the functional replicon from pMOL28 into pHCG3 may enable its successful conjugation into *C. necator*, allowing the organism to utilise CO as a sole carbon source.

Chapter 5: General discussion

5.1. Background

The need to tackle the problem of greenhouse gases is incumbent, given its connection to current environment damage (Fell et al. 2017, Hoegh-Guldberg et al. 2017, United Nations Framework Convention on Climate Change 2015). The development of new technologies for C1 gas fermentation to produce valuable bioproducts may be an important strategy to reduce greenhouse gas emissions in the future. *C. necator* H16 is a particularly well suited organism for C1 fermentation due to its ability to utilise CO_2 + H_2 as a source of carbon and energy (Cook & Schlegel 1978, Müller *et al.* 2013, Lu & Yu 2017). C1 feedstocks such as syngas typically contain high concentrations of CO, which can comprise of up to 60% of the gas mixture (National Energy Technology Laboratory 2018). CO has been shown to significantly inhibit the growth of numerous bacteria including *C. necator* H16 (Cypionka and Meyer 1982, King 2003, Shen *et al.* 2014). Therefore, the adaptation of *C. necator* to CO formed an important part of this project, as this is an essential step for the organism to grow efficiently using CO rich industrial waste gases.

Another important strategy to achieve this goal was to enable CO utilisation in *C. necator*. Certain aerobic bacteria possess the unique ability to utilise CO as sole source of carbon and energy via CODH (King & Weber 2007, King 2003, Mörsdorf *et al.* 1992). The CODH of *O. carboxidovorans* has been well studied and harbours a well-characterised cluster of *cox* genes (Rovaletti *et al.* 2019, Kreß *et al.* 2014, Resch *et al.* 2005, Kraut & Meyer 1988). Operons within the *cox* cluster have been well characterised, including *cosMSL*, which encode proteins forming the subunits of CODH (Dobbek *et al.* 2000) and *coxDEFG* and *coxI* which encode posttranslational maturation factors which modify the bimetallic [MoO2S-Cu] cluster in CODH (Kaufmann *et al.* 2018, Pelzmann *et al.* 2009, Dobbek *et al.* 2002). A subset of the *cox* operon genes have been expressed in *C. necator*, enabling utilisation of CO, although with much reduced efficiency in comparison to wild type *O. carboxidovorans* and requiring additional H_2 as a reducing equivalent (Heinrich *et al.* 2018). The *cox* genes are harboured on the megaplasmid pHCG3, which additionally contains *cbb* and *hox* genes for the utilisation of CO_2 + H_2 (Volland *et al.* 2011, Fuhrmann *et al.* 2003). Examining the conjugative

capacity of pHCG3 became an essential part of this work, as its transfer into *C. necator* may potentially enable utilisation of CO as a sole source of carbon and energy. Conjugation into closely related strains was also an important step to achieve, to establish the conjugative capacity of pHCG3 and perform additional genetic modifications if needed.

5.2. Key findings and future prospects

5.2.1. Genotypic and phenotypic analysis of carbon monoxide resistant *C. necator*

Through continuous subculturing of *C. necator* in the presence of CO with an additional carbon source of fructose, high tolerance to CO was successfully achieved. Ten separately evolving lineages all displayed a fast growing phenotype in the presence of CO when compared to the wild type starter culture. Individual isolates from these lineages also displayed the same level of CO resistance with a significantly higher growth rate than the wild type. To ascertain what the genetic differences were between these isolates and the wild type, which may cause this phenotype, Illumina whole genome sequencing was performed. This revealed a unique profile of mutations across each of the ten isolates, including SNPs within and upstream of *phcA* and a 12 base pair deletion within *phcR.* These genes are involved in the Phc cell density sensing system in *C. necator*, which shows a high degree of similarity to the homologous *R. solanacearum* system (Clough *et al.* 1997). The histidine kinase response regulator PhcR/PhcS regulates PhcA in this system, the latter sensing endogenous 3-OH-PAME signal molecules produced by PhcB (Clough *et al.* 1997, Flavier *et al.* 1997, Garg *et al.* 2002). Additionally two SNPs were found in the RNA polymerase sigma-54 factor gene, which in this organism is thought to be involved with activating denitrification genes, hydrogenase activity in *C. necator* (Merrick 1993, Schwartz *et al.* 1998) and controls motility and growth mode in *R. solanacerum* (Ray *et al.* 2015) However, the only mutation that was ubiquitous to all ten isolates was a SNP upstream of the cytochrome *bd* ubiquinol oxidase II gene *cydA2B2.* This respiratory quinol:O_2 oxidoreductase, is known to be active in *E. coli* during oxygen-limiting circumstances and other environmentally stressful conditions (Bristov *et al.* 2009, Trutko *et al.* 2003, Cotter *et al.* 1990, Avetisyan *et al.* 1992).

Future work may include further evolution of strains in the presence of syngas, i.e. when growing on CO_2 and H_2 in the presence of CO. This would present conditions more akin to those present in industrial gas fermentations. Although known to naturally exhibit a high degree of CO resistance (Van der Linden *et al.* 2004, Bernard *et al.* 1974), hydrogenases in *C. necator* and their expression may be further improved under these conditions, together with other components of the respiratory chain utilised under these conditions.

5.2.2. Characterisation of selected mutations

A mutant library was generated in *C. necator* wild type background in order to characterise the chosen mutations, in a similar manner to experiments conducted in *E. coli* (Mundhada *et al.* 2017, Mohamed *et al.* 2019). A series of knock-out / knock-in plasmids was created, which enabled the deletion of around 100 bp upstream and downstream of the mutation site and the subsequent integration of each mutation along with deleted wild type sequence. This method was repeated to generate double and triple knock-in strains, based on the mutational profiles of the evolved isolates. The growth of *C. necator cydA2B2*, *C. necator phcR* 12bp, *C. necator phcA* US, *C. necator* RNApol *4574* and *C. necator* RNApol *5744* was compared to the wild type. In the absence of CO, mutations *phcR* 12 bp, *phcA* US, RNApol *4574* and RNApol *5744* all enhanced the growth of *C. necator*, leading to a higher peak OD_{600}. Excluding *phcR*, these mutations slightly increased the growth of *C. necator* in the presence of CO. However, this increase was not statistically significant. Only the *cydA2B2* SNP mutation generated an evolved phenotype displaying significantly faster growth than the wild type. Interestingly this mutation had no effect on growth when *C. necator* was cultured in the absence of CO. When the other mutations were generated in combination with *cydA2B2*, the growth of *C. necator* improved slightly in comparison to the *cydA2B2* mutant in the presence and absence of CO, though this was not significant.

5.2.3. Characterisation of cytochrome *bd* ubiquinol oxidase expression

Since the *cydA2B2* SNP mutation generated the evolved fast growing phenotype, the mutation and gene were characterised. From the mutation a promoter site upstream of *cydA2B2* was predicted. Therefore, it was presumed that this mutation was causing a

regulatory effect on this gene. Two RFP reporter constructs were created; one harboured the wild type upstream region of *cydA2B2* and one harbouring the same region containing the altered nucleotide. The latter induced fluorescence in the reporter whereas the former failed to produce any signal above the wild type negative control background, indicating that the SNP increases gene expression. The operon was also expressed in *C. necator* via an arabinose inducible plasmid, which gave cells harbouring it a growth advantage when grown in the presence of CO. Following an in frame knock-out of *cydA2B2*, heightened sensitivity to CO was observed in comparison to the wild type. These results imply that cytochrome *bd* ubiquinol oxidase plays an important role in CO tolerance. A possible hypothesis is that the natural pattern of cytochrome expression in the wild type (Kömen et al, 1991) leads to decreased CO tolerance which is altered in the *cydA2B2* mutant and evolved isolates due to the increase in *cyd2B2* expression. This overexpression reduced the highest OD_{600} reached by *C. necator,* a trait that has also been observed in *E. coli* and *Corynebacterium glutamicum* (Calhoun *et al.* 1993, Kabus *et al.* 2007). It would be informative to further investigate overexpression and deletion at the protein level as well as the resulting impact on the respiratory chain. For instance, changes in presence and level of specific cytochromes in such strains might be investigated in the presence and absence of CO, e.g. by extraction of cell membranes and analysis of the associated cytochromes by determining their specific absorption spectra (Borisov *et al.* 2011). Additionally, RNAseq or qPCR could be performed on *C. necator cydA2B2* SNP and the deletion mutant, to elucidate whether upregulation of this gene is observed in the presence of CO. This may also reveal other direct or indirect consequences resulting from overexpression or absence of the gene.

5.2.4. Elucidation of the *O. carboxidovorans* quorum sensing systems

Examination of certain components of the pHCG3-encoded quorum sensing systems gave insights into the type of signalling molecules produced and how they are regulated in *O. carboxidovorans.* The AHL synthase genes of pHCG3, *traI* and *solI* were found to produce specific AHLs when expressed via pBBR1MCS-2 in *E. coli* and analysed by LC-MS/MS. Following the inactivation of the negative regulator of the quorum sensing system, TraM, AHL production was observed in *O. carboxidovorans*. By contrast, the wild type failed to produce any detectable AHLs. The AHLs produced by

the *traM* mutant, OM5_*traM*, included the two AHLs produced by TraI, suggesting that TraM interacts with the transcriptional regulator of *traI*, TraR. This mechanism of regulation is analogous to that used by *A. tumefaciens* (Fuqua *et al.* 1995, Hwang *et al.* 1995, Luo *et al.* 2000, Churchill & Chen 2011). Four AHLs produced by SolI in *E. coli* matched the pHCG3_*traM* AHL profile, suggesting that TraM possibly interacts with the putative regulator SolR. An alternative explanation could be that the *solRI* genes are controlled by the TraRI QS system in a regulatory QS hierarchy, examples of which have been observed for several other bacteria (Lazdunski *et al.* 2004), including members of the genus *Rhizobium* and its close relatives (Zheng *et al.* 2015, Federick *et al.* 2011, Wisniewski-Dye & Downie 2002)

Further research into the other genes of this QS system should be performed to elucidate the precise mechanisms governing the synthesis of AHLs in *O. carboxidovorans* and genes controlled by the SolR and TraR regulators. This may involve the inactivation of QS individual genes or combinations thereof, analysis of the resulting phenotypes, and identification of differentially regulated genes using transcriptomic techniques. It would also be interesting to use such mutants to investigate the potential cell-density (i.e. AHL) dependent transfer of pHCG3 in *O. carboxidovorans* populations. In *A. tumefaciens*, conjugal transfer of the Ti plasmid is stimulated at high donor densities, potentially resulting in frequent donor-donor matings (Alt-Mörbe *et al.* 1996, Fuqua & Winans 1996).

Whether this is a wasteful process, prevented through entry exclusion mechanisms, or desirable to enable recombination and potentially repair between different plasmid variants could then be investigated. It should be noted that some bacteria, for example *Enterococcus faecalis*, are able to sense recipient cells due to their production of chromosomally encoded signals, and to prevent plasmid transfer if donor concentrations are too high (Bandyopadhyay *et al.* 2016, Koraimann & Wagner 2014).

5.2.5. Analysis of the conjugative capacity of the megaplasmid pHCG3 and expression of encoded functions

OM5_*traM* was successfully conjugated into two plasmid-less derivatives of *A. tumefaciens* C58 (strains A676 and A677) for which acceptance of rhizobial megaplasmids has already been proved (Danino *et al.* 2003). The presence of pHCG3 was confirmed using PCR and whole genome sequencing, which upon further analysis using CLC bio software eliminated the possibility of *O. carboxidovorans* donor contamination. RepABC plasmids such as pHCG3 are typically found in α-proteobacteria (Pinto *et al.* 2012), therefore transfer into A676 and A677 was expected. Additionally, their parental strain C58 houses two RepABC-family plasmids which bear similarity to pHCG3 (Wetzel *et al.* 2015).

Unexpectedly, a species of *Pseudomonas* (*P*. cws) already containing pHCG3_*traM* was accidentally isolated during conjugation experiments, representing a genus *repABC* plasmids have never been reported in. A plasmid-less isolate of this strain was then successfully transformed with pHCG3_*traM*. However, the plasmid could not be transferred into *C. necator, E. coli* or *Pseudomonas* species closely matching the 16S rRNA sequence of *P.* cws, suggesting that pHCG3_*traM* and possibly other *repABC* plasmids can only be stably maintained in certain proteobacteria potentially due to plasmid instability or incompatibility (Döhlemann *et al.* 2017, Pérez-Oseguera & Cevallos 2013). Following its transfer into *A. tumefaciens,* evidence of AHL production could be seen, implying that the responsible genes on pHCG3 were being actively transcribed in this organism. However, *P.* cws utilisation of CO and CO_2 +H_2 was not observed under the employed conditions, suggesting only certain genes may be transcribed or active enzymes encoded by *cox* and *cbb* genes cannot be produced. For future analysis, RNA-seq or qPCR could be performed on these strains to ascertain which genes are being actively transcribed. In addition, enzyme assays, e.g. for CODH or enzymes of the Calvin-Benson-Bassham cycle may be employed to establish absence or presence of required activities. One possible explanation for the observed lack of CO and CO_2 utilisation is poor mRNA translation due to differences in codon usage. It may therefore become necessary to harmonise the codons of key genes to achieve the required levels of expression (Quax *et al.* 2015, Kudla *et al.* 2009, Corredu *et al.* 2019, Angov *et al.* 2008). Regarding the observed conjugation frequencies, these are in the

range observed for other α-proteobacterial megaplasmids (Danino *et al.* 2003, Pretorius-Güth *et al.* 1990). However, they were obtained under conditions of, presumably, maximum *tra* gene expression, due to inactivation of the TraM regulator. Hence, it would be of interest to observe transfer frequencies in presence of the native regulatory system. For this, a differently marked version of pHCG3 needs to be created, to allow conjugation frequencies of pHCG3 to be determined, with and without TraM being mutated. Such constructs could also be used to analyse maintenance and stability of pHCG3 in the heterologous hosts. Additionally, conjugation frequency could be monitored in the presence and absence CO or CO+H_2, since the transfer genes *trbL* and *trbI* are upreguated in the presence of syngas (Paul *et al.* 2010). In the current study, antibiotic selection was used so that transconjugants maintained the megaplasmid throughout the different experiments undertaken and there is currently no information on how quickly it would be lost in its absence. One of the key aims of this study, the transfer of the pHCG3 into *C. necator* could not be achieved. However, given the achieved transfer into *A. tumefaciens*, a genetically more amenable host, efforts should be made to equip pHCG3 with the pMOL28 replication operon, which was proven to function in *C. necator* (4.2.5.3). This may enable stable maintenance of pHCG3 and possibly utilisation of CO in this organism.

5.2.6. Analysis of CO regulation

Since utilisation of CO was not possible following the integration of the pHCG3-derived *cox* genes (Alagesan & Malys, unpublished findings) but was possible via plasmid expression (Heinrich *et al.* 2018), investigation into possible CO regulation of these genes was undertaken. Four different putative TFBSs upstream of *coxM* were therefore cloned into an RFP-expressing, inducible plasmid and transferred into *O. carboxidovorans*, *E. coli* and *C. necator* to monitor gene expression. Three of these plasmids encoded the putative transcriptional regulator CoxC (Fuhrmann *et al.* 2003), which was postulated to active the *cox* operon. None of the constructs gave RFP expression above that of the negative control in any of the organisms, in the presence or absence of CO. The RFP reporter system was tested in all organisms, which indicated that it could function in *E. coli* and *C. necator*, consistent with its use in previous studies (Hanko *et al.* 2018, Hanko *et al.* 2017), but could not in *O. carboxidovorans*. Therefore, it was concluded that none of the putative TFBSs functioned in *C. necator* or *E. coli*

but further investigation would be needed to decipher whether they were responsive in the host organism. This result potentially explained why *cox* gene-integrated *C. necator* was unable to utilise CO (Alagesan & Malys unpublished findings) and led to the conclusion that a strong, heterologous promoter may be required for sufficient *cox* gene expression as demonstrated by Heinrich *et al.* (2018). Future work could therefore aim to integrate a suitable promoter upstream of the *cox* operon on pHCG3 prior to transfer into *C. necator*.

