Quaternary amine metabolism in gut microbiota

Tiantian Fu

A thesis submitted to the School of Life Sciences in fulfilment of the requirements for the degree of Doctor of Philosophy

August 2017

The University of Warwick
Coventry, UK

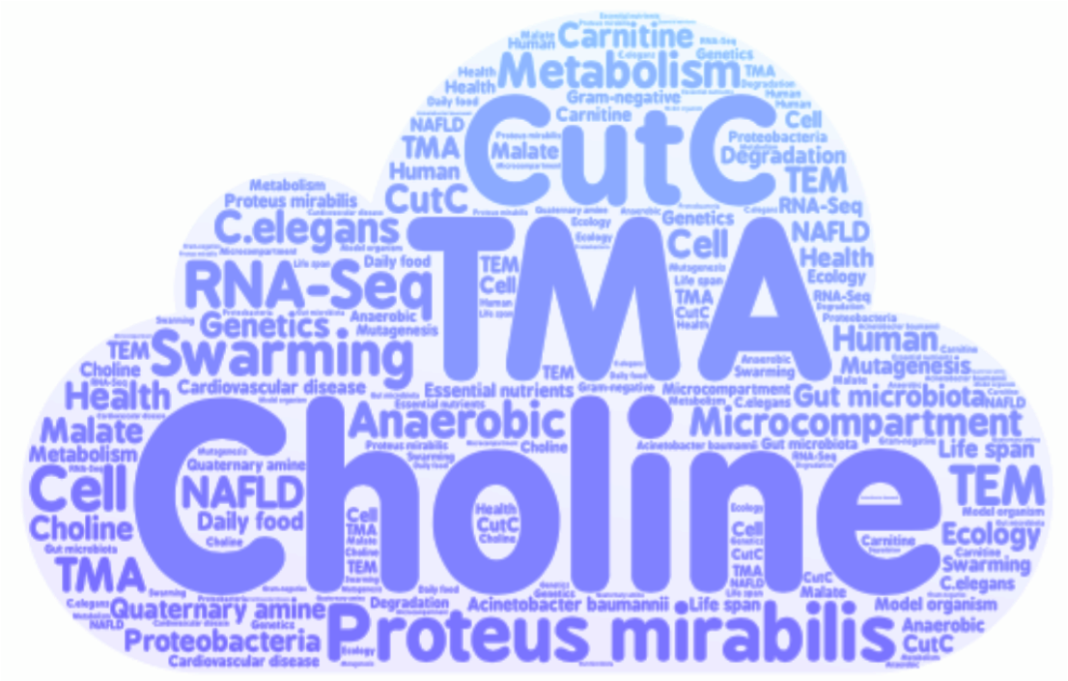


Table of contents

Acknowledgement……I

Declaration……..II

Abstract……III

Abbreviation……IV

List of Figures……..VII

List of Tables……XI

Chapter 1 Introduction……...1

1.1 The ecology and function of the microbial community in the human gut……..2

1.1.1 The ecology of the microbial community in the human gut……2

1.1.2 The function of the microbial community in the human gut……5

1.2 Quaternary amines and their degradation to methylated amines in the human gut……..9

1.2.1 Choline and its role in human health and disease……9

1.2.1.1 Sources of choline ……...11

1.2.1.2 Choline metabolism and cardiovascular disease……13

1.2.1.3 Choline metabolism and liver disease……13

1.2.2 Gut microbial degradation of quaternary amines ……15

1.3 *Proteus mirabilis*……...26

1.4 Comparative transcriptomics RNA-Seq analysis……...……...29

1.5 Bacterial microcompartments…….……..…….33

1.6 *Caenorhabditis elegans*……..……...37

1.6.1 *C. elegans* anatomy and life cycle……...……...37

1.6.2 Importance of *C. elegans* as a model organism for studying host-bacteria interactions…….40

1.6.3 Use of *C. elegans* to study *Acinetobacter baumanni* infection…….42

1.7 Aims and hypotheses…….45

Chapter 2 Materials and Methods……...46

2.1 *Proteus mirabilis*……...47

2.1.1 Maintenance of *Proteus mirabilis* strains……..47

2.1.2 Antibiotics used for *Proteus mirabilis* strains…….……..48

2.1.3 Preparation and transformation of eletrocompetent *P. mirabilis* cells…..48

2.1.4 Construction and complementation of *cutC* mutant in *P. mirabilis*……..48
2.1.5 Growth of *P. mirabilis* on choline and quantification for choline and TMA……...49
2.1.6 *P. mirabilis* swarming assays……49
2.1.7 Heterologous expression of *P. mirabilis* choline-TMA lyase CutC and CutD……49
2.1.8 Transmission eletron microscopy (TEM) imaging of *P. mirabilis* microcompartments……50
2.1.9 Purification of *P. mirabilis* microcompartments…….……..52
2.2 *Escherichia coli*…….52
2.2.1 Maintenance of *Escherichia coli*…….……...52
2.2.2 Antibiotics used for *Escherichia coli*……..……...53
2.2.3 Transformation of chemically competent *E. coli* cells……...53
2.3 *Acinetobacter baumannii*……...……...53
2.4 Extraction of nucleic acids…….……...53
2.5 RNA-Seq……...54
2.6 Nucleic acid manipulation techniques……...54
2.7 Analytical methods…….……...57
2.8 Protein manipulation methods……...59
2.9 Quantification of β-galactosidase activity (Miller assay) …….59
2.10 *Caenorhabditis elegans* life span assays……62
2.10.1 Maintenance of *C. elegans*…….……..62
2.10.2 Bleach synchronization of *C. elegans*……..62
2.10.3 The role of carnitine degradation on *C. elegans* life span…..……..63
2.10.3.1 Carnitine degradation in *Acinetobactor baumannii* on *C. elegans* life span……...63
2.10.3.2 Carnitine degradation in *Escherichia coli* SE11 on *C. elegans* life span……66
2.10.4 Survival analysis……..66
Chapter 3 Developing a method for genetic manipulation of *Proteus mirabilis*……..72
3.1 Introduction……...73
3.2 *Proteus mirabilis* strains and their growth on choline……..75
3.3 Minimum inhibitory concentration test on *P. mirabilis* strains…….75

3.4 Establishment of the protocol for plasmid transfer in *P. mirabilis* strain DMSZ4479……77
3.5 Generating a *cutC* knockout mutant in *P. mirabilis* using a TargeTron targeted transposon mutagenesis system……...79
3.5.1 Construction of the pACD4K-*cutC*……..……..79
3.5.2 Electroporation of *P. mirabilis* DSMZ4479 [pAR1219] competent cells with pACD4K-*cutC*……81
3.5.3 Complementation of the *cutC* mutant…….……...83
3.6 Conclusions……...85
Chapter 4 Comparative transcriptomic analysis of *Proteus mirabilis* by RNA-Seq……...86
4.1 Introduction……...87
4.2 Results……...88
4.2.1 Experimental setup……88
4.2.2 RNA-Seq raw read counts data summary……..91
4.2.3 Assessment of biological replicates……...96
4.2.4 Comparative transcriptome analysis by RNA-Seq in *Proteus mirabilis*...97
4.2.4.1 Differential expression of genes in response to choline degradation to TMA in *P. mirabilis*……97
4.2.4.2 Re-analysis of differential gene expression by RNA-Seq in *P. mirabilis*……100
4.2.5 Identification a choline-TMA lyase gene cluster in *Proteus mirabilis*…108
4.3 Discussion……...123
Chapter 5 Anaerobic choline degradation in *Proteus mirabilis*……..127
5.1 Introduction…….……128
5.2 Results……129
5.2.1 Heterologous expression of the putative *cutCD* from *P. mirabilis* to *E. coli*……129
5.2.2 Marker exchange mutagenesis of the essential *cutC* in TMA formation from choline in *P. mirabilis*……..131
5.2.3 Identification of the physiological role of choline degradation in *P. mirabilis*……133
5.2.3.1 Choline metabolism promotes anaerobic growth of *P. mirabilis* in liquid broth…..……...133

5.2.3.2 Choline metabolism promotes anaerobic swarming of *P. mirabilis*.133
5.2.3.3 Choline degradation to TMA enhances anaerobic swarming motility in *P. mirabilis*. ……..135
5.2.4 Choline induces microcompartment protein formation in anaerobic liquid cultures or on swarming plates of *P. mirabilis*……139
5.2.4.1 Microcompartment protein formation in anaerobic liquid cultures supplemented with/without choline……...139
5.2.4.2 Microcompartment protein formation on anaerobic swarming plates supplemented with/without choline……...140
5.2.4.3 Microcompartment protein formation in the wild type, the *cutC::kan* mutant and the complemented mutant…….…….....140
5.2.4.4 Microcompartment protein purification…….……...144
5.2.4.5 Microcompartment protein in recombinant *E. coli*……...146
5.2.5 Investigate the role of the N-terminal domain in CutC of *P. mirabilis*...159
5.2.6 Investigation of the role of a putative choline-responsive *tetR*-type transcriptional regulator in choline-TMA lyase expression…….163
5.3 Discussion……...166
5.3.1 The role of CutC and the role of choline degradation in TMA formation in *P. mirabilis*……...166
5.3.2 The role of microcompartment proteins in choline degradation……..169
5.3.3 The role of the N-terminal domain in CutC of *P. mirabilis*.. ……..172
5.3.4 The role of the *tetR* transcriptional regulator in choline-TMA lyase expression…….174
Chapter 6 The impact of gut microbial TMA formation on *Caenorhabditis elegans*……..175
6.1 Introduction…….176
6.2 Results…….177
6.2.1 The impact of TMA formation by *A. baumannii* on the life span of *C. elegans*……..177
6.2.2 The role of TMA formation in *E. coli* SE11 on the life span of *C. elegans*……..180
6.2.2.1 Feeding of *E. coli* SE11 by *C. elegans* strain N2……..……...180
6.2.2.2 Feeding of *E. coli* SE11 by *C. elegans* strain CF512….……...186
6.3 Discussion……...189

6.3.1 Compare *C. elegans* N2 life span during carnitine degradation in *A. baumannii* wild type, *cntA* and *cntB* mutants……...189
6.3.2 Compare *C. elegans* N2 life span fed on *E. coli* SE11 grown on carnitine and other carbon sources……..190
6.3.3 Comparison of *C. elegans* CF512 life span during carnitine degradation in *E. coli* SE11 wild type……..196
Chapter 7 Summary and future perspectives……...199
7.1 Chapter 3 - Developing a method for genetic manipulation of *Proteus mirabilis*……200
7.2 Chapter 4 - Comparative transcriptomic analysis of *Proteus mirabilis* by RNA-Seq……201
7.3 Chapter 5 - Anaerobic choline degradation in *Proteus mirabilis*……203
7.4 Chapter 6 - The impact of gut microbial TMA formation on *Caenorhabditis elegans*……...206
References…….……...210

Acknowledgments

I would like to acknowledge the expert guidance, both technical and academic, and the generous support of my supervisor, Dr. Yin Chen, as well as the time he has dedicated towards thought provoking discussions, and the invaluable expertise in writing and reviewing scientific reports. But most importantly, I would like to acknowledge him for the immeasurable influence on me of the highly self-disciplined attitude towards work and life.

I would like to pay tribute to my PhD committee Prof. Chris Dowson and Dr. Elizabeth Fullam from University of Warwick for their advice and helpful discussions. A special mention to Dr. Stefanie Frank from University of Kent for her genius support on my microcompartment work. Many thanks to Dr. Andre Pires da Silva from University of Warwick and Dr. Kumar Rajakumar from University of Leicester and their research groups for their great guidance and help on my *C.elegans* work.

I would like to thank all the members, both present and past, of Chen Lab and Schäfer Lab and departmental support staff for their support throughout the project, especially Dr. Ellie Jameson, who is warranted for her fundamental work and expert advice in this PhD project. Special mention to Julie for her great technical support, and Dr. Hendrik Schäfer, for his humor words and delicious cake. Thanks to my great colleagues ever, Ian, Yijun, Alastair, Michi, Eileen and Jason, for all the friendship and enjoyable moments we spent together in D14.

I would also like to thank all the members of my hockey team & basketball team, especially Mishan and Algis, who always show great support to my event organization, and have made my writing period delightful and inspiring.

Great thanks to my family for encouraging me and my work. Special mention to my elder brother cousin, who passed away during my PhD, but was responsible for bringing me into the world of microbiology and research, my great memorial to him forever! Last but not least, special tribute to my mum, who has given me this incredible life and great support for my PhD, not only the financial support, but also the influence on my whole life. Without whom, I will not have the chance to study abroad and fulfill this PhD.

Declaration

I declare that the work presented in this thesis was conducted by me under the direct supervision of Dr. Yin Chen, with the exception of those instances where as follows: **1)** Chapter 4, extraction and purification of RNA was carried out by Dr. Eleanor Jameson from University of Warwick, RNA enrichment, quality and quantity assessment and cDNA library preparation were carried out by Dr. Konrad Paszkiewicz from University of Exeter, and the raw read counts and the later RNA-Seq analysis were carried out by Dr. Eleanor Jameson and Dr. Konrad Paszkiewicz **2)** Chapter 5, fixing of samples for TEM and imaging of microcompartment protein were carried out by Ian Brown from University of Kent.

None of the work presented has been previously submitted for any other degree. Some of the data presented in Chapters 4 and 5, have been published, Jameson, E., Fu, T., Brown, I.R., Paszkiewicz, K., Purdy, K.J., Frank, S., Chen, Y. (2015) Anaerobic choline metabolism in microcompartments promotes growth and swarming of *Proteus mirabilis*. *Environ Microbiol* **18:** 2886-2898.

Tiantian Fu

Abstract

Quaternary amines such as choline and carnitine are essential nutrients for humans supplied from daily food; however, quaternary amines metabolism by gut microbiota can lead to the development of various diseases, including non-alcoholic fatty liver disease and cardiovascular disease. It is hypothesized that both diseases are promoted by microbial catabolism of choline and carnitine to trimethylamine (TMA).

Proteus mirabilis is a Gram-negative gut proteobacterium, which can metabolize choline anaerobically to form TMA. I demonstrated that the identified *cutC* gene is essential for choline degradation and subsequent TMA production in this bacterium. Using *P. mirabilis* as the model, I investigated the physiological role of quaternary amine metabolism from the bacterial perspective and demonstrated that *P. mirabilis* can rapidly uptake and degrade choline to enhance growth rate, cell yield and swarming speed under anaerobic and microaerophilic conditions. I also provide the first evidence of a novel choline-metabolizing microcompartment, which is present in both vegetative and swarming cells supplemented with choline.

Another important dietary source of TMA in human gut is carnitine. I used two model proteobacteria *Acinetobacter baumannii* and *Escherichia coli* in this project to investigate the role of carnitine metabolism to TMA in health and disease. *A. baumannii* and *E. coli* can use carnitine as a growth substrate to produce TMA. To better understand the role of quaternary amine metabolism in host health and disease, I used *Caenorhabditis elegans* model to investigate carnitine metabolism on the life span of the worm. My data suggest that malate, the degradation product of carnitine, extends the life span of *C. elegans* fed on *A. baumannii* or *E. coli*. Together, my study reveals that choline and carnitine metabolism as an adaptation strategy for gut proteobacteria and contributes to better understand the ecology of these TMA-forming gut bacteria in health and disease.

Abbreviations

AmpR	ampicillin (resistance)
BHMT	betaine homocysteine methyltransferase
BlastP	basic local alignment search tool using protein database for protein query
BMC	bacterial microcompartments
bp	base pairs
BSA	bovine serum albumin
C	carbon
cDNA	complementary DNA
CDP-choline	cytidine diphosphocholine
CFU	colony-forming units
CGC	Caenorhabiditis Genetic Centre
CntA/B	carnitine oxygenase
CutC	choline trimethylamine-lyase
CVD	cardiovascular disease
Da	Dalton
DMA	dimethylamine
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dH_2O	distilled water
dNTP	deoxynucleotide triphosphate
EC number	enzyme commission number
EDTA	ethylendiaminetetraacetic acid
EPR	electron paramagnetic resonance
EPS	extracellular polysaccharides
Eut	ethanolamine utilization protein
FAD	flavin adenine dinucleotide
FDR	false discovery rate
FMO	flavin-containing monooxygenase
g	gram / acceleration due to gravity
GBT	glycine betaine
GRE	glycyl radical enzyme
h	hour
HDL	high-density lipoprotein
IMG	Integrated microbial genomes
IPTG	Isopropyl β–D-1-thiogalactopyranoside
JGI	Joint Genome Institute
Kan^R	kanamycin (resistance)
l	litre
LacZ	β-galactosidase
LB	Luria-Bertani
LV	low viscosity
M	molar
MA	methylated amine
MA plot	A plot of M-values (log fold-change) as y axis against A-values (average of log expression value) as x axis

MALDI-TOF	matrix-assisted laser desorption/ionization-time of flight mass spectrometry
MAPK	mitogen-activated protein kinase
MCS	multiple cloning site
MIC	minimum inhibitory concentration
mg	milligram
µg	microgram
min	minute
ml	millilitre
mM	millimolar
MMA	monomethylamine
mol	mole
mRNA	messenger RNA
ms	millisecond
NADH	nicotinamide adenine dinucleotide (reduced form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
NAFLD	nonalcoholic fatty liver disease
NGM	nematode growth medium
ng	nanogram
NH_3	ammonia
$NH4^+$	ammonium
OD_{420}	optical density at 420 nm
OD_{600}	optical density at 600 nm
ONPG	O-nitrophenyl-ß-D-galactoside
PCR	polymerase chain reaction
Pdu	propanediol utilization protein
PEMT	phosphatidylethanolamine *N*-methyltransferase
QA	quaternary amine
qRT-PCR	quantitative real-time polymerase chain reaction
RNA	ribonucleic acid
RPKM	reads per kilobase million
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcriptase PCR
RuBisCo	ribulose-1,5-bisphosphate carboxylase/oxygenase
s	seconds
SAM	S-adenosylmethionine
SDS	sodium dodecyl sulphate
SNP	single nucleotide polymorphism
SpecR	Spectinomycin (resistance)
TBE	tris borate EDTA
TetR	Tet repressor protein
TEM	transmission electron microscopy
TMA	trimethylamine
TMAO	trimethylamine *N*-oxide
Tmm	TMA monooxygenase
UTI	urinary tract infection
VLDL	very-low-density lipoprotein
v/v	volume to volume

w/v	weight to volume

List of Figures

Figure 1.1 The proposed ecology and function of key human gut microbiome.....4
Figure 1.2 The chemical structures of choline, glycine betaine, carnitine and methylated amines.....16
Figure 1.3 The super metabolic pathways involved in the degradation of quaternary amines to methylated amines by the gut microbiota or the host.....19
Figure 1.4 Swimming and swarming cells of *P. mirabilis* under transmission electron microscope (TEM) and swarming pattern on agar plate.....27
Figure 1.5 Bacterial RNA-Seq work flow.....30
Figure 1.6 The polyhedral structure of bacterial microcompartment and the metabolic pathway associated with propanediol metabolosome such as PduA, PduB and PduJ.....34
Figure 1.7 *C. elegans* anatomy.....38
Figure 1.8 *C. elegans* life cycle.....39
Figure 1.9 Carnitine degradation and TMA formation in *A. baumanni* and *E. coli* SE11 and the predicted pathway of carnitine catabolism via CntA/B..43
Figure 1.10 The gene cluster of *cntAB* for carnitine degradation in *A. baumanni*, the knock-out genes to create mutants, and the quantification of TMA and carnitine for wild-type and *cntA/B* mutants.....44
Figure 2.1 The preparation of TEM sample for swarming cells on agar plate.....51
Figure 2.2 An example chromatograph showing the respective peaks and retention times for choline and TMA.....58
Figure 2.3 An example of the β-galactosidase activity assay.....61
Figure 2.4 Schematic representation of the impact of carnitine degradation in *A. baumannii* on *C. elegans* life span.....64
Figure 2.5 Different stages of *C. elegans* during life cycle and different observations of *C. elegans* during life span.....65
Figure 2.6 Schematic representation of the impact of carnitine degradation in *E. coli* SE11 on *C. elegans* life span.....67
Figure 3.1 One-step assembly PCR for amplifing the (re-target) intron RNA (A) gel electrophoresis shows the amplified retargeted intron (B)…….....80
Figure 3.2 The plasmid map of pACD4K-C and the retargeted intron RNA

site.....80
Figure 3.3 The primer binding positions for confirming the intron retargeting sites (A) gel electrophoresis shows the bands of wild type and the *cutC* mutant (B) 82
Figure 3.4 A plasmid map showing the insertion of *cutC-cutD* gene cluster cloned into pGEM-T vector.....84
Figure 3.5 Gel electrophoresis showing the amplified *cutC-cutD* gene cluster (A) purified *cutC-cutD* gene cluster (B), and digested pGEM-T vector and the *cutC-cutD* genecluster (C).....84
Figure 4.1 Growth curve of *P. mirabilis* in the defined medium supplemented with glucose or glucose plus choline anaerobically and cells were harvested for RNA extraction when OD_{600} reached 0.4.....89
Figure 4.2 RNA-Seq pipeline for the whole experiment.....90
Figure 4.3 Bioanalyzer analysis of *E. coli* RNA before and after mRNA enrichment.....92
Figure 4.4 The quality report of the sample C1rep3.....93
Figure 4.5 MDS plot of the variation of the RNA-Seq sample replicates.....96
Figure 4.6 MA plot of the RNA-Seq samples.....102
Figure 4.7 Parallel coordinates plot of the RNA-Seq samples.....105
Figure 4.8 The candidate gene clusters involved in choline degradation in the genome of *D. desulfuricans* and *P. mirabilis* (A) and the highly expressed microcompartment proteins in the *cutC* gene cluster with the presence of choline (B).....110
Figure 4.9 The metabolic pathway of choline degradation to TMA and acetaldehyde via the proposed enzymes.....118
Figure 4.10 Neighbour-joining phylogenetic tree constructed from amino acid sequences of glycyl radical enzymes.....120
Figure 4.11 The CutC sequence alignment of *P. mirabilis* and *D. desulfuricans* and the N-terminus of 300 amino acids in *P. mirabilis*.....122
Figure 5.1 Initial growth test of choline degradation to TMA in *P. mirabilis* anaerobically (A) quantification of TMA produced from recombinant *E. coli* carrying *P. mirabilis* wild type *cutC* and a glycyl radical site CutC mutant (B).....130

Figure 5.2 Quantification of choline and TMA concentrations and measurement of optical density (at 600 nm) of *P. mirabilis* wild type (A) *cutC::kan* mutant (B) and the complemented mutant (C).....132
Figure 5.3 Anaerobic liquid growth curves and anaerobic swarm radiuses curves of *P. mirabilis* in different carbon sources (A & B) Anaerobic swarming pattern of *P. mirabilis* on agar plates supplemented with different carbon sources (C).....134
Figure 5.4 Cumulative anaerobic swarming radiuses of *P. mirabilis* grown on glycerol or glycerol plus choline for wild-type (A) *cutC::kan* mutant (B) and complemented mutant (C).....137
Figure 5.5 Anaerobic liquid growth of *P. mirabilis* grown on glycerol, glycerol plus choline or choline alone for wild-type (A) *cutC::kan* mutant (B) and complemented mutant (C).....137
Figure 5.6 Anaerobic swarming for *P. mirabilis* wild type in 10 mM of glycerol supplemented with different concentrations of choline (1, 10, 50 mM) (A) different concentrations of glycerol (32, 50, 130 mM) (B) and different concentrations of TMA (1, 10, 50 mM) (C).....138
Figure 5.7 TEM images for microcompartment protein formation in anaerobic liquid culture of wild type *P. mirabilis* supplemented with glucose (A) choline (B), and glucose plus choline (C).....142
Figure 5.8 TEM images for microcompartment protein formation on anaerobic swarming plates of *P. mirabilis* wild type supplemented with glycerol (A) choline (B), and glycerol plus choline (C).....143
Figure 5.9 TEM images for wild type (A) *cutC* mutant (B) and complemented *cutC* mutant (C) MCPs grown on glucose supplemented with choline.....143
Figure 5.10 TEM image for the purified microcompartment in liquid culture.....145
Figure 5.11 The candidate *cutC* gene clusters involved in choline degradation in the genome of *P. mirabilis* including the five MCP shell proteins...149
Figure 5.12 The map of plasmid pET28a.....149
Figure 5.13 The gel pictures of the three amplified microcompartment proteins and the cloning strategy of the three fragments into pET28a.....151
Figure 5.14 SDS-PAGE gel picture of the overexpression of PMI2722-20 (three

shell proteins) and PMI2722-14 (five shell proteins).....154
Figure 5.15 TEM images for the recombinant microcompartments in *E. coli*.....156
Figure 5.16 The whole *cutC* operon from PMI2722 to PMI2710 (A) gel pictures of the amplified whole *cutC* operon PMI2722-10 (B).....158
Figure 5.17 A plasmid map showing the N-terminus truncated *cutC* cloned into the multiple cloning site MCS 1 using the NcoI & PstI restriction sites and *cutD* cloned into MCS 2 under the sites of NdeI & KpnI..161
Figure 5.18 The PCR products using primers to generate the N-terminus truncated *cutC* (*cutC'*).....161
Figure 5.19 SDS-PAGE gel showing the overexpression results of the total cell protein and supernatant for wild type and CutC truncated mutants..162
Figure 5.20 The map of the plasmid pBIO1878 (*lac*Z) with unique restriction sites and the cloning scheme to construct the pBIO1878 (P_{tetR}-*lac*Z).....164
Figure 5.21 The β-galactosidase assay for the promoter activity for *P. mirabilis* [pBIO1878 (P_{tetR}-*lac*Z)].....165
Figure 6.1 Lifespan analyses of *C. elegans* N2 strain at 20 °C fed on *A. baumannii* wild type, *cntA* or *cntB* mutant.....177
Figure 6.2 Lifespan analysis of *C. elegans* N2 at 25 °C fed on *E. coli* SE11 cultivated on different substrates in carnitine degradation.....180
Figure 6.3 Life span analysis of *C. elegans* N2 at 20 °C fed on *E. coli* SE11 cultivated on different substrates in carnitine degradation.....183
Figure 6.4 Lifespan analysis of *C. elegans* CF512 strains at 25 °C fed on *E. coli* SE11 cultivated on different substrates in carnitine degradation.....186
Figure 6.5 Survival curves (multiple group comparisons) of *C. elegans* N2 at 25 °C growth on different substrates in *E. coli* SE11.....191
Figure 6.6 Survival curves (multiple group comparisons) of *C. elegans* N2 at 20 °C growth on different substrates in *E. coli* SE11.....193
Figure 6.7 Survival curves (multiple group comparisons) of *C. elegans* CF512 at 25 °C growth on different substrates in *E. coli* SE11.....196

List of Tables

Table 1.1 Health impacts of the microbiota.....7
Table 1.2 The major roles of choline in host metabolism.....10
Table 1.3 The major sources of choline.....12
Table 1.4 The key enzymes involved in the degradation of quaternary amines to methylated amines associated with each pathway.....20
Table 1.5 Representative microorganism involved in QA metabolism to TMA…22
Table 1.6 Transcriptomic analysis methods for RNA-Seq in microbial metabolism.....32
Table 1.7 Metabolic pathways involving microcompartments and their functions35
Table 1.8 The applications of *C. elegans* as a model organism to investigate host-bacteria metabolism.....41
Table 2.1 The amount of each component in a standard PCR mix.....56
Table 2.2 The general procedure of a standard PCR cycle.....56
Table 2.3 A list of strains and plasmids used throughout the course of this study.68
Table 2.4 A list of primers used throughout the course of this study.....70
Table 3.1 MIC for four different strains of *P. mirabilis* on non-swarming agar plates.....76
Table 3.2 MIC for *P. mirabilis* DSMZ4479 on non-swarming agar plates.....76
Table 4.1 Raw data counts of RNA-Seq of the two conditions, glucose plus choline and glucose only.....95
Table 4.2 Top 20 differentially expressed genes in RNA-Seq in glucose plus choline and glucose only.....98
Table 4.3 Top 18 differentially expressed genes in glucose plus choline and glucose only in this study.....103
Table 4.4 The predicted functions of the choline-TMA lyase gene cluster in *D. desulfuricans*.....112
Table 4.5 The predicted functions of the choline-TMA lyase gene cluster in *P. mirabilis*.....114
Table 4.6 The read counts data of the highly expressed genes in the *cutC* gene cluster.....116
Table 5.1 The predicted function for the five shell proteins and the BlastP results.....147

Table 6.1 Survival analysis (log rank test/multiple comparison) of *C. elegans* N2 strain fed on *A. baumannii* wild type, *cntA* or *cntB* mutant.....178
Table 6.2 Survival analysis (log rank test) of *C. elegans* N2 fed on *E. coli* SE11 grown on various carbon sources at 25 °C.....181
Table 6.3 Survival analysis (log rank test) of *C. elegans* N2 fed on *E. coli* SE11 grown on various carbon sources at 20 °C.....184
Table 6.4 Survival analysis (log rank test) of *C. elegans* CF512 fed on *E. coli* SE11 grown on difference carbon sources at 25 °C.....187
Table 6.5 Survival analysis (multiple group comparisons) of *C. elegans* N2 fed on *E. coli* SE11 in different substrates at 25 °C.....192
Table 6.6 Survival analysis (multiple group comparisons) of *C. elegans* N2 fed on *E. coli* SE11 in different substrates at 20 °C.....194
Table 6.7 Survival analysis (multiple group comparisons) of *C. elegans* CF512 fed on *E. coli* SE11 in different substrates at 25 °C.....197

Chapter 1

Introduction

1.1 The ecology and function of the microbial community in the human gut

1.1.1 The ecology of the microbial community in the human gut

The human gut microbiome is a diverse community, which more than 1000 bacterial species inhabit; these encode about 5 million genes (D'Argenio and Salvatore, 2015). Thanks to the rapid development of next-generation sequencing technologies and its capacity for metagenome analysis, our understanding of the composition and ecosystem function of the human gut microbiota has been significantly advanced over the past decade (Shreiner *et al.*, 2015).

It is well known that the gut microbiota plays a fundamental role in health and disease. However, certain microbes have started to colonize the human gut as early as birth, and the gut microbial community continues developing during the first few years, in early childhood (Bull and Plummer, 2014). In the healthy human gut, the microbial community is composed of Bacteria, Archaea, viruses and some unicellular Eukaryotes (**Fig 1.1,** D'Argenio and Salvatore, 2015**)**. Among the bacterial species, the most dominant are *Bacteroidetes*, *Firmicutes*, Actinobacteria and *Proteobacteria*. *Bacteroidetes* and *Firmicutes* make up more than 90 % of the total microbial biomass in the healthy gut (Arumugam *et al.*, 2011; Yatsunenko *et al.*, 2012). It has been reported that the composition of the gut microbial community varies among individuals and is associated with numerous factors such as diet. *Bacteroidetes* and *Firmicutes* have been found to be responsible for the utilisation of fiber and peptide from our daily food (Ravcheev *et al.*, 2013; Schoberth and Gottschalk, 1969). Interestingly, the excess fat and energy intakes from diet might lead to the increased population of *Firmicutes* and decreased population of *Bacteroidetes* in the gut, which could result in obesity (Conlon and Bird, 2015).