5.3. Concluding statements

This study has contributed to the improvement of *C. necator* as an aerobic chassis for its potential future usage in larger scale fermentations using CO rich waste gases.

Adaptive evolution to CO has produced strains of *C. necator* with heightened CO tolerance and determination of the genes responsible has enabled the generation of artificially engineered strains which mimic these evolved traits. Exploration of the quorum sensing system of *O. carboxidovorans* megaplasmid has resulted in elucidation of the mechanisms governing AHL synthesis and its transfer into different species both closely and distantly related. Although the transfer of pHCG3 into *C. necator* to enable CO utilisation was not possible, the expansion of its host range and the prospect of future modifications to the megaplasmid provide confidence that this is still feasible.

Chapter 6: Appendix

6.1. Read coverage distribution for pHCG3_*traM* containing strains

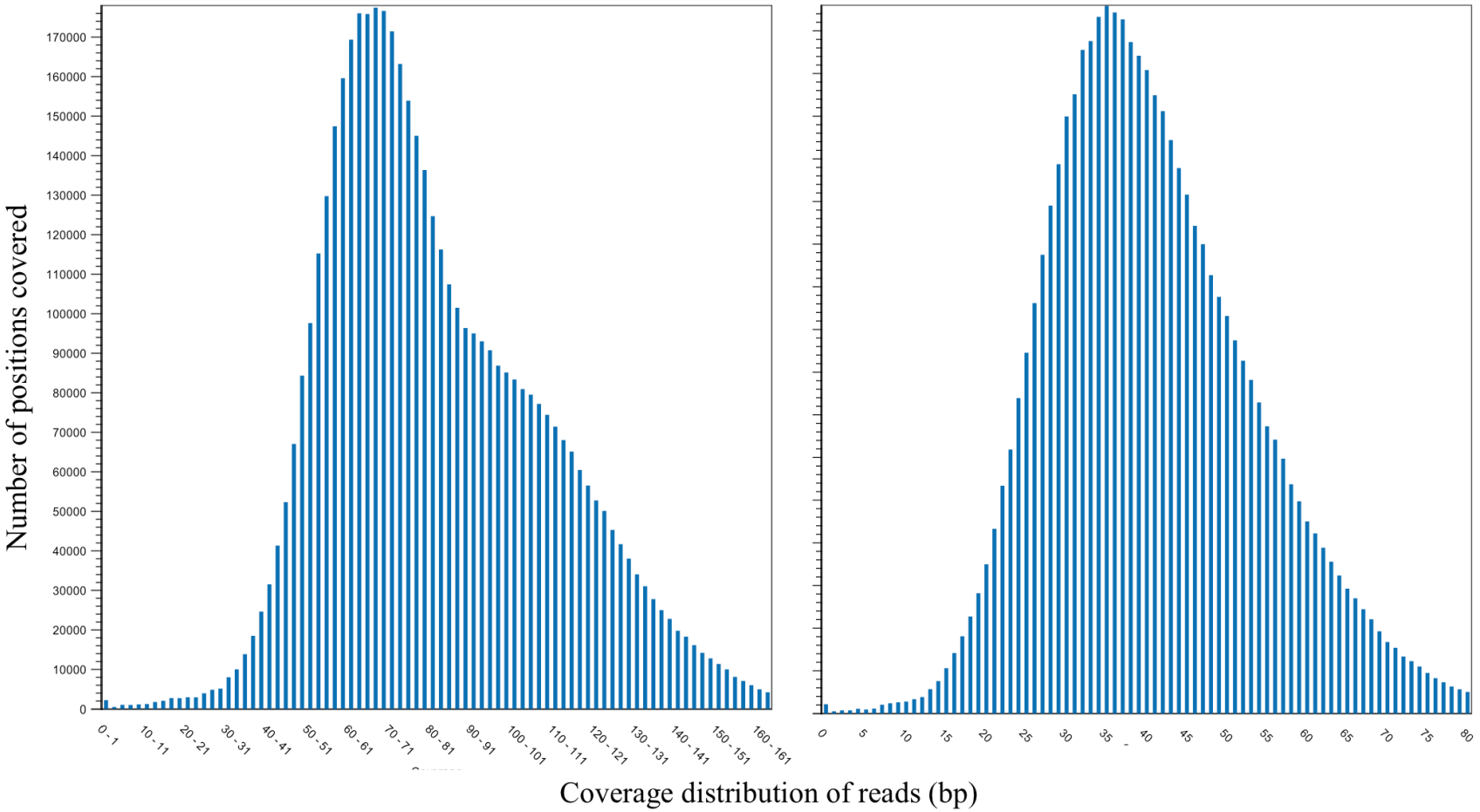


Figure 6.1. Read coverage distribution for A676 pHCG3_*traM* (left) and A677 pHCG3_*traM* mapped to the published *A. tumefacines* circular chromosome (Slater *et al.* 2008) ((right).

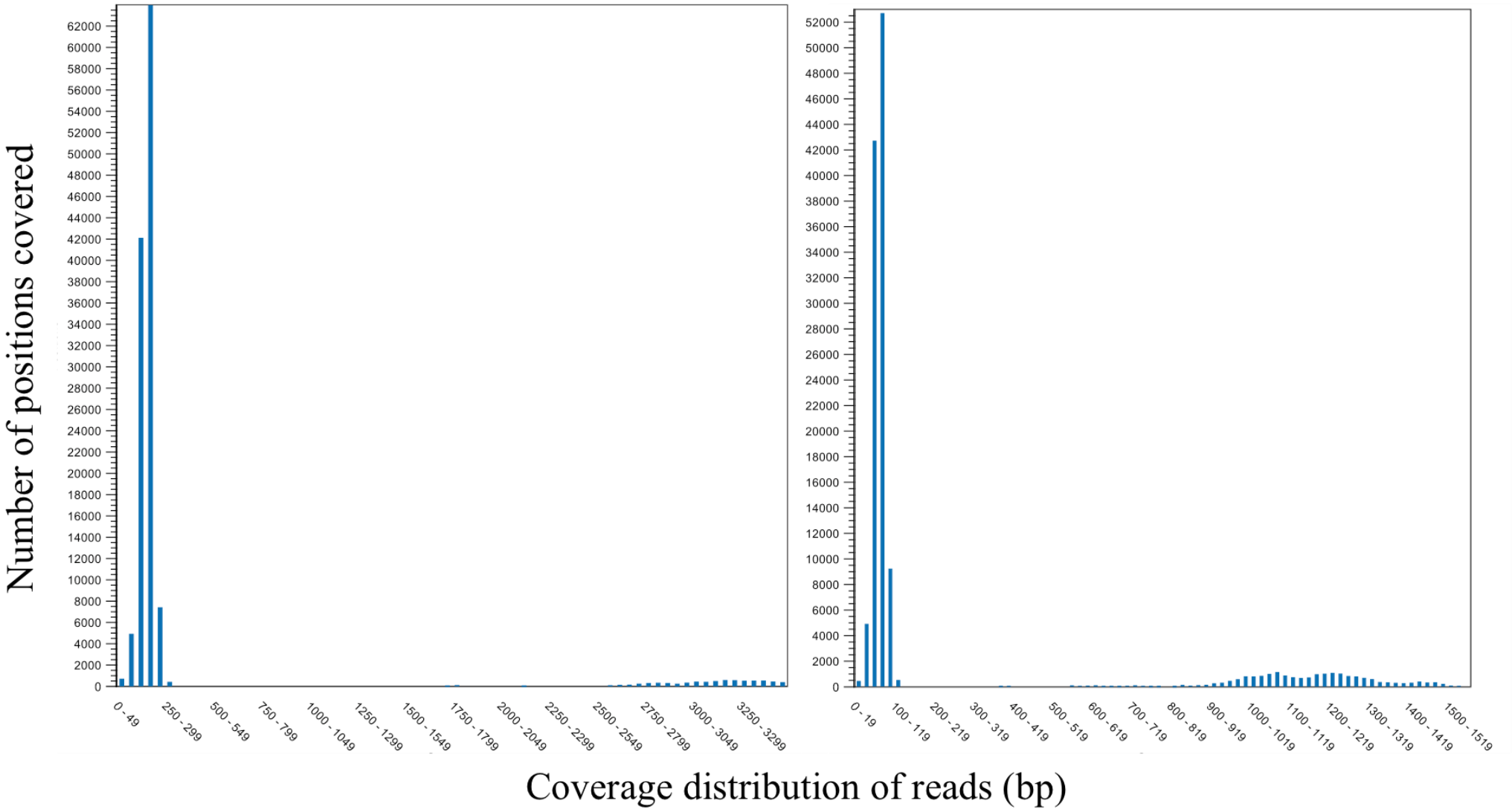


Figure 6.2. Read coverage distribution for A676 pHCG3_*traM* (left) and A677 pHCG3_*traM* (right) mapped to the published OM5 pHCG3 sequence (Volland *et al.* 2011).

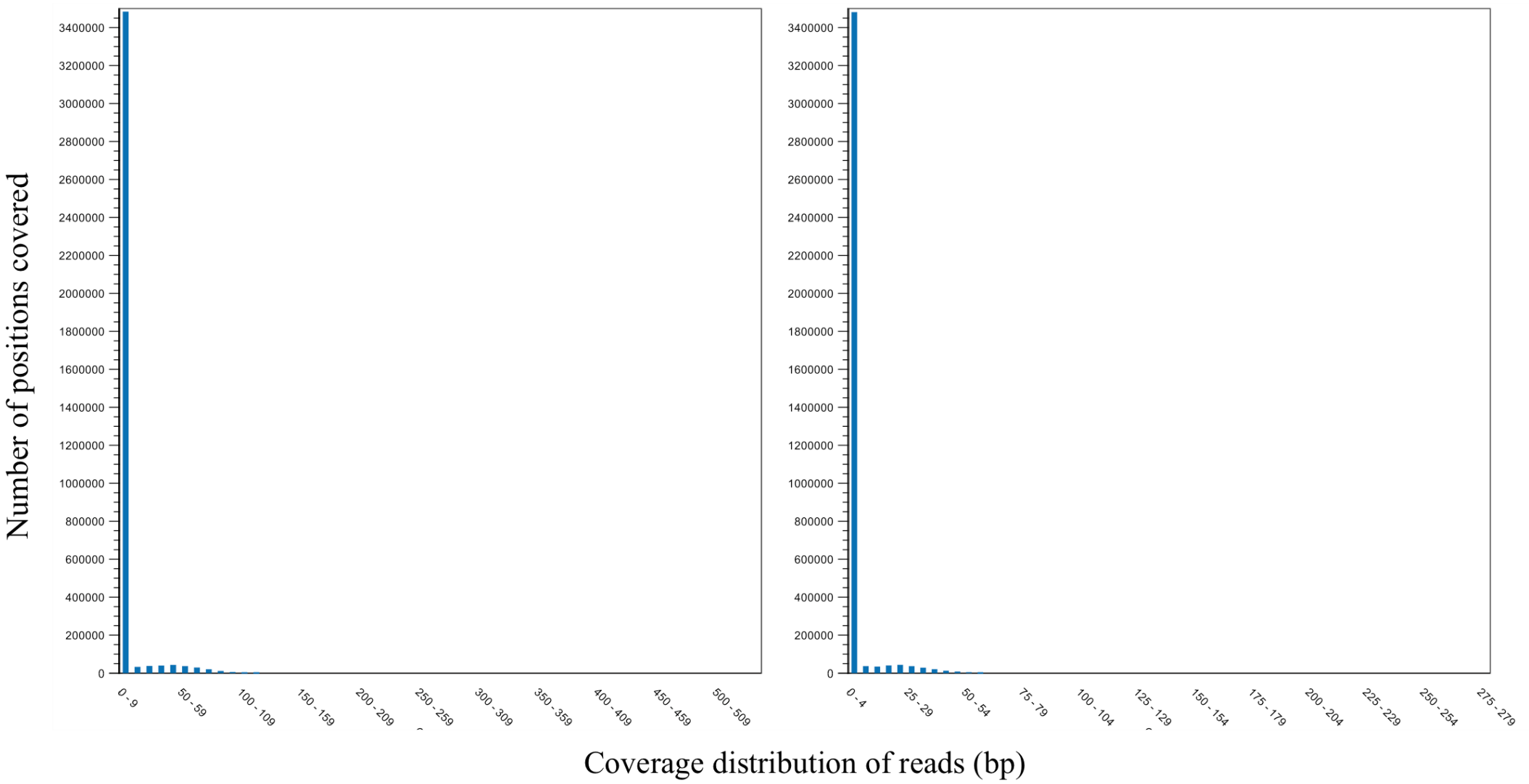


Figure 6.3. Read coverage distribution for A676 pHCG3_*traM* (left) and A677 pHCG3_*traM* (right) mapped to the published OM5 circular chromosome (Volland *et al.* 2011).

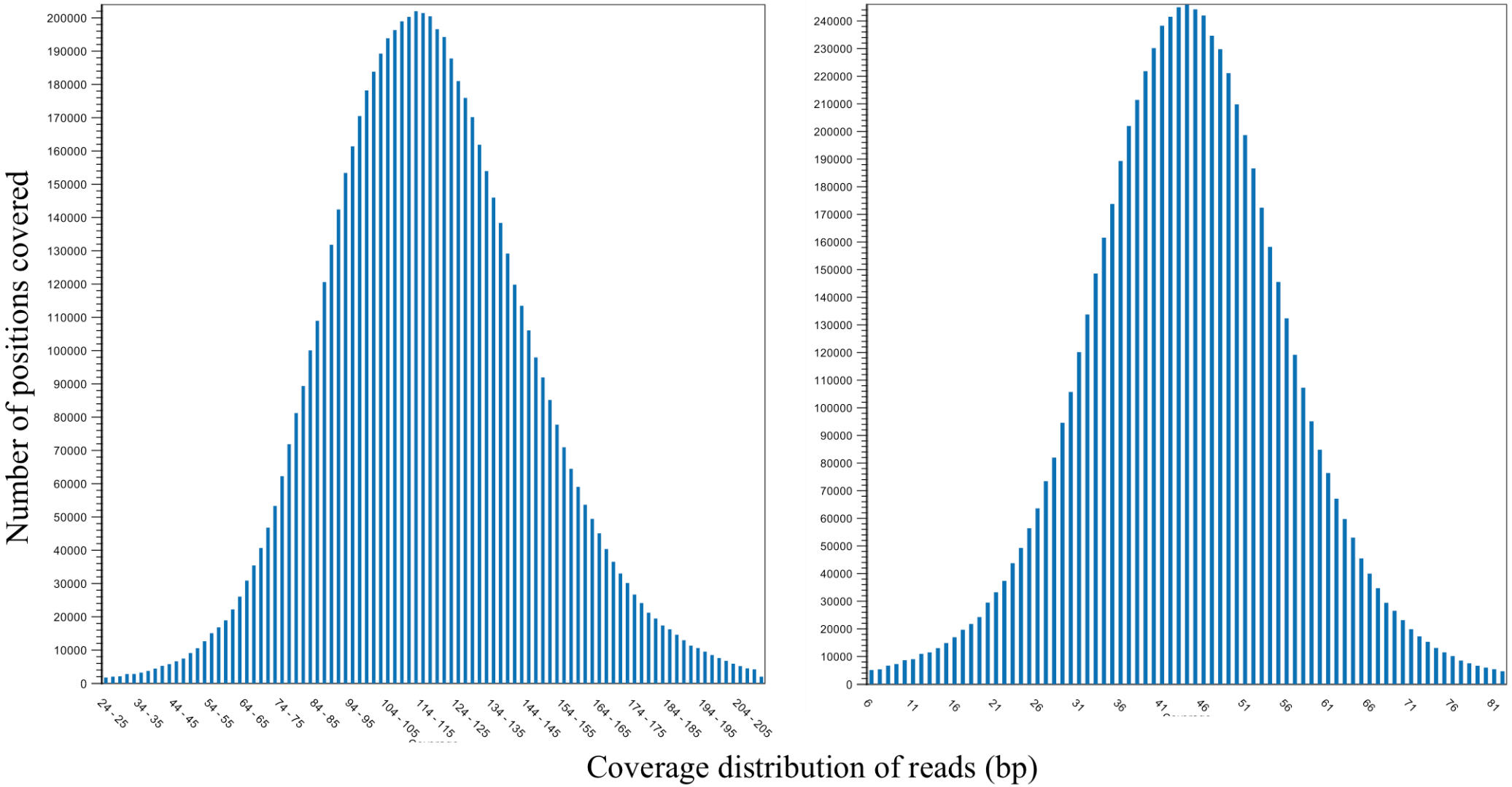


Figure 6.4. Read coverage distribution for *P.* cws pHCG3_*traM*_A (left) and *P.* cws pHCG3_*traM*_B (right) mapped to the assembled *P.* cws parental genome.

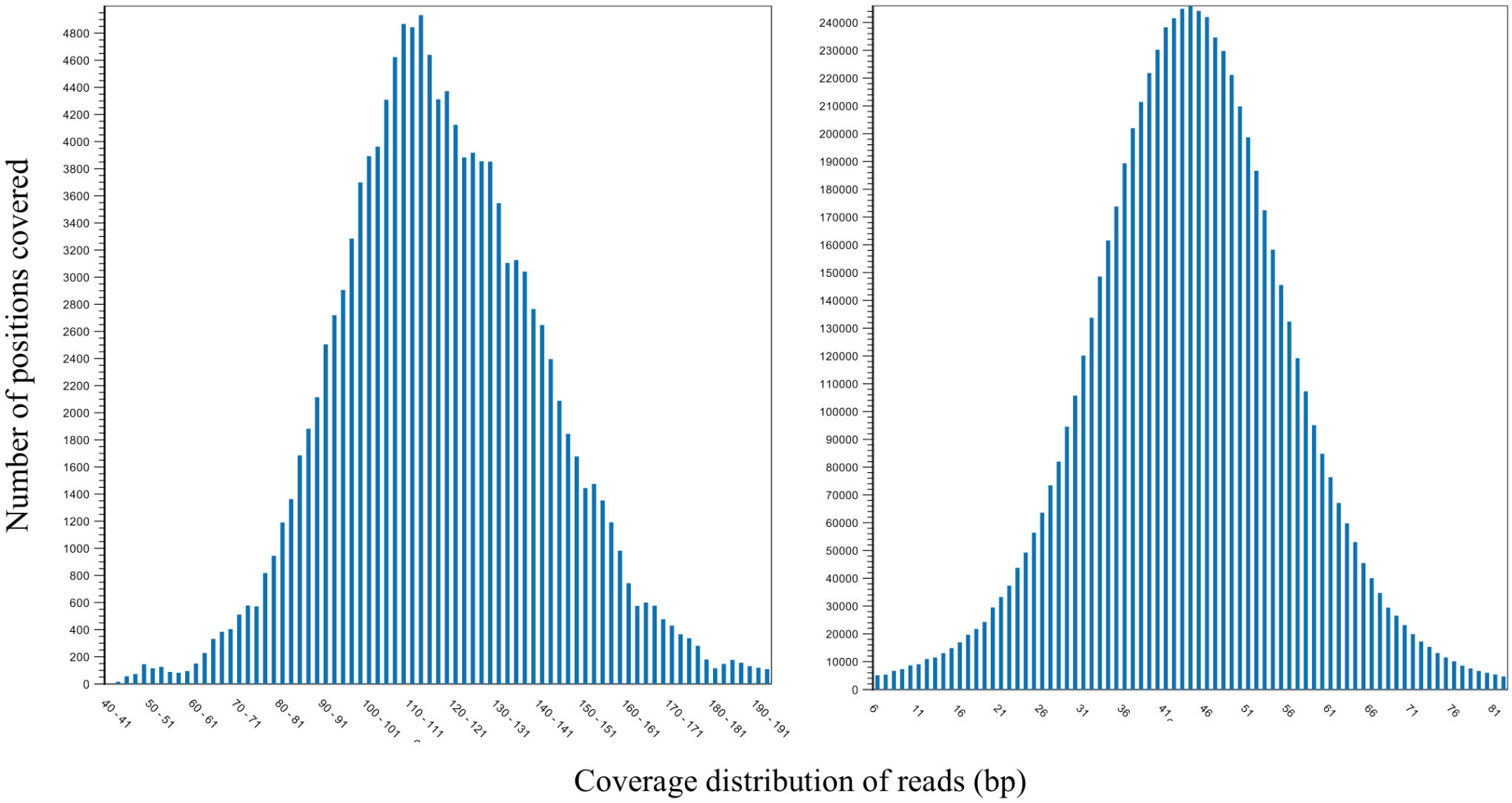


Figure 6.5. Read coverage distribution for *P.* cws pHCG3_*traM*_A (left) and *P.* cws pHCG3_*traM*_B (right) mapped to the OM5 pHCG3 sequence (Volland *et al.* 2011).

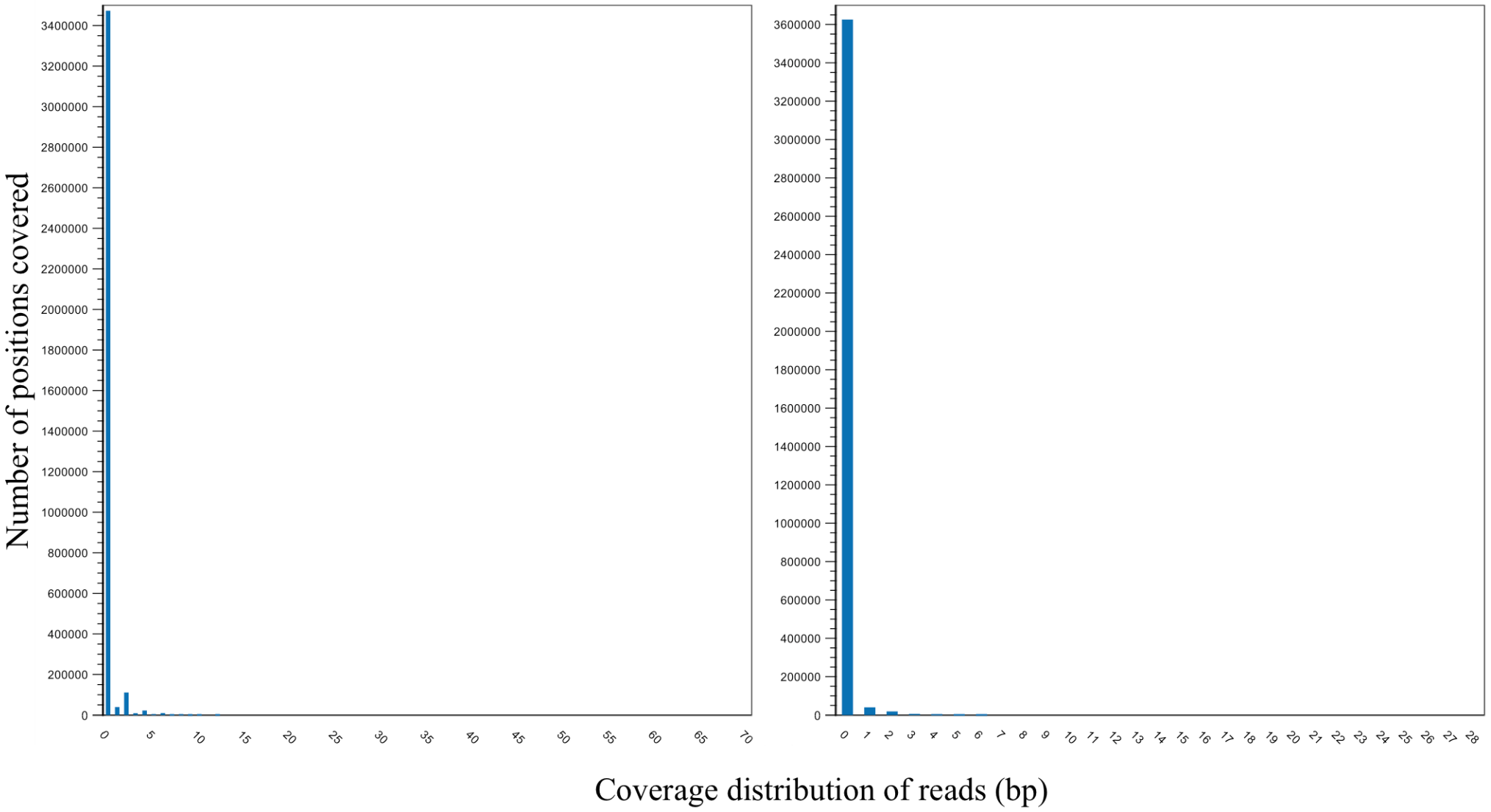


Figure 6.6. Read coverage distribution for *P.* cws pHCG3_*traM*_A (left) and *P.* cws pHCG3_*traM*_B (right) mapped to the published OM5 circular chromosome (Volland *et al.* 2011).

6.2. 16S *rRNA* sequences

Pseudomonas cws

16S *rRNA*

CTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTC
GAGCGGTAGAGAGAAGCTTGCTTCTCTTGAGAGCGGCGGACGGGTGAGTAATGCCTAGGA
ATCTGCCTGGTAGTGGGGGATAACGTTCGGAAACGGACGCTAATACCGCATACGTCCTAC
GGGAGAAAGCAGGGGACCTTCGGGCCTTGCGCTATCAGATGAGCCTAGGTCGGATTAGCT
AGTTGGTGGGGTAATGGCTCACCAAGGCGACGATCCGTAACTGGTCTGAGAGGATGATCA
GTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTG
GACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGATTGT
AAAGCACTTTAAGTTGGGAGGAAGGGTTGTAGATTAATACTCTGCAATTTTGACGTTACCG
ACAGAATAAGCACCGGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGGTGCAAGC
GTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTTGTTAAGTTGGATGTGAAA
TCCCCGGGCTCAACCTGGGAACTGCATTCAAAACTGACTGACTAGAGTATGGTAGAGGGT
GGTGGAATTTCCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAA
GGCGACCACCTGGACTAATACTGACACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGAT
TAGATACCCTGGTAGTCCACGCCGTAAACGATGTCAACTAGCCGTTGGAAGCCTTGAGCTT
TTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAA
CTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAA
CGCGAAGAACCTTACCAGGCCTTGACATCCAATGAACTTTCCAGAGATGGATTGGTGCCTT
CGGGAACATTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGT
TAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGTTACCAGCACGTAATGGTGGGCACTCT
AAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCC
CTTACGGCCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCCAAGCCGCGA
GGTGGAGCTAATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCG
TGAAGTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCC
TTGTACACACCGCCCGTCACACCATGGGAGTGGGTTGCACCAGAAGTAGCTAGTCTAACC
TTCGGGAGGACGGTTACCACGGTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAGC
CGTAGGGGAACCTGCGGCTGGATCACCTCCTT

6.2.1. 16S *rRNA* and *rpoB* from GenBank

Pseudomonas sp. J380

16S *rRNA*

AAGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCA
TGAATCACACCGTGGTAACCGTCCTCCCGAAGGTTAGACTAGCTACTTCTGGTGCAACCCA
CTCCCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGACATTCT
GATTCGCGATTACTAGCGATTCCGACTTCACGCAGTCGAGTTGCAGACTGCGATCCGGACT
ACGATCGGTTTTATGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATT
GTAGCACGTGTGTAGCCCAGGCCGTAAGGGCCATGATGACTTGACGTCATCCCCACCTTCC
TCCGGTTTGTCACCGGCAGTCTCCTTAGAGTGCCCACCATTACGTGCTGGTAACTAAGGAC
AAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCC
ATGCAGCACCTGTCTCAATGTTCCCGAAGGCACCAATCCATCTCTGGAAAGTTCATTGGAT
GTCAAGGCCTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGT
GCGGGCCCCCGTCAATTCATTTGAGTTTTAACCTTGCGGCCGTACTCCCCAGGCGGTCAAC
TTAATGCGTTAGCTGCGCCACTAAAAGCTCAAGGCTTCCAACGGCTAGTTGACATCGTTTA
CGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGCACCTCAGTGTC
AGTATTAGTCCAGGTGGTCGCCTTCGCCACTGGTGTTCCTTCCTATATCTACGCATTTCACC
GCTACACAGGAAATTCCACCACCCTCTACCATACTCTAGTCAGTCAGTTTTGAATGCAGTT
CCCAGGTTGAGCCCGGGGATTTCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGC
CCAGTAATTCCGATTAACGCTTGCACCCTCTGTATTACCGCGGCTGCTGGCACAGAGTTAG
CCGGTGCTTATTCTGTCGGTAACGTCAAAATTGCAGAGTATTAATCTACAACCCTTCCTCC
CAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTCACACACGCGGCATGGCTGGATCAG

GCTTTCGCCCATTGTCCAATATTCCCCACTGCTGCCTCCCGTAGGAGTCTGGACCGTGTCTC
AGTTCCAGTGTGACTGATCATCCTCTCAGACCAGTTACGGATCGTCGCCTTGGTGAGCCAT
TACCCCACCAACTAGCTAATCCGACCTAGGCTCATCTGATAGCGCAAGGCCCGAAGGTCC
CCTGCTTTCTCCCGTAGGACGTATGCGGTATTAGCGTCCGTTTCCGAACGTTATCCCCCACT
ACCAGGCAGATTCCTAGGCATTACTCACCCGTCCGCCGCTCTCAAGAGAAGCAAGCTTCTC
TCTACCGCTCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCCATGATCAA
ACTCTTCAG

Pseudomonas sp. MYb193

AAGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCA
TGAATCACACCGTGGTAACCGTCCTCCCGAAGGTTAGACTAGCTACTTCTGGTGCAACCCA
CTCCCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGACATTCT
GATTCGCGATTACTAGCGATTCCGACTTCACGCAGTCGAGTTGCAGACTGCGATCCGGACT
ACGATCGGTTTTATGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATT
GTAGCACGTGTGTAGCCCAGGCCGTAAGGGCCATGATGACTTGACGTCATCCCCACCTTCC
TCCGGTTTGTCACCGGCAGTCTCCTTAGAGTGCCCACCATAACGTGCTGGTAACTAAGGAC
AAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCC
ATGCAGCACCTGTCTCAATGTTCCCGAAGGCACCAATCTATCTCTAGAAAGTTCATTGGAT
GTCAAGGCCTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGT
GCGGGCCCCCGTCAATTCATTTGAGTTTTAACCTTGCGGCCGTACTCCCCAGGCGGTCAAC
TTAATGCGTTAGCTGCGCCACTAAAAGCTCAAGGCTTCCAACGGCTAGTTGACATCGTTTA
CGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGCACCTCAGTGTC
AGTATTAGTCCAGGTGGTCGCCTTCGCCACTGGTGTTCCTTCCTATATCTACGCATTTCACC
GCTACACAGGAAATTCCACCACCCTCTACCATACTCTAGTCAGTCAGTTTTGAATGCAGTT
CCCAGGTTGAGCCCGGGGATTTCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGC
CCAGTAATTCCGATTAACGCTTGCACCCTCTGTATTACCGCGGCTGCTGGCACAGAGTTAG
CCGGTGCTTATTCTGTCGGTAACGTCAAAATTGCAGAGTATTAATCTACAACCCTTCCTCC
CAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTCACACACGCGGCATGGCTGGATCAG
GCTTTCGCCCATTGTCCAATATTCCCCACTGCTGCCTCCCGTAGGAGTCTGGACCGTGTCTC
AGTTCCAGTGTGACTGATCATCCTCTCAGACCAGTTACGGATCGTCGCCTTGGTGAGCCAT
TACCCCACCAACTAGCTAATCCGACCTAGGCTCATCTGATAGCGCAAGGCCCGAAGGTCC
CCTGCTTTCTCCCGTAGGACGTATGCGGTATTAGCGTCCGTTTCCGAACGTTATCCCCCACT
ACCAGGCAGATTCCTAGGCATTACTCACCCGTCCGCCGCTCTCAAGAGAAGCAAGCTTCTC
TCTACCGCTCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCCATGATCAA
ACTCTTCAG