Proteobacteria, particularly those *Gammaproteobacteria* of the *Enterobacteriaceae* family, are mainly composed of facultative anaerobes, which are thought to persist in close proximity to the mucosa of the gut lumen (Eckburg *et al.*, 2005; El Kaoutari *et al.*, 2013; Winter *et al.*, 2013b). Recent studies have found a key role of *Proteobacteria* in quaternary amine metabolism in the human gut. For example, the

conversion of dietary choline to trimethylamine (TMA) by certain *Proteobacteria* may increase the risk of various diseases (Wang *et al.*, 2011; Dumas *et al.*, 2006). Interestingly, some methanogenic archaeal species, which are also present in human gut, may play an important role for the conversion of TMA to methane (Gaci *et al.*, 2014).

Moreover, it has been reported that in the healthy human gut, although *Gammaproteobacteria* are a key component of the gut microbiome, they only persist at low levels. Indeed, a sudden increase in the abundance of *Gammaproteobacteria*, particularly *Enterobacteriaceae*, has been reported to be associated with an imbalance in gut microbiota, named dysbiosis, which is a key indicator for gut malfunction and intestinal inflammation (Winter *et al.*, 2013b). Investigating the mechanisms underpinning the microbial community changes in *Gammaproteobacteria* during dysbiosis will advance our understanding of the variations between health and unbalanced gut microbial communities, and the relationship of gut dysbiosis and microbial metabolism. This would ultimately contribute to novel clinical treatment procedures to improve human health.

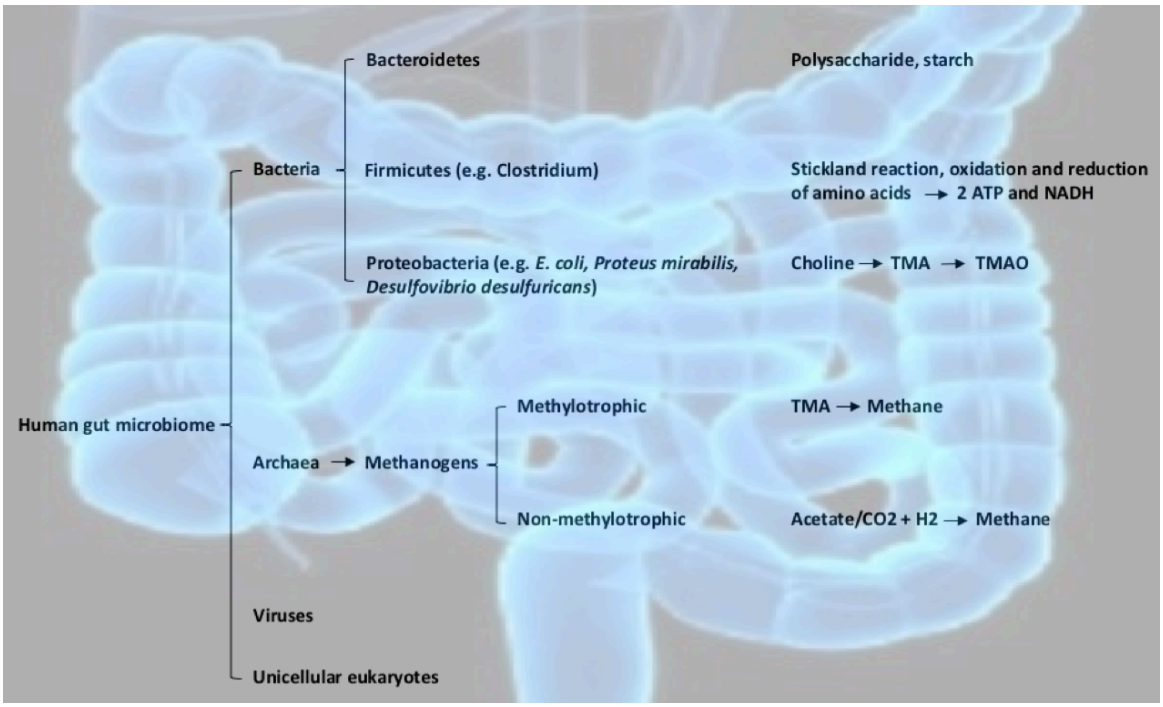


Figure 1.1 The proposed ecology and function of key human gut microbiome (D'Argenio and Salvatore, 2015, background picture modified from wed source: http://www.nutraingredients.com/Research/Microbiota-may-communicate-with-human-gut-cells-through-enzyme-signalling, copyright license: William Reed Business Media Ltd registered in England no. 2883992).

1.1.2 The function of the microbial community in the human gut

Due to the impact of rapid advances in next generation sequencing on the gut microbiome, there is a renewed interest in better understanding the function of key gut microbiome species. Large amount of scientific data has been accumulated over the past decade, highlighting a significant role of the microbial community in the human gut, not only in maintaining a healthy human gut, but also playing a pivotal role in disease development and progression. The activity of the gut microbiota can be directly associated with a number of health benefits in the human body, including biological metabolite synthesis, host-derived immune function, colonization resistance, cell differentiation, as well as regulating neural systems. **Table 1.1** summarises some of the most well studied impacts that are associated with the gut microbiome.

One of the best-known examples of gut microbiota in human health is arguably the synthesis of essential vitamins, which is largely dependent on the activity of gut microbiota. Experiments using mice model have confirmed that the host metabolism requires the supplementation with vitamin K and some B vitamins, which are primarily synthesized by certain gut microbiota including *Bacteroides*, *Eubacterium*, *Fusobacterium* and *Propionibacterium*n (Wostmann, 1981; Bentley, 1982).

Gut microbiota are also responsible for carbohydrates fermentation. The non-digestible carbohydrates in the diet are largely polysaccharides, undigested oligosaccharides and unabsorbed sugars. Furthermore, the product of carbohydrates fermentation is primarily short-chain fatty acids. Among them, propionate, acetate and butyrate are the three predominant short-chain fatty acid products, which are essential in epithelial cell proliferation and differentiation (Frankel, 1994). It is therefore not surprising that gut microbiota are essential for colonic physiology through their influence on epithelial cell regulation (Canny, 2008).

Important metabolic processes, which are associated with the gut microbiota, include the activation of bioactive food components, the conversion of pro-drugs to bioactive forms, xenobiotics and bile acids transformation (Marchesi, 2007; Blaut, 2007). These associations are important clues that reveal the relationship between the microbiome and drug metabolism, but also to uncover the impact of microbiome

on host processes of great importance for developing clinical therapeutics (Blaser, 2013).

One groundbreaking line of investigation into gut microbiota in recent years is the association between gut microbial metabolites and cardiovascular disease. Wang *et al*. has revealed that a diet rich in choline and carnitine can significantly increase the risk of atherosclerosis (Wang *et al*. 2011). Understanding the underlying mechanisms for the catabolism of choline by gut microbiota and the subsequent impacts on the host is the key aim of this thesis **(See Section 1.7)**.

Table 1.1 Health impacts of the microbiota (modified from MSc Module assignment, BS942 Project)

Function in host	Microbial contribution	Reference
Metabolism		
Vitamin synthesis	Synthesis of vitamin K and some B vitamins by *Bacteroides*, *Eubacterium*, *Fusobacterium* and *Propionibacterium*	Wostmann, 1981; Bentley, 1982
Carbohydrate fermentation	Fermentation of large polysaccharides, undigested oligosaccharides, unabsorbed sugars and alcohols	Canny, 2008
Cardiovascular diseases	Increasing risk of atherosclerosis from diet rich in choline and carnitine	Wang *et al*., 2011
Activation of bioactive food components and xenobiotic transformation	Metabolism of xenobiotic compounds by microbiota	Marchesi, 2007; Blaut, 2007
Immunity		
Development of mucosal immune system	Interact with host cells at mucosal interface; influence structure of gut-associated lymphoid tissue	Linares, 2016
Enhancement of immunomodulatory responses	Aid activation of macrophages, cytotoxic T-lymphocytes, cytokines; contribute to memory mechanisms e.g. oral tolerance	Canny, 2008
Other functions		

Colonization resistance	Attachment by commensals to epithelial cells targeted by pathogens; out-compete pathogenic bacteria for nutrients	Canny, 2008
Epithelial cell proliferation and differentiation	Directed by short-chain fatty acid produced from microbial fermentation	Frankel, 1994
Regulation of neuronal gene expression	Production of the short chain fatty acid regulator such as butyrate by gut microbiota	Linares, 2016

1.2 Quaternary amines and their degradation to methylated amines in the human gut

1.2.1 Choline and its role in human health and disease

Choline plays several key roles in human health via host metabolism. Apart from the common role as central nutrient, choline also contributes to the synthesis of phospholipids (phosphatidylcholine, PC and sphingomyelin), which are the essential components of cell membrane lipids. Phosphatidylcholine can be generated from dietary choline through the so-called cytidine 5'-diphosphocholine (CDP-choline) pathway or via the sequential methylation of phosphatidylethanolamine (Ueland, 2011; Gibellini and Smith, 2010). Choline is also involved in cell signaling, nerve transmission, lipid transport and osmoregulation (Gibellini and Smith, 2010; Li and Vance, 2008; Noga and Vance, 2003; Noga *et al.*, 2002; Ueland, 2011). Choline-derived glycine betaine is the major methyl donor in human and other mammals, which is used for the conversion of homocysteine to methionine through betaine: homocysteine methyltransferase (BHMT) (Pellanda, 2013). Current known functions of choline in human metabolism are summarized in **Table 1.2**.

Table 1.2 The major roles of choline in host metabolism

Role	Description	Reference
Maintaining structural integrity of cell membrane	Construction of the cell membrane via synthesis of phosphatidylcholine and sphingomyelin	Ueland, 2011; Gibellini and Smith, 2010
Cell signaling	Generating intracellular messages and cell-signaling molecules such as diacylglycerol and ceramide	Gibellini and Smith, 2010
Nerve impulse transmission	Synthesis of acetylcholine as a neurotransmitter involved in muscle control, circadian rhythm and memory	Li and Vance, 2008
Lipid (fat) transport and metabolism	Assisting very-low-density lipoproteins (VLDL, mainly fat and cholesterol) assembly and secretion from liver	Noga and Vance, 2003; Noga *et al.*, 2002
Major source of methyl donor	Methyl donor from choline-converted glycine betaine for methylation of homocysteine	Pellanda, 2013
Osmoregulation	Regulation of cell volume and protecting cell integrity against osmotic stress by choline-converted glycine betaine	Ueland, 2011
Integration with other nutrients	Metabolism of nucleic acids and amino acids and generation of methyl group donor together with other B-vitamins	Zeisel and Corbin, 2012

1.2.1.1 Sources of choline

Although humans can synthesise choline via a dedicated *de novo* synthesis pathway, our daily requirement of choline is in fact mainly met from dietary intake, such as fish, eggs, meat and milk. *De novo* choline synthesis is produced via the methylation of phosphatidylethanolamine to PC by phosphatidylethanolamine *N*-methyltransferase (PEMT) using S-adenosylmethionine (SAM) as the methyl group donor (Vance, 2014). In addition to the *de novo* choline pathway, choline can also be released as a byproduct during the synthesis of phosphatidylserine (Li and Vance, 2008), therefore it was not classified as an essential vitamin (Secades, 2011). However, *de novo* synthesis of choline is far from sufficient to meet the metabolic needs of a healthy human, and choline supplements in the form of choline salts or lecithin from soybean and egg yolk are required (Koc *et al.*, 2002; Hendler and Rorvik, 2008) **(Table 1.3)**.

Table 1.3 The major sources of choline

Sources	Description	Reference
De novo synthesis (biosynthesis)	Conversion of phosphatidylethanolamine to PC; substitution of choline by serine from PC to phosphatidylserine;	Li and Vance, 2008
Food sources	Eggs, liver, peanuts, fish, pasta, rice, meat, poultry and dairy food; homocysteine remethylation; PC in dietary food; Lecithin extracts	Food and Nutrition Board, 1998; Chester *et al.*, 2007-2008; Craig, 2004; Zeisel and Corbin, 2012; Koc *et al.*, 2002
Supplements	CDP-choline and choline salts, such as choline chloride and choline bitartrate; PC supplements; commercial lecithins prepared from soybean, sunflower and rapeseed; egg yolk lecithin	Koc *et al.*, 2002; Hendler and Rorvik, 2008

1.2.1.2 Choline metabolism and cardiovascular disease

As discussed previously, choline plays various fundamental roles through metabolism in healthy humans. One of the most important roles is that choline can generate glycine betaine, which is essential for the conversion of homocysteine to methionine by betaine:homocysteine methyltransferase (BHMT) (Pellanda, 2013). Disordered choline metabolism to glycine betaine, however, will cause increased concentrations of homocysteine in the blood. It has been reported that elevated homocysteine levels increase the risk of cardiovascular disease (CVD). Homocysteine-induced injury to the arterial walls initiate the process of atherosclerosis and eventually to strokes and heart attacks (Gerhard and Duell, 1999). Moreover, elevated homocysteine concentrations may trigger the development of atherosclerosis and thrombogenesis by the mechanisms of inflammation, abnormal blood coagulation, oxidative stress and disordered lipid metabolism endothelial (Zhou and Austin, 2009).

It is well-known that plasma choline levels are related to cardiovascular risk factors, such as the percentage of body fat, serum triglycerides and high-density lipoprotein (HDL) (Konstantinova *et al.*, 2008). However, an intriguing recent observation was that trimethylamine *N*-oxide (TMAO) in the blood, rather than choline, is the key biomarker in the development of cardiovascular events (Wang *et al.*, 2011). Understanding how TMAO is formed from dietary choline through microbial metabolism in representative gut microbiota is one of the key aims of this PhD thesis **(See Chapters 4&5)**.

1.2.1.3 Choline metabolism and liver diseases

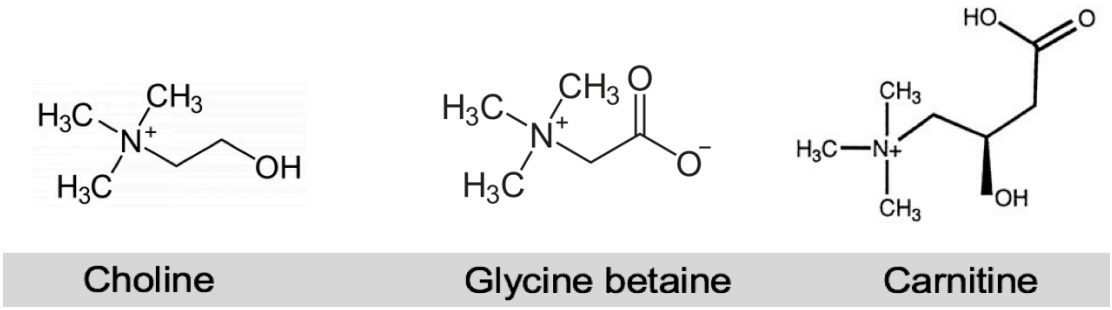
Choline metabolism in the gut is also associated to the development of liver diseases. Dumas *et al.* (2006) have demonstrated a remarkable finding that disruption of choline metabolism by gut microbiota is associated with nonalcoholic fatty liver disease (NAFLD). It is hypothesized that NAFLD is directly linked to gut microbial transformation of choline to TMA. This microbial bypass resulted in a reduction of choline bioavailability to the host, leading to a low level of plasma PC and a high level of urinary methylamines, triggering choline-deficiency and subsequent NAFLD.

In addition, animal experiments suggest that dietary choline deficiency caused an increased risk of spontaneous liver cancer (hepatocellular carcinoma) and increased sensitivity to carcinogenic chemicals (Pellanda, 2013). Several mechanisms have been reported to reveal the effect of choline deficiency on the increased risk of cancer, such as enhanced liver cell regeneration and tissue sensitivity to chemical insults, altered gene expressions and regulation (cell proliferation, differentiation, DNA repair, and improper DNA methylation caused apoptosis), damaged DNA by mitochondrial dysfunction-induced oxidative stress, and increased liver cell apoptosis caused by activated protein kinase C-mediated cell-signaling cascade (Ueland, 2011).

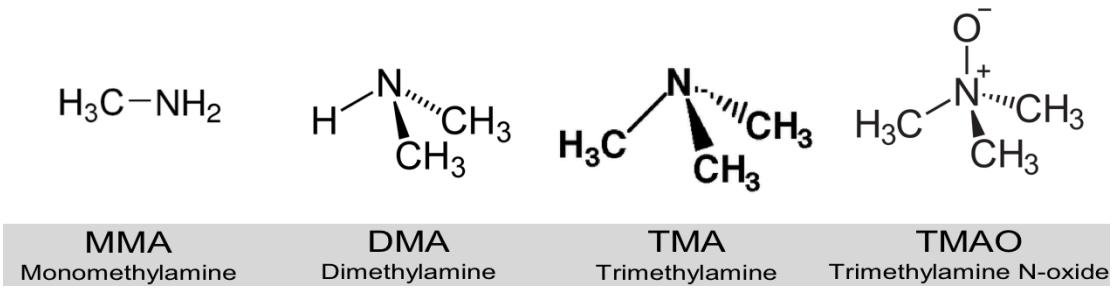
1.2.2 Gut microbial degradation of quaternary amines

Quaternary amines (QAs) play vital roles in both eukaryotic and prokaryotic organisms. They are characterized as various structures, involved in signaling, osmoregulatory and metabolic roles in human body. The most common QAs include choline, glycine betaine (GBT) and carnitine (Ikawa and Taylor, 1973; Landfald and Strøm, 1986; Graham and Wilkinson, 1992) (modified from the PhD thesis of Lidbury, 2015).

Choline and GBT are both important quaternary amine compounds, which can generate methylamines (MAs) anaerobically. They share a special feature that they both contain a tertiary amine group with a two-carbon (C2) side chain (acetyl group), which can be cleaved off during the degradation to trimethylamine (TMA) (King, 1984). Carnitine, another important QA, can be converted to TMA aerobically via the cleavage of the C4 side group by gut microbiota (Zhu *et al.*, 2014) (modified from the PhD thesis of Lidbury, 2015). The chemical structures of choline, glycine betaine, carnitine and methylamines are shown in **Fig 1.2**.



(A)



(B)

Figure 1.2 The chemical structures of **(A)** quaternary amines involved in this study: choline, glycine betaine and carnitine **(B)** methylated amines involved in this study: trimethylamine (TMA), dimethylamine (DMA), monomethylamine (MMA) and trimethylamine *N*-oxide (TMAO) (chemical structures modified from web source: http://www.wikiwand.com, free online open access).

Fig 1.3 summarizes the super metabolic pathways involved in the transformation of QAs to MAs by the gut microbiota or the host via the key functional enzymes, including choline-TMA lyase, which is the focus of this project (modified from Griffin *et al.*, 2015 and Fennema *et al.*, 2016). The key enzymes responsible for each step are shown in **Table 1.4** with corresponding EC numbers (http://www.brenda-enzymes.org). **Table 1.5** lists the representative microorganisms involved in QA metabolism to MA (Fennema *et al.*, 2016).

Choline conversion by gut microbiota to TMA has been reported as early as in the 1950s (Zeisel *et al.*, 1989; Chalmers *et al.*, 2006; Hayward and Stadtman, 1959). Craciun and Balskus (2012) have recently identified the choline-TMA lyase (CutC), which is a glycyl radical containing protein. CutC and associated proteins involved in choline metabolism by microorganisms is located on the so-called choline utilization gene cluster (*cut*) **(Fig. 1.3, Table 1.4 & Table 1.5)**. The essential role of *cutC* was unequivocally confirmed by the inactivation of *cutC* in *Desulfovibrio desulfuricans*, which completely abolished the ability of the bacterium to produce TMA from choline. A detailed bioinformatics analysis was carried out in **Chapter 4** to reveal *cutC* homologs in genome-sequenced gut bacteria.

It has also been reported that carnitine can be converted to TMA by gut bacteria, with the cleavage of the 3-hydroperoxybutyryl moiety (Meadows and Wargo, 2015). There are several candidate gut bacterial species, which may encode this pathway, for example, *Serratia marascescens* and *Acinetobacter calcoaceticus* (Meadows and Wargo, 2015). Zhu *et al.* (2014) has recently identified a two-component oxygenase/reductase Rieske-type enzyme encoded by *cntAB* using bioinformatics. This enzyme is responsible for catalyzing carnitine degradation to TMA **(Fig 1.3, Table 1.4 & Table 1.5)**.

TMAO can be produced from the oxygenation of TMA either via a bacterial TMA monooxygenase (Tmm) or by a class of flavin-containing monooxygenases (FMOs) in the human liver. The reaction uses NADPH as an electron donor and oxygen as the oxidizing agent (Chen *et al.*, 2011). TMAO can be reduced to TMA by bacterial TMAO reductase anaerobically (Kwan and Barrett, 1983) **(Fig 1.3 & Table 1.4).** It is thought that high amount of TMAO present in circulation may cause heart attack, stroke, arteriosclerosis or other cardiovascular disease although the underlying

biochemical mechanisms are not well-understood (Wang *et al.*, 2011).

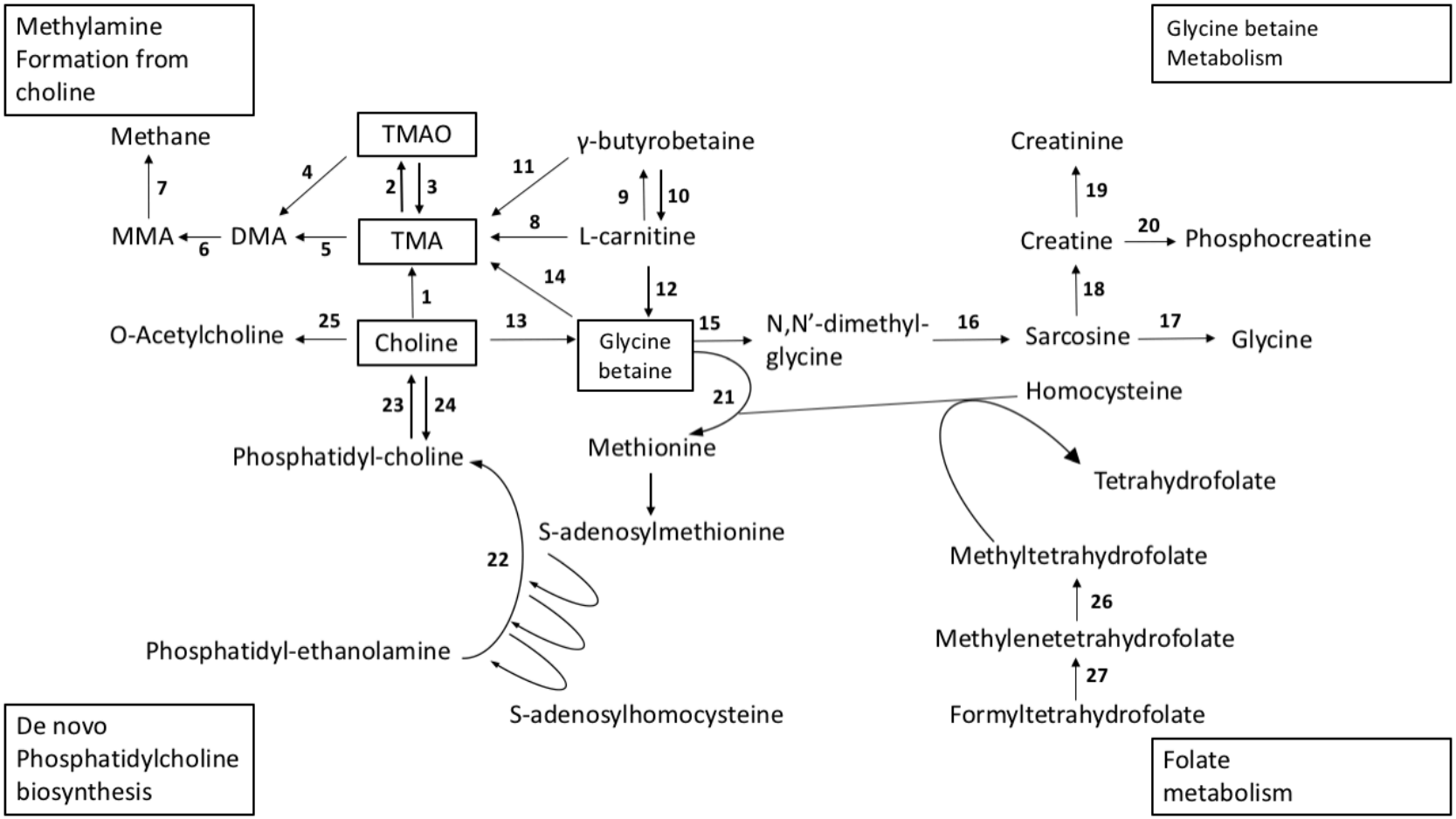


Figure 1.3 The super metabolic pathways involved in the degradation of quaternary amines (QA) to methylated amines (MA) by the gut microbiota or the host via the key functional enzymes including choline-TMA lyase (No.1), which is the focus of this project (modified from Griffin *et al.*, 2015 and Fennema *et al.*, 2016).

Table 1.4 The key enzymes involved in the degradation of quaternary amines to methylated amines associated with each pathway

No.	**Enzyme**	**EC number (**http://www.brenda-enzymes.org)
1	Choline-TMA lyase	EC 4.3.99.4
2	Flavin-containing TMA monooxygenase (FMO3)	EC 1.14.13.148
3	TMAO reductase	EC 1.7.2.3
4	Trimethylamine-oxide aldolase (TMAO demethylase)	EC 4.1.2.32
5	Trimethylamine dehydrogenase	EC 1.5.8.2
6	Dimethylamine dehydrogenase	EC 1.5.8.1
7	Methylamine dehydrogenase (amicyanin)	EC 1.4.9.1
8	Carnitine reductase/oxidase	/
9	γ -butyrobetainyl-CoA: carnitine-CoA transferase	EC 2.8.3.21
10	γ-butyrobetaine hydroxylase	EC 1.14.11.1
11	Carnitine-TMA lyase	/
12	L-carnitine dehydrogenase	EC 1.1.1.108
13	Choline dehydrogenase, betaine aldehyde dehydrogenase	EC 1.1.99.1, 1.2.1.8
14	betaine reductase	EC 1.21.4.4

15	Glycine betaine transmethylase	EC 2.1.1.5
16	Dimethylglycine dehydrogenase	EC 1.5.8.4
17	Sarcosine dehydrogenase	EC 1.5.8.3
18	Creatinase	EC 3.5.3.3
19	Creatininase	EC 3.5.2.10
20	Creatine kinase	EC 2.7.3.2
21	Betaine-homocysteine methyl transferase	EC 2.1.1.5
22	Phosphatidylethanolamine methyltransferase	EC 2.1.1.17, 2.1.1.71
23	Phospholipase D	EC 3.1.4.4
24	Choline kinase	EC 2.7.1.32
25	Choline O-acetyltransferase	EC 2.3.1.6
26	Methylene-tetrahydrofolate reductase	EC 1.5.1.20
27	Methylene-tetrahydrofolate dehydrogenase	EC 1.5.1.5

Table 1.5 Representative microorganism involved in QA metabolism to TMA (modified from Fennema *et al.*, 2016)

Phylum	Genus or Species	Reference
Choline		
Actinobacteria	*Mobiluncus*	Cruden and Galask, 1988
	Olsenella	Craciun and Balskus, 2012; Martinez-del Campo *et al.*, 2015
Bacteroidetes	*Bacteroides*	Cruden and Galask, 1988
Firmicutes	*Anaerococcus*	Craciun and Balskus, 2012; Romano *et al.*, 2015
	Clostridium	Robinson *et al.*, 1952; Bradbeer, 1965; Fiebig and Gottschalk, 1983; Moller *et al.*, 1986; Craciun and Balskus, 2012; Martinez-del Campo *et al.*, 2015; Romano *et al.*, 2015
	Desulfitobacterium	Craciun and Balskus, 2012
	Enterococcus	Simenhoff *et al.*, 1976
	Streptococcus	Robinson *et al.*, 1952; Simenhoff *et al.*, 1976; Chao and Zeisel, 1990; Craciun and Balskus, 2012; Martinez-del Campo *et al.*, 2015
Proteobacteria	*Desulfovibrio*	Hayward and Stadtman, 1959, 1960; Baker *et al.*, 1962; Bradbeer, 1965; Fiebig and Gottschalk, 1983; Craciun and Balskus, 2012
	Edwardsiella	Romano *et al.*, 2015
	Enterobacter	Eddy, 1953; Craciun and Balskus, 2012
	Escherichia	Craciun and Balskus, 2012; Martinez-del Campo *et al.*, 2015; Romano *et al.*, 2015

	Klebsiella	Eddy, 1953; Craciun and Balskus, 2012; Kuka *et al.*, 2014; Kalnins *et al.*, 2015; Martinez-del Campo *et al.*, 2015
	Proteus	Seim *et al.*, 1982a; Craciun and Balskus, 2012; Kuka *et al.*, 2014; Martinez-del Campo *et al.*, 2015; Romano *et al.*, 2015
	Providencia	Craciun and Balskus, 2012; Romano *et al.*, 2015
	Pseudomonas	Robinson *et al.*, 1952; Kleber *et al.*, 1978
	Yokenella	Craciun and Balskus, 2012
Carnitine		
Proteobacteria	*Acinetobacter*	Kleber *et al.*, 1977; Seim *et al.*, 1982b; Miura-Fraboni *et al.*, 1982; Ditullio *et al.*, 1994; Zhu *et al.*, 2014
	Citrobacter	Zhu *et al.*, 2014
	Escherichia	Zhu *et al.*, 2014
	Klebsiella	Kuka *et al.*, 2014; Zhu *et al.*, 2014
	Proteus	Seim *et al.*, 1982a
	Pseudomonas	Kleber *et al.*, 1978; Miura-Fraboni *et al.*, 1982
Glycine Betaine		
Firmicutes	*Clostridium*	Naumann *et al.*, 1983; Moller *et al.*, 1986
	Eubacterium	Zindel *et al.*, 1988; Hormann and Andreesen, 1989
	Sporomusa	Moller *et al.*, 1986; Hormann and Andreesen, 1989
TMAO		

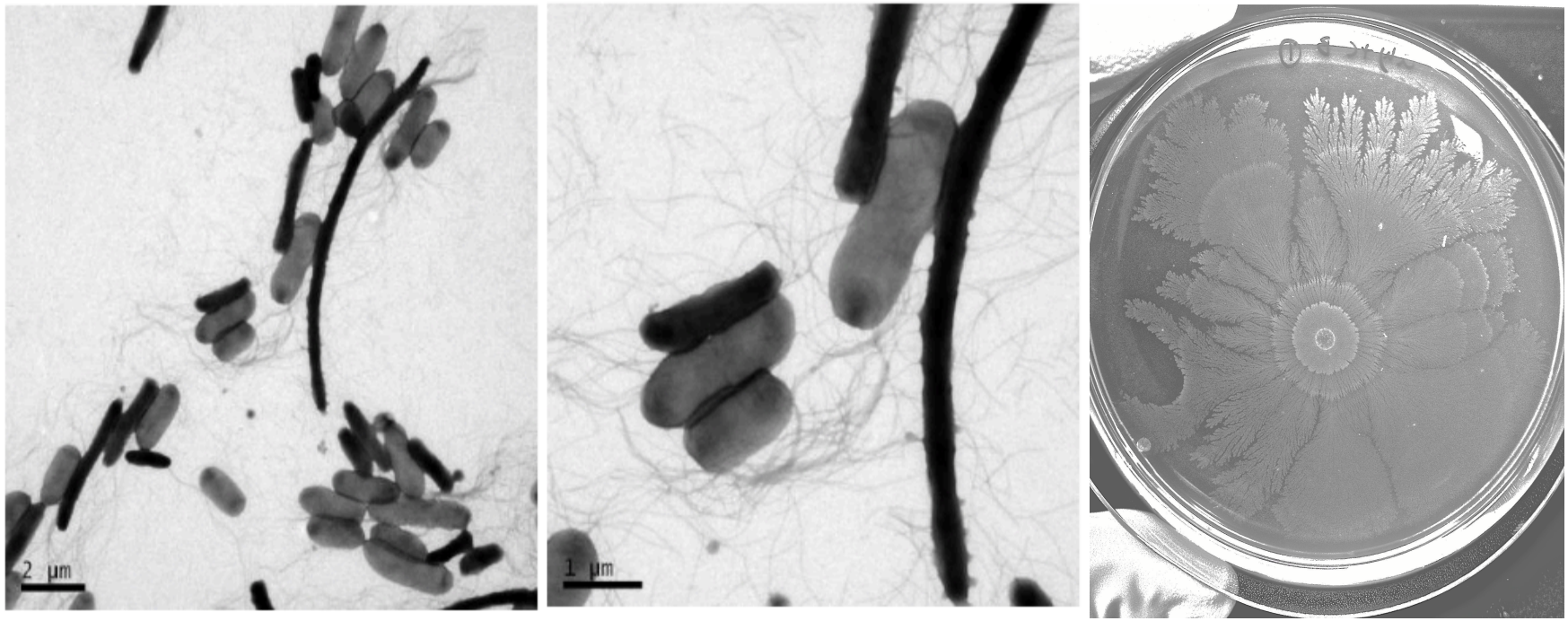
Actinobacteria	*Micrococcus*	Robinson *et al.*, 1952
	Mobiluncus	Cruden and Galask, 1988
Firmicutes	*Bacillus*	Robinson *et al.*, 1952
	Clostridium	Robinson *et al.*, 1952
	Staphylococcus	Robinson *et al.*, 1952
	Sarcina	Robinson *et al.*, 1952
	Streptococcus	Robinson *et al.*, 1952
Proteobacteria	*Alcaligenes*	Robinson *et al.*, 1952
	Campylobacter	Sellars *et al.*, 2002
	Citrobacter	Lin and Hurng, 1989
	Escherichia	Robinson *et al.*, 1952; Ishimoto and Shimokawa, 1978; Cox and Knight, 1981; Takagi *et al.*, 1981; Easter *et al.*, 1982; Lin and Hurng, 1989; Denby *et al.*, 2015
	Proteus	Robinson *et al.*, 1952; Strom *et al.*, 1979; Stenberg *et al.*, 1982
	Pseudomonas	Robinson *et al.*, 1952; Lee *et al.*, 1977; Easter *et al.*, 1982; Chen *et al.*, 2011
Ergothioneine		
Proteobacteria	*Alcaligenes*	Yanasugondha and Appleman, 1957; Kelly and Appleman, 1961
	Escherichia	Wolff, 1962

Although it has been known for over a century that anaerobic degradation of choline can produce TMA, the underpinning molecular and biochemical mechanisms remained largely un-established when the PhD project started. While my investigation was on-going, research carried out at Harvard University has revealed a gene cluster responsible for anaerobic choline utilization by a sulfate-reducing bacterium *Desulfovibrio desulfuricans* (Craciun and Balskus, 2012), which has been previously reported to metabolize choline to TMA (Chao and Zeisel, 1990). This gene cluster was analysed by bioinformatics and found to contain a set of genes to encode a glycyl radical enzyme encoded by *cutC* and a glycyl activating protein encoded by *cutD*. The function of the gene cluster in *D. desulfuricans* was confirmed by genetic knockout strategy. The *cutC* gene was disrupted by transposon mutagenesis and the generated mutant strain was found unable to degrade choline any more. In addition, the *cutC* and *cutD* genes were heterologously expressed in *Escherichia coli,* which enabled *E. coli* to utilize choline and produce TMA. The mechanism of choline degradation by this novel enzyme was studied by electron paramagnetic resonance (EPR) spectroscopy and it was found that *cutC* belongs to the member of C-N bond cleaving glycyl radical enzyme (GRE) family (Craciun and Balskus, 2012).