Pseudomonas sp. FDAARGOS_380

AAGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCA
TGAATCACACCGTGGTAACCGTCCTCCCGAAGGTTAGACTAGCTACTTCTGGTGCAACCCA
CTCCCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGACATTCT
GATTCGCGATTACTAGCGATTCCGACTTCACGCAGTCGAGTTGCAGACTGCGATCCGGACT
ACGATCGGTTTTATGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATT
GTAGCACGTGTGTAGCCCAGGCCGTAAGGGCCATGATGACTTGACGTCATCCCCACCTTCC
TCCGGTTTGTCACCGGCAGTCTCCTTAGAGTGCCCACCATTACGTGCTGGTAACTAAGGAC
AAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCC
ATGCAGCACCTGTCTCAATGTTCCCGAAGGCACCAATCTATCTCTAGAAAGTTCATTGGAT
GTCAAGGCCTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGT
GCGGGCCCCCGTCAATTCATTTGAGTTTTAACCTTGCGGCCGTACTCCCCAGGCGGTCAAC
TTAATGCGTTAGCTGCGCCACTAAAAGCTCAAGGCTTCCAACGGCTAGTTGACATCGTTTA
CGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGCACCTCAGTGTC
AGTATTAGTCCAGGTGGTCGCCTTCGCCACTGGTGTTCCTTCCTATATCTACGCATTTCACC
GCTACACAGGAAATTCCACCACCCTCTACCATACTCTAGTCAGTCAGTTTTGAATGCAGTT
CCCAGGTTGAGCCCGGGGATTTCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGC
CCAGTAATTCCGATTAACGCTTGCACCCTCTGTATTACCGCGGCTGCTGGCACAGAGTTAG
CCGGTGCTTATTCTGTCGGTAACGTCAAAATTGCAGAGTATTAATCTACAACCCTTCCTCC
CAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTCACACACGCGGCATGGCTGGATCAG
GCTTTCGCCCATTGTCCAATATTCCCCACTGCTGCCTCCCGTAGGAGTCTGGACCGTGTCTC

AGTTCCAGTGTGACTGATCATCCTCTCAGACCAGTTACGGATCGTCGCCTTGGTGAGCCAT
TACCTCACCAACTAGCTAATCCGACCTAGGCTCATCTGATAGCGCAAGGCCCGAAGGTCC
CCTGCTTTCTCCCGTAGGACGTATGCGGTATTAGCGTCCGTTTCCGAACGTTATCCCCCACT
ACCAGGCAGATTCCTAGGCATTACTCACCCGTCCGCCGCTCTCAAGAGAAGCAAGCTTCTC
TCTACCGCTCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCCATGATCAA
ACTCTTCAG

Pseudomonas fluorescens A506

AAGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCA
TGAATCACACCGTGGTAACCGTCCTCCCGAAGGTTAGACTAGCTACTTCTGGTGCAACCCA
CTCCCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGACATTCT
GATTCGCGATTACTAGCGATTCCGACTTCACGCAGTCGAGTTGCAGACTGCGATCCGGACT
ACGATCGGTTTTATGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATT
GTAGCACGTGTGTAGCCCAGGCCGTAAGGGCCATGATGACTTGACGTCATCCCCACCTTCC
TCCGGTTTGTCACCGGCAGTCTCCTTAGAGTGCCCACCATAACGTGCTGGTAACTAAGGAC
AAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCC
ATGCAGCACCTGTCTCAATGTTCCCGAAGGCACCAATCTATCTCTAGAAAGTTCATTGGAT
GTCAAGGCCTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGT
GCGGGCCCCCGTCAATTCATTTGAGTTTTAACCTTGCGGCCGTACTCCCCAGGCGGTCAAC
TTAATGCGTTAGCTGCGCCACTAAAAGCTCAAGGCTTCCAACGGCTAGTTGACATCGTTTA
CGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGCACCTCAGTGTC
AGTATTAGTCCAGGTGGTCGCCTTCGCCACTGGTGTTCCTTCCTATATCTACGCATTTCACC
GCTACACAGGAAATTCCACCACCCTCTACCATACTCTAGTCAGTCAGTTTTGAATGCAGTT
CCCAGGTTGAGCCCGGGGATTTCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGC
CCAGTAATTCCGATTAACGCTTGCACCCTCTGTATTACCGCGGCTGCTGGCACAGAGTTAG
CCGGTGCTTATTCTGTCGGTAACGTCAAAATTGCAGAGTATTAATCTACAACCCTTCCTCC
CAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTCACACACGCGGCATGGCTGGATCAG
GCTTTCGCCCATTGTCCAATATTCCCCACTGCTGCCTCCCGTAGGAGTCTGGACCGTGTCTC
AGTTCCAGTGTGACTGATCATCCTCTCAGACCAGTTACGGATCGTCGCCTTGGTGAGCCAT
TACCCCACCAACTAGCTAATCCGACCTAGGCTCATCTGATAGCGCAAGGCCCGAAGGTCC
CCTGCTTTCTCCCGTAGGACGTATGCGGTATTAGCGTCCGTTTCCGAACGTTATCCCCCACT
ACCAGGCAGATTCCTAGGCATTACTCACCCGTCCGCCGCTCTCAAGAGAAGCAAGCTTCTC
TCTACCGCTCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCCATGATCAA
ACTCTTCAG

Pseudomonas fluorescens strain Pt14

CTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTC
GAGCGGTAGAGAGAAGCTTGCTTCTCTTGAGAGCGGCGGACGGGTGAGTAATGCCTAGGA
ATCTGCCTGGTAGTGGGGGATAACGTTCGGAAACGGACGCTAATACCGCATACGTCCTAC
GGGAGAAAGCAGGGGACCTTCGGGCCTTGCGCTATCAGATGAGCCTAGGTCGGATTAGCT
AGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCGTAACTGGTCTGAGAGGATGATCA
GTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTG
GACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGATTGT
AAAGCACTTTAAGTTGGGAGGAAGGGTTGTAGATTAATACTCTGCAATTTTGACGTTACCG
ACAGAATAAGCACCGGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGGTGCAAGC
GTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTTGTTAAGTTGGATGTGAAA
TCCCCGGGCTCAACCTGGGAACTGCATTCAAAACTGACTGACTAGAGTATGGTAGAGGGT
GGTGGAATTTCCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAA
GGCGACCACCTGGACTAATACTGACACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGAT
TAGATACCCTGGTAGTCCACGCCGTAAACGATGTCAACTAGCCGTTGGAAGCCTTGAGCTT
TTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAA
CTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAA
CGCGAAGAACCTTACCAGGCCTTGACATCCAATGAACTTTCTAGAGATAGATTGGTGCCTT
CGGGAACATTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGT
TAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGTTACCAGCACGTTATGGTGGGCACTCT
AAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCC
CTTACGGCCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCCAAGCCGCGA

GGTGGAGCTAATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCG
TGAAGTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCC
TTGTACACACCGCCCGTCACACCATGGGAGTGGGTTGCACCAGAAGTAGCTAGTCTAACC
TTCGGGAGGACGGTTACCACGGTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAGC
CGTAGGGGAACCTGCGGCTGGATCACCTCCTT

Pseudomonas libanensis strain DMSP-1

CTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTC
GAGCGGTAGAGAGAAGCTTGCTTCTCTTGAGAGCGGCGGACGGGTGAGTAATGCCTAGGA
ATCTGCCTGGTAGTGGGGGATAACGTTCGGAAACGGACGCTAATACCGCATACGTCCTAC
GGGAGAAAGCAGGGGACCTTCGGGCCTTGCGCTATCAGATGAGCCTAGGTCGGATTAGCT
AGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCGTAACTGGTCTGAGAGGATGATCA
GTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTG
GACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGATTGT
AAAGCACTTTAAGTTGGGAGGAAGGGTTGTAGATTAATACTCTGCAATTTTGACGTTACCG
ACAGAATAAGCACCGGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGGTGCAAGC
GTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTTGTTAAGTTGGATGTGAAA
TCCCCGGGCTCAACCTGGGAACTGCATTCAAAACTGACTGACTAGAGTATGGTAGAGGGT
GGTGGAATTTCCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAA
GGCGACCACCTGGACTAATACTGACACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGAT
TAGATACCCTGGTAGTCCACGCCGTAAACGATGTCAACTAGCCGTTGGAAGCCTTGAGCTT
TTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAA
CTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAA
CGCGAAGAACCTTACCAGGCCTTGACATCCAATGAACTTTCTAGAGATAGATTGGTGCCTT
CGGGAACATTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGT
TAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGTTACCAGCACGTCATGGTGGGCACTCT
AAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCC
CTTACGGCCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCCAAGCCGCGA
GGTGGAGCTAATCCCACAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCG
TGAAGTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCC
TTGTACACACCGCCCGTCACACCATGGGAGTGGGTTGCACCAGAAGTAGCTAGTCTAACC
TTCGGGAGGACGGTTACCACGGTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAGC
CGTAGGGGAACCTGCGGCTGGATCACCTCCTT

Pseudomonas azotoformans strain S4

AAGGAGGTGATCCAGCCGCAGGTTCCCCTACGGCTACCTTGTTACGACTTCACCCCAGTCA
TGAATCACACCGTGGTAACCGTCCTCCCGAAGGTTAGACTAGCTACTTCTGGTGCAACCCA
CTCCCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGACATTCT
GATTCGCGATTACTAGCGATTCCGACTTCACGCAGTCGAGTTGCAGACTGCGATCCGGACT
ACGATCGGTTTTATGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATT
GTAGCACGTGTGTAGCCCAGGCCGTAAGGGCCATGATGACTTGACGTCATCCCCACCTTCC
TCCGGTTTGTCACCGGCAGTCTCCTTAGAGTGCCCACCATTACGTGCTGGTAACTAAGGAC
AAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCC
ATGCAGCACCTGTCTCAATGTTCCCGAAGGCACCAATCTATCTCTAGAAAGTTCATTGGAT
GTCAAGGCCTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGT
GCGGGCCCCCGTCAATTCATTTGAGTTTTAACCTTGCGGCCGTACTCCCCAGGCGGTCAAC
TTAATGCGTTAGCTGCGCCACTAAAAGCTCAAGGCTTCCAACGGCTAGTTGACATCGTTTA
CGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGCACCTCAGTGTC
AGTATTAGTCCAGGTGGTCGCCTTCGCCACTGGTGTTCCTTCCTATATCTACGCATTTCACC
GCTACACAGGAAATTCCACCACCCTCTACCATACTCTAGTCAGTCAGTTTTGAATGCAGTT
CCCAGGTTGAGCCCGGGGATTTCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGC
CCAGTAATTCCGATTAACGCTTGCACCCTCTGTATTACCGCGGCTGCTGGCACAGAGTTAG
CCGGTGCTTATTCTGTCGGTAACGTCAAAATTGCAGAGTATTAATCTACAACCCTTCCTCC
CAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTCACACACGCGGCATGGCTGGATCAG
GCTTTCGCCCATTGTCCAATATTCCCCACTGCTGCCTCCCGTAGGAGTCTGGACCGTGTCTC
AGTTCCAGTGTGACTGATCATCCTCTCAGACCAGTTACGGATCGTCGCCTTGGTGAGCCAT
TACCTCACCAACTAGCTAATCCGACCTAGGCTCATCTGATAGCGCAAGGCCCGAAGGTCC

CCTGCTTTCTCCCGTAGGACGTATGCGGTATTAGCGTCCGTTTCCGAACGTTATCCCCCACT
ACCAGGCAGATTCCTAGGCATTACTCACCCGTCCGCCGCTCTCAAGAGAAGCAAGCTTCTC
TCTACCGCTCGACTTGCATGTGTTAGGCCTGCCGCCAGCGTTCAATCTGAGCCATGATCAA
ACTCTTCAG

Biblography

Angov E, Hillier CJ, Kincaid RL & Lyon JA. Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. PloS one. 2008; 3: 2189.

Alagesan S, Hanko EKR, Malys N, Ehsaan M, Winzer K, Minton NP. Functional Genetic Elements for Controlling Gene Expression in *Cupriavidus necator* H16. Appl Environ Microbiol. 2018; 84(19):e00878-18.

Alt-Mörbe J, Stryker JL, Fuqua C, Li PL, Farrand SK, Winans SC. The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. *J Bacteriol*. 1996; 178(14):4248–4257.

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ. "Basic local alignment search tool." J. Mol. Biol. 1990; 215:403-410

An SS, Kim YM. Purification and characterization of a manganese-containing superoxide dismutase from a carboxydobacterium, *Pseudomonas carboxydohydrogena*. Mol Cells. 1997; 7(6):730–737

Arenas-López C, Locker J, Orol D *et al.* The genetic basis of 3-hydroxypropanoate metabolism in *Cupriavidus necator* H16. Biotechnol Biofuels. 2019; 12:150.

Atsumi S, Wu TJ, and Machado JM *et al.* Evo lution, genomic analysis and reconstruction of isobu tanol tolerance in *Escherichia coli*, Mol. Syst. Biol. 2010; 6: 449–453.

Avetisyan AV, Bogachev AV, Murtasina RA, Skulachev VP. Involvement of a *d*-type oxidase in the Na^+-motive respiratory chain of *Escherichia coli* growing under low $\Delta\mu_H^+$ conditions. FEBS Lett. 1992; 306:199–202.

Bandyopadhyay A, O'Brien S, Frank KL, Dunny GM, Hu W-S. Antagonistic Donor Density Effect Conserved in Multiple Enterococcal Conjugative Plasmids. Applied and Environmental Microbiology. 2016; 82 (15): 4537-4545

Bartosik D, Baj J, Wlodarczyk M. Molecular and functional analysis of pTAV320, a repABC-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. Microbiology. 1998; 144, 3149–3157.

Behjati S, Tarpey PS. What is next generation sequencing? *Arch Dis Child Educ Pract Ed*. 2013; 98(6):236-8.

Behrendt U, Ulrich A & Schumann P. Fluorescent pseudomonads associated with the phyllosphere of grasses; *Pseudomonas trivialis* sp. nov., *Pseudomonas poae* sp. nov. and *Pseudomonas congelans* sp. nov. Int. J. Syst. Evol. Microbiol. 2003; 53: 1461-9.

Belitsky BR, Wray LV Jr, Fisher SH, Bohannon DE, Sonenshein AL. Role of TnrA in nitrogen source-dependent repression of Bacillus subtilis glutamate synthase gene expression. *J Bacteriol*. 2000;182(21):5939–5947.

Berg IA, Kockelkorn D, Buckel W & Fuchs G. A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in archaea. Science. 2007; 318:1782–1786.

Bergamaschi P, Hein R, Heimann M & Crutzen PJ. Inverse modeling of the global CO cycle. J. Geophys. Res. 2000; 105: 1909–1927.

Bernard U, Probst I, and Schlegel HG. The cytochromes of some hydrogen bacteria. *Arch Microbiol.* 1974; 95: 29–37

Berney M & Cook GM. Unique flexibility in energy metabolism allows mycobacteria to combat starvation and hypoxia. PLoS One 2010; 5: e8614.

Bertsova YV, Bogachev AV, Skulachev VP. Generation of protonic potential by the bd-type quinol oxidase of *Azotobacter vinelandii*. FEBS Lett. 1997;414:369–372

Bi C, Su P, Müller J, et al. Development of a broad- host synthetic biology toolbox for Ralstonia eutropha and its application to engineering hydrocarbon bio- fuel production. Microb Cell Fact. 2013;12:107.

Blanca-Ordóñez H, Oliva-García JJ, Pérez-Mendoza D, et al. pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. J Bacteriol. 2010;192(23):6309–6312.

Borisov VB. Interaction of bd-type quinol oxidase from Escherichia coli and carbon monoxide: heme d binds CO with high affinity. Biochemistry-Moscow. 2008;73(1):14–22.

Borisov VB, Forte E, Davletshin A, Mastronicola AD, Sarti P, Giuffrè A. Cytochrome *bd* oxidase from *Escherichia coli* displays high catalase activity: an additional defence against oxidative stress. FEBS Lett. 2013; 587: 2214-2218.

Borisov VB, Forte E, Giuffrè A, Konstantinov A, Sarti P. Reaction of nitric oxide with the oxidized di-heme and heme-copper oxygen-reducing centres of terminal oxidases: Different reaction pathways and end-products. J Inorg Biochem. 2009;103:1185–1187

Borisov VB, Forte E, Konstantinov AA, Poole RK, Sarti P, Giuffre A. Interaction of the bacterial terminal oxidase cytochrome *bd* with nitric oxide. FEBS Lett. 2004;576:201–204.

Borisov VB, Gennis RB, Hemp J, Verkhovsky MI. The cytochrome bd respiratory oxygen reductases. Biochim Biophys Acta. 2011;1807(11):1398–413.

Bowien B, Kusian B. Genetics and control of CO_2 assimilation in the chemoautotroph Ralstonia eutropha. Arch. Microbiol. 2002; 178; 85-93

Bresan S, Jendrossek D. New Insights into PhaM-PhaC-Mediated Localization of Polyhydroxybutyrate Granules in *Ralstonia eutropha* H16. Appl Environ Microbiol. 2017;83(12):e00505-17.

Brigham C. Perspectives for the biotechnological production of biofuels from CO_2 and H_2 using *Ralstonia eutropha* and other 'Knallgas' bacteria. Appl Microbiol Biotechnol. 2019;103(5):2113–2120

Brigham CJ, Budde CF, Holder JW, Zeng Q, Mahan AE, Rha C, *et al*. Elucidation of β-oxidation pathways in *Ralstonia eutropha* H16 by examination of global gene expression. J Bacteriol. 2010; 192(20):5454–64.

Brigham CJ, Zhila N, Shishatskaya E, Volova TG, Sinskey AJ. Manipulation of Ralstonia eutropha carbon storage pathways to produce useful bio-based products. Subcell Biochem. 2012;64:343–366.

Butler CD. Climate Change, Health and Existential Risks to Civilization: A Comprehensive Review (1989-2013). Int J Environ Res Public Health. 2018; 15(10):2266.

Byrom D. Polymer synthesis by microorganisms: technology and economics. Trends Biotechnol. 1987;5:246–250.

Calhoun MW, Oden KL, Gennis RB, Teixeira-de-Matto MJs & Neijssel OM. Energetic efficiency of *Escherichia coli*: effects of mutations in components of the aerobic respiratory chain. J. Bacteriol. 1993; 175**:**3020-3025.

Cao H, Wei D, Yang Y, *et al.* Systems-level understanding of ethanol-induced stresses and adaptation in E. coli. *Sci Rep*. 2017;7:44150.

Castillo-Ramirez S, Vazquez-Castellanos JF, Gonzalez V, Cevallos MA. Horizontal gene transfer and diverse functional constraints within a common replication-

partitioning system in Alphaproteobacteria: the repABC operon. BMC Genomics. 2009; 10:536.

Cervantes-Rivera R, Pedraza-Lopez F, Perez-Segura G, Cevallos MA. The replication origin of a repABC plasmid. BMC Microbiol. 2011; 11:158.

Cevallos MA, Cervantes-Rivera R, Gutiérrez-Ríos RM. The repABC plasmid family. Plasmid. 2008; 60:19–37.

Chai Y, Winans SC. A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. Mol Microbiol. 2005; 56:1574–1585.

Chai Y, Zhu J, Winans SC. 2001. TrlR, a defective TraR-like protein of *Agrobacterium tumefaciens*, blocks TraR function in vitro by forming inactive TrlR:TraR dimers. Mol. Microbiol. 2001; 40:414–421.

Chen GQ. A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. Chem Soc Rev. 2009;38:2434–2446.

Churchill ME, Chen L. Structural basis of acyl-homoserine lactone-dependent signaling. *Chem Rev*. 2011;111(1):68–85.

Clomburg JM, Crumbley AM and Gonzalez R, Industrial biomanufacturing: The future of chemical production. Science 2017; 355(6320): 0804

Clough SJ, Lee K-M, Schell MA, Denny TP. A two-component system in *Ralstonia (Pseudomonas) solanacearum* modulates production of PhcA-regulated virulence factors in response to 3-hydroxypalmitic acid methyl ester. *J Bacteriol.* 1997; 179: 3639–3648.

Commichau FM, Herzberg C, Tripal P, Valerius O & Stülke, J. A regulatory protein–protein interaction governs glutamate biosynthesis in *Bacillus subtilis*: the glutamate

dehydrogenase RocG moonlights in controlling the transcription factor GltC. Molecular Microbiology. 2007; 65: 642-654.

Cook DM, Li P-L, Ruchaud F, Padden S, Farrand SK. Ti plasmid conjugation is independent of vir: reconstitution of the tra functions from pTiC58 as a binary system. J. Bacteriol. 1997; 179:1291–1297

Cook AM & Schlegel HG. Metabolite concentrations in *Alcaligenes eutrophus* H16 and a mutant defective in poly-(β-hydroxybutyrate synthesis. Arch. Microbiol. 1978; 119: 231–235.

Cordero PRF, Bayly K, Leung PM, Huang C, Islam ZF, Schittenhelm RB, King GM, Greening C. Carbon monoxide dehydrogenases enhance bacterial survival by oxidising atmospheric CO. ISME J. 2019.

Correddu D, Montaño López JDJ, Angermayr SA, Middleditch MJ, Payne LS and Leung IK. Effect of consecutive rare codons on the recombinant production of human proteins in Escherichia coli. *IUBMB life*. 2019; 1– 9.

Cotter PA, Chepuri V, Gennis RB, Gunsalus RP. Cytochrome *o* (*cyoABCDE*) and *d* (*cydAB*) oxidase gene expression in *Escherichia coli* is regulated by oxygen, pH, and the *fnr* gene product. J Bacteriol. 1990; 172:6333–6338.

Cramm R. Genomic view of energy metabolism in Ralstonia eutropha H16. J Mol Microbiol Biotechnol. 2009; 16: 38–52.

Crutzen PJ & Gidel LT. A two-dimensional photochemical model of the atmosphere: the tropospheric budgets of the anthropogenic chlorocarbons, CO, CH4, CH3Cl and the effect of various NOx sources on tropospheric ozone. J. Geophys. Res. 1983; 88: 6641–6661.

Cunliffe M. Physiological and metabolic effects of carbon monoxide oxidation in the model marine bacterioplankton *Ruegeria pomeroyi* DSS-3. Appl Environ Microbiol 2013; 79: 738–740.

Cypionka H & Meyer O. Influence of carbon monoxide on growth and respiration of carboxydotrophic and other aerobic organisms. FEMS Microbiol Lett. 1982; 15:209–214.

Cypionka H & Meyer O. Carbon monoxide-insensitive respiratory chain of *Pseudomonas carboxydovorans.* J Bacteriol. 1983;156(3):1178–1187.

Cypionka H, van Verseveld HW & Stouthamer AH. Proton translocation coupled to carbon monoxide- insensitive and sensitive electron transport in *Pseudomonas carboxydovorans.* FEMS Microbiol. Lett. 1984; 22, 209–213.

Dabboussi F, Hamze M, Elomari M, Verhille S, Baida N, Izard D & Leclerc H. Pseudomonas libanensis sp. nov., a new species isolated from Lebanese spring waters. Int. J. Syst. Bacteriol. 1999; 49: 1091-101.

Danino VE, Wilkinson A, Edwards A, Downie JA. Recipient-induced transfer of the symbiotic plasmid pRL1JI in *Rhizobium leguminosarum bv. viciae* is regulated by a quorum-sensing relay. Mol. Microbiol. 2003; 50: 511–525.

Debabov VG. Modern approaches to the creation of industrial microorganism strains. Russ. J. Genet. 2015; 51: 365–376.

Dernedde J, Eitinger T, Patenge N, Friedrich B. *hyp* gene products in *Alcaligenes eutrophus* are part of a hydrogenase-maturation system. Eur. J. Biochem., 1996; 235 , pp. 351-358

Diaz Ricci JC, Hernández ME. Plasmid effects on *Escherichia coli* metabolism. Crit. Rev. Biotechnol. 2000;20:79–108.

diCenzo GC, Finan TM. The divided bacterial genome: structure, function, and evolution. Microbiol Mol Biol Rev . 2017; 81: 19–17.

Diels L, Faelen M, Mergeay M, Nies D. Mercury transposons from plasmids governing multiple resistance to heavy metals in *Alcaligenes eutrophus* CH34. Arch Intern Physiol Bioch. 1985;93:27–28

Ditta G, Schmidhauser T, Yakobson E, et al. Plasmids related to the broad host range vector, pRK290, use- ful for gene cloning and for monitoring gene expres- sion. Plasmid. 1985;13:149–153.

Dobbek H, Gremer L, Meyer O, Huber R. Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcysteine. Proc Natl Acad Sci U S A. 1999;96(16):8884–8889.

Dobbek H, Gremer L, Kiefersauer R, Huber R, Meyer O. Catalysis at a dinuclear [CuSMo(==O)OH] cluster in a CO dehydrogenase resolved at 1.1-A resolution. Proc Natl Acad Sci U S A. 2002; 99(25):15971–15976.

Döhlemann J, Wagner M, Happel C, Carrillo M, Sobetzko P, Erb TJ *et al.* A Family of Single Copy repABC -Type Shuttle Vectors Stably Maintained in the Alpha-Proteobacterium *Sinorhizobium meliloti*. ACS Synth Biol. American Chemical Society; 2017; 6: 968–984.

Dormeyer M, Lübke AL, Müller P, Lentes S, Reuß DR, Thürmer A, Stülke J, Daniel R, Brantl S, Commichau FM. Hierarchical mutational events compensate for glutamate auxotrophy of a *Bacillus subtilis* gltC mutant. Environ Microbiol Rep. 2017; 9:279–289.

Dougherty K, Smith BA, Moore AF *et al.* Multiple phenotypic changes associated with large-scale horizontal gene transfer. PLoS One. 2014; 9(7):e102170.

Drzyzga O, Revelles O, Durante‐Rodríguez G, Díaz E, García JL & Prieto A. New challenges for syngas fermentation: towards production of biopolymers. J Chem Technol Biotechnol. 2015; 90: 1735‐1751.

Dürre P & Eikmanns BJ. C1-carbon sources for chemical and fuel production by microbial gas fermentation. Curr Opin Biotechnol. 2015; 35:63–72.

Environmental Protection Agency (EPA), GHGRP (Greenhouse Gas Reporting Program) Overview Report. [Online]. Available:https://www.epa.gov/sites/production/files/2017-09/documents/ry16_ghgrp_overview_report.pdf [21 April 2018]. [Accessed: 21 June 2018]

Esquivel-Elizondo S, Delgado AG, Krajmalnik-Brown R. Evolution of microbial communities growing with carbon monoxide, hydrogen, and carbon dioxide, FEMS Microbiology Ecology. 2017; 93:1-12

Ewering C, Heuser F, Benölken JK *et al.* Metabolic engineering of strains of *Ralstonia eutropha* and Pseudomonas putida for biotechnological production of 2-methylcitric acid. Metab Eng. 2006; 8:587–602.

Fei Q, Guarnieri MT, Tao L, Laurens LML, Dowe N and Pienkos. PT, Bioconversion of natural gas to liquid fuel: Opportunities and challenges. Biotechnol Adv. 2014; 32(3):596–614.

Fell SC, Carrivick JL, Brown LE. The Multitrophic Effects of Climate Change and Glacier Retreat in Mountain Rivers. Bioscience. 2017;67(10):897–911.

Finan TM, Weidner S, Wong K, Buhrmester J, Chain P, Vorholter FJ, Hernandez-Lucas I, Becker A, Cowie A, Gouzy J, Golding B, Puhler A. The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont *Sinorhizobium meliloti.* Proc Natl Acad Sci USA. 2001; 98:9889–9894.

Flavier AB, Clough SJ, Schell MA, Denny TP. Identification of 3-OH-hydroxypalmitic acid methyl ester as a novel autoregulator controlling virulence in *Ralstonia solanacearum*. *Mol Microbiol.* 1997; **26**: 251–259.

Frederix M, Edwards A, McAnulla C and Downie JA. Co-ordination of quorum-sensing regulation in *Rhizobium leguminosarum* by induction of an anti-repressor. Molecular Microbiology. 2011; 81: 994-1007.

Fricke WF, Kusian B, Bowien B. The genome organization of *Ralstonia eutropha* strain H16 and related species of the Burkholderiaceae. J Mol Microbiol Biotechnol. 2009; 16:124–135

Friedrich CG, Friedrich B, Bowien B. Formation of enzymes of autotrophic metabolism during hetero- trophic growth of *Alcaligenes eutrophus*. Microbiology. 1981;122:69–78.

Friedrich B & Schwartz E. Molecular biology of hydrogen utilization in aerobic chemolithotrophs. Annu Rev Microbiol. 1993; 47: 351–383.

Frommeyer M, Wiefel L, Steinbüchel A. Features of the biotechnologically relevant polyamide family “cyanophycins” and their biosynthesis in prokaryotes and eukaryotes. Crit Rev Biotechnol. 2016;36: 153–164.

Fuhrmann S *et al.* Complete nucleotide sequence of the circular megaplasmid pHCG3 of *Oligotropha carboxidovorans*: function in the chemolithoautotrophic utilization of CO, H2 and CO_2. Gene. 2003; 322:67–75.

Fuqua WC & Winans SC. A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. J. Bacteriol. 1994; 176:2796–2806.

Fuqua C & Winans SC. Conserved cis-acting promoter elements are required for density-dependent transcription of *Agrobacterium tumefaciens* conjugal transfer genes. J. Bacteriol. 1996; 178:435–440.

Fuqua C, Burbea M, Winans S. Activity of the *Agrobacterium* Ti plasmid conjugal transfer regulator TraR is inhibited by the product of the traM gene. J. Bacteriol. 1995; 177:1367–1373.

Füzik T, Ulbrich P, and Ruml T. Efficient mutagenesis independent of legation (EMILI), J. Microbiol. Methods, 2014; 106: 67–71

Gay P, Le Coq D, Steinmetz M, Berkelman T, Kado CI. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J Bacteriol. 1985;164(2):918–921.

Garcia SR, Karhumac KK, Fonseca C *et al.* Improved xylose and arabinose utilization: on by an industrial recombinant S. cerevisiae strain using evolutionary engineering, Biotechnol. Biofuel, 2010; 3: 15.

Gardan L, Bella P, Meyer JM, Christen R, Rott P, Achouak W & Samson R. *Pseudomonas salomonii* sp. nov., pathogenic on garlic, and *Pseudomonas palleroniana* sp. nov., isolated from rice. Int. J. Syst. Evol. Microbiol. 2002; 52, 2065-74.

Garg RP, Yindeeyoungyeon W, Gilis A, Denny TP, Van Der Lelie D & Schell MA. Evidence that *Ralstonia eutropha (Alcaligenes eutrophus)* contains a functional homologue of the *Ralstonia solanacearum* Phc cell density sensing system. Molecular Microbiology. 2000; 38: 359-367.

Gee DL & Brown WD. The effect of carbon monoxide on bacterial growth. Meat Sci. 1981; 5:215–222.

Giuffrè Alessandro, Borisov Vitaliy B., Arese Marzia, Sarti Paolo, Forte Elena. Cytochrome bd oxidase and bacterial tolerance to oxidative and nitrosative stress. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2014;1837(7):1178–1187.

Giusti MdlÁ *et al.* Genetic and functional characterization of a yet-unclassified rhizobial Dtr (DNA-transfer-and-replication) region from a ubiquitous plasmid conjugal system present in *Sinorhizobium meliloti*, in *Sinorhizobium medicae*, and in other nonrhizobial Gram-negative bacteria. Plasmid. 2012; 67:199–210.