In contrast, in this PhD project, I started the investigation of choline metabolism in a model gut microbiota organism, namely *Proteus mirabilis*, which is known to be capable of producing TMA from choline (Sandhu and Chase, 1986). In order to identify the key genes involved in choline-to-TMA transformation, a comparative transcriptomics method was employed **(See Chapter 4)**.

1.3 *Proteus mirabilis*

In this study, the Gram-negative enteric *Proteus mirabilis*, a bacterium capable of colonizing the gut and the urogenital tract, was used as a model to investigate the role of choline metabolism. *P. mirabilis* is a facultative anaerobic, rod shaped bacterium which is found in the gut microbiota in healthy individuals. It is also widely distributed in soil and water (Burall *et al.*, 2004). The name of *P. mirabilis* is suggested from its typical swarming motility, which refers to the flagellum-dependent movement across a surface **(Fig 1.4)**, in contrast to swimming through liquid or soft agar. This form of motility is thought to allow *P. mirabilis* to migrate across catheters, gaining entry to the urinary tract and causes infection (Armbruster and Mobley, 2012).



(A) **(B)** **(C)**

Figure 1.4 (A&B) Swimming and swarming cells of *P. mirabilis* under transmission electron microscope, rod shaped cells are swimming cells and elongated cells with flagella are swarming cells. **(C)** Swarming pattern of *P. mirabilis* on solid agar plate (images taken during the course of this study, TEM pictures obtained in collaboration with Dr Stefanie Frank from University of Kent).

P. mirabilis can causes severe infections in humans. The most commonly infected body site is the urinary tract, which is characterised by an alkaline urine with a very distinct fishy odor (the smell of TMA). *P. mirabilis* can be diagnosed in the lab according to its characteristic swarming motility, and the incapability to metabolize lactose using a MacConkey agar plate assay. If left untreated, the increased alkalinity can lead to the formation of calcium carbonate crystals of struvite, which can subsequently lead to kidney stones. Once the stones develop and if they are allowed to grow large enough, it will cause obstruction and renal failure. Although rare, some *Proteus* species can also cause septicaemia, wound infections and pneumonia (Armbruster and Mobley, 2012).

P. mirabilis is also known to have probably the ability for quaternary amine metabolism. It has been reported, in the 1980s, that *Proteus mirabilis* can also produce TMA from choline (Sandhu and Chase, 1986). It has been suggested that the increasing concentration of TMAO in the plasma, which results from dietary intake of choline or carnitine followed by gut microbial degradation to TMA and subsequent oxygenation to TMAO by FMO3 in the liver, increases the risk of cardiovascular disease (Wang *et al.*, 2011; Bennett *et al.*, 2013; Koeth *et al.*, 2013; Tang *et al.*, 2013; Obeid *et al.*, 2016). Therefore, I aim to use this strain as the model organism in this study to explore molecular genetics and biochemistry underpinning choline degradation to TMA.

1.4 Comparative transcriptomics RNA-Seq analysis

Transcriptome sequencing technologies have been developed and widely adopted over the last decades to facilitate the genome-wide analyses of gene expression in various conditions. Microbial transcriptome and meta-transcriptome information can be used for predicting resistance to specific antibiotics, understanding host-pathogen immune interactions, tracking disease progression and quantifying gene expression changes during microbial metabolism (Han *et al.*, 2015). Of particular interest is the RNA-Seq technique, which was significantly advanced thanks to the improvement and reduced cost of next-generation sequencing platforms. There are many advantages of RNA-Seq enabled transcriptomic analyses compared to hybridization-based methods such as microarrays, which require gene-specific probes and large quantity of mRNA (Nookaew *et al.*, 2012). For example, RNA-Seq identifies common and novel transcripts in an unbiased strand-specific way. It allows for confident identification for both high and low expression of gene transcripts at a dynamic range, and makes it possible for streamlined workflow with multiple samples. It offers low-cost, large and complex datasets at single nucleotide resolution (Han *et al.*, 2015). However, the methods for data-mining, interpretation, and analysis of RNA-Seq for the metabolic mechanism is the most important part of transcriptomics. **Fig 1.5** illustrates the general workflow of RNA-Seq in bacteriology, including the key steps such as RNA extraction, mRNA sequence enrichment, qualification and quantification, cDNA library preparation, RNA-Seq alignment, read counts comparison, and finally gene expression and statistical analysis. **Table 1.6** lists several commonly used methods for transcriptomic data analyses.

In this project, RNA-Seq and the aforementioned bioinformatics tools are employed to identify the genes involved in choline-to-TMA transformation in the model bacterium, *Proteus mirabilis* **(Chapter 4)**.

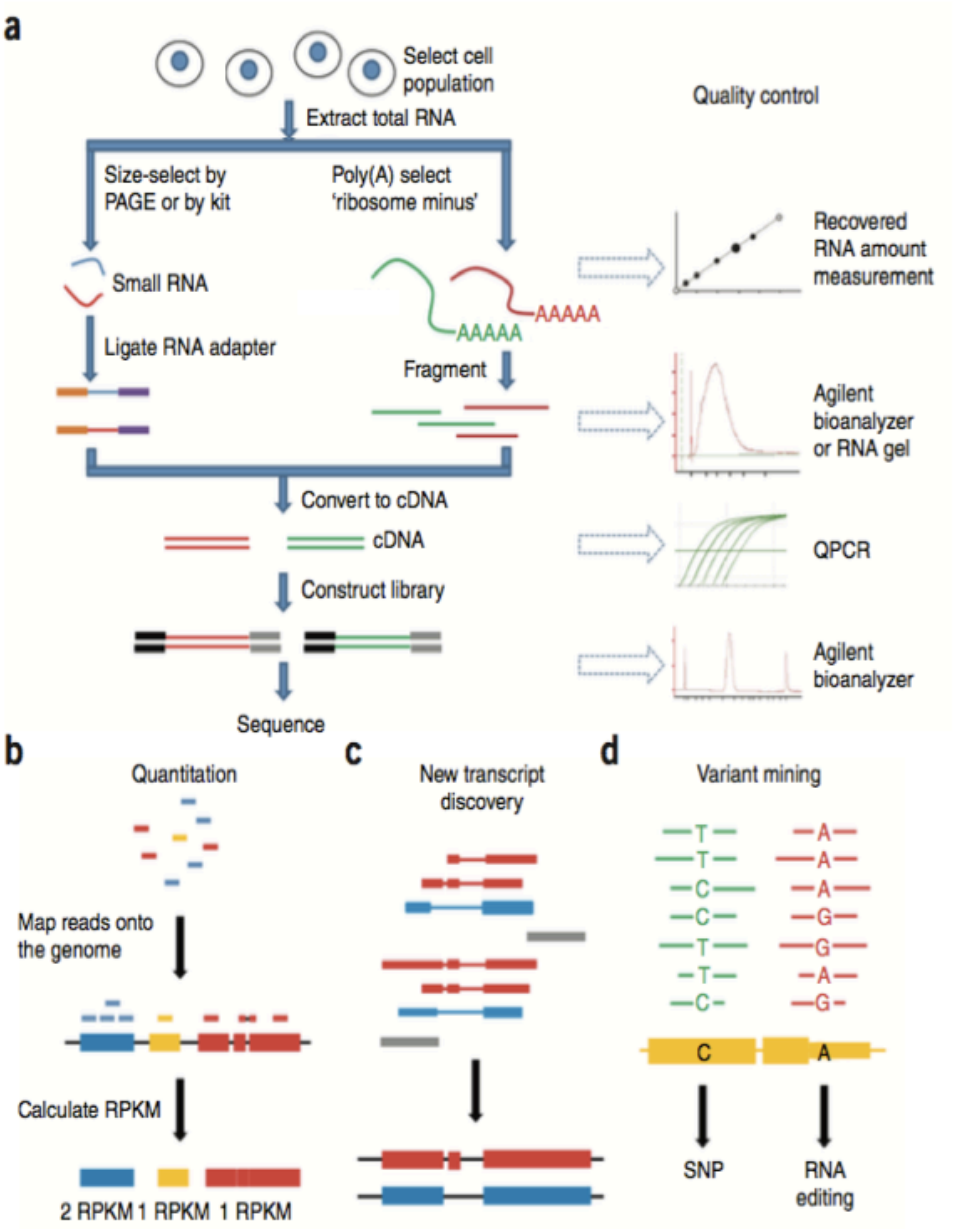

a
Select cell population
Extract total RNA
Size-select by PAGE or by kit
Poly(A) select 'ribosome minus'
Small RNA
Ligate RNA adapter
AAAAA
AAAAA
Fragment
Convert to cDNA
cDNA
Construct library
Sequence
Quality control
Recovered RNA amount measurement
Agilent bioanalyzer or RNA gel
QPCR
Agilent bioanalyzer
b
Quantitation
Map reads onto the genome
Calculate RPKM
2 RPKM 1 RPKM 1 RPKM
c
New transcript discovery
d
Variant mining
T T C C T T C
A A A G G A G
C A
SNP
RNA editing

Figure 1.5 The schematic workflow of bacterial RNA-Seq work. **(a)** RNA-Seq library construction. Firstly, 300,000 to 3 million cells are used to extract the total RNA. Secondly, mRNA enrichment is carried out to remove any rRNA thus only mRNA is retained. mRNA is fragmented into a uniform size distribution and RNA gel electrophoresis or Agilent Bioanalyzer is used to monitor the size and the quality of the fragment. Finally, cDNA library is built and Agilent Bioanalyzer is used to check the size distribution pattern of the library. **(b)** RNA-Seq alignment. Sequence read counts are aligned to the reference genome and splice junctions are mapped via mapping programs. The gene expression level is quantified as absolute read counts or normalized values such as RPKM (reads per kilo base per million mapped reads). **(c)** RNA-Seq *de novo* sequence assembly. RNA transcripts can be generated by assembling mapped reads. **(d)** RNA-Seq analysis. Transcriptome reads are compared with the reference genome, and the sequence variants are mined and identified as genomic variants (for example, SNPs) or candidates for RNA editing (picture adapted from Zeng and Mortazavi, 2012, copyright license number: 4123720380187).

Table 1.6 Transcriptomic analysis methods for RNA-Seq in microbial metabolism

Methods	**Description**	**Reference**
DGE analysis	Assess differences in transcript level across samples	Han *et al.*, 2015
RPKM method	Account for expression and normalized read counts with respect to overall mapped read number and gene length, but not sequencing depth, gene length, and isoforms abundance.	Han *et al.*, 2015
Expectation Maximization (RSEM)	Accurate estimates for gene and isoform expression levels and can be used even for species without a reference genome assembly	Han *et al.*, 2015
EdgeR function in the Degust Interactive Web-based RNA-Seq visualization software	Comparison of read counts between different conditions, to calculate log-fold values for count data so that comparison of two conditions can be visualized and significance can be tested statistically.	Zhao *et al.*, 2016

1.5 Bacterial microcompartments

During the metabolic activities of some bacteria, an unusual metabolic compartment called bacterial microcompartments (BMCs) are formed. These organelle-like structures are characterized by being enclosed with a protein shell rather than a lipid bilayer seen in eukaryotic organelles (Armbruster and Mobley, 2012; Davanloo *et al.*, 1984; Hampel *et al.*, 2014). The role of BMCs has been well documented and it is suggested that they are formed to optimize pathways with toxic or volatile intermediates. It has been revealed that the protein shell of the BMC usually acts as a diffusion barrier that helps channel an intermediate to the next pathway enzyme (Chowdhury *et al.*, 2014).

The most well-known BMC is the carboxysome, which is mostly found in cyanobacteria and some chemoautotrophic bacteria. It is reported that this 100 nm in diameter shell protein plays a role in the fixation of carbon dioxide (Iancu *et al.*, 2007; Schmid *et al.*, 2006; Tanaka *et al.*, 2008). Another important BMC is propanediol BMC, which has polyhedral shapes with proteinaceous structures within the cell. The metabolism associated with 1,2-propanediol is shown schematically below **(Fig 1.6)**. The faces of the structure are formed with the association of hexagonal shell proteins, which are the most abundant proteins in the metabolosome, such as PduA, PduB, and PduJ. The vertex of the assembly is formed from a pentameric shell protein, representing PduN (highlighted in red in **Fig 1.6**). **Table 1.7** summaries the most commonly reported BMC and their functional pathways.

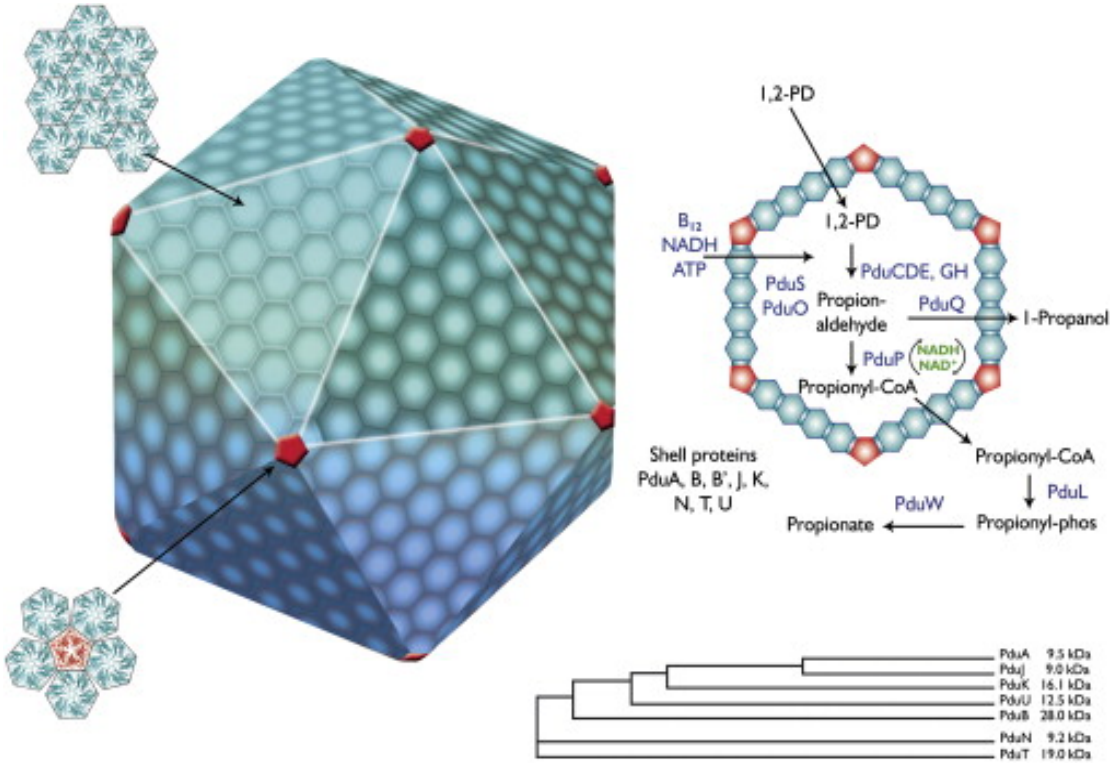


Figure 1.6 The polyhedral-shaped proteinaceous structures of propanediol BMC, with faces of the structure formed with the association of hexagonal shell proteins PduA, PduB, PduJ, PduK and PduU, and the vertex formed with pentameric shell protein PduN (highlighted in red). The propanediol BMC is associated with 1,2-propanediol utilization pathway (pictures adapted from Parsons *et al*., 2010, copyright license number 4123720753402).

Table 1.7 Metabolic pathways involving microcompartments and their functions (Yeates *et al.*, 2010)

Names	**Pathways**	**Descriptions**
Carboxysome	Carbon dioxide fixation	House both the Calvin cycle enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and carbonic anhydrase; capsules the inter-conversion of carbon dioxide and bicarbonate; concentrates carbon dioxide and overcome the well-documented inefficiency of RuBisCO.
Propanediol (Pdu) BMCs	1,2-propanediol utilization pathway	Organize enzymes for forming and consuming propionaldehyde within the protein shell; channel the aldehyde to the next enzyme to limit its cellular toxicity.
Ethanolamine (Eut) BMCs	Ethanolamine utilization pathway	Glutathione-based detoxification pathways and DNA polymerase repair functions; channel toxic or volatile intermediates
Pyruvate to ethanol	N/A	Metabolize Pyruvate to ethanol in *Vibrio furnissii* M1
Ethanol to acetate	N/A	Butyrate synthesis in *Clostridium kluyveri*
Glycerol	N/A	Metabolic pathway through a 3- hydroxylpropionaldehyde intermediate

One of the key findings from this PhD project was the identification and observation of BMC formation during choline-to-TMA transformation in the model bacterium, *Proteus mirabilis* **(Chapters 4 and 5)**. The proposed microcompartments in choline utilization in *P. mirabilis* comprise five different proteins in the *cut* gene cluster: PMI2714, PMI2718, PMI2720, PMI2721 and PMI2722 **(See Chapter 4)**. In the RNA-Seq experiment shown in **Chapter 4**, it was found that these microcompartment shell proteins were heavily induced by choline. The role of microcompartments is hypothesized either to help retain the volatile intermediates such as aldehyde and carbon dioxide or to protect the cell from the potential toxic effects of the intermediates (e.g. acetaldehyde) (Hampel *et al.*, 2014). Since one of the products of choline metabolism is acetaldehyde, I hypothesized that the microcompartments are essential in preventing the release of this toxic compound into the cell.

In an attempt to investigate the role of each of the five microcompartment genes and to identify which are essential for the assembly of the shell, these genes were amplified and introduced into *E. coli* for overexpression **(See Chapter 5)**.

1.6 *Caenorhabditis elegans*

Caenorhabditis elegans, also referred here as nematode or worm, was first described as a free-living organism in the 1900s and classified in the *Nematoda* phylum (Riddle *et al.*, 1997). It feeds on bacteria and is usually found in soil or on vegetation of temperate zones (Altun and Hall, 2017).

1.6.1 *C. elegans* anatomy and life cycle

As can be seen from **Fig 1.7,** the length of an adult *C. elegans* is ~1 mm (Altun and Hall, 2017) and its body can be divided in three parts: the head, the mid-body and the tail. The head of *C. elegans* also contains three parts: the grinder (to crush food), the pharynx (to push food into the grinder), and the pharyngeal-intestinal valve (to allow the passage of crushed food into the intestine) (McGhee, 2007). The mid-body section is the longest part of *C. elegans* body. It is surrounded by the hypodermis that secretes cuticles functioning as exoskeleton (Altun and Hall, 2017). The intestine is the most important place, where degradation of bacteria, absorption and storage of nutrients occurs (McGhee 2007). The gonads, the pseudocoelom (or body cavity) and the muscles present beneath the hypodermis (Altun and Hall, 2017). The last part is the tail, which contains the anus where defecation occurs (Altun and Hall, 2017).

The most significant character of *C. elegans* is its predominant population of hermaphrodites, which produce both sperm and oocytes and lay fertilized eggs in which early embryogenesis takes place **(Fig 1.7)**. The life cycle of nematodes is usually divided into 4 stages, starting after embryo-hatching, it undergoes four moulting states (larval stages named L1 to L4) before becoming an adult able to lay eggs (Riddle *et al.*, 1997; Altun and Hall, 2017). Besides, the nematodes are very sensitive to the environment such as temperature, food and growth conditions. The change of temperature could significantly affect the life cycle of *C. elegans* - it takes ~2.5 days at 25 °C, ~3 days at 20 °C or ~6 days at 15 °C to complete the entire life cycle (Byerly *et al.*, 1976) **(Fig 1.8)**. Interestingly, under food deficiency, a special form of nematodes, named dauer, will develop directly after L1 larval and go through a completely different life cycle to L4 larval, skipping the L2 and L3 stages **(Fig 1.8)**.

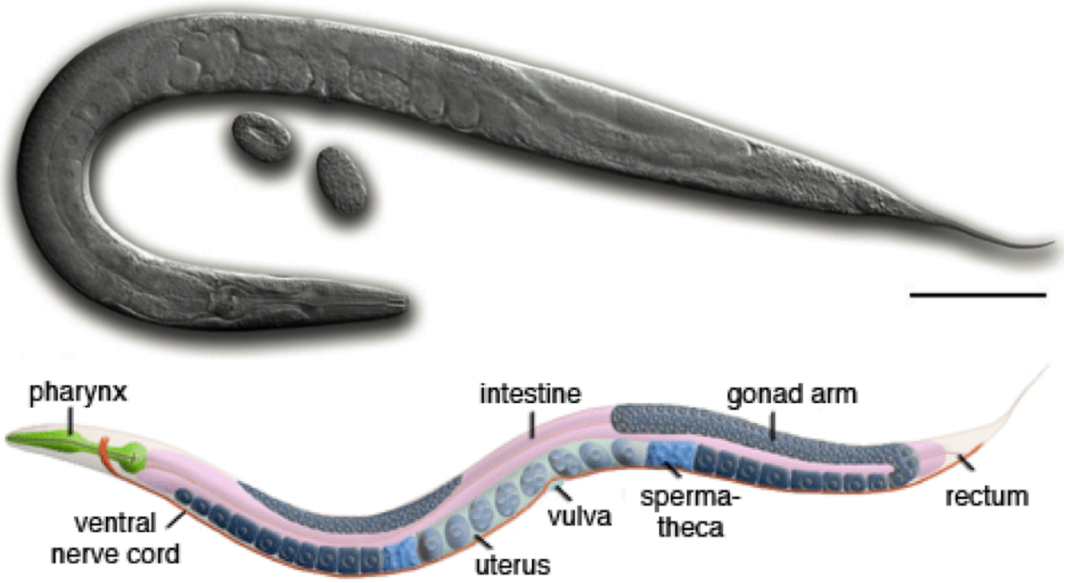


Figure 1.7 Anatomy of an adult *C. elegans*. Upper: Image of *C. elegans* obtained by differential interference contrast microscopy, left lateral side, scale bar 0.1mm. Lower: Schematic drawing of *C. elegans* anatomical structures, left lateral side (picture modified from Altun and Hall, 2017, free online open access).

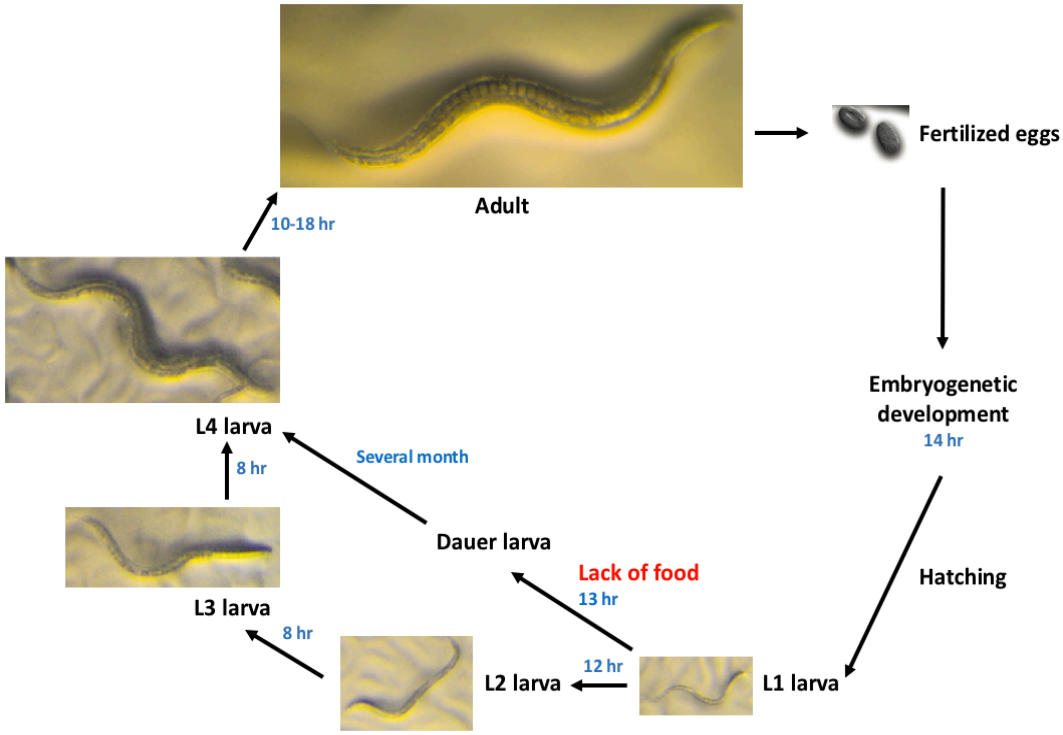


Figure 1.8 Schematic representation of *C. elegans* life cycle at 22°C. Numbers in blue indicates the time required for each stage and L1, L2, L3 and L4 represent the nematode larval stages (picture modified from Altum and Hall, 2017 and images taken during the course of this study).

1.6.2 Importance of *C. elegans* as a model organism for studying host-bacteria interactions

C. elegans was first employed as a model organism by Sydney Brenner who used the N2 (Bristol) strain to study its development and nervous system (Brenner, 1973; Brenner, 1974; Brenner, 1988). There are many important characteristics of *C. elegans* that set it apart from other animal models and make it a unique model organism for research: 1) easy maintenance. Nematodes can be stored stably for years and maintained in large numbers in the lab by feeding on an *E. coli* strain OP50, which is also cheap and straightforward to grow (Stiernagle, 2006); 2) high reproducibility. ~1000 offspring can be generated by nematodes every day, which are identical to the progenitor (Altun and Hall, 2017); 3) easy to study. It has a short life cycle and a transparent body (Altun and Hall, 2017), allowing study of its development under a microscope and the adoption of any florescent compounds (Irazoqui and Ausubel, 2010); 4) suitable for genetic manipulation as *C. elegans* genome was completely sequenced in 1998 and mutants can be readily made (Fay, 2013).

There are some similarities between *C. elegans* model and murine models of infection, for example, they both have innate immunity, however, *C. elegans* doesn't have an adaptive immune system, circulatory system and professional phagocytes (Sifri *et al.*, 2005) and cannot be reared at 37°C, at which temperature some virulence factors are expressed (Wurtzel *et al.*, 2012). Despite dissimilarities, a lot of evidence have shown the advantage of *C. elegans* model over mouse model for host-bacteria metabolism. **Table 1.8** summarizes key studies using *C. elegans* to investigate bacteria-host interactions.