Gnida M, Ferner R, Gremer L., Meyer O & Meyer-Klaucke W. A novel binuclear [CuSMo] cluster at the active site of carbon monoxide dehydrogenase: characterization by X-ray absorption spectroscopy. Biochemistry. 2003; 42: 222-230.

Goodarzi H, Bennett BD, Amini S, Reaves ML, Hottes AK, Rabinowitz JD, Tavazoie S. Regulatory and metabolic rewiring during laboratory evolution of ethanol tolerance in *E. coli*. Mol Syst Biol. 2010, 6:378.

Goodner B, Hinkle G, Gattung S, Miller N, Blanchard M, Qurollo B *et al*. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science. 2001; 294: 2323–2328.

González V *et al.* The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc Natl Acad Sci U S A. 2006; 103:3834–3839.

Grant SG, Jessee J, Bloom FR, Hanahan D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc Natl Acad Sci USA. 1990; 87(12):4645-9.

Greening C, Constant P, Hards K, Morales SE, Oakeshott JG, Russell RJ, et al. Atmospheric hydrogen scavenging: from enzymes to ecosystems. Appl Environ Microbiol. 2015; 81: 1190–1199.

Grethlein AJ, Worden RM, Jain MK *et al*. Evidence for production of n-butanol from carbon monoxide by *Butyribacterium methylotrophicum*. J Ferment Bioeng 1991;72:58–60.

Gruber S, Hagen J, Schwab H, et al. Versatile and stable vectors for efficient gene expression in *Ralstonia eutropha* H16. J Biotechnol. 2014;186:74–82.

Gruber S, Schwab H, Koefinger P. Versatile plasmid-based expression systems for Gram-negative bacteria – general essentials exemplified with the bacterium *Ralstonia eutropha* H16. N Biotechnol. 2015;32: 552–558.

Ha SM, Kim CK, Roh J, Byun JH, Yang SJ, Choi SB, Chun J & Yong D. Application of the whole genome-based bacterial identfication system, TrueBac ID, in clinical isolates which were not identified with three MALDI-TOF M/S systems. Ann Lab Med. 2019; 39(6):530-536

Halvorsen BM, Adhikari U, Eikeland M.S. Gasification of Biomass for Production of Syngas for Biofuel. Proceedings of the 56th SIMS. 2015; 255-260.

Handrick R, Reinhardt S, Jendrossek D. Mobilization of poly(3-hydroxybutyrate) in *Ralstonia eutropha*. J Bacteriol. 2000;182(20):5916–5918.

Hanko EKR, Minton NP, Malys N. (2017) Characterisation of a 3-hydroxypropionic acid-inducible system from *Pseudomonas putida* for orthogonal gene expression control in Escherichia coli and *Cupriavidus necator*. Sci. Rep. 7(1), 1724.10.1038/s41598-017-01850-w.

Hanko EKR, Minton NP, Malys N. A Transcription Factor-Based Biosensor for Detection of Itaconic Acid. ACS Synth Biol. 2018;7(5):1436–1446.

Hanson RS & Hanson TE. Methanotrophic bacteria. Microbiol Rev. 1996; 60(2):439–471.

Hanzelka BL & Greenberg EP. Quorum sensing in *Vibrio fischeri*: Evidence that S-adenosylmethionine is the amino acid substrate for autoinducer synthesis. J. Bacteriol. 1996; 178:5291–5294.

He X, Chang W, Pierce DL, Seib LO, Wagner J, Fuqua C. Quorum sensing in *Rhizobium sp*. strain NGR234 regulates conjugal transfer (tra) gene expression and influences growth rate. J. Bacteriol. 2003;185:809–822

Heinrich D, Raberg M, Steinbüchel A. Studies on the aerobic utilization of synthesis gas (syngas) by wild type and recombinant strains of *Ralstonia eutropha* H16. Microb Biotechnol. 2018;11(4):647–656.

Hill S, Viollet S, Smith AT, Anthony C. Roles for enteric *d*-type cytochrome oxidase in N_2 fixation and microaerobiosis. J Bacteriol. 1990;172:2071–2078.

Hille R, Dingwall S & Wilcoxen The aerobic CO dehydrogenase from *Oligotropha carboxidovorans*. JJ Biol Inorg Chem. 2015; 20: 243.

Hoegh-Guldberg O, Poloczanska ES, Skirving W and Dove S. Coral Reef Ecosystems under Climate Change and Ocean Acidification. Front. Mar. Sci. 2017; 4:158.

Hooykaas PJJ, Van Brussel, AAN, Dendulk-Ras H, Van slogteren, GMS and Schilperoort RA. Sym plasmid of Rhizobium trifolii expressed in different rhizobial species and *Agrobacterium tumefaciens*. Nature. 1981; 291: 351-353.

Hooykaas PJ, Snijdewint FG, Schilperoort RA. Identification of the Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its transfer to and expression in other rhizobia and Agrobacterium tumefaciens. Plasmid. 1982;8(1):73–82

Horinouchi T, Tamaoka K, Furusawa C, Ono N, Suzuki S, Hirasawa T, Yomo T, Shimizu H. Transcriptome analysis of parallel-evolved *Escherichia coli* strains under ethanol stress. BMC Genomics. 2010; 11:579.

Hess WT, Kurtz A and Stanton D. Methane, ethane, and propane, in Kirk-Othmer Encyclopedia of Chemical Technology. 1993; 13(4): 398–400.

Hücker SM, Ardern Z, Goldberg T, et al. Discovery of numerous novel small genes in the intergenic regions of the *Escherichia coli* O157:H7 Sakai genome. PLoS One. 2017;12(9):e0184119.

Hwang I, Li PL, Zhang L, Piper KR, Cook DM, Tate ME, Farrand SK. TraI, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid N-acylhomoserine lactone autoinducer. Proc. Natl. Acad. Sci. U. S. A. 1994; 91:4639–4643.

Hwang I, Cook D, Farrand S. A new regulatory element modulates homoserine lactone-mediated autoinduction of Ti plasmid conjugal transfer. J. Bacteriol. 1995; 177:449–458

Hwang I, Smyth AJ, Luo Z-Q, Farrand SK. Modulating quorum sensing by antiactivation: TraM interacts with TraR to inhibit activation of Ti plasmid conjugal transfer genes. Mol. Microbiol. 1999; 34:282–294.

Hyeon JE, Kim SW, Park C, and Han SO. Efficient biological conversion of carbon monoxide (CO) to carbon dioxide (CO_2) and for utilization in bioplastic production by *Ralstonia eutropha* through the display of an enzyme complex on the cell surface. Chem Commun. 2015; 51: 10202–10205

Iizuka H & Komagata K. New species of *Pseudomonas* belonged to fluorescent group (Studies on the microorganisms of cereal grains. Part V). J. Agric. Chem. Soc. Jpn. 1963; 37, 137-141.

Intergovernmental Panel on Climate Change (IPCC), Carbon Monoxide (CO) and Hydrogen (H2). [Online]. Available: https:// www.ipcc.ch/ipccreports/tar/wg1/139.htm [12 April 2018].[Accessed 19 June 2019]

Ingledew WJ, Poole RK. The respiratory chains of *Escherichia coli* Microbiol. Rev. 1984; 48: 222-271

Ishizaki A, Tanaka K, Taga N. Microbial production of poly-d-3-hydroxybutyrate from CO_2. Appl. Microbiol. Biotechnol. 2001; **57**:6–12.

Islam ZF, Cordero PRF, Feng J, Chen Y-J, Bay S, Gleadow RM, et al. Two *Chloroflexi* classes independently evolved the ability to persist on atmospheric hydrogen and carbon monoxide. *ISME J.* 2019; in press.

Iwata S. Structure and function of bacterial cytochrome c oxidase. J Biochem. 1998; 123:369–375.

Jajesniak P, Ali HEMO and Wong TS. Carbon dioxide capture and utilization using biological systems: Opportunities and challenges. J Bioprocess Biotech. 2014; 4(3):1–15.

Jensen TO, Kvist T, Mikkelsen MJ, Westermann P. Production of 1,3-PDO and butanol by a mutant strain of Clostridium pasteurianum with increased tolerance towards crude glycerol. AMB Express. 2012; 2:44.

Jesse HE, Nye TL, McLean S, Green J, Mann BE, Poole RK. Cytochrome bd-I in Escherichia coli is less sensitive than cytochromes *bd*-II or *bo*" to inhibition by the carbon monoxide-releasing molecule, CORM-3: *N*-acetylcysteine reduces CO-RM uptake and inhibition of respiration. Biochim Biophys Acta. 2013; 1834(9):1693–1703.

Ji M, Greening C, Vanwonterghem I, Carere CR, Bay SK, Steen JA, *et al.* Atmospheric trace gases support primary production in Antarctic desert surface soil. Nature 2017; 552: 400–403.

Jiang, Y, Chen B, Duan C, Sun B, Yang J & Yang S. Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. Appl. Environ. Microbiol. 2015; 81: 2506–2514.

Johnson BF & Stanier RY. Dissimilation of aromatic compounds by *Alcaligenes eutrophus*. J Bacteriol. 1971; 107(2):468–75.

Juengert JR, Patterson C, Jendrossek D. Poly(3-Hydroxybutyrate) (PHB) Polymerase PhaC1 and PHB Depolymerase PhaZa1 of Ralstonia eutropha Are Phosphorylated In Vivo. Appl Environ Microbiol. 2018;84(13):e00604-18.

Juhnke S, Peitzsch N, Hübener N, Große C, Nies DH. New genes involved in chromate resistance in *Ralstonia metallidurans* strain CH34. Archives of Microbiology. 2002; 179(1):15–25.

Jünemann S. Cytochrome *bd* terminal oxidase. Biochim. Biophys. Acta. 1997;1321:107–127

Khalil MAK & Rasmussen RA. Carbon monoxide in the Earth's atmosphere: indications of a global increase. Nature. 1988; 332: 242–245.

Kabus A, Niebisch A & Bott M. Role of cytochrome *bd* oxidase from *Corynebacterium glutamicum* in growth and lysine production. *Appl Environ Microbiol*. 2007;73(3):861–868. doi:10.1128/AEM.01818-06

Karl TR & Trenberth KE. Modern Global Climate Change. Science. 2003; 302: 1719–1723.

Kaufmann P, Duffus B, Teutloff C & Leimkühler S. Functional Studies on *Oligotropha carboxidovorans* Molybdenum–Copper CO Dehydrogenase Produced in Escherichia coli. Biochemistry, 2018; 57(19):2889-2901.

Kleihues L, Lenz O, Brnhard M, Buhrke T & Friedrich B. The H_2 sensor of *Ralstonia eutropha* is a member of the subclass of regulatory (NiFe) hydrogenases J. Bacteriol. 2000: 182: 2716-2724

Keeling PJ. Functional and ecological impacts of horizontal gene transfer in eukaryotes. Curr Opin Genet Dev. 2009; 19: 613–619.

Keller MW, Schut GJ, Lipscomb GL, Menon AL, Iwuchukwu IJ, Leuko TT *et al.* Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. Proc Natl Acad Sci. 2013; 110(15):5840–5845.

Kelly MJS, Poole RK, Yates MG & Kennedy C. Cloning and mutagenesis of genes encoding the cytochrome *bd* terminal oxidase complex in *Azotobacter vinelandii*: Mutants deficient in the cytochrome *d* complex are unable to fix nitrogen in air. J Bacteriol. 1990;172:6010–6019.

Khalil MAK, Pinto JP & Shearer MJ. Atmospheric carbon monoxide. Chemosph. Glob. Change Sci. 1999; 1: 9-11.

Kirk RE. Carbon monoxide. Kirk–Othmer Encyclopedia of Chemical Technology 1998; 5(4):48–61.

King GM. Attributes of atmospheric carbon monoxide oxidation by Maine forest soils. Appl. Environ. Microbiol. 1999; 65:5257-5264

King GM. Molecular and culture based analyses of aerobic carbon monoxide oxidizer diversity. Appl. Environ. Microbiol. 2003; 69, 7257–7265.

King GM, Weber CF. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. Nat Rev Microbiol 2007; 5: 107–118.

King GM. Carbon monoxide as a metabolic energy source for extremely halophilic microbes: implications for microbial activity in Mars regolith. Proc Natl Acad Sci. 2015; 4465–4470.

King GM, Weber CF. Interactions between bacterial carbon monoxide and hydrogen consumption and plant development on recent volcanic deposits. ISME J 2008; 2: 195–203

King GM, Weber CF, Nanba K, Sato Y, Ohta H. Atmospheric CO and hydrogen uptake and CO oxidizer phylogeny for Miyake-jima, Japan Volcanic Deposits. Microbes Environ 2008; 23: 299–305

Klenk H-P *et al*. The complete genome sequence of the 41. hyperthermophilic sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature. 1997; 390: 364–370.

Kolonay JF & Jr Maier RJ. Formation of pH and potential gradients by the reconstituted *Azotobacter vinelandii* cytochrome bd respiratory protection oxidase. J Bacteriol. 1997;179:3813–3817

Kömen R, Zannoni D, Ingledew WJ & Schmidt K. The electron transport system of *Alcaligenes eutrophus* . Arch Microbiol. 1991; 155: 382–390.

Koraimann G, Wagner MA. Social behavior and decision making in bacterial conjugation. *Front Cell Infect Microbiol*. 2014;4:54.

Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM II, Peterson KM. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene. 1995; 166 (1): 175-176.

Kreß O, Gnida M, Pelzmann AM, Marx C, Meyer-Klaucke W, Meyer O. Reversible inactivation of CO dehydrogenase with thiol compounds. Biochem. Biophys. Res. Commun. 2014; 447: 413–418.

Krausse D, Hunold K, Kusian B, Lenz O, Stulke J, Bowien B, Deutscher J. Essential role of the hprk gene in ralstonia eutropha h16. J Mol Microbiol Biotechnol. 2009;17:146–152.

Krieg T, Sydow A, Faust S, Huth I, Holtmann D. CO_2 to terpenes: autotrophic and electroautotrophic alpha-humulene production with *Cupriavidus necator*. Angew Chem Int Ed Engl. 2018; 57:1879–1882

Kudla G, Murray AW, Tollervey D & Plotkin JB. Coding-sequence determinants of gene expression in Escherichia coli. *Science, 324*(5924), pp.255-258.

Kuyper M, Toirkens MJ, Diderich JA, Winkler AA, van Dijken JP, Pronk JT, Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. FEMS Yeast Research. 2005; 5: 925–934.

LaCroix RA, Sandberg TE, O'Brien EJ, et al. Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of *Escherichia coli* K-12 MG1655 on glucose minimal medium. Appl Environ Microbiol. 2015;81(1):17–30.

Landt O, Grunert HP & Hahu U. A general method for rapid sitedirected mutagenesis using the polymerase chain reaction, Gene, 1990; 96: 125–128.

Lazdunski AM, Ventre I, Sturgis JN. Regulatory circuits and communication in Gram-negative bacteria. *Nat. Rev. Microbiol.* 2004; 2: 581–592.

Lee SH, Kim MS, Lee JH *et al.* Adaptive engineering of a hyperthermophilic archaeon on CO and discovering the underlying mechanism by multi-omics analysis. Sci Rep. 2016;6:22896.

Lenski RE. What is adaptation by natural selection? Perspectives of an experimental microbiologist. PLoS Genet. 2017;13(4):e1006668.

Lenz O & Friedrich B. A novel multicomponent regulatory system mediates H2 sensing in Alcaligenes eutrophus. Proc Natl Acad Sci U S A. 1998;95(21):12474–12479.

Letunic I & Bork P. Interactive tree of life (itol): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2007;23:127–128.

Letunic Ivica, Bork Peer. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Research. 2019;47:256–259.

Levine M & Anderson DQ. Two New Species of Bacteria Causing Mustiness in Eggs. J. Bacteriol. 1932; 23, 337-47.

Li P-L, Everhart DM, Farrand SK. Genetic and sequence analysis of the trb locus on pTiC58, a mating-pair formation system related to members of the type IV secretion family. J. Bacteriol. 1998; 180: 6164–6172.