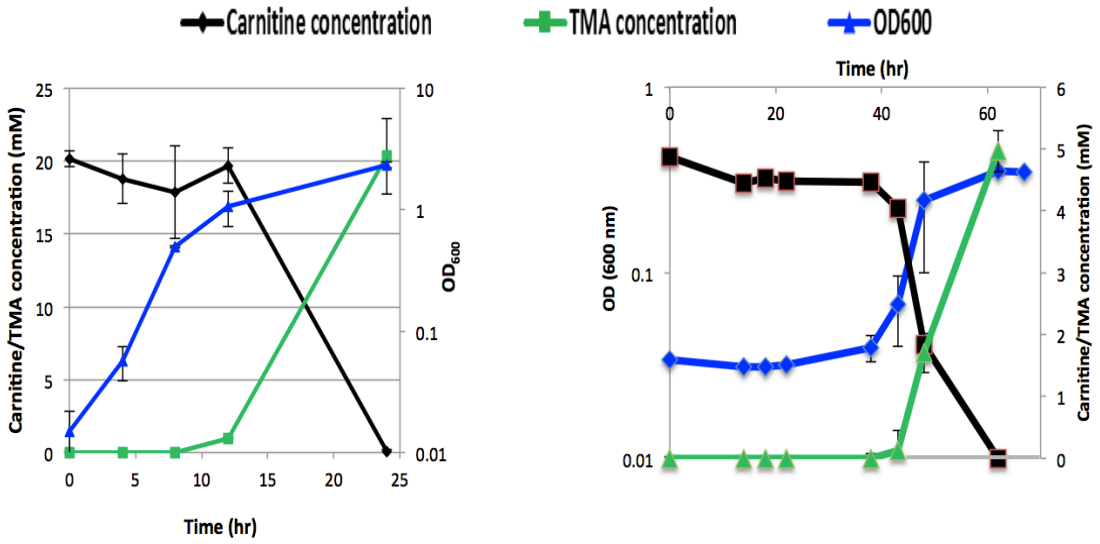
Table 1.8 The applications of *C. elegans* as a model organism to investigate host-bacteria interactions

Application	**Description**	**Reference**
Cell biology	Study programmed-cell death; signal transduction	Metzstein *et al.*, 1998; Kornfeld, 1997
Neuroscience	Nervous and muscle systems development	Hekimi *et al.*, 1998
Molecular microbiology	Control of gene expression without genetic manipulation by RNA interference	Ahringer, 2006; Billi *et al.*, 2014; Fire *et al.*, 1998
Physiology	Sex determination; embryogenesis; ageing;	Marin and Baker, 1998; Hekimi *et al.*, 1998
Immunology	Employing innate immunity as defense against infection with conserved innate immune pathways compared to mammalian cells such as the p38 MAPK and the insulin-like pathways.	Ewbank 2006
Microorganism infections	Nematodes can be fed on different bacterial strains by growing worms on media prepared in the lab	Kurz and Ewbank, 2007; Sifri *et al.*, 2005
Host-pathogen interactions	Monitoring pathology by following nematode survival; changes in behavior; changes in appearance; accumulation of bacteria in the intestine; monitoring gene expression via qRT-PCR or reporter transgenic systems	Darby, 2005; Tan *et al.*, 1999; Zhang *et al.*, 2005; Hodgkin *et al.*, 2000; Aballay *et al.*, 2000; Troemel *et al.,* 2006; Irazoqui *et al.,* 2008

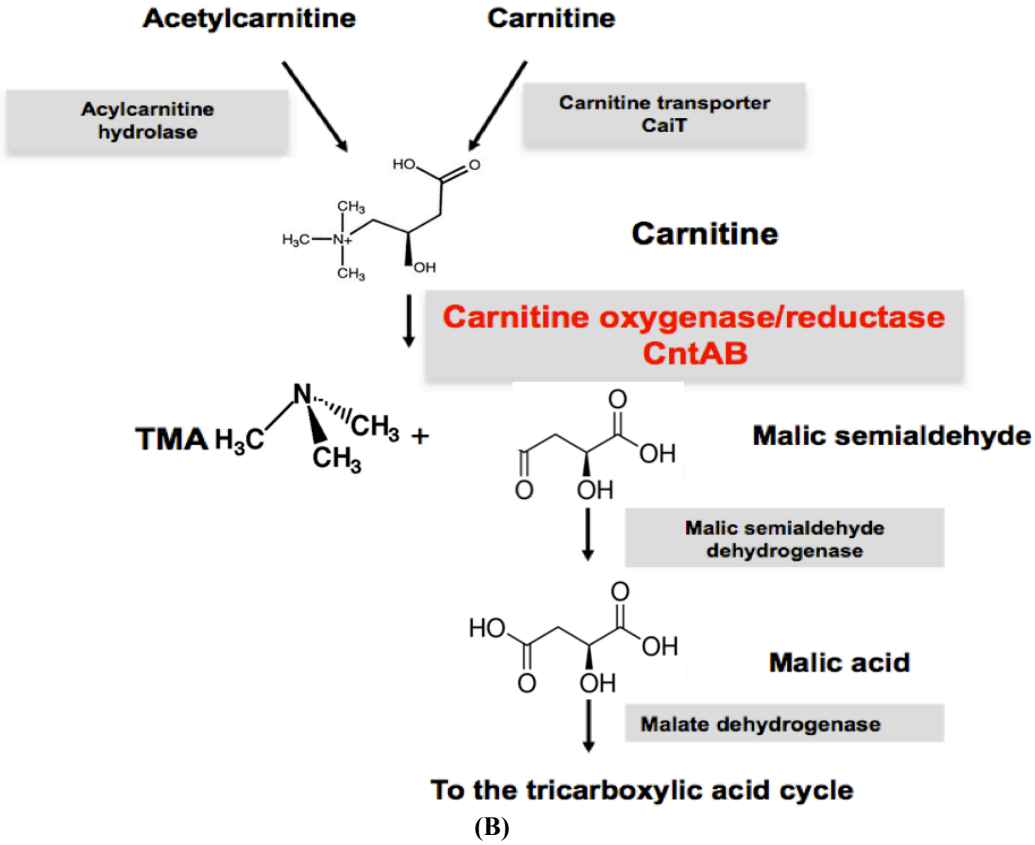
1.6.3 Use of *C. elegans* to study *Acinetobacter baumanni* infection

As I have discussed earlier, dietary intake of carnitine can promote cardiovascular diseases in humans through microbial production of TMA and its subsequent oxidation to TMAO by hepatic flavin-containing monooxygenases (Koeth *et al.*, 2013; Benett *et al.*, 2013; Wang *et al.*, 2011; Tang *et al.*, 2013). Previously studies in the lab have shown that both *A. baumanni* and *Escherichi coli* SE11 are excellent bacterial models for the investigation of biochemical pathways and mechanisms of carnitine degradation to TMA **(Fig 1.9A&B)**. Using bioinformatics approaches, Zhu *et al.* first identified a two-component Rieske-type oxygenase/reductase (CntAB) and associated gene cluster proposed to be involved in carnitine metabolism in these bacteria. CntA belongs to a group of previously uncharacterized Rieske-type proteins and has an unusual "bridging" glutamate but not the aspartate residue, which is believed to facilitate inter-subunit electron transfer between the Rieske centre and the catalytic mononuclear iron centre. Using *Acinetobacter baumannii* as the model, *cntAB* is confirmed to have an essential role in carnitine degradation to TMA. Heterologous overexpression of *cntAB* enables *Escherichia coli* to produce TMA, confirming that these genes are sufficient in TMA formation. Site-directed mutagenesis experiments have confirmed that this unusual "bridging glutamate" residue in CntA is essential in catalysis and neither mutant (E205D, E205A) is able to produce TMA **(Fig 1.10)**.

In an attempt to reveal the role of TMA formation in *A. baumannii* and *E. coli* infection, *C. elegans* was chosen as the model to compare carnitine degradation of the wild type and mutants in TMA formation and the subsequent impact on the life span of the worm. This offers the opportunity to better understand the human gut microbiota and their dynamics in human health and disease in large-scale epidemiological and dietary intervention studies.



(A)



(B)

Figure 1.9 (A) The quantification of carnitine degradation and TMA formation in *A. baumanni* (left) and *E. coli* SE11 (right) **(B)** The predicted pathway of carnitine degradation to TMA via CntA/B (modified from Zhu *et al.*, 2014, freely available online through the PNAS open access option).

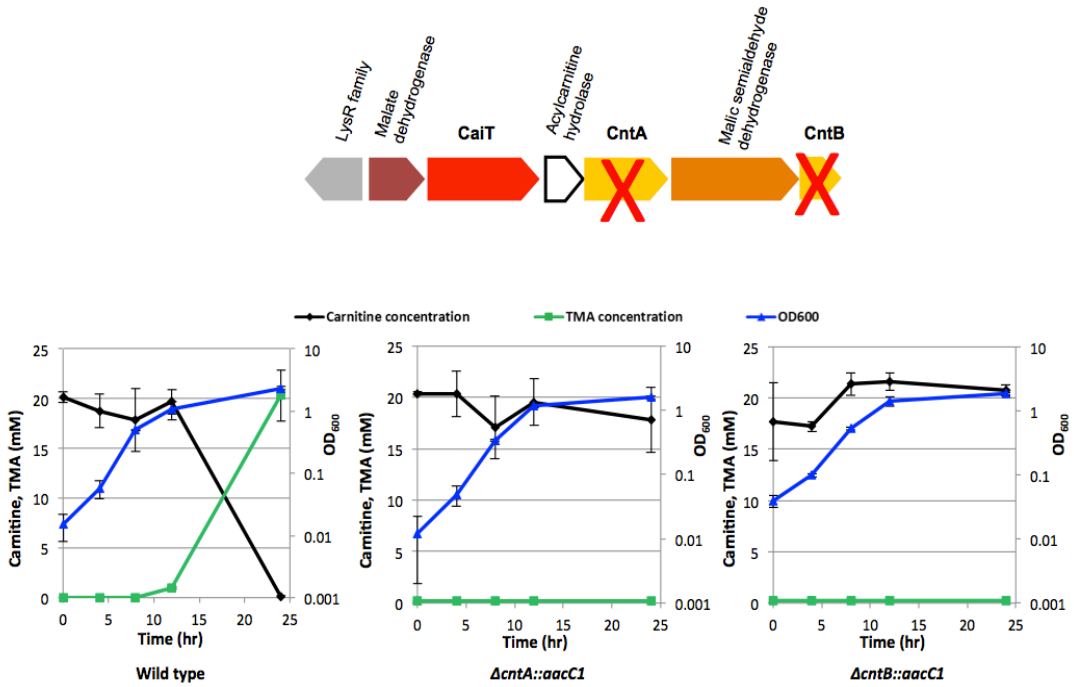


Figure 1.10 Carnitine degradation gene cluster of *cntAB* for *A. baumanni* and the knock-out genes are shown with red cross which create the mutants (ΔcntA::aacC1 and ΔcntB::aacC1, respectively) (upper); the quantification of carnitine and TMA for wild-type and the mutants in the culture medium supplemented with carnitine and succinate. The experiments run in triplicates and standard deviation was shown as error bars (lower) (modified from Zhu *et al.*, 2014, freely available online through the PNAS open access option).

1.7 Aims and hypotheses

The overall aim of this project is to establish the function of choline metabolism and TMA formation in representative gut microbiota. Specifically, I aim to

1) use RNA-Seq to identify the key genes responsible for choline degradation to TMA in *P. mirabilis* and confirm the essential role of choline-TMA lyase (*cutC*) in TMA formation (Chapter 4).
2) investigate the physiological role of choline degradation in *P. mirabilis* from the bacterial perspective (Chapter 5).
3) investigate the microcompartment protein (BMC) formation in choline degradation in *P. mirabilis* (Chapter 5).
4) investigate the impact of gut microbiota formation of TMA from dietary carnitine on the life span of the host using *C. elegans* as the model (Chapter 6).

Chapter 2

Materials and Methods

2.1 *Proteus mirabilis*

All reagents used in this thesis were prepared using Milli-Q H_2O and sterilized by autoclaving at 121 °C for 15 min or filter-sterilized through a 0.2 μm membrane (Millipore, Darmstadt, Germany). pH of the reagents was adjusted with the addition of NaOH (10 M) or HCl (10 M). The vitamins solution, containing biotin (0.4 mg l^{-1}), folic acid (0.4 mg l^{-1}), thiamine hydrochloride (1 mg l^{-1}), riboflavin (1 mg l^{-1}), nicotinic acid (1 mg l^{-1}), pantothenic acid (1 mg l^{-1}), 4-aminobenzoic acid (1 mg l^{-1}), lipoic acid (1 mg l^{-1}), pyridoxine hydrochloride (2 mg l^{-1}) and vitamin B_{12} (20 mg l^{-1}), was filter-sterilized using a 50 ml syringe.

2.1.1 Maintenance of *Proteus mirabilis* strains

All the maintenance and growth experiments of *Proteus mirabilis* strains were carried out by myself. All *P. mirabilis* strains listed in **Table 2.3** were stored as glycerol stocks (Luria-Bertani containing 10 % (v/v) glycerol) at -80 °C. For all experiments, cells were first revived from the glycerol stocks on non-swarming (see receipt below) agar plates (1.5 %, w/v). The plates were incubated at room temperature around 20 °C, and for short-term storage at 4 °C (See details described below in this section).

Growth experiments using *P. mirabilis* in this thesis were carried out anaerobically using a defined NH_4^+ medium containing (per litre) 7 g K_2HPO_4, 3 g KH_2PO_4, 0.1 g $MgSO_4 \cdot 7H_2O$, 1 g $(NH_4)_2SO_4$, 5 ml micronutrients (per litre: 3 g each L-histidine, L-tryptophan, L-nicotinamide and L-isoleucine), 50 μM ferric citrate, 50 mM sodium fumarate. The growth medium was supplemented with choline and other carbon sources (10 mM glucose, glycerol or choline).

For anaerobic growth in liquid, 90 ml liquid cultures were grown anaerobically in crimp-sealed 125 ml serum vials until OD_{600} >0.6 at 37 °C with no shaking. Starter cultures were inoculated by taking 1 ml of overnight *P. mirabilis* liquid culture grown in LB broth, and then injecting into the sealed serum vials. Each serum vial was de-gassed for 30 min with O_2-free N_2 gas to establish anaerobic conditions.

For growth on solid agar plates, aerobic non-swarming agar plates were prepared with 10 g tryptone, 5 g yeast extract, 0.5 g NaCl, and 5 ml glycerol and 1.5 % (w/v) agar. Anaerobic swarming agar plates were prepared using the same defined NH_4^+ medium described above solidified with 1.5 % (w/v) agar and supplemented with yeast extract (0.025 %, w/v).

2.1.2 Antibiotics used for *Proteus mirabilis* strains

All antibiotics used in this study were prepared by the in-house media preparation team at the University of Warwick. The following concentrations were required to isolate and maintain *P. mirabilis* mutant strains unless otherwise stated: ampicillin (150 μg ml^{-1}), kanamycin (150 μg ml^{-1}), and chloramphenicol (200 μg ml^{-1}).

2.1.3 Preparation and transformation of electrocompetent *P. mirabilis* cells

The method to prepare electrocompetent *P. mirabilis* cells was adapted from Visalli *et al.* (2003) and detailed in Jameson *et al* (2015). Briefly, LB broth without NaCl was used to grow *P. mirabilis* cells aerobically at 37 °C, until mid-log phase (OD_{600} of 0.3-0.6) before the cultures were harvested. Cells were washed four times with ice-cold, sterile 10 % (v/v) glycerol to remove any LB medium and then re-suspended in a final volume of 2 mL 10 % (v/v) glycerol. 50 μl aliquots were rapidly frozen in dry ice/ethanol and stored at −80 °C for a maximum for 4 months.

2.1.4 Construction and complementation of c*utC* mutant in *P. mirabilis*

The c*utC* mutant was constructed using the TargeTron targeted transposon mutagenesis system (Sigma-Aldrich Company Ltd., Dorset), which inserted a kanamycin-resistance gene into the *cutC* gene following the manufacturer's instructions. Non-swarming agar plates containing kanamycin (150 μg ml^{-1}) were used for the selection of c*utC::kan* mutant. The c*utC* complemented mutant was constructed by cloning the native *cutCD* (PMI2716/15) of *P. mirabilis* into pGEM-T. The resulting vector pGEM-T-*cutCD* was electroporated into the *P. mirabilis cutC::kan* mutant. The complimented mutant was then selected on non-swarming

agar plates containing ampicillin (150 μg ml^{-1}). All primers used for the construction and complementation of the *cutC::kan* mutant are shown in **Table 2.4**.

2.1.5 Growth of *P. mirabilis* on choline and quantification of choline and TMA

The aforementioned NH_4^+ defined medium was used to investigate the growth of *P. mirabilis* on choline. The degradation of choline and production of TMA were quantified using ion chromatography on an 881 Compact IC Pro (Metrohm, Herisau, Switzerland).

2.1.6 *P. mirabilis* swarming assays

For the swarming assays, the same NH_4^+ defined medium was used. The protocol was detailed in Jameson *et al* 2015. Briefly, 2 μl *P. mirabilis* culture was incubated on agar plates which were then left at 30 °C in an anaerobic growth cabinet (Don Whitley Scientific). Swarm radius was measured for triplicate plates supplemented with glycerol, choline or glycerol plus choline as the carbon sources.

2.1.7 Heterologous expression of *P. mirabilis* choline-TMA lyase CutC and CutD

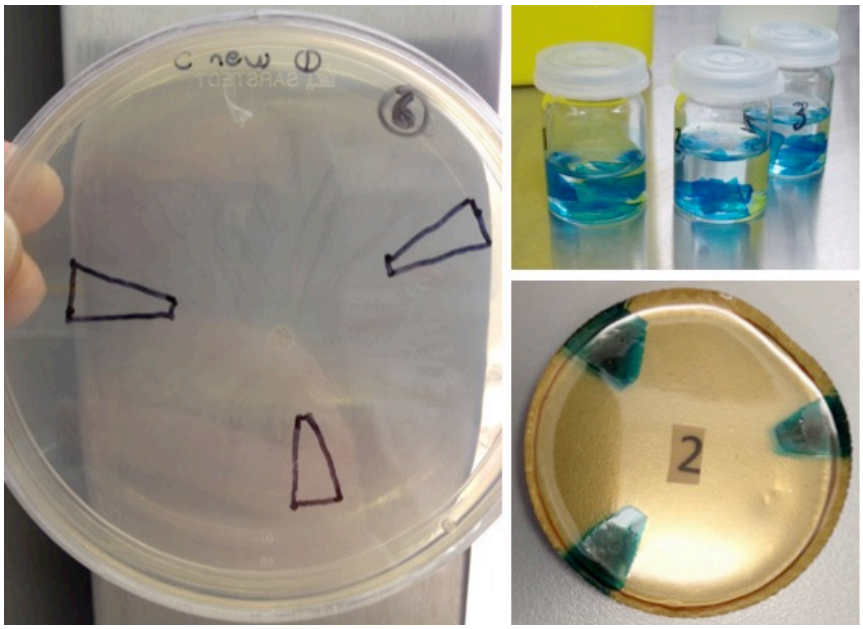
Codon-optimised *cutC* (PMI2716) and *cutD* (PMI2715) of *P. mirabilis* were ordered from GenScript (Piscatsaway, USA) to make sure that the putative *cutCD* genes are compatible for expression in *E. coli* before heterologous expression was carried out. The detailed cloning protocol was adapted from Jameson *et al.* (2015) using the expression vector pCOLADuet-1 (Novagen, Merck, Darmstadt, Germany). A modified M9 medium was used to grow the *E. coli* expressing *P. mirabilis* proteins anaerobically. The medium was supplemented with 40 mM $NaNO_3$, 50 μM $FeCl_3$, 30 mM glucose, yeast extract (0.05 %, w/v) and 400 mM NaCl. A high concentration of salt (400 mM) was used here in order to enhance the choline-uptake by *E. coli* via increased choline transporter *betT* expression (Lamark *et al.*, 1991). Expression of the recombinant protein was confirmed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of whole extract. Individual bands corresponding to the expected size of CutC and CutD were removed from the gel and analysed by

matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF).

2.1.8 Transmission electron microscopy (TEM) imaging of *P. mirabilis* microcompartments

To observe microcompartments by TEM, anaerobic cultures of *P. mirabilis* were grown in 80 ml NH_4^+ defined medium supplemented with glucose, choline or glucose plus choline. Liquid cultures were fixed for 14 hours by injection with fixative to a final concentration of 2.5 % (w/v) glutaraldehyde in 100 mM sodium cacodylate (pH 7.2). The cells were harvested by centrifugation at 4,000 rpm for 15 min and the cell pellet was re-suspended in 4 ml of the sodium cacodylate buffer. 2 ml of cells was centrifuged at 10,000 rpm for 2 mins, and then dehydrated twice in 100 mM sodium cacodylate (pH 7.2) for 5 mins, 1 % osmium tetroxide for 2 hours, three times in 100 mM sodium cacodylate (pH 7.2) for 5 mins, 50 % (v/v) ethanol for 15 mins and 70 % (v/v) ethanol overnight. The cells were further dehydrated in 90 % (v/v) ethanol for 15 mins, three times with 100 % ethanol for 15 mins, twice with propylene oxide for 15 mins, propylene oxide: agar low viscosity (LV) resin (1:1) for 30 mins, 2x LV resin for 90 mins. The samples were placed in 0.5 ml BEEM® capsules, centrifuged for 5 min at 900 × g to concentrate the cells to the tip and incubated at 60 °C for 24 hr to polymerize.

To observe the microcompartments of *P. mirabilis* grown on agar plates, agar plus bacteria were cut into a wedge shape (wider at the edge of the colony and narrower towards the middle of the colony) approximately 10 mm long, 5-6 mm (at wide end of wedge) and 2-3 mm (at narrow end of wedge). Three samples per plate were prepared. The wedge was cut with sloping sides-narrower at the surface of the agar and wider at the base of the agar so that the location of the bacteria relative to where they were on the plate was easy to identify **(Fig 2.1A)**. After the agar wedges were stained with 0.1 % alcian blue (w/v) in 0.1 % acetic acid (v/v) to make it easy for sample handling and processing as described above. Aluminium dish moulds were used to place the processed wedges with the *P. mirabilis* swarming culture facing down in LV resin. All the samples were incubated at 60 °C for 24 hr to polymerize **(Fig 2.1B)**.



(A) **(B)**

Figure 2.1 Swarming plates used for TEM sample preparation: **(A)** Agar plus bacteria of swarming cells were cut in a wedge shape, three samples per plate for each conditions, choline, glycerol and glycerol plus choline. **(B)** Agar wedges were stained with 0.1 % (w/v) alcian blue in 0.1 % (v/v) acetic acid processed for TEM, as described above.

2.1.9 Purification of *P. mirabilis* microcompartments

An attempt was made to purify microcompartments (BMCs) from choline-induced *P. mirabilis* culture. 2 L of liquid culture of *P. mirabilis* was grown on choline and glucose plus choline respectively until a certain cell density was achieved (OD_{600} 0.4-0.6). The BPER II purification method was carried out to purify the microcompartments (Sinha *et al.*, 2012). Briefly, cells were harvested by centrifugation and washed twice with 40 ml of buffer A (12.5 mM $MgCl_2$, 50 mM Tris-HCl [pH 8.0], 500 mM KCl, and 1.5 % 1,2-PD). After that, 2 g wet weight of cells were re-suspended in a mixture of 20 ml of buffer A and 30 ml of BPER-II containing 3 EDTA-free protease inhibitor tablets, 5 mM β-mercaptoethanol, 20 µl Benzonase and 50 mg of lysozyme. The re-suspended sample was incubated with 60 rpm of rotator shaking for 30 mins at room temperature, followed by 5 mins incubation on ice. Cell debris was removed by centrifugation twice at 12,000 × g for 5 mins at 4 °C. The remained microcompartment proteins were then centrifuged at 20,000 × g for 20 min at 4 °C. The pellet was washed once with a mixture of 8 ml of buffer A and 12 ml of BPER-II supplemented with 1 ml of buffer B (5 mM $MgCl_2$, 50 mM Tris-HCl [pH 8.0], 50 mM KCl, 1 % 1,2-PD) containing protease inhibitor. Finally, the remaining cell debris was removed by centrifugation three times at 12,000 × g for 1 min at 4 °C. The purified BMCs were stored at 4 °C until used.

2.2 *Escherichia coli*

2.2.1 Maintenance of *Escherichia coli*

For growth in liquid medium, Luria-Bertani (LB) medium: 10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 5 g l^{-1} NaCl, 40 µl 5 M NaOH (Sambrook *et al*., 2001) was used to cultivate *E. coli* strains, listed in **Table 2.3**. For growth on agar plates, Bacto agar (Difco) was added (1.5 % w/v). Liquid cultures were incubated at 37 °C in an orbital shaker (150 rpm). Stock cultures were prepared by adding 10 % (v/v) glycerol into LB culture and then stored at -80 °C.

There were four *E. coli* strains used throughout this thesis **(Table 2.3)**. *E. coli* JM109 (Sigma-Aldrich Company Ltd., Dorset) was used for cloning. *E. coli* BLR (DE3) pLysS (Novagen, Merck, Darmstadt, Germany) was used for protein overexpression.

E. coli SE11 was used to investigate carnitine degradation to TMA, which was cultivated either in LB medium or in the same defined medium for *A. baumannii* **(See Chapter 6)**. *E. coli* OP50 was used to maintain *C. elegans* (see below).

2.2.2 Antibiotics used for *Escherichia coli*

The antibiotics used for *E. coli* in this study and their appropriate concentrations were prepared as follows: kanamycin (25-50 μg ml^{-1}), gentamicin (10 μg ml^{-1}), spectinomycin (25 μg ml^{-1}), tetracycline (10 μg ml^{-1}), ampicillin (100 μg ml^{-1}), and streptomycin (50 μg ml^{-1}).

2.2.3 Transformation of chemically competent *E. coli* cells

E. coli JM109 high competency cells (Promega, Fitchburg, WI, USA) were used for all the routine cloning steps performed in this thesis. Transformation was carried out following the protocol from Sambrook *et al*. (Sambrook *et al*., 2001). Transformed cells were plated onto LB agar plates containing the appropriate antibiotic(s).

2.3 *Acinetobacter baumannii*

A. baumannii ATCC19606 wild-type and the mutants (ΔcntA::aacC1, ΔcntB::aacC1) were cultivated in either Luria-Bertani (LB) medium or M9 defined medium. M9 defined medium was prepared following the protocol from Zhu *et al.* (2012): NH_4Cl (1 g l^{-1}), NaCl (0.5 g l^{-1}), KH_2PO_4 (3 g l^{-1}), $Na_2HPO_4 \cdot 7H_2O$ (12.8 g l^{-1}), $MgSO_4 \cdot 7H_2O$ (0.5 g l^{-1}), $CaCl_2 \cdot 2H_2O$ (0.15 g l^{-1}), $Na_2MoO_4 \cdot 2H_2O$ (0.5 mg l^{-1}), $FeCl_3$ (50 μM), supplemented with a mix of the vitamins: biotin (0.4 mg l^{-1}), folic acid (0.4 mg l^{-1}), thiamine hydrochloride (1 mg l^{-1}), riboflavin (1 mg l^{-1}), nicotinic acid (1 mg l^{-1}), pantothenic acid (1 mg l^{-1}), 4-aminobenzoic acid (1 mg l^{-1}), lipoic acid (1 mg l^{-1}), pyridoxine hydrochloride (2 mg l^{-1}) and vitamin B_{12} (20 mg l^{-1}).

2.4 Extraction of nucleic acids

DNA extractions from bacteria in this study were carried out using the FastDNA® SPIN Kit for Soil (MP Biomedicals, LLC, CA, USA) according to the manufacturer's instructions. Normally, 1-10 ml of bacteria liquid culture was

centrifuged (8, 000 x g for 5 min) to generate a cell pellet before extraction was carried out.

RNA extraction and purification were performed following the protocol from Jameson *et al.* (2015). Anaerobic *P. mirabilis* cultures were grown in the NH_4^+ defined medium supplemented with glucose only or glucose plus choline, five replicates for each condition. Cells were grown until the OD_{600} reached 0.4 and were harvested by centrifugation. Total RNA was extracted using the TRI reagent (Sigma-Aldrich Company Ltd., Dorset). DNA was removed by treating with on-column DNAse (QIAgen, Crawley, UK). Total RNA was purified using the RNeasy Mini Kit (QIAgen, Crawley, UK) according to the manufacturer's protocols.

For plasmid mini-prep, 1.5-5 ml overnight *E. coli* culture containing the appropriate plasmid was used. Plasmid extraction was carried out using the QIAprep Miniprep Kit (QIAgen) or GeneJET kit (Fermentas) according to the manufacturer's protocols.

For *E. coli* carrying a low-copy number plasmid (*e.g.* pCOLADuet), the QIApep Midiprep Kit (QIAgen) was applied according to the manufacturer's instructions.

2.5 RNA-Seq

60 µl RNase-free water was used to re-suspend the purified RNA. Messenger RNA was enriched using the Ribo-Zero rRNA Removal Kit (Gram-negative Bacteria, Epicentre, Cambio, UK). An Agilent bioanalyser (Agilent, Edinburgh, UK) was used to assess the quality and quantity of RNA. ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre, Cambio, UK) was used to prepare complementary DNA libraries. Illumina HiSeq 2500 platform (Illumina, Little Chesterfield, UK) was then applied to run each sample of one hundred bp paired-end sequencing. FASTX Toolkit was used to trim and filter the raw Illumina RNA-Seq reads to 55 bp based on read quality. Only high-quality sequences, with Illumina base quality score of 20 for either 96 % or 100 % of bases, were retained. Finally, the RNA-Seq reads were aligned to the *P. mirabilis* HI4320 genome sequence via the TOPHAT program. A more detailed method for subsequent bioinformatics analysis of RNA-Seq data is presented in **Chapter 4**.

2.6 Nucleic acid manipulation techniques

An ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA) was used to quantify nucleic acids. A 1 % (w/v) agarose gel made with TBE buffer was used to check the integrity and purity of DNA.

PCR in this study was carried out in a total reaction volume of 50 µl. A T-100 thermocycler (Bio-rad Laboratories Inc., Hercules, CA, USA) was used to perform the reaction, with either KAPATaq (KAPA biosystems, Willimgton, MA, USA), Dreamtaq (Fermentas) or *Pfu* polymerase (Promega). Each component for a general PCR is shown in **Table 2.1** and a typical reaction procedure is shown in **Table 2.2**.

For colony PCR, DMSO (4 %, v/v) and BSA (0.04 %, w/v) were added to the reaction to help permeabilise cell membranes and stabilise enzymes respectively. Negative controls (with no template DNA) were always run in parallel to check for any DNA contamination. For overnight running, the reaction is usually set to be held at 4 °C.

Table 2.1 The amount of each component in a standard PCR mix

Component	**50 µl reaction**	**Final concentration**
Nuclease-free sterile dH_2O	Up to 50 µl	N/A
5 x Buffer	10 µl	1 x
$MgCl_2$ (usually present in the 5 x buffer)	N/A	N/A
dNTPs (10 mM each nucleotide)	1 µl	0.2 mM each nucleotide
10 µM forward primer	2 µl	0.4 µM
10 µM reverse primer	2 µl	0.4 µM
Template DNA	Dependent on the manufacturer's protocol	<0.5 µg/50 µl
DNA polymerase	Dependent on the manufacturer's supply	2.5 unites

Table 2.2 The general procedure of a standard PCR cycle

Step	**Temperature**	**Duration**	**Cycles**
Initial denaturation	95 °C	3 min (5 min for colony PCR)	1
Denaturation	95 °C	1 min	
Annealing	45-60 °C (dependent on primers)	30 sec	30
Elongation	72 °C	1 min/kb	
Final elongation	72 °C	5 min	1

DNA restriction digests were carried out using enzymes from Promega or Fermentas according to the manufacturer's recommendations.

DNA fragments from agarose gels were purified using QIAquick (QIAgen) or Nucleospin (Macherey-Nagel, Düren, Germany) gel extraction kits according to the manufacturers' instructions.

DNA ligations were routinely carried out using T4 DNA ligase (Promega, or Fermentas) with varying amounts of DNA and varying ratios of vector: insert, according to the general guidance of Molecular Cloning (Green & Sambrook).

PCR products were routinely cloned into a pGEM-T Easy vector (Promega) according to the manufacturer's instructions.