Li X, Griffin D, Li X & Henson MA. Incorporating hydrodynamics into spatiotemporal metabolic models of bubble column gas fermentation. Biotechnology and Bioengineering. 2018; 116: 28– 40.

Li Z & Nair SK. Quorum sensing: how bacteria can coordinate activity and synchronize their response to external signals?. Protein Sci. 2012;21(10):1403–1417.

Li H, Opgenorth PH, Wernick DG, Rogers S, Wu TY, Higashide W, Malati P, Huo YX, Cho KM & Liao JC. Integrated electromicrobial conversion of CO_2 to higher alcohols. Science. 2012; 335: 1596.

Lick S, Kröckel L, Wibberg D, Winkler A, Blom J, Bantleon A, Goesmann A & Kalinowski J. *Pseudomonas carnis* sp. nov., isolated from meat. Int. J. Syst. Evol. Microbiol. 2020; DOI 10.1099/ijsem.0.0039282020

Liu K, Atiyeh HK, Tanner RS, Wilkins MR & Huhnke RL. Fermentative production of ethanol from syngas using novel moderately alkaliphilic strains of *Alkalibaculum bacchi*. *Bioresour. Technol.* 2012; 104: 336–341

Long CP, Gonzalez JE, Feist AM, Palsson BO & Antoniewicz MR. Fast growth phenotype of *E. coli* K-12 from adaptive laboratory evolution does not require intracellular flux rewiring. Metab Eng. 2017; 44:100–7.

Loper JE, Hassan KA, Mavrodi DV, et al. Comparative genomics of plant-associated Pseudomonas spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS Genet. 2012; 8(7):e1002784. doi:10.1371/journal.pgen.1002784

Lorite MJ, Tachil J, Sanjuan J, Meyer O & Bedmar EJ. Carbon monoxide dehydrogenase activity in *Bradyrhizobium japonicum.* Appl. Environ. Microbiol. 2000; 68: 1871–1876.

Lu Y & Khalil MAK. Methane and carbon monoxide in OH chemistry: the effects of feedbacks and reservoirs generated by the reactive products. Chemosphere. 1993; 26: 641–655 (1993).

Lu Y & Yu J. Gas mass transfer with microbial CO_2 fixation and poly(3‐hydroxybutyrate) synthesis in a packed bed bioreactor. Biochem Eng J. 2017; 122: 13‐21.

Luo Z-Q, Qin Y & Farrand SK. The antiactivator TraM interferes with the autoinducer-dependent binding of TraR to DNA by interacting with the C-terminal region of the quorum-sensing activator. J. Biol. Chem. 2000; 275:7713–7722.

Luo G, Wang W & Angelidaki I. Anaerobic digestion for simultaneous sewage sludge treatment and CO biomethanation: process performance and microbial ecology. Environ Sci Technol 2013;47:10685–93.

Lynd L, Kerby R & Zeikus JG. Carbon monoxide metabolism of the methylotrophic acidogen *Butyribacterium methylotrophicum*. J Bacteriol 1982; 149:255–63.

MacKellar D, Lieber L, Norman JS *et al.* Streptomyces thermoautotrophicus does not fix nitrogen. Sci Rep. 2016;6:20086.

Malaviya A, Jang YS & Lee SY. Continuous butanol production with reduced byproducts formation from glycerol by a hyper producing mutant of *Clostridium pasteurianum*. Appl. Microbiol. Biotechnol. 2012;93:1485–1494

Marco SA, Shakun JD, Clark PU and Mix AC. A reconstruction of regional and global temperature for the past 11,300 years. Science. 2013; 339(6124): 1198–201.

McKinlay JB & Harwood CS. Carbon dioxide fixation as a central redox cofactor recycling mechanism in bacteria. Proc. Natl Acad. Sci. USA. 2010; 107, 11669–11675.

Mergeay M, Nies D, Schlegel HG, Gerits J, Charles P, Van Gijsegem F. *Alcaligenes eutrophus* CH34 is a facultative chemolithotroph with plasmid-bound resistance to heavy metals. J Bacteriol. 1985;162(1):328–334.

Merrick MA, Knight KL, Ingersoll CD, Potteiger JA. The effects of ice and compression wraps on intramuscular temperatures at various depths. *J Athl Train*. 1993;28(3):236–245.

Meyer O & Schlegel HG. Reisolation of the carbon monoxide utilizing hydrogen bacterium *Pseudomonas carboxydovorans* (Kistner) comb. nov. Arch Microbiol. 1978; 118: 35–43.

Meyer O & Schlegel H G. Oxidation of carbon monoxide in cell extracts of *Pseudomonas carboxydovorans.* J Bacteriol. 1979;137:811–817.

Meyer O & Schlegel HG. Carbon monoxide: methylene blue oxidoreductase from *Pseudomonas carboxydovorans*. J Bacteriol. 1980; 141: 74–80.

Meyer O & Schlegel HG. Biology of aerobic carbon monoxide-oxidizing bacteria. Annu. Rev. Microbiol. 1983; 37:277–310.

Meyer O, Stackebrandt E & Auling G. Reclassification of ubiquinone Q-10 containing carboxidotrophic bacteria: transfer of "[*Pseudomonas*] *carboxydovorans*" OM5T to *Oligotropha*, gen. nov., as *Oligotropha carboxidovorans*, comb. nov., transfer of "[*Alcaligenes*] *carboxydus*" DSM 1086T to *Carbophilus*, gen. nov., as *Carbophilus carboxidus*, comb. nov., transfer of "[*Pseudomonas*] compransoris" DSM 1231T to Zavarzinia, gen. nov., as Zavarzinia compransoris, comb. nov., and amended descriptions of the new genera. Syst. Appl. Microbiol. 1993; 16:390-395

Meyer O, Gremer L, Ferner R, Ferner M, Dobbek H, Gnida M, Meyer-Klaucke W & Huber R. The role of Se, Mo and Fe in the structure and function of carbon monoxide dehydrogenase. Biol. Chem. 2000;381:865–876.

Miura K. An Overview of Current Methods to Confirm Protein-Protein Interactions. *Protein Pept Lett*. 2018;25(8):728–733.

Morsdorf G, Frunzke K, Gadkari D & Meyer O. Microbial growth on carbon monoxide. Biodegradation. 1992; 3, 61-82.

Morton ER, Fuqua C. Genetic manipulation of *Agrobacterium*. Curr Protoc Microbiol. 2012; 3:Unit–3D.2.

Mueller-Cajar O, *et al*. Structure and function of the AAA. Nature. 2012; 479:194–198.

Müller J, MacEachran D, Burd H *et al*. Engineering of *Ralstonia eutropha* H16 for autotrophic and heterotrophic production of methyl ketones. Appl Environ Microbiol. 2013;79(14):4433–4439.

Mundhada H, Seoane JM, Schneider K, Christensen HB, Koza A, Herrgard M, ... Nielsen AT. Increased production of L-serine in *Escherichia coli* through Adaptive Laboratory Evolution. Metabolic Engineering, 2017; 39: 141–150.

Mundhada H, Seoane JM, Schneider K, Koza A, Christensen HB, Klein T, Phaneuf PV, Herrgard M, Feist AM, Nielsen AT. Increased production of L-serine in *Escherichia coli* through adaptive laboratory evolution. Metab Eng. 2017;39:141–150. doi: 10.1016/j.ymben.2016.11.008

Muthusamy S, Lundin D, Mamede Branca RM, Baltar F, González JM, Lehtiö J, et al. Comparative proteomics reveals signature metabolisms of exponentially growing and stationary phase marine bacteria. Environ Microbiol 2017; 19: 2301–2319

Nikolskaya AN, Galperin MY. A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. Nucleic Acids Res. 2002;30(11):2453–2459.

Norman RO *et al.* Progress towards platform chemical production using *Clostridium autoethanogenum*. Biochem. Soc. Trans. 2018; 46, 523–535.

Nybo SE, Khan NE, Woolston BM *et al*. Metabolic engineering in chemolithoautotrophic hosts for the production of fuels and chemicals. Metab Eng. 2015;30:105–120

Netl.doe.gov. 5.1.5. Syngas Composition | netl.doe.gov. [online] Available at: https://www.netl.doe.gov/research/coal/energy-systems/gasification/gasifipedia/syngas-composition. 2018 [Accessed 12 Aug. 2019].

O'Brien JM, Wolkin RH, Moench TT *et al.* Association of hydrogen metabolism with unitrophic or mixotrophic growth of *Methanosarcina barkeri* on carbon monoxide. J Bacteriol 1984;158:373–5.

Oelgeschläger E & Rother M. Carbon monoxide-dependent energy metabolism in anaerobic bacteria and archaea. Arch Microbiol 2008; 190:257–69.

Pappas KM. Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. Plasmid. 2008; 60, 89–107.

Pappas KM & Winans SC. The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid repABC operon. Mol Microbiol. 2003;49(2):441–55.

Park JM, Kim TY & Lee SY. Genome-scale reconstruction and in silico analysis of the *Ralstonia eutropha* H16 for polyhydroxyalkanoate synthesis, lithoautotrophic growth, and 2-methyl citric acid production. BMC Syst Biol. 2011;5:101.

Park C & Zhang J. High expression hampers horizontal gene transfer. Genome Biol Evol. 2012;4(4):523–532.

Parsek MR & Greenberg EP. Acyl-homoserine lactone quorum sensing in gram-negative bacteria: a signaling mechanism involved in associations with higher organisms. Proc Natl Acad Sci U S A. 2000;97(16):8789–8793.

Patrauchan MA, Miyazawa D, LeBlanc JC, Aiga C, Florizone C, Dosanjh M *et al.* Proteomic analysis of survival of *Rhodococcus jostii* RHA1 during carbon starvation. Appl Environ Microbiol. 2012; 78: 6714–6725.

Paul D, Bridges S, Burgess SC, Dandass Y & Lawrence ML. Genome sequence of the chemolithoautotrophic bacterium *Oligotropha carboxidovorans* OM5T. J. Bacteriol. 2008; 190:5531–5532

Paul D, Bridges S, Burgess S, Dandass Y & Lawrence M. Complete genome and comparative analysis of the chemolithoautotrophic bacterium *Oligotropha carboxidovorans* OM5. BMC Genomics. 2011; 11:511.

Pelzmann A, Ferner M, Gnida M, Meyer-Klaucke W, Maisel T & Meyer O. The CoxD protein of *Oligotropha carboxidovorans* is a predicted AAA+ ATPase chaperone involved in the biogenesis of the CO dehydrogenase [CuSMoO2] cluster. J Biol Chem. 2009;284(14):9578–9586.

Pelzmann AM, Mickoleit F & Meyer OJ, Insights into the posttranslational assembly of the Mo-, S- and Cu-containing cluster in the active site of CO dehydrogenase of *Oligotropha carboxidovorans*. Biol Inorg Chem. 2014; 19: 1399.

Pérez -Mendoza D, Sepúlveda E, Pando V, Muñoz S, Nogales J, Olivares J, Soto MJ, Herrera-Cervera JA, Romero D, Brom S, and Sanjuán J. Identification of the *rctA* gene, which is required for repression of conjugative transfer of rhizobial symbiotic megaplasmids. J. Bacteriol. 2005; 187: 7341–7350.

Pérez-Oseguera Á & Cevallos MA. RepA and RepB exert plasmid incompatibility repressing the transcription of the repABC operon. Plasmid. 2013; 70:362–376.

Pfeiffer D & Jendrossek D. Interaction between poly(3-hydroxybutyrate) granule-associated proteins as revealed by two-hybrid analysis and identification of a new phasin in *Ralstonia eutropha* H16. Microbiology. 2011; 157:2795–2807.

Phillips JR, Atiyeh HK, Tanner RS, Torres JR, Saxena J, Wilkins MR & Huhnke RL. Butanol and hexanol production in *Clostridium carboxidivorans* syngas fermentation: Medium development and culture techniques. *Bioresour. Technol.* 2015; 190, 114–121.

Piper KR & Farrand SK. Quorum sensing but not autoinduction of Ti plasmid conjugal transfer requires control by the opine regulon and the antiactivator TraM. J. Bacteriol. 2000;182:1080–1088.

Piper KR, von Bodman SB & Farrand SK. Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. 1993; Nature 362:448–450.

Pinto UM, Flores-Mireles AL, Costa ED & Winans SC. RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within repC and functions only in cis. Mol. Microbiol. 2011; 81 1593–1606.

Pinto UM, Pappas KM & Winans SC. The ABCs of plasmid replication and segregation. Nat Rev Microbiol. 2012; 10:755–765.

Petersen J, Brinkmann H, Pradella S. Diversity and evolution of repABC type plasmids in Rhodobacterales. Environ. Microbiol. 2009; 11 2627–2638.

Poehlein A, Kusian B, Friedrich B, Daniel R, Bowien B. Complete genome sequence of the type strain Cupriavidus necator N-1. J Bacteriol. 2011; 193(18):5017. doi:10.1128/JB.05660-11.

Poole RK & Cook GM. Redundancy of aerobic respiratory chains in bacteria? Routes, reasons and regulation. Adv Microb Physiol. 2000;43:165–224.

Pretorius-Güth IM, Pühler A & Simon R. Conjugal Transfer of Megaplasmid 2 between Rhizobium meliloti Strains in Alfalfa Nodules. Appl Environ Microbiol. 1990;56(8):2354–2359.

Probst I & Schlegel HG. Respiratory components and oxidase activities in Alcaligenes eutrophus . Biochim Biohys Acta. 1976; 440: 412–428.

Qin Y, Luo ZQ, Smyth AJ, Gao P, Beck von Bodman S & Farrand SK. Quorum-sensing signal binding results in dimerization of TraR and its release from membranes into the cytoplasm. EMBO J. 2000;19:5212–5221

Quandt J & Hynes MF. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene. 1993;127(1):15–21.

Quax TE, Claassens NJ, Söll D & van der Oost J. Codon bias as a means to fine-tune gene expression. *Molecular cell, 59*(2), pp.149-161.

Raberg M, Volodina E, Lin K & Steinbüchel A. *Ralstonia eutropha* H16 in progress: applications beside PHAs and establishment as production platform by advanced genetic tools. Crit Rev Biotechnol. 2018;38(4):494–510.

Ramírez-Romero MA, Soberón N, Pérez-Oseguera A, Téllez-Sosa J & Cevallos MA. 2000. Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. J Bacteriol. 182:3117–3124.

Rauch R, Hrbek J, & Hofbauer H. Biomass gasification for synthesis gas production and applications of the syngas. WIREs Energy Environ. (2014); 3: 343-362.

Ray SK, Kumar R, Peeters N, Boucher C, Genin S. *rpoN1*, but not *rpoN2*, is required for twitching motility, natural competence, growth on nitrate, and virulence of *Ralstonia solanacearum. Front Microbiol*. 2015;6:229.

Resch M, Dobbek H & Meyer O. Structural and functional reconstruction in situ of the [CuSMoO2] active site of carbon monoxide dehydrogenase from the carbon monoxide oxidizing eubacterium *Oligotropha carboxidovorans*. J Biol Inorg Chem. 2005; 10: 518–528.

Richter M & Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. USA. 2009; 106, 19126–19131.

Romanchuk A, Jones CD, Karkare K, Moore A, Smith BA, et al. Bigger is not always better: Transmission and fitness burden of a ∼1 MB *Pseudomonas syringae* megaplasmid pMPPla107. Plasmid. 2014; 73: 16–25.

Rovaletti A, Bruschi M, Moro G, Cosentino U, Greco C. The challenging in silico description of carbon monoxide oxidation as catalyzed by molybdenum-copper CO dehydrogenase. Front Chem. 2019;6:630.

Sandberg TE, Lloyd CJ, Palsson BO, Feist AM. Laboratory Evolution to Alternating Substrate Environments Yields Distinct Phenotypic and Genetic Adaptive Strategies. Appl Environ Microbiol. 2017;83(13):e00410-17.

Sandoval-Espinola WJ, Chinn MS, Thon MR & Bruno-Bárcena JM. Evidence of mixotrophic carbon-capture by n-butanol-producer *Clostridium beijerinckii*. Sci Rep. 2017;7(1):12759.