Sanger sequencing of PCR products or purified plasmids was carried out by GATC, Germany. DNA sequence analysis was carried out using SeqMan Pro, part of the Lasergene software suite (DNASTAR, Madison, WA, USA).

For electrophoresis of DNA, 0.5-2 % (w/v) agarose gels were made with 1 × TBE buffer, with the addition of ethidium bromide (0.5 μg ml^{-1}) prior to casting to visualize the DNA. DNA fragments were run and separated on the gel alongside 1kb plus DNA ladder (Fermentas) as marker to estimate the sizes of DNA fragments. Gels were visualised on a Gene Genius transilluminator (Syngene, Cambridge, UK).

2.7 Analytical methods

Cation-exchange ion chromatograph coupled with a conductivity detector was used in this study to quantify choline and TMA, retention times 18.34 min and 14.46 min respectively **(Fig 2.2)**. A Metrosep C4/250-mm separation column was used to separate methylated amines and quaternary amines from other cations.

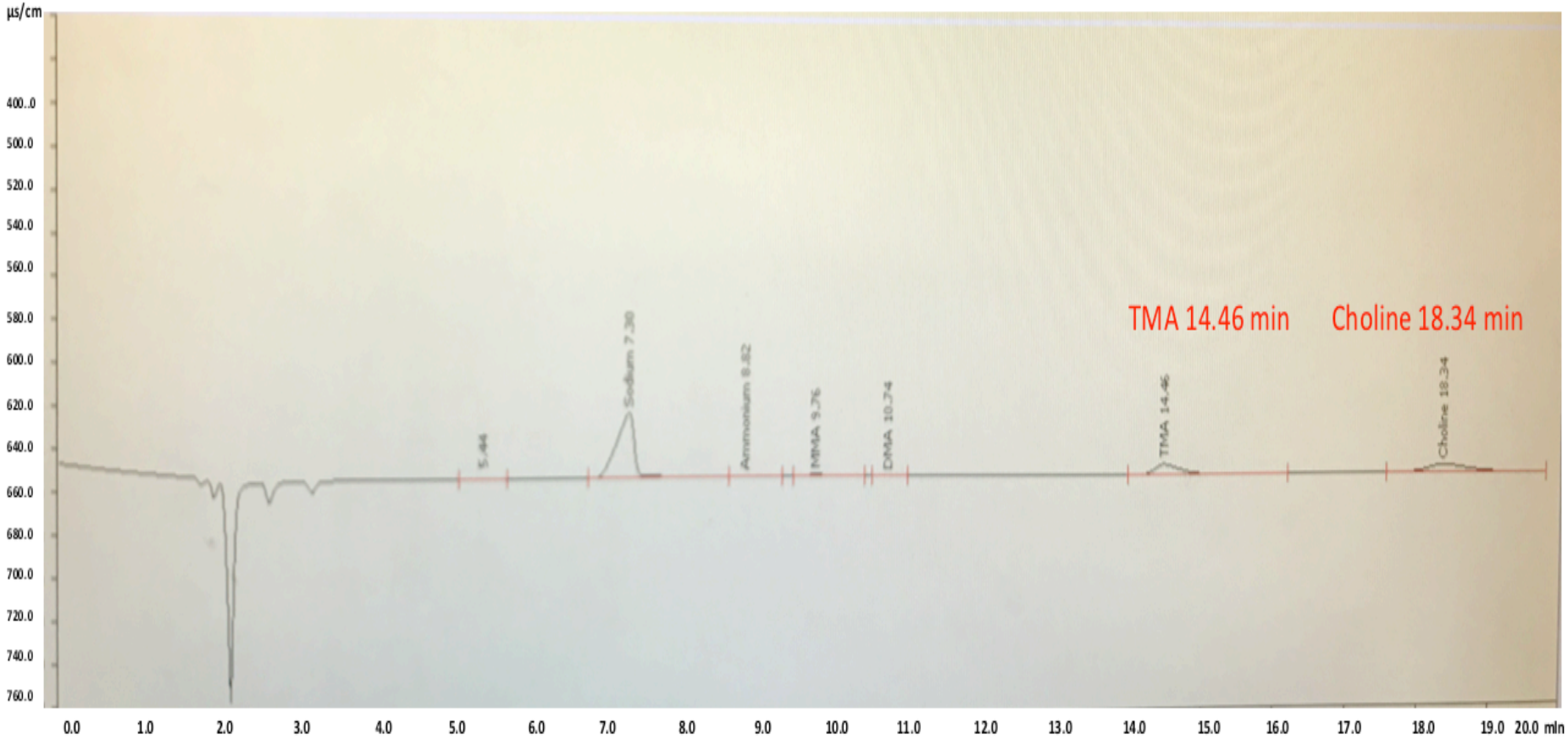


Figure 2.2 An example chromatograph showing the respective peaks and retention times for choline and TMA.

2.8 Protein manipulation methods

Bacterial cells were broken for protein work by passing through a French Press unit (American Instrument) three times at 110 megapascals at 4 °C.

A Bio-Rad Protein Assay Kit (Bio-Rad) was used to determine the total protein concentration, according to the manufacturer's instructions. A standard curve was plotted using bovine serum albumin (BSA, 0.25, 0.5, 0.75, 1 mg ml^{-1}).

Sodium dodecyl sulphate (SDS)-PAGE was applied in this study for routine analyses of purified proteins or over-expressed, IPTG-induced recombinant proteins, with PageRuler Plus pre-stained protein ladder (Fermentas) as a molecular mass reference.

2.9 Quantification of β-galactosidase activity (Miller assay)

Z-buffer stock solution (4.27 g Na_2HPO_4, 2.75 g $NaH_2PO_4 \cdot H_2O$, 0.375 g KCl, 0.125 g $MgSO_4 \cdot 7H_2O$) was prepared according to the protocol of Lidbury, 2015. 500 ml of stock solution was stored at 4 °C with pH adjusted to 7.0. Before use, 50 ml of Z-buffer was mixed with 0.14 ml ß-mercaptoethanol. Fresh O-nitrophenyl-ß-D-galactoside (ONPG) (4 mg ml^{-1}) was also prepared prior to use. 1 M Na_2CO_3 was prepared and stored at 4 °C.

Transconjugants, carrying the appropriate *lac* fusion plasmid, were stored on *P. mirabilis* non-swarming agar plates with spectinomycin (150 μg ml^{-1}). Fresh cells were prepared by re-streaking onto new agar a week prior to the enzyme assays. Colonies were picked and placed into M9 defined medium containing glucose and spectinomycin (150 μg ml^{-1}). This culture was then grown overnight and used as an inoculum (2.5 %, v/v). Cells were grown overnight in M9 with either choline or glucose (10 mM) as the sole carbon source. For these two groups, equal amounts of cells (10-30 ml, according to OD_{600}) were harvested (centrifugation for 10 min at 8,000 x g) and washed with Z Buffer (5 ml) prior to re-suspension in 1 ml Z buffer.

The enzyme assay was initiated by mixing cells with the artificial substrate OPNG, resulting in the formation of the yellow product O-nitrophenol, as can be seen from **Fig 2.3**. The following equation was used to calculate the promotor activity by converting into Miller Units:

$$\text{Activity} = \frac{OD_{420}}{OD_{650} \times \text{time} \times \text{vol}} \times 1000$$

Time = time of reaction (min)

Vol = volume of cells used (ml)

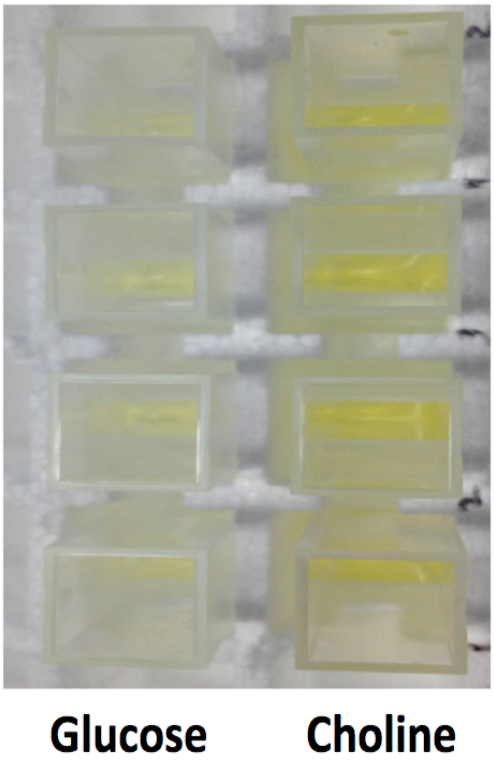


Figure 2.3 An example of the β-galactosidase activity assay, showing the choline-TMA lyase promoter activity of *P. mirabilis* in choline medium (right), in contrast to in glucose medium (left) **(See Chapter 5)**.

2.10 *Caenorhabditis elegans* life span assays

2.10.1 Maintenance of *C. elegans*

C. elegans N2 strain was obtained from Dr Andre Pires da Silva (University of Warwick, UK). *C. elegans* strain CF512 fer-15(b26) II; fem-1(hc17) IV was purchased from *Caenorhabditis* Genetic Center (CGC, University of Minnesota, USA). Modified nematode growth medium (NGM) plates were used to maintain nematodes: 3 g l^{-1} NaCl, 2.5 g l^{-1} bacteriological peptone, 25 mM potassium phosphate buffer (1 M KH_2PO_4, 1 M K_2HPO_4, pH= 6), 1 mM $CaCl_2$, 1 mM $MgSO_4$, 5 mg ml^{-1} cholesterol, 25 ug ml^{-1} nystatin, 17 g l^{-1} Bacto agar (Difco), and 50 ug ml^{-1} Streptomycin (Tan *et al.*, 1999). The plates were seeded with *Escherichia coli* OP50 strain as previously described by Stiernagle (2006). Briefly, 50 µl of *E. coli* OP50 overnight LB culture was seeded onto NGM agar. The plates were then incubated for 8 h at 37 °C or at room temperature for 24 h. OP50 plates were kept at 4 °C in the fridge until needed. *C. elegans* were then transferred to a new plate seeded with OP50 by transferring a chunk of agar from an old plate. Plates with nematodes were incubated at 16 °C for 4-7 days until the *E. coli* (nematode food) was almost finished. A new agar chunk was then transferred to seed a new plate.

2.10.2 Bleach synchronization of *C. elegans*

This method is applied to obtain age-synchronized *C. elegans* N2 or CF512 for the life span assay. The nematodes were washed off from 3-4 day old OP50 plates with M9 buffer (3 g l^{-1} of KH_2PO_4, 6 g l^{-1} of Na_2HPO_4, 5 g l^{-1} NaCl and 100 mM $MgSO_4$), and collected into a 50 ml falcon tube. The worms were then treated with 0.5 M NaOH and 1 % NaClO for 5 min. The aim of this step is to kill the nematodes at both adult and larval stages with only eggs intact, as reported by Stiernagle (2006). After that, the falcon tube was filled with PBS 10X (80 g l^{-1} of NaCl. 2 g l^{-1} of KCl. 14.4 g l^{-1} of Na_2HPO_4 2.4 g l^{-1} of KH_2PO_4, pH=7.4) with gentle shaking for a few seconds. This step dilutes and neutralizes the NaOH and bleach. Eggs were collected by centrifugation at 1,500 x g for 5 min and washed twice with the aforementioned M9 buffer. Eggs were then re-suspended with M9 buffer and transferred to a new 50 ml falcon tube and incubated for 16-24 h at 20 °C while spinning on a rotating

carousel at 10 rpm. During this time, the eggs hatched but remained at the first larval stage (L1), due to a lack of food, as reported by Stiernagle (2006). After incubation, L1 larvae were gained by spinning at 1,000 *g* for 5 min and removing the top layer of medium. L1 larvae were then plated on NGM plates seeded with OP50 for 3 days at 20 °C to reach the young adult stage.

2.10.3 The role of carnitine degradation on *C. elegans* life span

2.10.3.1 Carnitine degradation in *Acinetobactor baumannii* on *C. elegans* life span

All work on *A. baumannii* described in this thesis was carried out by myself in the laboratory of Dr Kumar Rajakumar at The University of Leicester. As shown in **Fig 2.4**, *A. baumannii* wild type, *cntA* and *cntB* mutant strains were recovered from frozen stocks by streaking onto LB plates and incubating overnight at 37 °C. One colony of each strain was cultivated in 5 mL LB for 16 h at 37 °C with shaking (200 rpm). 1 % (v/v) LB inoculum of wild type, the *cntA* and the *cntB* mutants were added to a defined M9 medium supplemented with 20 mM of carnitine or 20 mM of succinate with gentamycin (80 μg ml^{-1}), incubated at 37 °C while shaking at 200 rpm until OD_{600} reached 0.6. The same amount of cells of wild type or mutants were collected and washed with NH_4^+ free M9 medium, containing carnitine or succinate. 100 μl of M9-washed culture was spotted on NH_4^+ free M9 medium agar plates supplemented with carnitine or succinate without gentamycin. The plates were incubated overnight at room temperature before use.

Ten age-synchronized adult nematodes were manually transferred onto a plate containing *A. baumannii* wild type, *cntA*, or *cntB* mutant. Ten plates were used for each bacterial strain, resulting in 100 worms in total for each experiment. The nematode plates were observed at 20 °C every day and the numbers of nematodes found alive, dead or censored were recorded. 'Censored' worms indicate those that had dried on a plate wall, bagged with progenies or missing during the life span experiment **(Fig 2.5)**. Nematodes were observed as dead if they did not react or move when touched. The percentage of surviving worms was plotted over time and log-rank tests were performed using the Kaplan-Meier survival analysis in Sigmaplot (version 11.0).

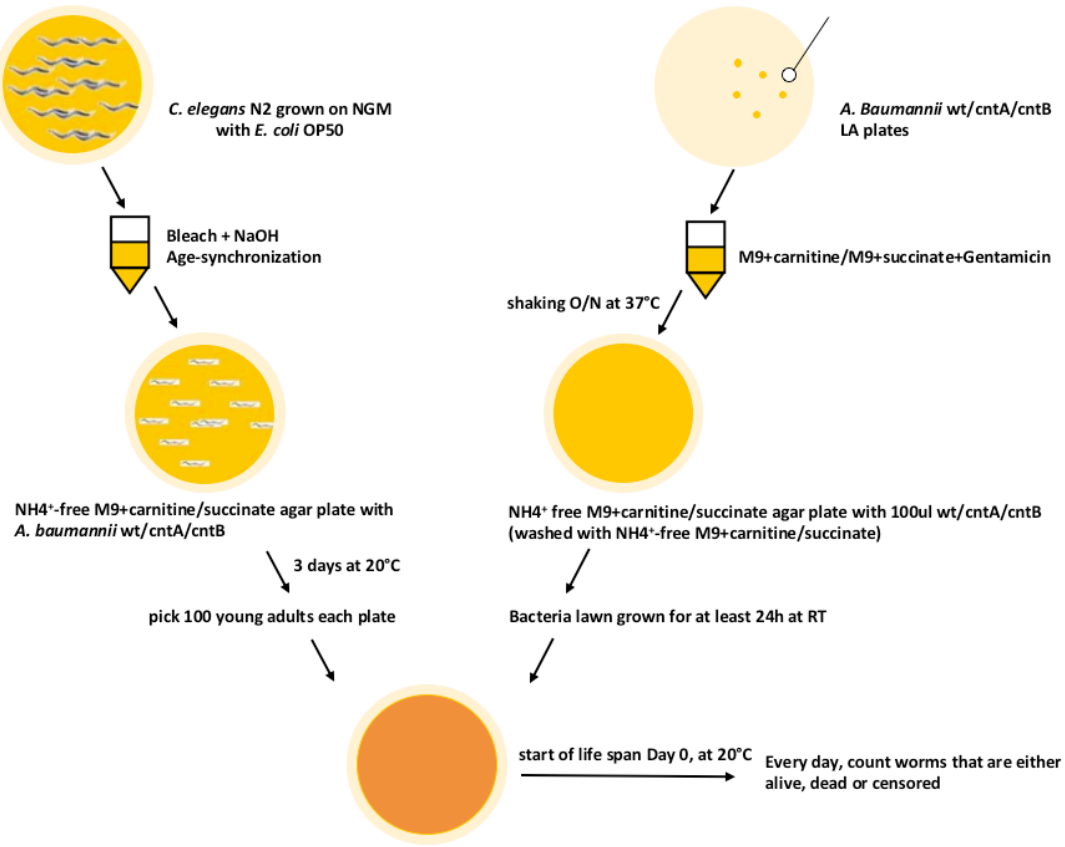
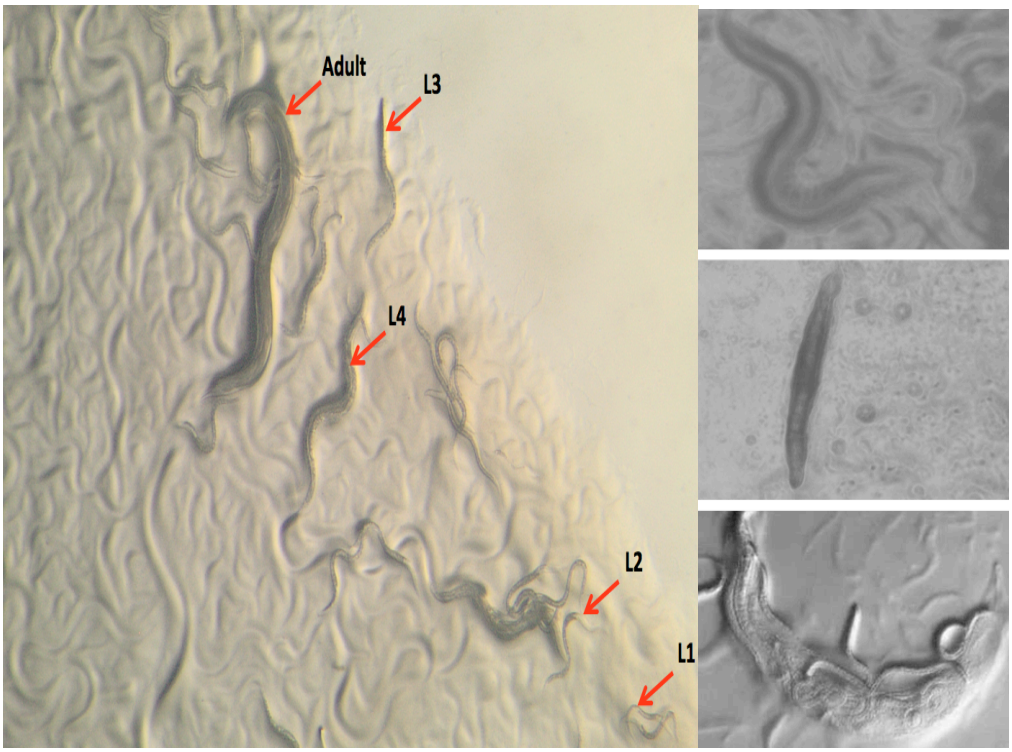


Figure 2.4 Schematic representation of the impact of carnitine degradation in *A. baumannii* on *C. elegans* life span (modified from the protocol of Crosatti, 2014).



(A) **(B)**

Figure 2.5 (A) Different stages of *C. elegans* on un-synchronized NGM plates, from L1 to L4 lava and to adult during their life cycle, as indicated by red arrows. **(B)** Observation of *C. elegans* during its life span was recorded as, upper: alive; middle: dead; lower: censored (bagged).

2.10.3.2 Carnitine degradation in *Escherichia coli* SE11 on *C. elegans* life span

The same protocol **(See section 2.10.3.1)** was used to investigate the impact of carnitine degradation in *E. coli* SE11 on *C. elegans* life span **(Fig 2.6)**.

A temperature sensitive nematode strain was used. *C. elegans* CF512 strain was maintained on OP50 plates at 15 °C until needed. Age-synchronized worms were grown at 25 °C to adults and transferred onto five different groups of *E. coli* SE11 plates supplemented with: 20 mM of carnitine, malate, succinate, malate plus TMA, or succinate with TMA respectively (100 worms per group). The plates were kept at 25 °C throughout the life span experiment since CF512 is a strain unable to produce offspring at this temperature and thus remain sterile (Hsin, 2007).

Survival assays were carried out for the *C. elegans* N2 strain at both 20 °C and 25 °C to compare the difference of the impact of carnitine degradation in *E. coli* SE11 on *C. elegans* life span.

2.10.4 Survival analysis

At the end of each assay, the life span data collected were used to carry out the survival analysis, taking into account the missing nematodes during the experiments as censored. The life span data were analysed to investigate whether each worm survived, died or was censored, as well as the time at which any of these events occurred. Sigma-plot software was used to carry out this analysis using the Kaplan-Meier method. The comparisons of the survival of nematodes between different bacterial strains and between mutants and wild type were carried out using Log-rank statistical test.

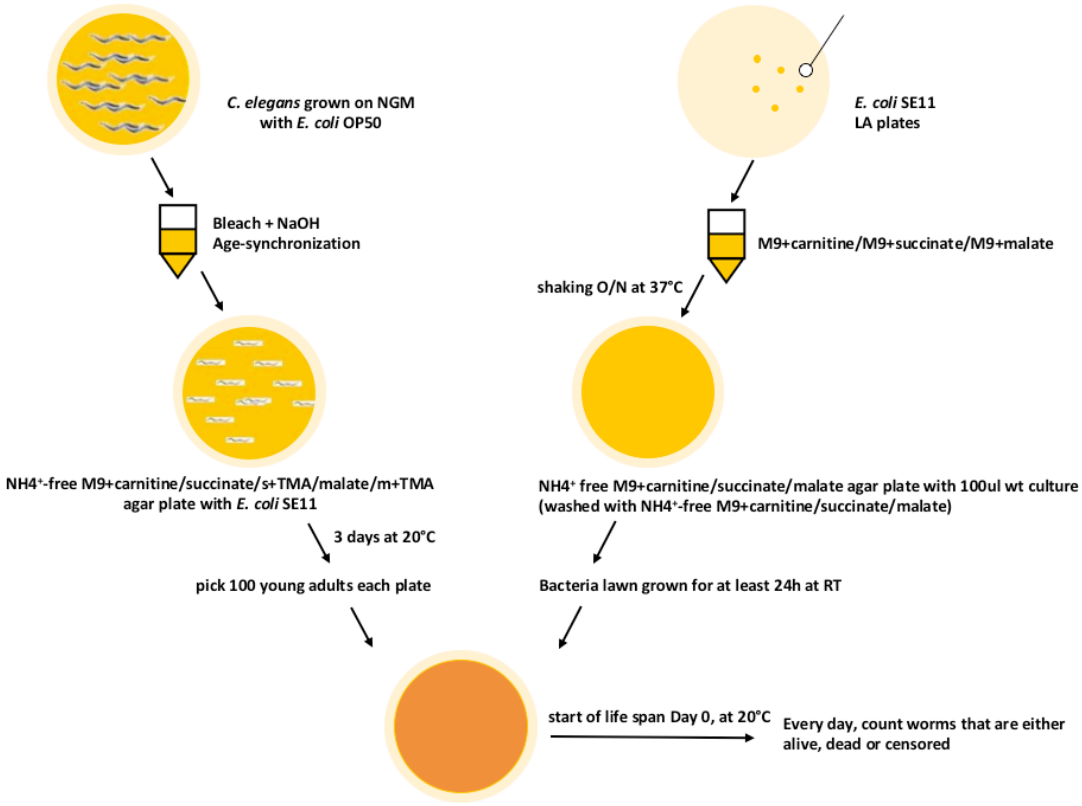


Figure 2.6 Schematic representation of the impact of carnitine degradation in *E. coli* SE11 on *C. elegans* life span (modified from the protocol of Crosatti, 2014).

Table 2.3 A complete list of strains and plasmids used throughout the course of this study

Plasmids/strains	**Description/use**	**Reference**
Bacterial strains		
Acinetobacter baumannii ATCC 19606	Wild type	University of Leicester laboratory collection
A. baumanii ΔcntA::aacC1	Wild type derivative with *cntA* deletion	Zhu *et al.*, 2014
A. baumanii ΔcntB::aacC1	Wild type derivative with *cntB* deletion	Zhu *et al.*, 2014
Escherichia coli BLR(DE3) pLysS	Host for heterologously protein expression	(Sambrook *et al.*, 2001)
E. coli JM109	Routine host for cloning	(Sambrook *et al.*, 2001)
E. coli SE11	Wild type	(Oshima *et al.*, 2008)
E. coli OP50	Wild type	(Stiernagle, 2006)
Proteus mirabilis DSMZ4479	Wild type	DSMZ culture collection, Braunschweig, Germany
P. mirabilis HI4320	Fully sequenced as genome reference	Pathology Department of Cambridge University
P. mirabilis P14	Clinical isolate	Pathology Department of University of Cambridge
P. mirabilis P19	Clinical isolate	Pathology Department of University of Cambridge
P. mirabilis ΔcutC::Km	Wild type with disrupted *cutC*	This thesis
P. mirabilis ΔcutC::Km+*cutCD*	*cutC* mutant complemented with pGEM-T-*cutC/D*	This thesis
Plasmids		
pACD4K-C	TargeTron vector, Kan^R	TargeTron gene knock out system, TA0100, Sigma
pACD4K-*cutC*	Disrupted *cutC* cloned into pACD4K-C	This thesis
pAR1219	Helping plasmid for TargeTron transposon mutagenesis, Amp^R	TargeTron gene knock out system, TA0100, Sigma
pBIO1878	Spc^R derivative of pMP220 with LacZ reporter gene	(Todd *et al.*, 2012)
pBIO1878 (P_{tetR}-*lacZ*)	The promoter of *cutC* gene cluster cloned into pBIO1878	This thesis
pCOLADuet	Dual expression vector, Kan^R	Novagen

pCOLADuet-*cutC/D*	*cutC* and *cutD* cloned into pCOLADuet	This thesis
pCOLADuet-*cutC'/cutD*	N-terminus truncated *cutC* cloned into pCOLADuet-*cutD*	This thesis
pGEM-T-*cutC/D*	*cutC* and *cutD* cloned into pGEM-T easy vector, Amp^R	This thesis
pk18mob*sacB*	Positive control vector for electroporation of *P. mirabilis*, Kan^R	(Schäfer *et al.*, 1994)
Animal strains		
Caenorhabditis elegans N2	Wild type, store at 15 °C, grow at 20 °C or 25 °C	University of Warwick laboratory collection
C. elegans CF512 fer-15(b26) II; fem-1(hc17) IV	Temperature sensitive strain, grow at 15 °C, sterile at 25 °C	*Caenorhabditis* Genetic Centre, University of Minnesota, USA

Table 2.4 A list of primers used throughout the course of this study

Primers	Sequence	Used for
CutC-IBS	AAAAAAGCTTATAATTATCCTTAAAATTCCATCTGGTGCG CCCAGATAGGGTG	*ΔcutC* by TargeTron transposon mutagenesis system
CutC -EBS2	TGAACGCAAGTTTCTAATTTCGGTTAATTTCCGATAGAGG AAAGTGTCT	*ΔcutC* by TargeTron transposon mutagenesis system
CutC -EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCCATCTG ACTAACTTACCTTTCTTTGT	*ΔcutC* by TargeTron transposon mutagenesis system
cdt mutant F	GCGGTTAATCCGCGTGTTTC	Confirmation of *ΔcutC* complement mutant
cdt mutant R	AGGCATTTGCCGGTGTATCA	Confirmation of *ΔcutC* complement mutant
T7 F	TAATACGACTCACTATAGGG	Confirmation of *ΔcutC* complement mutant
5'R kpnI	AAAAAGGTACCACGCCATGCGATATCAGGAG	Confirmation of *ΔcutC* complement mutant
5898-6656 R	AAATATTTTGCCGGTTATCAA	Confirmation of *ΔcutC* complement mutant
6656-7415 F	CCACCGTTAAGTTGATAAAGG	Confirmation of *ΔcutC* complement mutant
3' F kpnI	AAAAAGGTACCAGGGCACGAAGAGTTACTGC	Confirmation of *ΔcutC* complement mutant
M13 R	GGATAACAATTTCACACAGG	Confirmation of *ΔcutC* complement mutant
N-terminus truncated *cutC* F	CCATGGACATCAACGATCCGCGCGTGA	Cloning region of N-terminus truncated *cutC*
N-terminus truncated *cutC* R	GATTATGCGGCCGTGTACAA	Cloning region of N-terminus truncated *cutC*
F1 pGEMT	TCAGATCTGTCGTATCACCA	Confirmation of N-terminus truncated *cutC*
F2 pGEMT	TTCGTTAAAGTCTACCAGCC	Confirmation of N-terminus truncated *cutC*
F3 pGEMT	CTATCCGCAAACTGGTGTAT	Confirmation of N-terminus truncated *cutC*
P_{tetR} F	GAATTCAAATCCACATTACTTTCTCGCTCA	Cloning region of the choline-TMA lyase promoter
P_{tetR} R	GGTACCTTTTATTTCCTCCAGGAAGCCGTT	Cloning region of the choline-TMA lyase promoter

PMI2722-20 F	GGAATTCCATATGGGTGATGCATTAGGTCTG	Cloning region of microcompartment proteins
PMI2722-20 R	CTAGCTAGCTCAGGCTTTATGTTGCTCTG	Cloning region of microcompartment proteins
PMI2718 F	CATGCCATGGATGATCCTCGCAAAGGTAATC	Cloning region of microcompartment proteins
PMI2718 R	CTAGCTAGCTTACTCCTTACTGTCCCGATAAAC	Cloning region of microcompartment proteins
PMI2814 F	CATGCCATGGATGAACAGTTTAGGTGTGATTG	Cloning region of microcompartment proteins
PMI2714 R	CTAGCTAGCTTATTTTTTACCTTTACGGC	Cloning region of microcompartment proteins
PMI2722-10 F	CGCGGATCCATGGGTGATGCATTAGGTCTG	Cloning region of the *cutC* operon
PMI2722-10 R	AAAAAGCGGCCGCTTACATCGTTGAAAAGGTATTC	Cloning region of the *cutC* operon
PMI2725-09 F	CGCGGATCCATGAATACAACCGCACATCAC	Cloning region of *cutC* operon+*tetR* regulator+ protein-tyrosine phosphatase
PMI2725-09 R	AAAAAGCGGCCGCTTATTTTATTTCAATATTATC	Cloning region of *cutC* operon+*tetR* regulator+ protein-tyrosine phosphatase

Chapter 3

Developing a method for genetic manipulation of *Proteus mirabilis*

3.1 Introduction

Proteus mirabilis is a Gram-negative, facultative anaerobic, rod shaped bacterium, which is part of the gut microbiota in healthy individuals, but also a frequent cause of urinary tract infections (UTI) after *Escherichia coli* (Pearson *et al.*, 2008). *P. mirabilis* is unique in that this bacterium can differentiate from short vegetative swimmer cells to elongated highly-flagellated swarmer cells (Pearson *et al.*, 2008). This distinct feature was first observed by Hauser in 1885 (Rather, 2005). It has been well-documented since the 1980s that *P. mirabilis* can produce TMA from choline (Sandhu and Chase, 1986), therefore it was chosen in this project as a model to uncover the key genes and metabolic pathways for choline degradation to TMA. However, the method for plasmid transfer and gene knock-out mutagenesis for *P. mirabilis* has not been well reported.

Between January, 2004 and December, 2007, a total of 198 non-duplicate isolates of *P. mirabilis* were isolated from clinical specimens in hospitalized patients (Yu *et al.*, 2010). *P. mirabilis* strain HI4320 is one of the representatives of the species, which was isolated from the urine of an elderly female patient suffering from a long-term indwelling catheter (Pearson *et al.*, 2008). The whole genome of strain HI4320 was subsequently sequenced by the Mobley group. It is 4.063 Mb long with a relatively low GC content (39 %). 3,685 gene-coding ORFs were identified in the genome and it has seven rRNA loci. One plasmid of 36,289 nucleotides was also identified from the genome sequences. HI4320 has since become the most widely-used strain in recent studies because it is well-characterized and the whole genome is fully closed. Complete genome sequencing of HI4320 has significantly facilitated the study of *P. mirabilis* physiology and metabolism (Pearson *et al.*, 2008).

During the course of this PhD study, two clinical isolates were also used. Strain P14 and strain P19 (U6250) were isolated from hospital patients with chronic urinary tract infections (Liaw *et al.*, 2001).

Strain DSMZ4479 has also been widely investigated for urinary tract infection, genome sequencing analysis and clinic detection studies (Schaffer and Pearson, 2015; Gosiewski *et al.*, 2014; Bonkat *et al.*, 2012). It was identified by ion mobility spectrometry, which is used to determine microbial volatile organic compounds, and

also brings a new identification tool for human pathogenic bacteria via determination of their characteristic volatile metabolomes (Jünger *et al.*, 2012). Gene mutagenesis and homologous recombination studies in *E. coli* showed that this strain was the easiest and most preferred strain for genetic manipulation (Schaffer and Pearson, 2015), although its genome has not been fully sequenced. The aim of this chapter is therefore 1) to test several *P. mirabilis* isolates for their ability to degrade choline to TMA and 2) to establish a method for genetic manipulation in representative *P. mirabilis* strains.

3.2 *Proteus mirabilis* strains and their growth on choline

During my PhD study, I obtained the four strains of *P. mirabilis* and tested their ability to produce TMA from choline anaerobically. *P. mirabilis* strain DSMZ4479 was obtained from the DSMZ culture collection, Braunschweig, Germany. Strains HI4320, P14 and P19 were obtained from the Pathology Department of Cambridge University. These four strains were grown anaerobically at 37 °C in NH_4^+ medium on choline (10 mM) as the sole carbon source and choline was degraded completely within 5 days. Quantification of TMA using a cation exchange Ion chromatography (881 Compact IC Pro Metrohm, Herisau, Switzerland) demonstrated the production of TMA from choline degradation in all four isolates **(See Chapter 2 materials and Methods)**.

3.3 Minimum inhibitory concentration test on *P. mirabilis* strains

In order to establish a protocol for generating mutants in *P. mirabilis*, I firstly carried out minimum inhibitory concentration (MIC) tests. I tested the four strains on the non-swarming agar plates **(See Chapter 2 Materials and Methods** for medium compositions**)** using kanamycin, ampicillin and gentamycin. The MIC data for kanamycin, ampicillin and gentamycin showed no significant difference among the four strains **(Table 3.1)**. I chose the strain DSMZ4479 for further exploration because it has been reported previously that it was amenable for genetic manipulation (Schaffer and Pearson, 2015) although its genome sequence has not been reported **(Table 3.2)**.

Table 3.1 MIC (μg ml^{-1}) for four different strains of *P. mirabilis* on non-swarming agar plates

MIC / Strains	Kanamycin	Ampicillin	Gentamicin
DSMZ4479	150	150	50
HI4320	150	150	20
P14	150	100	50
P19	150	150	20

Table 3.2 MIC (μg ml^{-1}) for *P. mirabilis* DSMZ4479 on non-swarming agar plates

Antibiotics	MIC
Kanamycin	150
Ampicillin	150
Gentamicin	50
Spectinomycin	150
Streptomycin	250
Chloramphenicol	150

3.4 Establishment of the protocol for plasmid transfer in *P. mirabilis* strain DMSZ4479

I initially planned to generate marker-exchange knockout mutants using the vector pKNG101 through conjugation. pKNG101 is a suicide plasmid, suitable for positive selection of double recombination mutants in a range of Gram-negative bacteria (Kaniga *et al.*, 1991). It contains an oriR6k type of origin of replication, a mobRK2 type of origin of transfer and an antibiotics selection marker (streptomycin resistance, Sm^R) (Kaniga *et al.*, 1991). However, after several attempts, I did not manage to obtain any mutants in *P. mirabilis* DSMZ4479 using this approach. A previous deep investigation on the whole genome of *P. mirabilis* HI4320, however, has revealed the presence of an R6K-type of plasmid in this bacterium. It has been reported that this R6K-type plasmid in strain HI4320, named pHI4320, is highly related to the R6K plasmid in *E. coli*, which encodes the R6K replication proteins (encoded by PMIP37 to PMIP39) and the π protein (encoded by PMIP01) (Kaniga *et al.*, 1991; Pearson *et al.*, 2008). One of the necessary facts for successful bacterial conjugation is plasmid compatibility, which means the donor cell provides a conjugative genetic element, usually a plasmid, while the recipient cell does not already contain a similar plasmid. The genome of *P. mirabilis* strain DSMZ4479 has not been genome-sequenced but it is possible that a similar R6K type plasmid is present in DSMZ4479, therefore making it impossible for conjugating pKNG101 (which contains an oriR6k type of replication origin) into this bacterium due to plasmid incompatibility. This approach was therefore abandoned.

I then tried to use electroporation for plasmid transfer to DSMZ4479 using a vector that does not contain an oriR6k type of replication origin. I adapted the protocol from Visalli *et al.* (2003) to make electrocompetent cells. The protocol was further optimised to improve competency. The cells were grown in a LB broth medium (without NaCl) and harvested when the optical density (OD_{600}) reached between 0.3 and 0.6, which took 2.5-4 hours after inoculation. Immediately before electroporation, the competent cells were mixed with plasmid DNA and incubated on ice for 1 hour before electroporation to improve electroporation efficiency. During the experiment, I used a pGEMT-based construct containing a Kan^R cassette

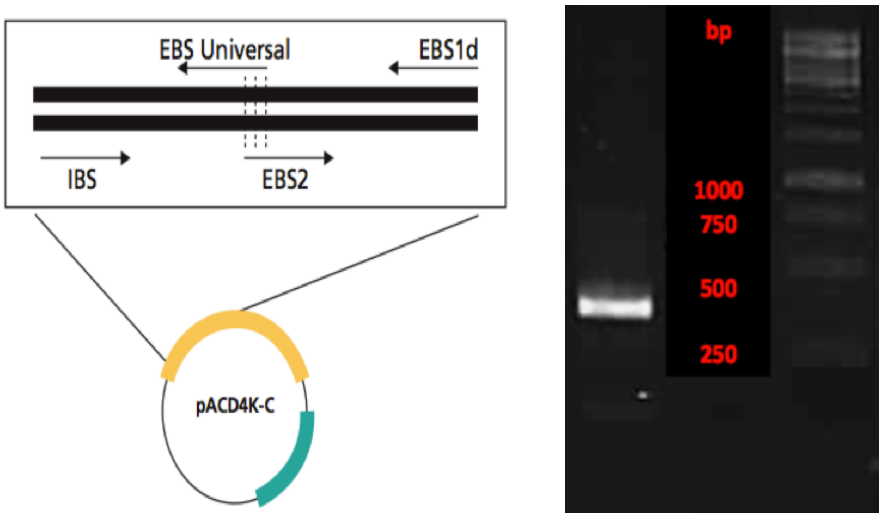
as a positive control to optimize the gene transfer protocol. The transformation efficiency of plasmid transfer using this established protocol was estimated to be ~ 4×10^5 CFU/μg.

Briefly, mid-log phase *P. mirabilis* cultures were harvested when they reached OD_{600} 0.3-0.6, approximately 2.5-4 hours in LB (minus NaCl) at 37 °C with shaking. Cells were chilled on ice and washed 4 times by centrifugation at 3,200 *g* with ice-cold, sterile, 10 % (v/v) glycerol solution. The culture was concentrated 500-fold and aliquots of 60 μl snap-frozen in a dry-ice ethanol bath and stored at -80 °C. Before electroporation thawed competent cells were combined with 500 ng of plasmid DNA and incubated on ice for 1 hour. The cells were then transferred to a 1 mm gap cuvette and electroporated at 1.8 kV and the time constant was 4.5-4.6 ms.

3.5 Generating a *cutC* knockout mutant in *P. mirabilis* using a TargeTron targeted transposon mutagenesis system

3.5.1 Construction of the pACD4K-*cutC*

In **Chapter 4**, a comparative transcriptomic approach is used to identify key genes involved in choline degradation to TMA. One of the key genes that is highly induced by choline is designated as *cutC* **(See later in Chapter 4)**. I therefore chose this gene for testing the gene knockout method. The *cutC* mutant (*ΔcutC*) was generated using the TargeTron gene knockout system (Sigma-Aldrich, TA0100-3EA) **(Fig 3.1 and Fig 3.2)**. The *cutC* gene sequence was loaded onto the TargeTron Design Website (sigma-aldrich.com/targetronacess) and three *cutC*-retarget specific primers IBS, EBS2, and EBS1d were designed by the software. The system provided a few sets of these three specific primers for the target site and the set with lowest E value (0.010) was used in this study **(Table 2.2, Materials and Methods)**.



(A) **(B)**

Figure 3.1 (A) One-step assembly PCR using the EBS universal primer and the three primers specific for *cutC* was carried out to mutate (re-target) the intron RNA. The PCR product ~ 350 bp **(B)** of the retargeted intron was cloned into pGEM-T which was then confirmed by DNA sequencing using the M13 primers. The retargeted intron was then sub-cloned into the plasmid pACD4K-C to create pACD4K-*cutC* (**Fig 3.2**).

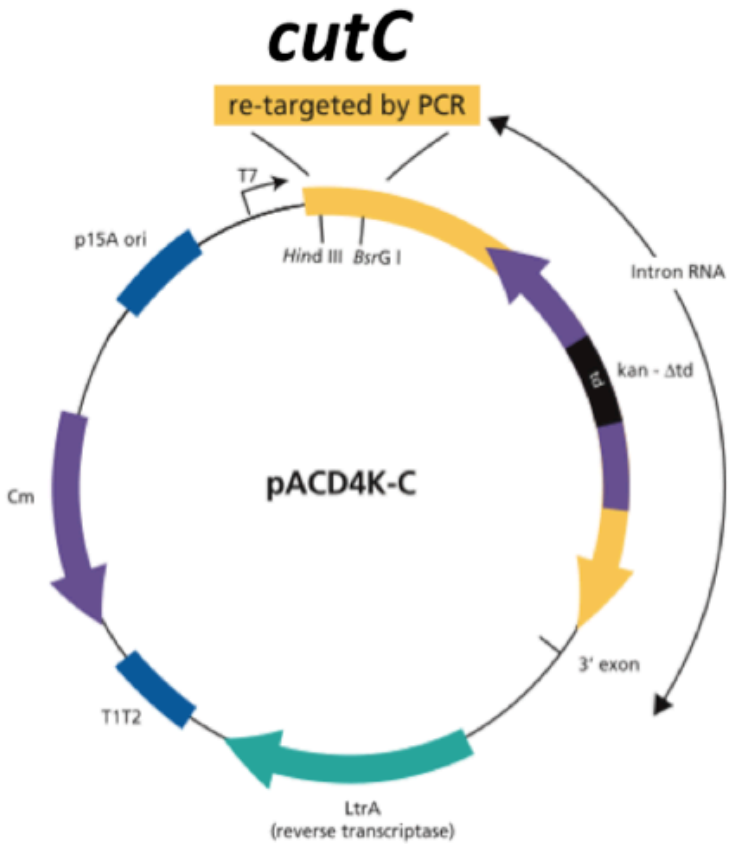
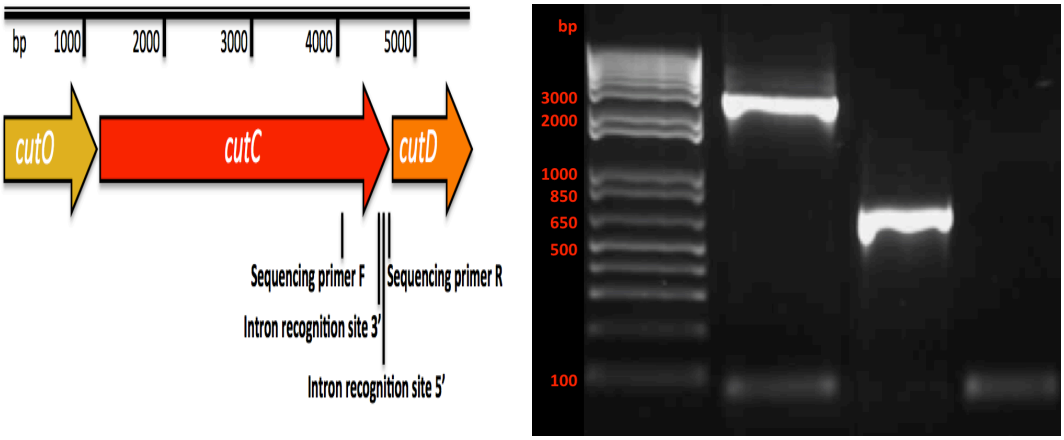


Figure 3.2 The plasmid map of pACD4K-C and the retargeted intron RNA site.

3.5.2 Electroporation of *P. mirabilis* DSMZ4479 [pAR1219] competent cells with pACD4K-*cutC*

P. mirabilis DSMZ4479 competent cells were first electroporated with the T7 helper plasmid pAR1219 (Davanloo *et al.*, 1984). The helper plasmid is a pBR322-based vector that expresses T7 RNA polymerase under control of the IPTG inducible lac UV51 promoter and is intended for use with the TargeTron Gene Knockout System (TA0100). Therefore, this step is essential because the helper plasmid would assist the late intron RNA transcription once the transformation of *P. mirabilis* DSMZ4479 and pACD4K-*cutC* was done. Transformants were selected on non-swarming agar plates containing 150 µl/ml ampicillin. Electrocompetent *P. mirabilis* DSMZ4479 containing pAR1219 were then prepared which were stored at -80 °C before use. The pAR1219 containing cells were electroporated with the construct pACD4K-*cutC*, and putative *cutC* mutants were selected for by overnight growth in LB containing 150 µl/ml ampicillin cultures, 200 µl/ml chloramphenicol and 1 % glucose. The addition of 0.5 mM IPTG to the chloramphenicol- and ampicillin-resistant isolate induced the intron to insert into the *cutC* gene in the *P. mirabilis* chromosome. Successful transformants were selected on non-swarming agar plates containing 150 µl/ml kanamycin. This mutant, *ΔcutC*, was confirmed by PCR using sequencing primers F and R **(Fig 3.3A)**. The resulting PCR product for the wild type is around 1 kb whereas the size of mutant is around 3 kb since the Kan^R cassette from the intron is about 2kb **(Fig 3.3B)**.



(A) **(B)**

Figure 3.3 (A) The primer binding positions for the sequencing primers F and R, and the intron retargeting sites (position 3359 bp in total 3429 bp of *cutC*). **(B)** Agarose gel electrophoresis of the PCR products of the wild type and the mutant. lane 1, 1kb plus DNA ladder (Invitrogen); lane 2, PCR product of the *ΔcutC* ; Lane 3, PCR product of the wild type; and Lane 4, negative control of PCR.

3.5.3 Complementation of the *cutC* mutant

To complement the *cutC* mutant, the intact *cutC* and *cutD* genes were amplified from the genomic DNA of *P. mirabilis* wild type DMSZ4479 using a proof-reading DNA polymerase (KAPA HiFi Tag). The PCR product was analysed by agarose gel electrophoresis in **Fig 3.5 (A)**, which was further purified before dA-tailing. In the same way, the dA-tailing product was purified and ligated with the pGEM-T vector (Promega) to obtain pGEM-T-*cutC/D* **(Fig 3.4 & Fig 3.5B)**. This construct was confirmed by digestion with the restriction enzyme, NotI **(Fig 3.5C)** and subsequent DNA sequencing analysis using the primers listed in **(See Table 2.2, Materials and Methods)**. The competent cells of *P. mirabilis cutC* mutant were electroporated with pGEM-T-*cutC/D,* and the complemented mutant was selected on an LB agar plate with ampicillin (150 μg/ml).

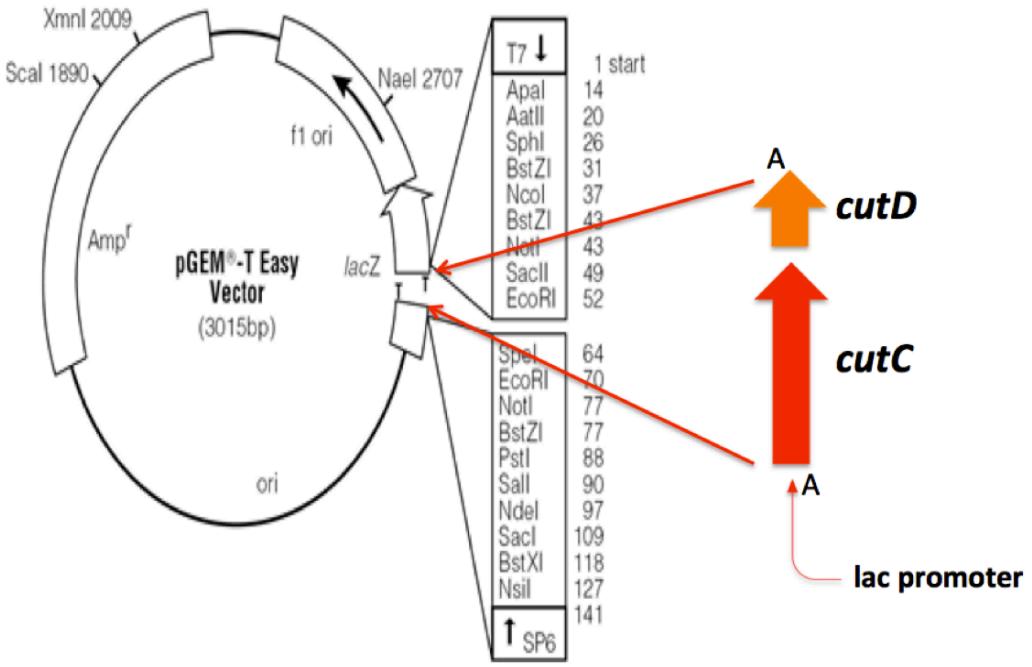
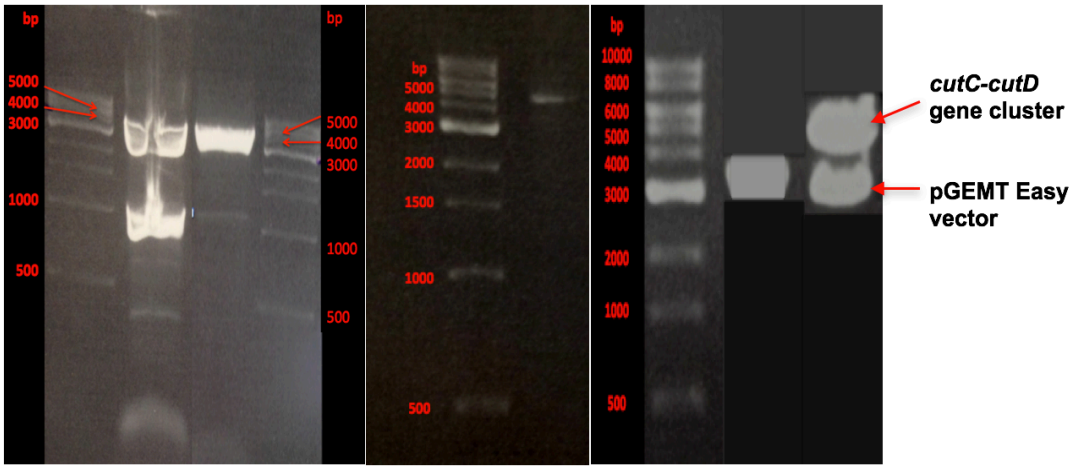


Figure 3.4 *cutC-cutD* genes (4.7 kb) were amplified from *P. mirabilis* that was cloned into the pGEM-T vector. Its expression was driven by the lac promoter in the upstream of *cutC* (picture modified from pGEMT Easy Vector System technical manual, Promega, free online open access).



(A) **(B)** **(C)**

Figure 3.5 (A) Agarose gel electrophoresis of the PCR product of the *cutC-cutD* gene cluster. From left to right, 1 kb DNA ladder (Invitrogen), PCR product of *cutC-cutD* (replicate 1), PCR product (replicate 2), 1 kb DNA ladder. The uppermost 4.7 kb band is the product of expected size. **(B)** The dA-tailing product of *cutC-cutD* was purified from the agarose gel, which was then ligated into the pGEM-T Easy vector (Promega). **(C)** Digested pGEM-T-*cutC/D* with the restriction enzyme, NotI to confirm the presence of the *cutC-cutD* gene cluster in this plasmid. From left to right, 1 kb DNA ladder (Invitrogen), linearized pGEM-T Easy vector control (3 kb), pGEM-T-*cutC/D* (the upper 4.7 kb band represents the *cutC-cutD* gene cluster and the lower 3 kb band was the linearized pGEM-T Easy vector).

3.6 Conclusions

This chapter describes the initial test of four *P. mirabilis* strains (strains HI4320, P14, P19 and DSMZ4479) and their ability to convert choline to TMA. In good agreement with previous studies (Sandhu and Chase, 1986) all four strains tested are capable of producing TMA from choline anaerobically. MIC tests were then carried out for the four stains of *P. mirabilis*, which demonstrated no difference in antibiotic resistant profile **(Table 3.1)**. DSMZ4479 was therefore used as the model strain in this study to establish the method of plasmid transfer and targeted gene knock-out mutagenesis, because it proved amenable for genetic manipulation although its genome sequence has not been reported. To validate the method, one of the highly expressed genes during growth on choline was chosen for further mutagenesis (See results presented in next chapter). This *cutC* gene was successfully deleted from *P. mirabilis* DSMZ4479. The *cutC::kan* mutant was further complemented by cloning and expression of the native *cutC/D* gene from DSMZ4479. Further growth tests of the wild-type, mutant and the complemented mutant are presented in **Chapter 5**.

Chapter 4

Comparative transcriptomic analysis of *Proteus mirabilis* by RNA-Seq

4.1. Introduction

The advent of next generation sequencing methods, and in particular comparative transcriptomics (RNA-Seq), has revolutionized the study of prokaryotic transcriptomes (Güell *et al.*, 2009; Sharma *et al.*, 2010b; Mitschke *et al.*, 2011). Despite this, few studies have employed this technology to study the transcriptomes of quaternary amine metabolism in bacteria.

One of the most important advantages of RNA-Seq is to identify the differential expression of genes under two or more conditions. The differentially expressed genes are identified by a combination of expression change threshold and score cutoff, which are usually based on *p* values calculated by statistical modeling. It is necessary to normalize the context of different analyses, as the presence of systematic variation between samples, as well as, differences in library composition will be taken into account via normalization.

It is known that *P. mirabilis* can grow on choline as a sole carbon and energy source (Sandhu and Chase, 1986), however it is unclear which genes, enzymes and metabolic pathways are involved. The aim of this chapter is therefore to provide a global view of whole cell response to choline metabolism in this bacterium, by comparing transcripts of choline-grown cultures with glucose-grown cultures.

In this chapter, RNA-Seq was carried out to investigate which genes are switched on and highly expressed under growth conditions with or without choline. Several modern bioinformatics tools have been used in this study: Bedtool-intersect toolset (Quinlan, 2015) for raw data comparison and calculation, to get read counts files, Degust interactive web-based RNA-Seq visualization software (Zhao *et al.*, 2016) for read counts analysis and R program (DESEQ package) for statistical analysis. The phylogenetic and genome analysis were subsequently carried out to identify the functional choline-TMA lyase and the *cutC* gene cluster in *P. mirabilis* and to compare the CutC between *P. mirabilis* and *D. desulfuricans*.

4.2 Results

4.2.1 Experimental setup

P. mirabilis cultures were grown anaerobically in defined medium with glucose only or glucose plus choline as a carbon source, with five replicates for each condition. Cells were grown and harvested at the late exponential phase, when OD_{600} reached 0.4, for glucose plus choline at 15 hours, and for glucose only at 20 hours, as indicated in black arrows **(Fig 4.1)**. The whole process of RNA-Seq was carried out using the pipeline indicated in **Fig 4.2**. Briefly, the total ribonucleic acid (RNA) was firstly extracted and purified, and then the messenger RNA was enriched using an rRNA-depletion kit. The quality and quantity of the enriched mRNA were assessed on an Agilent bioanalyser. cDNA library preparation, RNA Hi-Seq program, and raw Illumina RNA-Seq alignment were carried out following the procedures detailed in **Chapter 2 (Materials and Methods, Section 2.5 RNA-Seq)**. Note that the raw reads were mapped to *P. mirabilis* HI4320 genome sequence, because the genome sequence of the strain used in this study is not available. The raw read counts were gained from extracting via the HTSEQ PYTHON package.

Two independent methods were then used to analyze the raw read counts data. In the first instance, differential expressions of genes were determined by the R software package DESEQ for the cultures grown under conditions with or without choline. This was carried out by Dr Jameson and myself at the beginning of my PhD (Jameson *et al.*, 2015). **Fig 4.2** indicates the contributions for each one of us for the work presented in this chapter.

Additionally, the raw read counts of RNA-Seq were also processed using the transcriptomic Bedtool online program (Bedtool-intersect toolset, http://bedtools.readthedocs.io/en/latest/content/tools/intersect.html) by myself. The alignment of the data (file format: sample_hits.bam.bai) was carried out in Artemis (Sanger), using the reference strain *P. mirabilis* HI4320. The expression analysis and plots were carried out using the Degust website program (http://www.vicbioinformatics.com/degust/index.html), while the different expression levels of each gene were also investigated **(Fig 4.2)**.

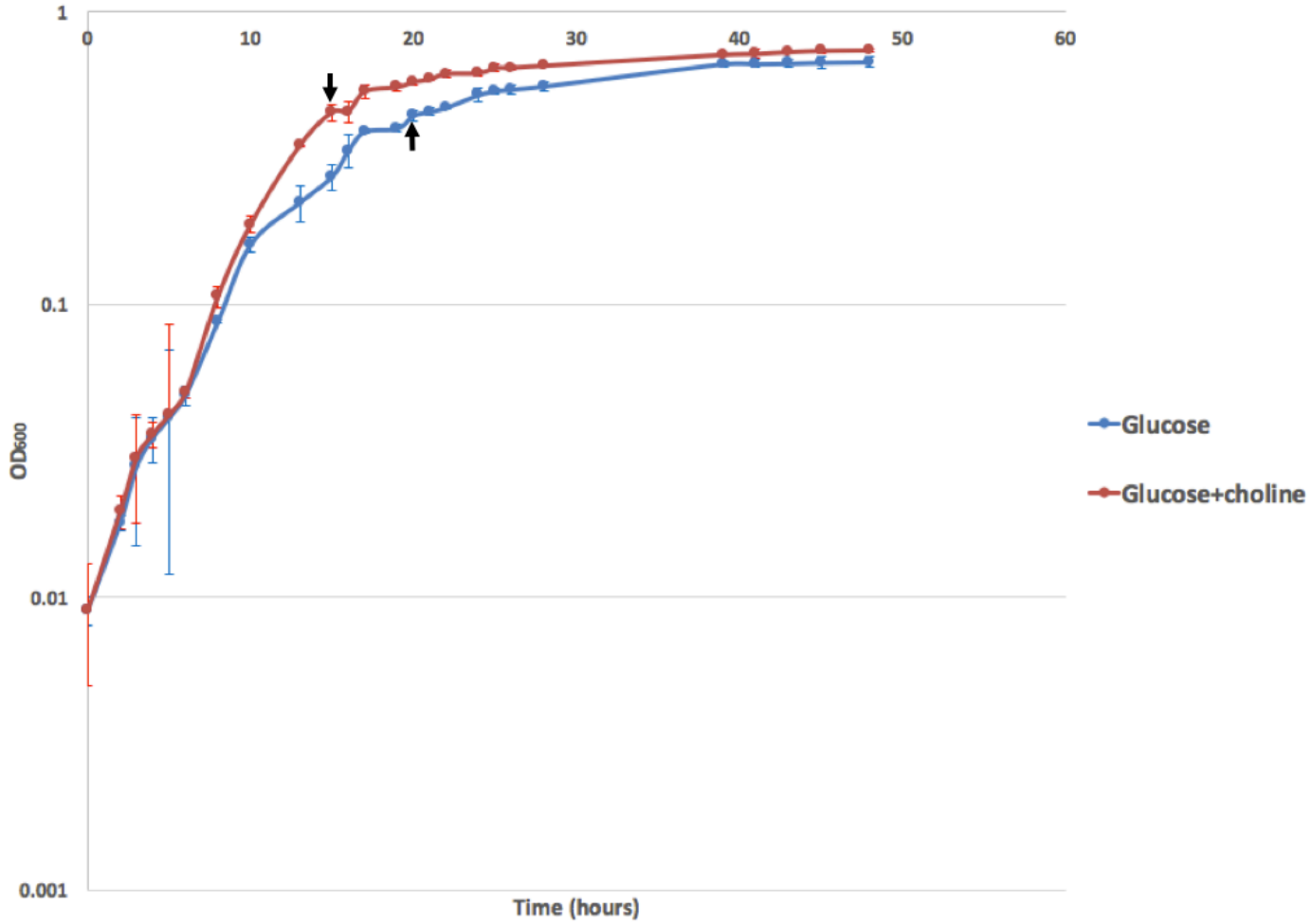


Figure 4.1 Growth curve of *P. mirabilis* in the defined medium supplemented with glucose or glucose plus choline anaerobically. Cells were harvested when OD_{600} reached 0.4, for glucose plus choline at 15 hours, and for glucose only at 20 hours, as shown in black arrows.