Santiago B & Meyer O. Purification and molecular characterization of the H_2 uptake membrane-bound NiFe-hydrogenase from the carboxidotrophic bacterium *Oligotropha carboxidovorans.* J. Bacteriol. 1997; 179: 6053-6060.

Schittkowski J, Ruland H, Laudenschleger D, Girod K, Kähler K, Kaluza S, Muhler M & Schlögl R. Methanol synthesis from steel mill exhaust gases: challenges for the industrial Cu/ZnO/Al2O3 catalyst. Chem Ing Tec. 2018; 90(10):1419–1429.

Schlegel HG & Lafferty R. Growth of Knallgas bacteria (*Hydrogenomonas*) using direct electrolysis of culture medium. Nature; 1965:205:308–309.

Schlegel HG & Lafferty R. Novel energy and carbon sources. A. The production of biomass from hydrogen and carbon dioxide. Adv Biochem Eng. 1971;1: 143–168.

Schneider K, Cammack R, Schlegel HG & Hall DO. Iron-sulfur centers of soluble hydrogenase from *Alcaligenes eutrophus*. Biochim Biophys Acta. 1979;578(2):445–461.

Schnürer A & Nordberg A. Ammonia, a selective agent for methane production by syntrophic acetate oxidation at mesophilic temperature. Water Sci. Technol. 2008; 57: 735–740.

Schuchmann K, Chowdhury NP & Müller V. Complex Multimeric [FeFe] Hydrogenases: Biochemistry, Physiology and New Opportunities for the Hydrogen Economy. *Front Microbiol*. 2018;9:2911.

Schübel U, Kraut M, Mörsdorf G & Meyer O. Molecular characterization of the gene cluster coxMSL encoding the molybdenum‐containing carbon monoxide dehydrogenase of *Oligotropha carboxidovorans* . J Bacteriol. 1995; 177: 2197‐2203.

Schwartz E, Gerischer U & Friedrich B. Transcriptional regulation of *Alcaligenes eutrophus* hydrogenase genes. Journal of Bacteriology. 1998;180:3197–3204.

Schwartz E, Henne A, Cramm R, Eitinger T, Friedrich B & Gottschalk G. Complete nucleotide sequence of pHG1: a *Ralstonia eutropha* H16 megaplasmid encoding key enzymes of H_2-based lithoautotrophy and anaerobiosis. J Mol Biol. 2003;332(2):369–83.

Shen Y, Chen X, Peng B *et al*. An efficient xylosefermenting recombinant *Saccharomyces cerevi* siae strain obtained through adaptive evolution and its global transcription profile, Appl. Microbiol. Biotech nol. 2012; 96: 1079–1091.

Shen GJ, Shieh JS, Grethlein AJ, Jain AK & Zeikus JK. Biochemical basis for carbon monoxide tolerance and butanol production by *Butyribacterium methylotrophicum* Appl Microbiol Biotechnol. 1999; 51: 827-832.

Shen WC, Wang X, Qin WT, Qiu XF & Sun BW. Exogenous carbon monoxide suppresses *Escherichia coli* vitality and improves survival in an *Escherichia coli*-induced murine sepsis model. Acta Pharmacol Sin. 2014; 35(12):1566–1576.

Shinde S, Cumming JR, Collart FR, Noirot PH, Larsen PE. *Pseudomonas fluorescens* transportome is linked to strain-specific plant growth promotion in aspen seedlings under nutrient stress. Front Plant Sci. 2017;8:348.

Shintani M, Sanchez ZK & Kimbara K. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Front. Microbiol.* 2015; 6:242.

Sievers F & Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. Protein Science. 2018; 27: 135-145.

Simon J & Hederstedt L. Composition and function of cytochrome *c* biogenesis System II. FEBS J. 2011; 278: 4179–4188.

Simon R, Priefer U & Pühler A. A broad host range mobilization system for invivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. Bio/Technology. 1983;1:784-91.

Skerman VBD, Sneath PHA & McGowan V. Approved lists of bacterial names. Int J Syst Bacteriol 1980; 30:225–420.

Slater SC, Voige WH & Dennis DE. Cloning and expression in *Escherichia coli* of the Alcaligenes eutrophus H16 poly-beta-hydroxybutyrate biosynthetic pathway. J Bacteriol. 1988;170:4431–4436.

Smith KM & Liao JC. An evolutionary strategy for isobutanol production strain development in Escherichia coli, Metab. Eng. 2011; 13: 674– 681.

Sokolova TG *et al.* The first evidence of anaerobic CO oxidation coupled with H_2 production by a 42. hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. Extremophiles. 2004; 8, 317–323.

Solomon S, Plattner G-K, Knutti R & Friedlingstein P. Irreversible climate change due to carbon dioxide emissions. Proc. Natl. Acad. Sci. U. S. A. 2009; 106: 1704–1709.

Sonderegger M & Sauer U. Evolutionary engineering of *Saccharomyces cerevisiae* for anaerobic growth on xylose. Appl Environ Microbiol. 2003;69(4):1990–1998.

Steinbuchel A & Fuchtenbusch B. Bacterial and other biological systems for polyester production. Trends Biotechnol. 1998;16(10):419–27.

Straub M, Demler M, Weuster-Botz D & Dürre P. Selective enhancement of autotrophic acetate production with genetically modified *Acetobacterium woodii*. J Biotechnol. 2014; 178:67– 72.

Su S, Khan SR & Farrand SK. Induction and loss of Ti plasmid conjugative competence in response to the acyl-homoserine lactone quorum-sensing signal. *J. Bacteriol.* 2008; 190, 4398–4407.

Sun F, Zhou D, Wang Q *et al.* Genetic characterization of a novel blaDIM-2-carrying megaplasmid p12969-DIM from clinical *Pseudomonas putida*. J. Antimicrob. Chemother. 2016; 71: 909–912.

Sun X, Atiyeh HK Kumar A, Zhang H & Tanner RS. Biochar enhanced ethanol and butanol production by *Clostridium carboxidivorans* from syngas. Bioresour. Technol. 2018; 265: 128–138.

Syvanen M. Evolutionary Implications of Horizontal Gene Transfer. Annu Rev Genet. 2012; 46: 341–358.

Tabata S, Hooykaas PJJ & Oka A. Sequence determination and characterization of the replicator region in the tumour-inducing plasmid pTiB6S3. J Bacteriol. 1989; 171, 1665-1672.

Taghavi S, Provoost A, Mergeay M & van der Lelie D. Identification of a partition and replication region in the *Alcaligenes eutrophus* megaplasmid pMOL28. Mol Gen Genet. 1996; 250:169–179.

Tamegai H, Kato C, Horikoshi K. Pressure-regulated respiratory system in barotolerant bacterium, *Shewanella* sp. strain DSS12. J Biochem Mol Biol Biophys. 1998; 1:213–220.

Tarkka MT, Feldhahn L, Buscot F & Wubet T. Genome Sequence of the *Mycorrhiza* Helper Bacterium *Streptomyces* sp. Strain AcH 505. *Genome Announc*. 2015; 3(2):e01386-14.

Tee K L, Grinham J, Othusitse AM, Gonzalez‐Villanueva M, Johnson AO & Wong TS. An efficient transformation method for the bioplastic‐producing "Knallgas" bacterium *Ralstonia eutropha* H16. Biotechnology Journal. 2017; 12: 1700081.

Torella JP, Gagliardi CJ, Chen JS, Bediako DK, Colon B, Way JC, Silver PA & Nocera DG. Efficient solar-to-fuels production from a hybrid microbial-water-splitting catalyst system. Proc Natl Acad Sci U S A. 2015; 112:2337–2342

Truchet G, Rosenberg C, Vasse J, Julliot JS, Camut S, Denarie J. Transfer of Rhizobium meliloti pSym genes into *Agrobacterium tumefaciens*: host-specific nodulation by atypical infection. J Bacteriol. 1984; 157(1):134–142.

Trutko SM, Evtushenko LI, Dorofeeva LV, Shlyapnikov MG, Gavrish EY, Suzina NE & Akimenko VK. Terminal Oxidases in Representatives of Different Genera of the Family Microbacteriaceae, Mikrobiologiya. 2003; 72:1–8.

Tully BJ, Wheat CG, Glazer BT & Huber JA. A dynamic microbial community with high functional redundancy inhabits the cold, oxic subseafloor aquifer. ISME J. 2018;12(1):1–16.

Van der Linden E, Burgdorf T, Bernhard M, Bleijlevens B, Friedrich B *et al.* The soluble [NiFe]-hydrogenase from *Ralstonia eutropha* contains four cyanides in its active site, one of which is responsible for the insensitivity towards oxygen. J Biol Inorg Chem. 2004; 9: 616–626.

Vega JL, Antorrena GM, Clausen EC *et al.* Study of gaseous substrate fermenations: carbon monoxide conversion to acetate. 2. Continuous culture. Biotechnol Bioeng 1989; 34:785–93.

Verhille S, Baïda N, Dabboussi F, Hamze M, Izard D & Leclerc H. *Pseudomonas gessardii* sp. nov. and *Pseudomonas migulae* sp. nov., two new species isolated from natural mineral waters. Int. J. Syst. Bacteriol. 1999; 49, 1559-72.

Verma A, Hirsch DJ, Glatt CE, Ronnett GV & Snyder SH. Carbon monoxide: a putative neural messanger. Science.1993; 259: 381–384.

Volland S, Rachinger M, Strittmatter A, Daniel R, Gottschalk G & Meyer O. Complete genome sequences of the chemolithoautotrophic *Oligotropha carboxidovorans* strains OM4 and OM5. J Bacteriol. 2011;193(18):5043.

Volova TG, Kalacheva GS & Altukhova OV. Autotrophic synthesis of polyhydroxyalkanoates by the bacteria *Ralstonia eutropha* in the presence of carbon monoxide. Appl Microbiol Biotechnol. 2002;58:675–678.

von Neubeck M, Huptas C, Glück C, Krewinkel M, Stoeckel M, Stressler T, Fischer L, Hinrichs J, Scherer S & Wenning M. *Pseudomonas helleri* sp. nov. and *Pseudomonas weihenstephanensis* sp. nov., isolated from raw cow's milk. Int J Syst Evol Microbiol. 2017; 66:1163–1173.

Wang X, Jin D, Zhou L, et al. Draft Genome Sequence of Halotolerant Polycyclic Aromatic Hydrocarbon-Degrading *Pseudomonas bauzanensis* Strain W13Z2. Genome Announc. 2014; 2(5):e01049-14.

Wall D, Delaney JM, Fayet O, Lipinska B, Yamamoto T & Georgopoulos C. *arc*-Dependent thermal regulation and extragenic suppression of the *Escherichia coli* cytochrome *d* operon. J Bacteriol. 1992;174:6554–6562.

Waterhouse AM, Procter JB, Martin DMA, Clamp M & Barton GJ. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009; 25: 1189-1191

Winkler J, Reyes LH, Kao KC. Adaptive Laboratory Evolution for Strain Engineering. Syst. Metab. Eng. 2013; 985:211–222.

Winslow CE, Broadhurst J, Buchanan RE, Krumwiede C, Rogers LA & Smith GH. The Families and Genera of the Bacteria: Final Report of the Committee of the Society of American Bacteriologists on Characterization and Classification of Bacterial Types. J. Bacteriol. 1920; 5, 191-229.

Weber CF & King GM. The phylogenetic distribution and ecological role of carbon monoxide oxidation in the genus *Burkholderia*. FEMS Microbiol Ecol. 2012; 79: 167–175

Wetzel ME, Olsen GJ, Chakravartty V & Farrand SK. The repABC Plasmids with Quorum-Regulated Transfer Systems in Members of the Rhizobiales Divide into Two Structurally and Separately Evolving Groups. Genome Biol Evol. 2015; 7:3337–3357.

Wilcoxen J, Zhang B & Hille R. Reaction of the molybdenum- and copper-containing carbon monoxide dehydrogenase from *Oligotropha carboxydovorans* with quinones. Biochemistry. 2011; 50(11):1910–1916.

Wilkins MR & Atiyeh HK. Microbial production of ethanol from carbon monoxide. *Curr. Opin. Biotechnol.* 2011; *22*: 326–330.

Windhorst C & Gescher J. Efficient biochemical production of acetoin from carbon dioxide using *Cupriavidus necator* H16. Biotechnology for Biofuels. 2019; 12(1):163.

Winson MK *et al*. Construction and analysis of *luxCDABE*-based plasmid sensors for investigating N-acyl homoserine lactone-mediated quorum sensing. FEMS Microbiol. Lett. 1998; 163:185–92.

Wisniewski-Dyé F & Downie J.A. Quorum-sensing in *Rhizobium*. Antonie Van Leeuwenhoek. 2002;81:397–407.

Wisselink HW, Toirkens MJ, del Rosario Franco Berriel M *et al.* Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. Appl Environ Microbiol. 2007;73(15):4881–4891.

World Energy Outlook 2016. [online] Available at: https://www.iea.org/newsroom/news/2016/november/world-energy-outlook-2016.html [Accessed 26 June 2019].

Wübbeler JH, Steinbüchel A. New pathways for bacterial polythioesters. Curr Opin Biotechnol. 2014; 29:85–92.

Xiong B, Li Z, Liu L, Zhao D, Zhang X & Bi C. Genome editing of *Ralstonia eutropha* using an electroporation-based CRISPR-Cas9 technique. Biotechnol Biofuels. 2018; 11:172.

Xue L, Farrugia G, Miller SM, Ferris CD, Snyder SH & Szurszewski JH. Carbon monoxide and nitric oxide as coneurotransmitters in the enteric nervous system: evidence from genomic deletion of biosynthetic enzymes. Proc Natl Acad Sci U S A. 2000; 97(4):1851–1855.

Yu T & Chen YG. Effects of elevated carbon dioxide on environmental microbes and its mechanisms: a review Sci. Total Environ. 2019; 655: 865-879.

Yu, J & Munasinghe P. Gas fermentation enhancement for chemolithotrophic growth of *Cupriavidus necator* on carbon dioxide. Fermentation. 2018; 4:63.

Zakhary R *et al*. Targeted gene deletion of heme oxygenase 2 reveals neural role for carbon monoxide. Proc. Natl Acad. Sci. USA. 1997; 94: 14848–14853.

Zavarzin GA, Nozhevnikova AN. Aerobic carboxydobacteria. Microb Ecol. 1977; 3:305-326.

Żebracki K, Koper P, Marczak M, Skorupska A & Mazur A. Plasmid-Encoded RepA Proteins Specifically Autorepress Individual repABC Operons in the Multipartite *Rhizobium leguminosarum* bv. trifolii Genome. PLoS One. 2015; 10(7):e0131907.

Zhang L, Murphy PJ, Kerr A & Tate ME. Agrobacterium conjugation and gene regulation by N-acyl-L-homoserine lactones Nature. 1993; 362:446.

Zhu J & Winans SC. Activity of the quorum-sensing regulator TraR of Agrobacterium tumefaciens is inhibited by a truncated, dominant defective TraR-like protein. Mol. Microbiol. 1998; 27:289–297.

Zhu J & Winans SC. Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters *in vitro* and decreases TraR turnover rates in whole cells. Proc. Natl Acad. Sci. USA. 1999; 96:4832–4837.

Zhu J & Winans SC. The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. Proc. Natl Acad. Sci. USA. 2001; 98:1507–1512.

Zhang J & Madden TL. "PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation." Genome Res. 1997; 7:649-656.

Zhang R, Wang M & Ren L. Long-term trends of carbon monoxide inferred using a two-dimensional model. Chemosphere. 2001; 3: 123–132.

Zheng H, Mao Y, Zhu Q *et al.* The quorum sensing regulator CinR hierarchically regulates two other quorum sensing pathways in ligand-dependent and -independent fashions in Rhizobium etli. *J Bacteriol*. 2015; 197(9):1573–1581.

Zheng X, Ying Z, Wang B & Chen C. Hydrogen and syngas production from municipal solid waste (MSW) gasification via reusing CO_2 .Appl Therm Eng. 2018; 144:242-247.