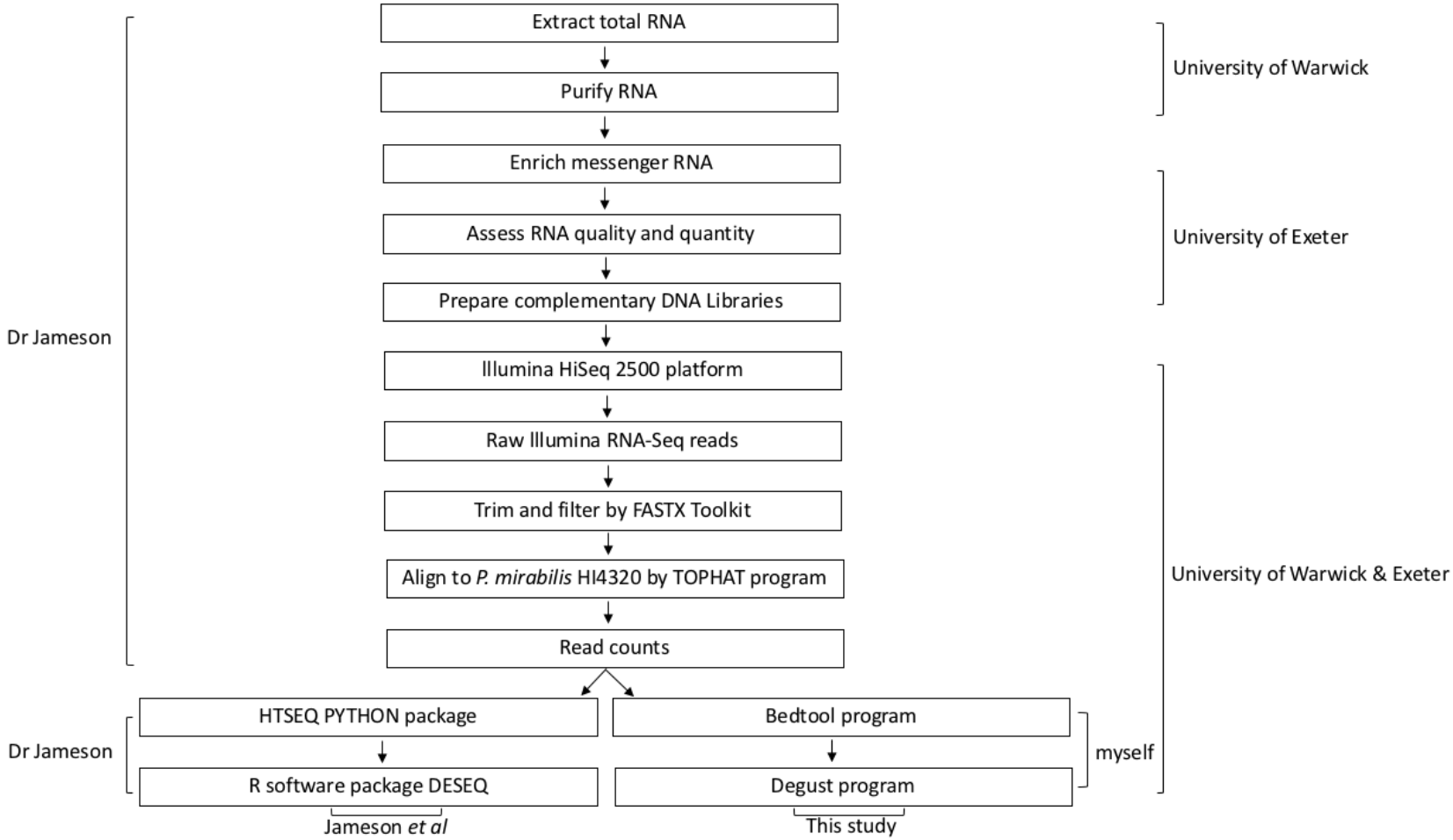


Figure 4.2 RNA-Seq pipeline for the whole experiment. The extraction and purification of RNA was carried out in Warwick; RNA enrichment, quality and quantity assessment and cDNA library preparation were carried out in Exeter, and the raw read counts and the later RNA-Seq analysis were carried out in Warwick and Exeter.

4.2.2 RNA-Seq raw read count data summary

After extraction and purification, messenger RNA was enriched using an rRNA-depletion kit. The quality of enriched mRNA and their quantities before and after the enrichment were assessed on an Agilent bioanalyser to assess the quality of the enrichment. **Fig 4.3** is a typical example showing the difference of *E. coli* mRNA before and after enrichment. In enriched mRNA samples, the main big peaks of 16S and 23S rRNA were depleted, and only small peaks of mRNA remained and thus were enriched (MICROB *Express*^{TM} Kit protocol, Part number AM1905). Exactly the same process of analysis was carried out for the RNA enrichment of *P. mirabilis* under choline and non-choline conditions. Among those five biological replicates in each condition, three of them had better quality and higher quantity. They are C1rep1, C1rep2 and C1rep3 represent choline-grown condition, and C2rep1, C2rep2 and C2rep3 are glucose-grown condition, respectively. Thus these six samples (three replicates for each condition) were used for later library preparation and sequencing analysis. **Fig 4.4** shows a typical QC report of one sample (C1rep3) after Hiseq runs. **Table 4.1** shows the raw data counts of RNA-Seq of the three replicates for each condition.

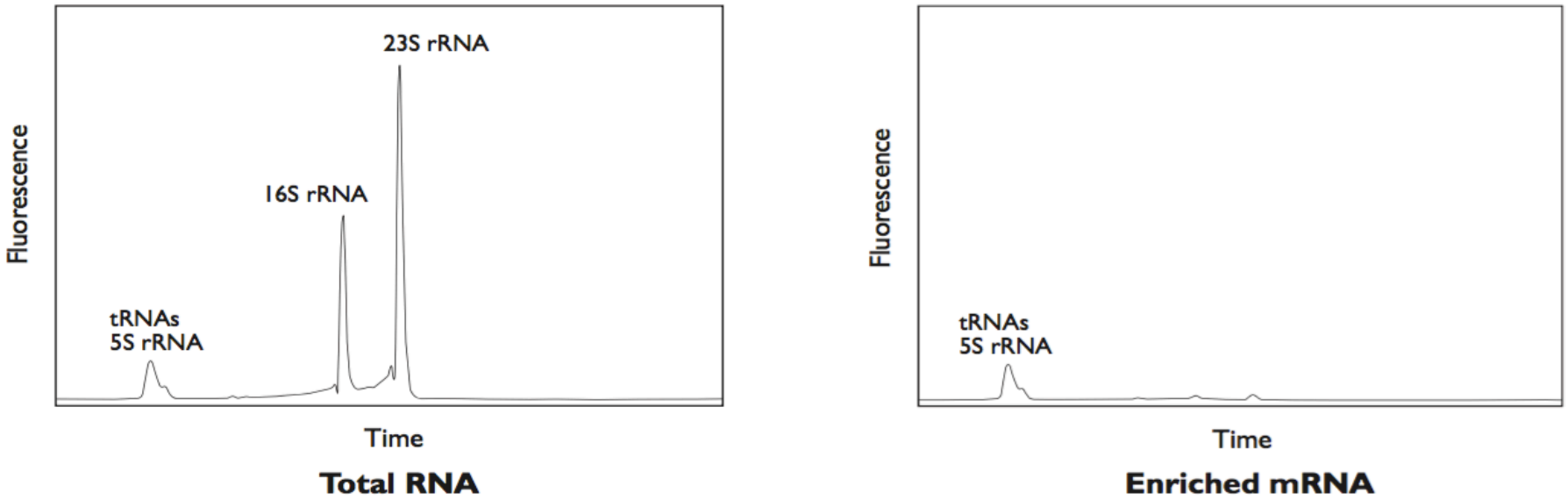
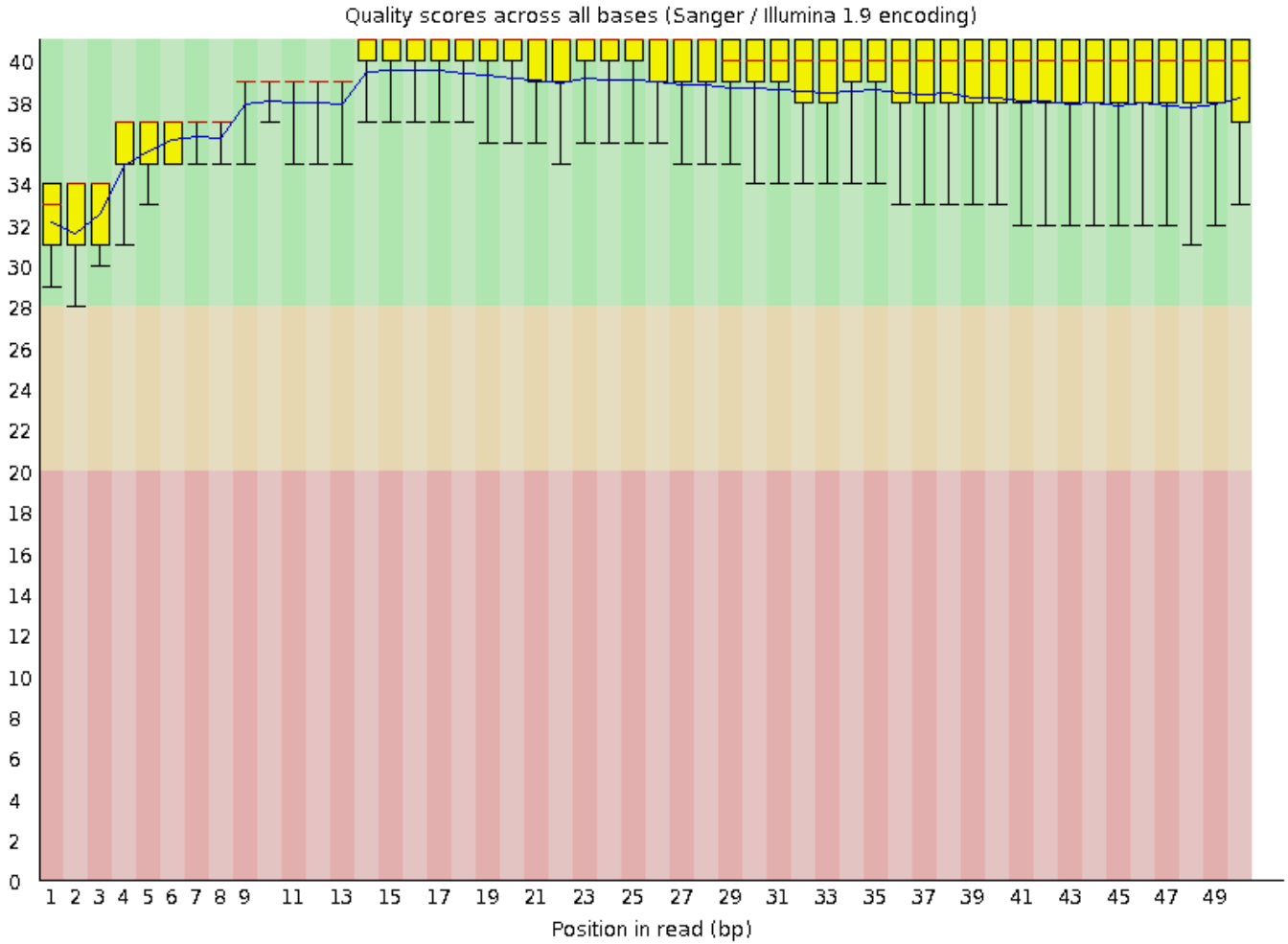
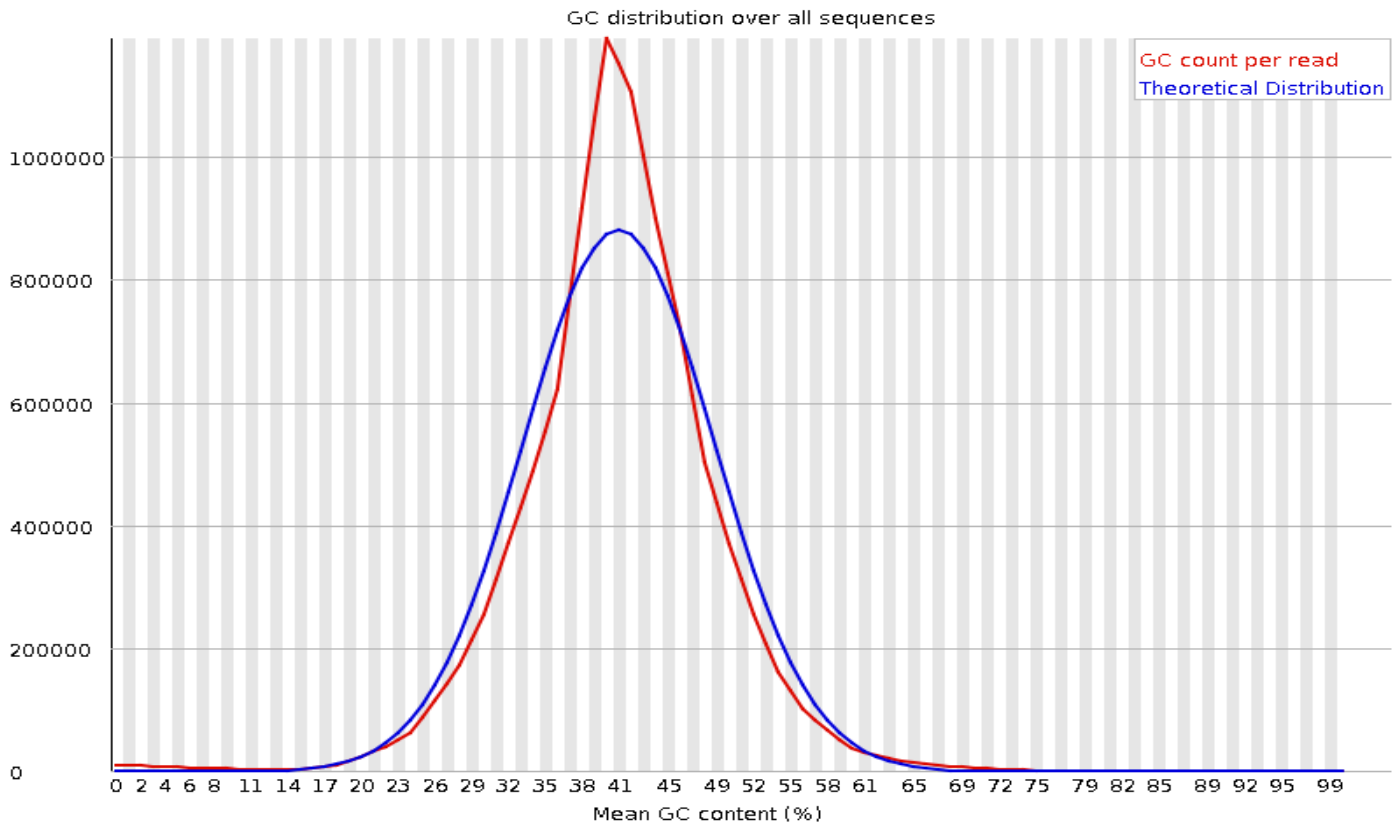


Figure 4.3 Bioanalyzer analysis of *E. coli* RNA before and after mRNA Enrichment. The *E. coli* mRNA was enriched using the MICROB*Express* kit from 10 µg of total RNA. 200 ng of RNA was used to generate the above electropherograms before and after mRNA enrichment by Agilent 2100 bioanalyzer using the RNA LabChip kit (MICROB *Express*™ Kit protocol, Part number AM1905).


Quality scores across all bases (Sanger / Illumina 1.9 encoding)
40
38
36
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
1 2 3 4 5 6 7 8 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49
Position in read (bp)

(A)



(B)

Figure 4.4 The quality report of C1rep3 replicate (sequence length 15-50bp, GC content 40 %) of the total six replicates under two conditions. **(A)** quality score across all base. Yellow bars show the quality scores of each base. **(B)** GC distribution over the whole 100bp sequence, the red line is the GC count per read and the blue line is the theoretical distribution.

Table 4.1 Raw data counts of RNA-Seq of the two conditions, glucose plus choline and glucose only

	Glucose + choline			Glucose		
	C1rep1	C1rep2	C1rep3	C2rep1	C2rep2	C2rep3
Total reads	15,674,358	13,332,183	13,830,281	9,629,506	8,887,177	14,357,080
Mean	14,278,941			10,957,921		
Standard Deviation	1,233,863			2,967,065		

4.2.3 Assessment of biological replicates

The repeatability of the three samples in each conditions was assessed by MDS plot. **Fig 4.5** gives an overview of the variation of the three biological replicates under each condition. Both conditions had good repeatability between samples.

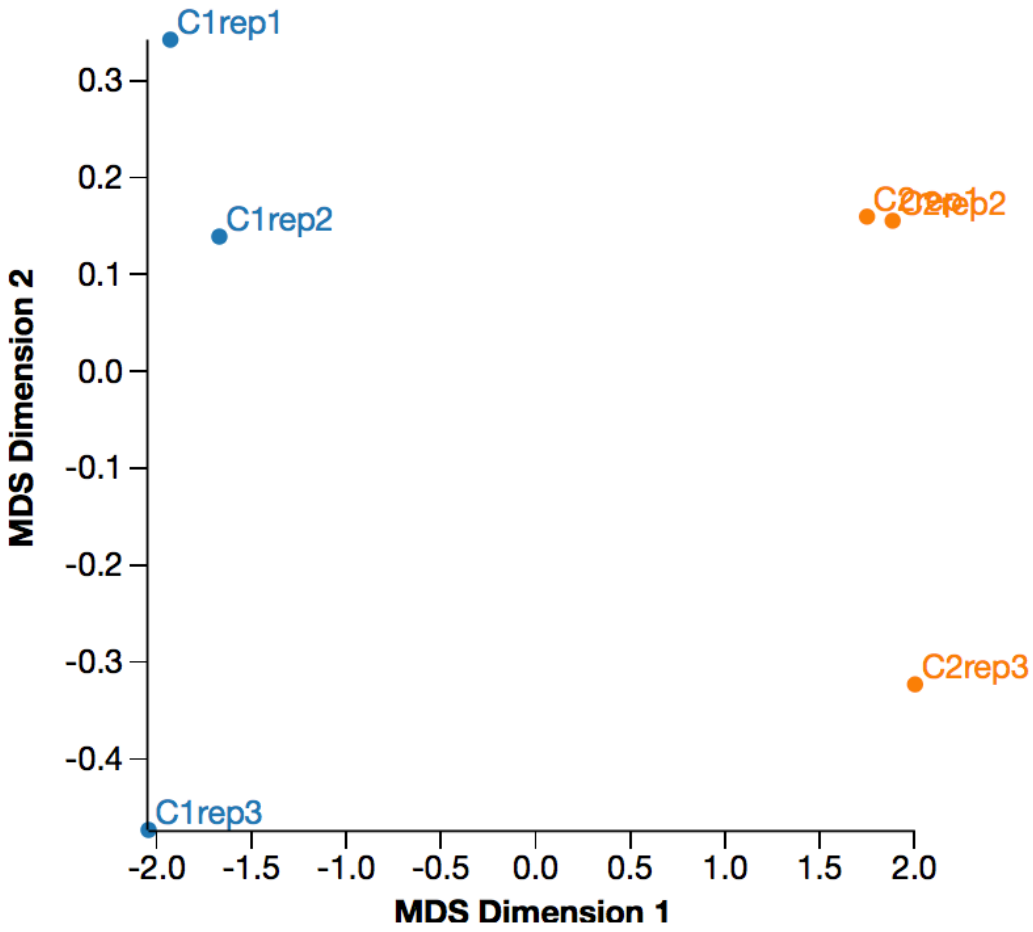


Figure 4.5 MDS plot of the variation of the RNA-Seq sample replicates for each condition. Replicates shown in blue with the prefix "C1" represent choline-grown cultures, and replicates shown in orange with the prefix "C2" represent are glucose-grown cultures. The analysis was carried out using the Degust program (http://www.vicbioinformatics.com/degust/index.html).

4.2.4 Comparative transcriptome analysis by RNA-Seq in *Proteus mirabilis*

4.2.4.1 Differential expression of genes in response to choline degradation to TMA in *P. mirabilis*

Although Sandhu and Chase (1986) have previously confirmed that *P. mirabilis* can metabolize choline to TMA anaerobically, the key enzymes responsible for this transformation remain unknown. Comparative transcriptomic analysis of *P. mirabilis* cultivated on glucose with/without additional choline provided an overview of the key genes that are regulated by choline addition. Initial analyses carried out by Jameson *et al.* showed that a number of genes are highly induced by choline. The top 20 highly expressed genes in *P. mirabilis* are presented in **Table 4.2**. Interestingly, several of these highly induced proteins (PMI2722, PIM2721, PMI2720, PMI2718 and PMI2714) were annotated as bacterial microcompartment shell proteins by the Integrated Microbial Genomes database at the Joint Genome Institute (http:// img.jgi.doe.gov/).

Table 4.2 Top 20 differentially expressed genes. Positive log_2-fold change indicates down regulation with the addition of choline and negative values indicate up regulation with choline (Jameson *et al.*, 2015).

No.	Locus tag	Gene name	Log_2-fold change	*p*-value	Gene description
1	PMI2721	*eutM*/*pduA*/*J* homologue	-10.853338	8.58E-38	Microcompartment protein (BMC domain, PF00936.14)
2	PMI2720	*eutM*/*pduA*/*J* homologue	-9.9694882	1.25E-36	Microcompartment protein (BMC domain, PF00936.14)
3	PMI2722	*eutM*/*pduA*/*J* homologue	-9.3293139	2.06E-27	Microcompartment protein (BMC domain, PF00936.14)
4	PMI2716	*cutC*	-7.3519216	3.39E-26	Choline-trimethylamine lyase, CutC
5	PMI2719	*cutF*	-7.9456261	1.66E-20	Aldehyde dehydrogenase, CutF
6	PMI2715	*cutD*	-6.0539258	8.64E-20	Choline-trimethylamine lyase, activating enzyme, CutD
7	PMI2714	*eutM*/*pduA*/*J* homologue	-5.2834574	7.49E-17	Microcompartment protein (BMC domain, PF00936.14)
8	PMI2718	*eutN*/*ccmL* homologue	-10.804654	1.71E-15	Microcompartment/carboxysome/ethanolamine utilizatin protein (EutN_CcmL, PF03319.8)
9	PMI2711	*emrE* homologue	-8.6639517	7.24E-13	Quaternary ammonium compound resistance protein
10	PMI2717	*cutO*	-9.6507023	1.08E-10	Alcohol dehydrogenase

11	PMI2710	*emrE* homologue	-9.2090116	2.81E-06	Multi-drug resistance protein, EmrE homologue
12	PMI0695	*trxB*	-1.9182025	0.000123	Thioredoxin reductase
13	PMI2713	*cutH*	-5.5647202	0.000259	Phosphate acetyltransferase
14	PMI0501		-1.8393886	0.000698	Uncharacterized conserved protein YjdB, phage protein
15	PMIt054		3.6411425	0.001508	tRNA-Met
16	PMI1739		-1.6050110	0.001628	Hypothetical protein YfaZ precursor, putative exporter protein
17	PMI3109	*fdoH*	-1.6200982	0.001691	Formate dehydrogenase-O subunit beta, transmembrane
18	PMI0297		-1.7078902	0.001728	Fimbrial subunit
19	PMI1288	*ydfG*	-1.7860287	0.001896	NADP-dependent L-serine/L-allo-threonine dehydrogenase
20	PMI1955	*eco*	-1.6974899	0.002404	Protease inhibitor ecotin; homodimeric protease inhibitor

BMC, bacterial microcompartment; NADP, nicotinamide adenine dinucleotide phosphate

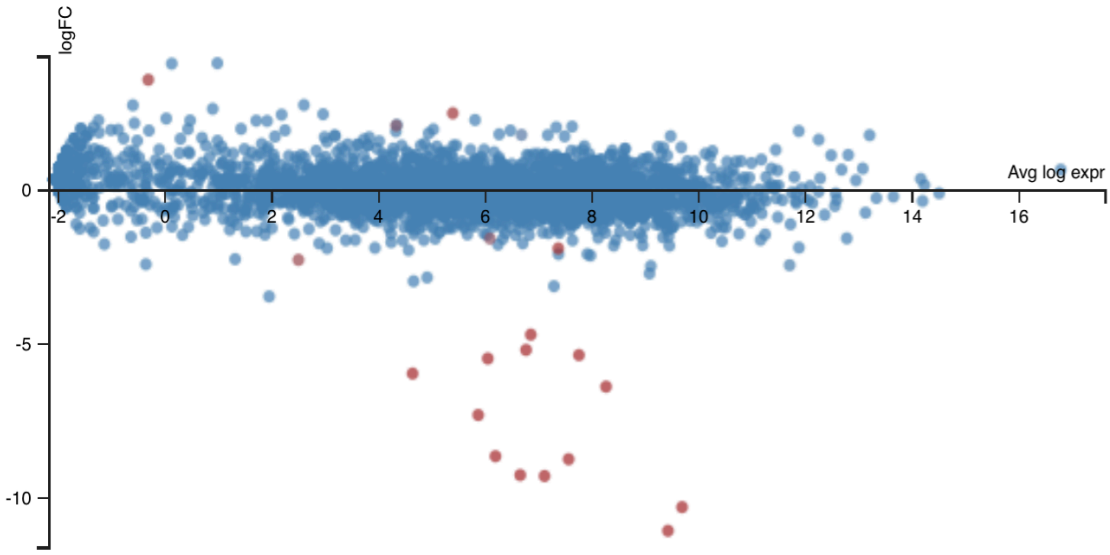
4.2.4.2 Re-analysis of differential gene expression by RNA-Seq in *P. mirabilis*

The initial analysis by Jameson *et al.* used the FASTX toolkit to trim the data and the TOPHAT program to align the RNA-Seq reads to the *P. mirabilis* HI4320 genome. Statistical analysis was performed using the R software package DESEQ. However, in the DESEQ experiment, $p < 0.05$ was applied rather than *p* adjust value, thus there is a likelihood of obtaining false positive hits. Therefore, a different bioinformatics approach was applied, aiming to re-analyze the RNA-Seq data to minimize the chance of false positives in data analyses. For this purpose, the user-friendly online program Bedtool-intersect (Quinlan, 2015) was used to re-run the raw read counts. Degust online program (Zhao *et al.*, 2016) was then used to analyse and visualize the new read counts and perform the statistical analysis. False discovery rate (FDR) was applied at 0.05 cut off.

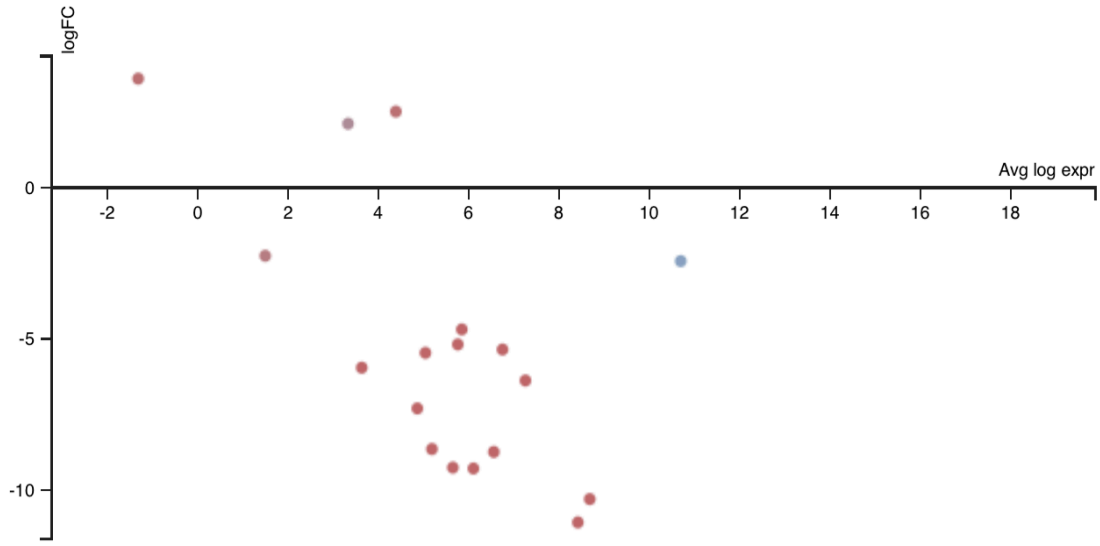
The MA plot **(Fig 4.6A&B)** visualized the different gene expression levels of the choline and non-choline conditions, by transforming the measured data of these two conditions onto M-values and A-values. M-values (y axis) are the log intensity ratio of the two conditions (log fold-change), whereas A-values (x axis) are the log intensity averages of the two conditions (average of log expression value). The red dots represent the FDR value < 0.05 while the blue indicates FDR value ≥ 0.05. **Fig 4.6A** shows the distribution of genes at FDR cut off of 1, and **Fig 4.6B** shows at $\log_2$-fold change with a more stringent cut off of FDR <0.05, there are 18 highly expressed genes that are statistically significant, 17 genes are indicated in red with FDR value less than 0.05, while only one gene is shown in blue with the FDR value 0.05.

The details of these 18 genes are listed in **Table 4.3**. Positive log-fold change indicates down regulation with the addition of choline and negative values indicate up regulation with choline. The highest log-fold change is -11.07 for PMI2721, a gene annotated as a *pduA/J* microcompartment protein, followed by-10.30 (PMI2720), -9.29 (PMI2722), -9.26 (PMI2718), and -5.35 (PMI2714). The high log-fold changes of the five putative bacterial microcompartment shell proteins suggest that the microcompartment formation is highly correlated to the presence of choline.

The parallel coordinates plot **(Fig 4.7)** was used to visualise the top 18 most highly differentially expressed genes and how expression varied between the choline supplemented culture and glucose only culture, at $\log_2$-fold change with FDR cut off of 0.05. Each condition has three replicates. Each line in **Fig 4.7A** shows the difference of expression in log-fold change values between the two conditions. Positive log-fold change values (above zero on the right axis) indicate up-regulation of glucose (down regulation with choline), and negative values (below zero on the right axis) indicate down-regulation of glucose (up regulation with choline). The red lines show the genes with FDR < 0.05 and only one blue line shows one gene with FDR ≥ 0.05. Gene numbers were shown on the right axis. **Fig 4.7B** is a heatmap indicating the different expression levels of the 18 genes between the two conditions. Each block represents one gene. The bar legend indicates the positive log-fold change with red colour and the negative log-fold change with blue colour. The top 18 blocks on the heatmap is choline grown condition, showing an inceseaing up-regulation of choline from left to right (gene no. 18 to 1), while the bottom 18 bolcks is glucose grown condition, showing a decreasing down-regulation of glucose from left to right (gene no. 18 to 1). Gene numbers were shown on top of the heatmap with gene annotations corresponding to **Table 4.3**.



(A)



(B)

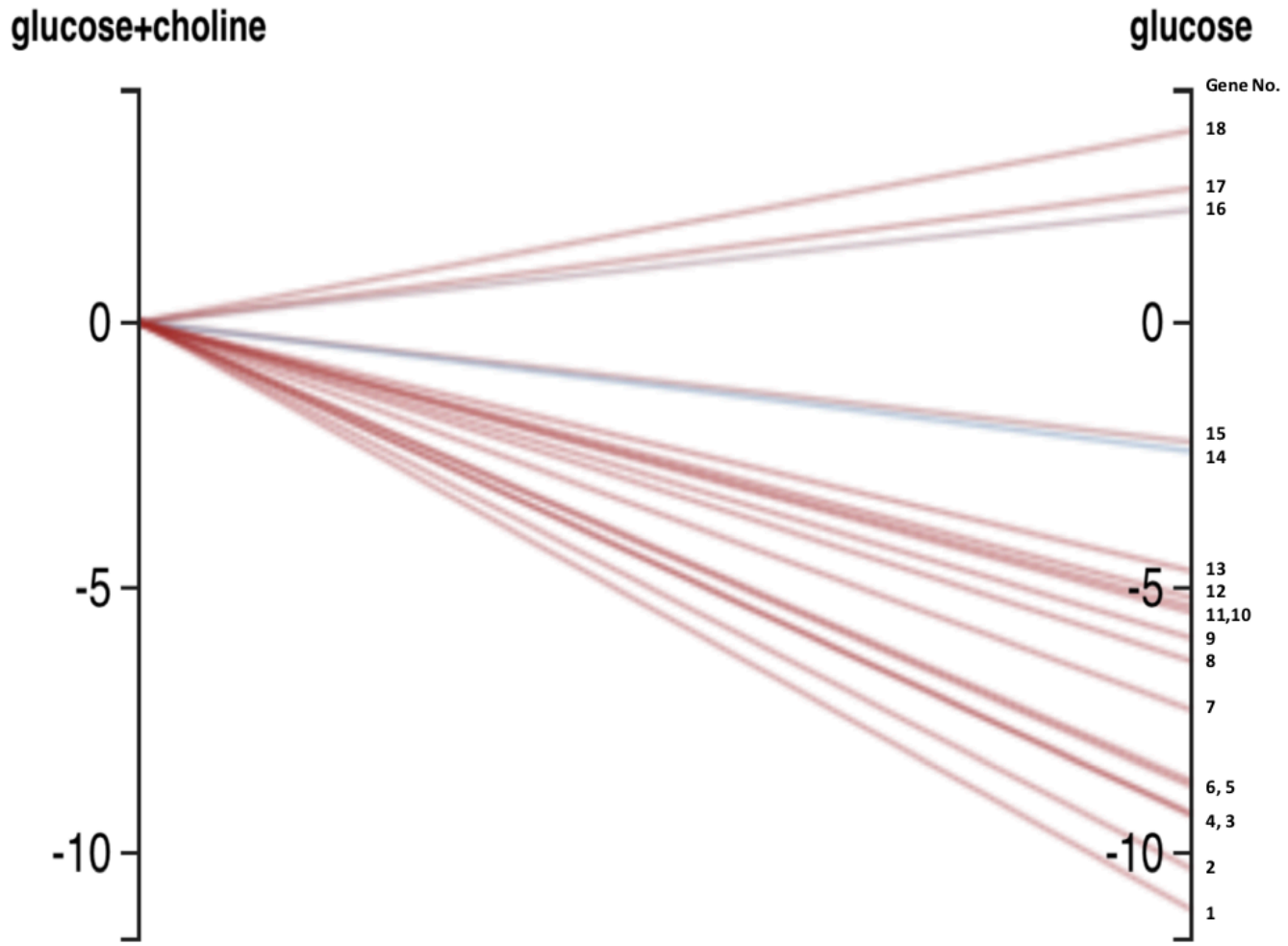
Figure 4.6 shows the MA plot of the RNA-Seq data for the two conditions **(A)** at FDR cut off of 1 **(B)** at FDR cut off of 0.05. The red dots represent the FDR value < 0.05 and the blue indicates FDR value ≥ 0.05. The x axis represents average log expression and the y axis represents log fold change. The analysis was carried out using the Degust program (http://www.vicbioinformatics.com/degust/index.html).

.

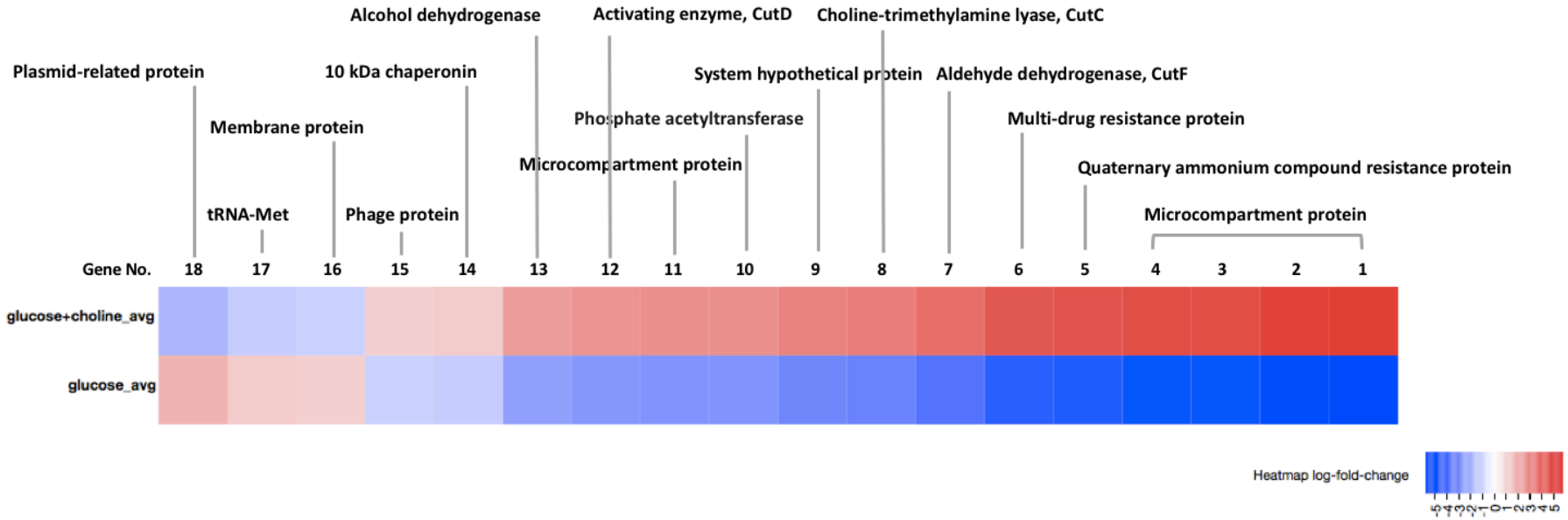
Table 4.3 Top 18 differentially expressed genes. Positive log_2-fold change indicates down regulation with the addition of choline and negative values indicate up regulation with choline (this study).

No.	Locus tag	Gene name	Log_2-fold change	FDR ($p < 0.05$)	Gene description
1	PMI2721	*eutM*/*pduA/J* homologue	-11.07	1.15E-47	Microcompartment protein (BMC domain, PF00936.14)
2	PMI2720	*eutM*/*pduA/J* homologue	-10.30	6.67E-31	Microcompartment protein (BMC domain, PF00936.14)
3	PMI2722	*eutM*/*pduA/J* homologue	-9.29	4.24E-41	Microcompartment protein (BMC domain, PF00936.14)
4	PMI2718	*eutN*/*ccmL* homologue	-9.26	9.61E-34	Microcompartment/carboxysome/ethanolamine utilizatin protein (EutN_CcmL, PF03319.8)
5	PMI2711	*emrE* homologue	-8.74	2.29E-33	Quaternary ammonium compound resistance protein
6	PMI2710	*emrE* homologue	-8.65	7.16E-26	Multi-drug resistance protein, EmrE homologue
7	PMI2719	*cutF*	-7.30	7.30E-37	Aldehyde dehydrogenase, CutF
8	PMI2716	*cutC*	-6.38	1.34E-28	Choline-trimethylamine lyase, CutC
9	PMI2712		-5.96	3.35E-14	System hypothetical protein
10	PMI2713	*cutH*	-5.47	2.10E-17	Phosphate acetyltransferase
11	PMI2714	*eutM*/*pduA/J* homologue	-5.35	1.91E-19	Microcompartment protein (BMC domain, PF00936.14)

12	PMI2715	*cutD*	-5.19	7.26E-27	Choline-trimethylamine lyase, activating enzyme, CutD
13	PMI2717	*cutO*	-4.69	1.08E-16	Alcohol dehydrogenase
14	PMI2544	*groS*	-2.43	0.05	10 kDa chaperonin
15	PMI0908		-2.26	0.01	Phage protein
16	PMI0178		2.11	0.03	Membrane protein
17	PMIt053		2.51	5.42E-03	tRNA-Met
18	PMI2552		3.60	5.42E-03	Plasmid-related protein


glucose+choline
glucose
Gene No.
18
17
16
15
14
13
12
11,10
9
8
7
6, 5
4, 3
2
1
0
-5
-10
0
-5
-10

(A)


Plasmid-related protein
Membrane protein
tRNA-Met
10 kDa chaperonin
Phage protein
Alcohol dehydrogenase
Activating enzyme, CutD
Microcompartment protein
Phosphate acetyltransferase
System hypothetical protein
Choline-trimethylamine lyase, CutC
Aldehyde dehydrogenase, CutF
Multi-drug resistance protein
Quaternary ammonium compound resistance protein
Microcompartment protein
Gene No. 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
glucose+choline_avg
glucose_avg
Heatmap log-fold-change
5 4 3 2 1 0 1 2 3 4 5

(B)

Figure 4.7 (A&B) Parallel coordinates plot of the RNA-Seq for two conditions, glucose with choline and without choline at log_2-fold change, indicating the top 18 highly expressed genes (down regulation with choline with positive log-fold change in purple colour, up regulation with choline with negative log-fold change in red colour). (**A**) Parallel coordinates plot, each line shows the change in expression in one gene, between choline and glucose only. (**B**) Heatmap, each block represents one gene. Each condition has three biological replicates. The analysis was carried out using the Degust program (http://www.vicbioinformatics.com/degust/index.html).

4.2.5 Identification a choline-TMA lyase gene cluster in *Proteus mirabilis*

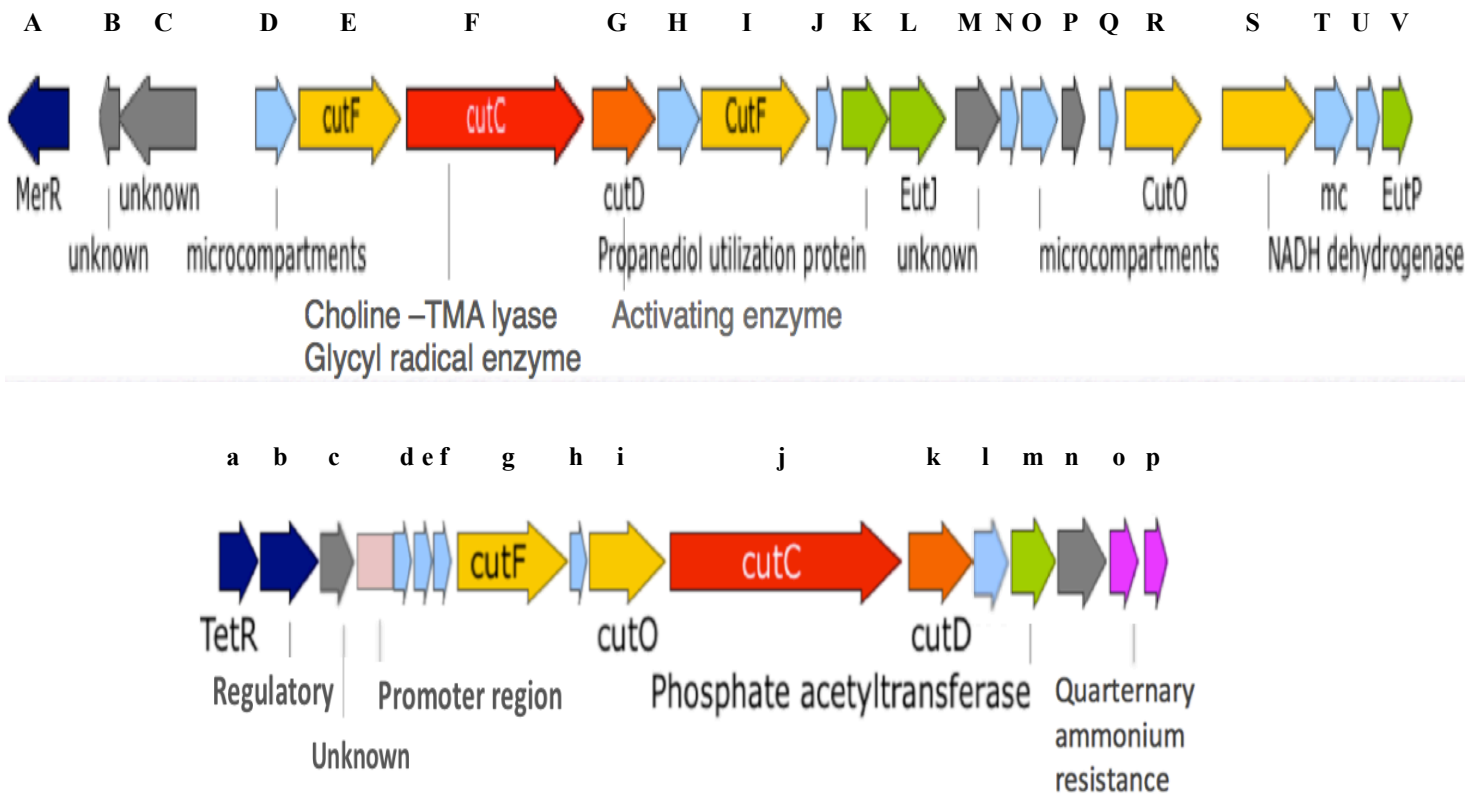
The top 13 genes identified by Degust **(Table 4.3)** were exactly the same as ones identified by DESEQ **(Table 4.2)**, confirming that these genes are highly expressed in the presence of choline. Many of these genes form a putative gene cluster in the genome, suggesting they have a key role in the conversion of choline to TMA. Therefore, further phylogenetic and genome analysis of this gene cluster were carried out.

One of the most highly induced genes from the RNA-Seq analyses is annotated as a glycyl radical containing protein (PMI2716). While this PhD project was on-going, the enzyme responsible for the conversion of choline to TMA was identified from *Desulfovibrio desulfuricans* (Cracium and Balskus, 2012). This choline-TMA lyase enzyme has high sequence similarity to PMI2716 and also belongs to the glycyl radical enzyme family. Using the online IMG program hosted by the Joint Genome Institute (http://img.jgi.doe.gov/), the two genes clusters were compared **(Fig 4.8A)**. Functional annotation of these two gene clusters are also carried out by IMG, which are summarized in **Table 4.4** and **Table 4.5**.

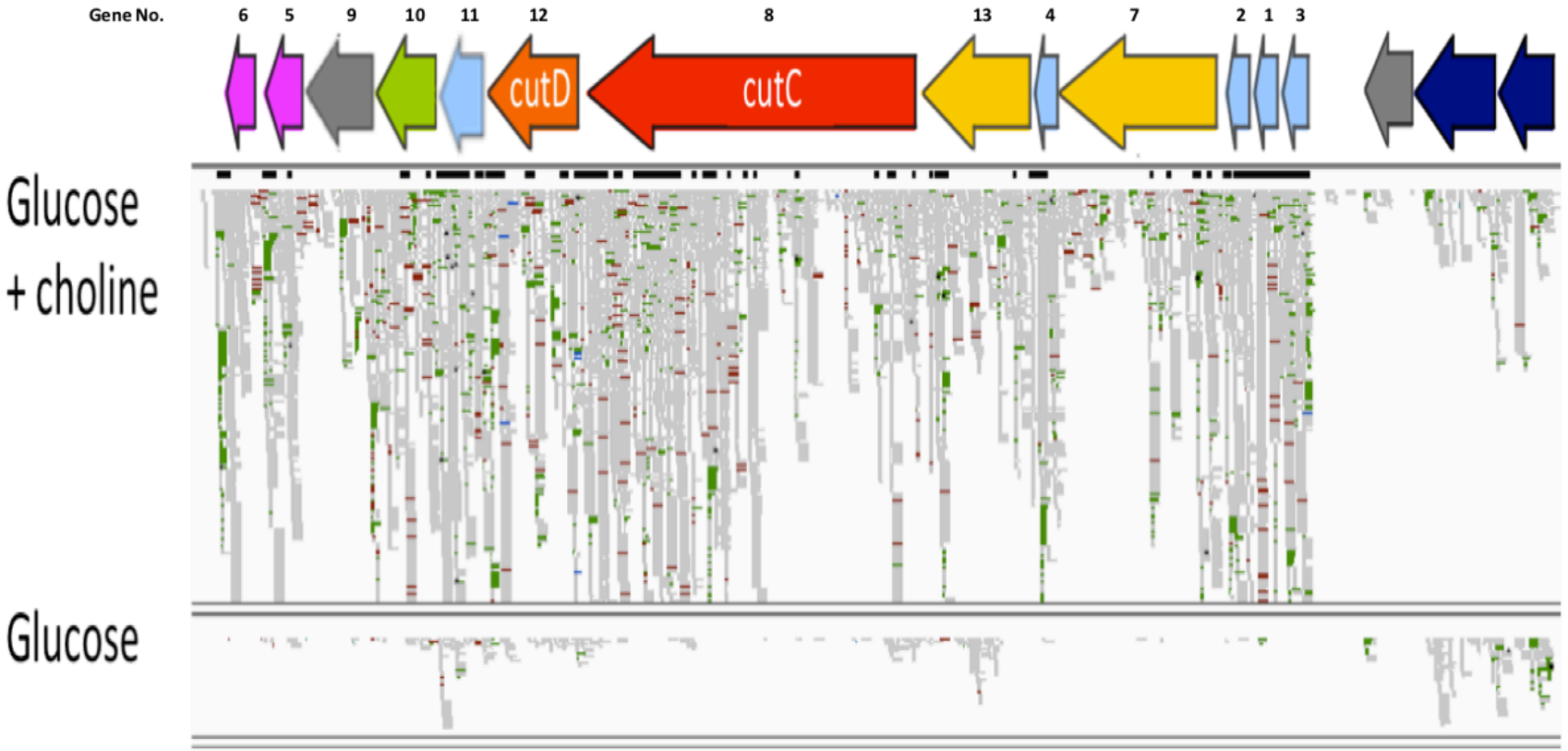
The 13 top expressed genes in the *cutC* gene cluster, and their gene number are visualised in **Fig 4.8B** and cross-referenced to **Table 4.3**. Among them, the microcompartment proteins (gene numbers 1, 2, 3 and 4) are the most highly expressed genes within the *cutC* gene cluster, in the presence of choline. There is a significant difference in the total read counts and expression level between the choline-grown and glucose-grown conditions for these shell proteins, with total reads 59860, 70485, 12249, 8116 in choline-grown culture, compared to 21, 42, 14, 10 in glucose-grown culture **(Table 4.6)**.

Based on the RNA-Seq data and the work on *D. desulfuricans*, the metabolic pathway of choline degradation to TMA is proposed **(Fig 4.9)**. The function of the bacterial microcompartments is proposed to optimize pathways with toxic or volatile intermediates, in this case, acetaldehyde, to channel it to the next pathway enzyme,

converting it to ethanol by an alcohol dehydrogenase (PMI2717, *cutO*) or to acetyl-CoA by an acetaldehyde dehydrogenase (PMI2719, *cutF*) (Max and Lidstrom, 2002).


A B C D E F G H I J K L M N O P Q R S T U V
cutF
cutC
CutF
MerR
unknown
cutD
EutJ
CutO
mc
EutP
unknown
microcompartments
Propanediol utilization protein
unknown
microcompartments
NADH dehydrogenase
Choline –TMA lyase
Glycyl radical enzyme
Activating enzyme
a b c d e f g h i j k l m n o p
cutF
cutC
TetR
cutO
cutD
Regulatory
Promoter region
Phosphate acetyltransferase
Quarternary ammonium resistance
Unknown

(A)



(B)

Figure 4.8 (A) The candidate gene clusters involved in choline degradation in the genome of *D. desulfuricans* (upper) and *P. mirabilis* (lower) predicted by the online IMG program at the Joint Genome Institute (http:// img.jgi.doe.gov/) **(B)** shows the putative microcompartment proteins (highlighted in light blue, gene No. 1, 2, 3, and 4) are most highly expressed on the *cutC* gene cluster with the presence of choline. Gene No. represents the top 13 genes in **Table 4.3 (**picture adapted from the Hi-Seq Illumina raw RNA-Seq from Dr Jameson during the course of this study).

Table 4.4 The predicted functions of the choline-TMA lyase gene cluster in *D. desulfuricans.*

Gene	Locus Tag	Homology-based function prediction	Proposed function in choline degradation
A	Ddes_1352	MerR-family transcriptional regulator	pfam06445- GyrI-like, pfam13411- MerR_1
B	Ddes_1353	system hypothetical protein	accession no.YP_002479933
C	Ddes_1354	system hypothetical protein	accession no.YP_002479934
D	Ddes_1355	microcompartments protein	microcompartment shell protein
E	Ddes_1356	aldehyde-alcohol dehydrogenase	aldehyde dehydrogenase
F	Ddes_1357	Formate C-acetyltransferase	Pyruvate-formate lyase
G	Ddes_1358	glycyl-radical enzyme activating protein family	Pyruvate-formate lyase-activating enzyme
H	Ddes_1359	microcompartments protein	microcompartment shell protein
I	Ddes_1360	acetaldehyde dehydrogenase (acetylating)	aldehyde dehydrogenase
J	Ddes_1361	microcompartments protein	microcompartment shell protein

K	Ddes_1362	Propanediol utilization protein	accession no.YP_002479942
L	Ddes_1363	ethanolamine utilization protein EutJ family protein	accession no.YP_002479943
M	Ddes_1364	system hypothetical protein	accession no.YP _002479944
N	Ddes_1365	Ethanolamine utilization protein EutN	accession no.YP _002479945
O	Ddes_1366	microcompartments protein	microcompartment shell protein
P	Ddes_1367	system hypothetical protein	accession no.YP_002479947
Q	Ddes_1368	microcompartments protein	microcompartment shell protein
R	Ddes_1369	iron-containing alcohol dehydrogenase	alcohol dehydrogenase
S	Ddes_1370	Respiratory-chain NADH dehydrogenase	accession no.YP_002479950
T	Ddes_1371	microcompartments protein	microcompartment shell protein
U	Ddes_1372	microcompartments protein	microcompartment shell protein
V	Ddes_1373	ethanolamine utilization protein EutP	accession no.YP_002479953

Table 4.5 The predicted functions of the choline-TMA lyase gene cluster in *P. mirabilis.*

Gene	Locus Tag	Homology-based function prediction	Proposed function in choline degradation
a	PMI2725	TetR-family transcriptional regulator	pfam00440 - TetR_N
b	PMI2724	regulatory protein	pfam00486 - Trans_reg_C
c	PMI2723	system hypothetical protein	accession no.YP_002152424
d	PMI2722	microcompartments protein	microcompartment shell protein
e	PMI2721	microcompartments protein	microcompartment shell protein
f	PMI2720	microcompartments protein	microcompartment shell protein
g	PMI2719	aldehyde-alcohol dehydrogenase	acetylaldehyde dehydrogenase (*cutF*)
h	PMI2718	system hypothetical protein	microcompartment shell protein
i	PMI2717	propanediol utilization protein (alcohol dehydrogenase)	alcohol dehydrogenase (*cutO*)
j	PMI2716	propanediol utilization protein (dehydratase)	choline-TMA lyase (*cutC*)

k	PMI2715	propanediol utilization protein (dehydratase activating enzyme)	choline-TMA lyase, activating enzyme (*cutD*)
l	PMI2714	propanediol utilization protein	microcompartment shell protein
m	PMI2713	propanediol utilization protein	Phosphate acetyltransferase, acetyl-CoA => acetate
n	PMI2712	system hypothetical protein	accession no.YP_002152413
o	PMI2711	quaternary ammonium compound resistance protein	pfam00893 - Multi_Drug_Res
p	PMI2710	multidrug resistance protein	pfam00893 - Multi_Drug_Res

Table 4.6 The read counts data of the highly expressed genes on the *cutC* gene cluster (Reads Per Kilobase Million, RPKM)

No.	**Locus Tag**	Glucose + choline				Glucose			
		Rep1	Rep2	Rep3	Total reads	Rep1	Rep2	Rep3	Total reads
1	PMI2721	26264	18171	15425	59860	7	3	11	21
2	PMI2720	26231	19631	24623	70485	29	3	10	42
3	PMI2722	6060	2342	3847	12249	9	1	4	14
4	PMI2718	1540	3256	3320	8116	6	3	1	10
5	PMI2711	6533	2473	7286	16292	10	11	6	27
6	PMI2710	2225	619	3367	6211	3	1	8	12
7	PMI2719	1692	1151	2061	4904	9	3	12	24
8	PMI2716	11786	5488	9285	26559	82	56	100	238
9	PMI2712	338	337	1285	1960	13	4	8	25
10	PMI2713	1574	1083	2729	5386	54	15	24	93
11	PMI2714	5950	4991	6876	17817	173	68	86	327

12	PMI2715	2798	2718	3355	8871	64	46	79	189
13	PMI2717	1767	2971	4280	9018	87	73	121	281

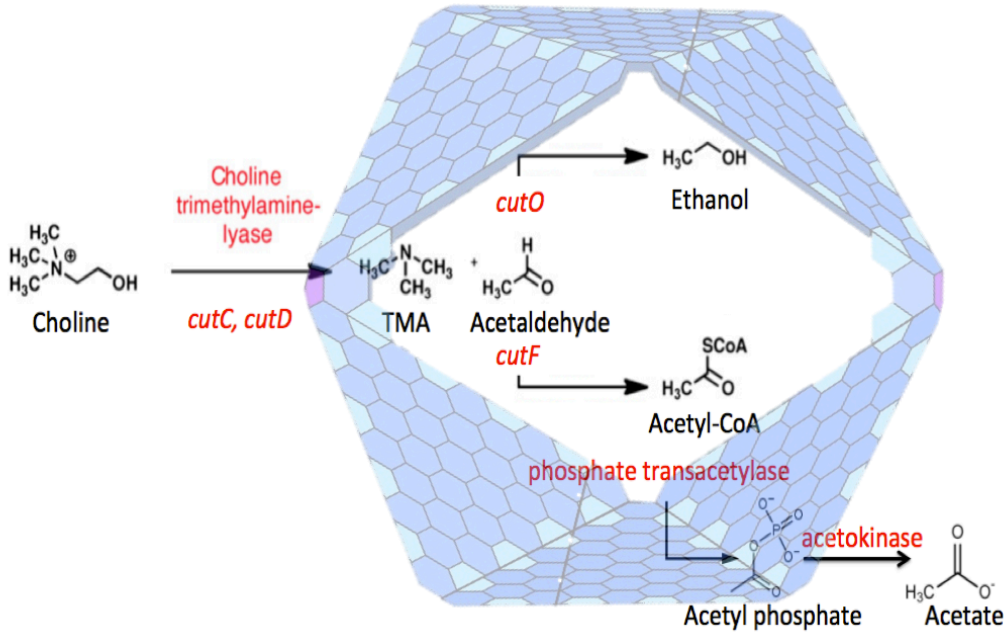


Figure 4.9 The metabolic pathway of choline degradation to TMA and acetaldehyde via the proposed enzymes. Choline is converted to TMA and acetaldehyde by the choline-TMA lyase and the activating protein (PMI2715, *cutC* and PMI2716, *cutD* respectively). Acetaldehyde is further metabolized inside the microcompartment to ethanol by an alcohol dehydrogenase (PMI2717, *cutO*) or to acetyl-CoA by an acetaldehyde dehydrogenase (PMI2719, *cutF*) (Max and Lidstrom, 2002) (picture modified from Craciun and Balskus, 2012, Chowdhury *et al.* 2014, free online open access).

The identification of choline TMA lyase (CutC) offers the opportunity to further understand the distribution of this novel metabolic pathway in bacteria. Genome analysis was therefore carried out to investigate the phylogeny and distribution of CutC in genome-sequenced bacteria using the IMG database. A neighbour-joining phylogenetic tree was constructed, which is shown in **Fig 4.10**. As reported by Martínez-del Campo *et al.* (2015) recently, phylogenetic analysis indicates that there are two distinct clusters of CutC proteins. The first cluster, designated as type I CutC, is mainly found in obligate anaerobes including *D. desulfuricans* (Craciun and Balskus, 2012). The second, type II CutC cluster, which contains facultative anaerobes including *P. mirabilis* and many isolates of the *Enterobacteriaceae* family, is further divided into two subclades, II.a containing *Gammaproteobacteria* and II.b including *Firmicutes* and *Deltaproteobacteria*.

Multiple sequence alignment of type I CutC versus type II CutC have revealed differences between these *cutC* genes in their length: the type I and II.b *cutC* gene are approximately 2.5 kb, while the type II.a is about 3.4 kb. The translated CutC of type II.a therefore is almost 300 amino acids longer **(Fig 4.11)**. Despite the notable difference in N terminus of CutC variants, the rest of the CutC sequence is highly conserved.


Benzylsuccinate synthase
1-methyl alkyl succinate synthase
4-Hydroxyphenylacetate decarboxylase
B12-independent glycerol dehydrogenase
Choline TMA-lyase type II.b (cluster 2.2 GRM1)
Uncharacterized CPFL
Ribonucleotide reductase
Pyruvate formate-lyase
Choline TMA-lyase type II.a (cluster 2.3 GRM2)
Choline TMA-lyase type I (cluster 2.2 GRM1)
C. difficile 630
T. forsythia ATCC 43037
H. foetida DSM 6591
D. tolulica DSM 7467
G-9B TRM1
Geobacter sp. 7TMJ1
Azoarcus sp. HxN1
E. coli 101-1
C. bolteae ATCC BAA-613
S. putrefaciens CN-32
R. obeum ATCC 29174
S. mutans UA159
S. suis 98HAH33
L. plantarum WCFS1
V. harveyi HY01
H. influenzae Rd KW20
B. coahuilensis m4-4
L. monocytogenes HPB2262
A. marina MBIC11017
S. thermophilum IAM 14863
C. aerofaciens ATCC 25986
A. caccae DSM 14662
B. plebeius DSM 17135
Desulfovibrio alaskensis G20
Pelobacter carbinolicus DSM 2380
Desulfotomaculum reducens MI-1
Desulfotomaculum ruminis DSM 2154
Desulfitobacterium hafniense DP7
Desulfitobacterium dehalogenans ATCC 51507
Desulfosporosinus youngiae DSM 17734
Vibrio furnissii NCTC 11218
Proteus penneri ATCC 35198
Providencia rettgeri DSM 1131
Providencia rustigianii DSM 4541
Providencia alcalifaciens DSM 30120
Yokenella regensburgei ATCC 43003
Escherichia fergusonii B253
Escherichia coli MS 200-1
Aeromonas hydrophila ATCC 7966
Klebsiella oxytoca 10-5246
Enterobacter blattae DSM 4481
Klebsiella pneumoniae ATCC 13884
Enterobacter aerogenes KCTC 2190
Proteus mirabilis HI4320
Paenibacillus dendritiformis C454
Streptococcus intermedius SK54
Streptococcus dysgalactiae ATCC 12394
Olsenella uli DSM 7084
Clostridium tanakaei YIT 12063
Anaerococcus vaginalis ATCC 51170
Anaerococcus hydrogenalis DSM 7454
Desulfovibrio desulfuricans G20
Desulfovibrio desulfuricans ATCC 27774
Desulfotalea psychrophila LSv54
Desulfovibrio salexigens DSM 2638
Alkaliphilus oremlandii OhILAs
Alkaliphilus metalliredigens
Clostridium tetani E88
Clostridium carboxidivorans P7
Clostridium sporogenes ATCC 15579
Clostridium botulinum ATCC 3502
Clostridium ljungdahlii DSM 13528
Clostridium saccharolyticum WM1
Clostridium phytofermentans ISDg
Clostridium hathewayi DSM 13479
Clostridium asparagiforme DSM 15981
Clostridium citroniae WAL-17108
Clostridium bacterium 1_7_47_FAA
100
100
100
80
100
0.1

Figure 4.10 shows the relationship of choline TMA-lyase relative to other glycyl-radical enzymes of the family. The phylogenetic tree was constructed using the amino acid sequences using the neighbour-joining method. Two distinct clusters of CutC were observed, which has also been reported by Martínez-del Campo *et al.* previously (2015). Type I and II.b sub-clusters are approximately 2.5 kb in length whereas Type II.a sub-cluster is approximately 3.4 kb in length. The scale bar presented the evolutionary distance in mutations per residue. 1000 replicates were used with bootstrap values > 70 % (figure taken from Jameson *et al.*, 2015).

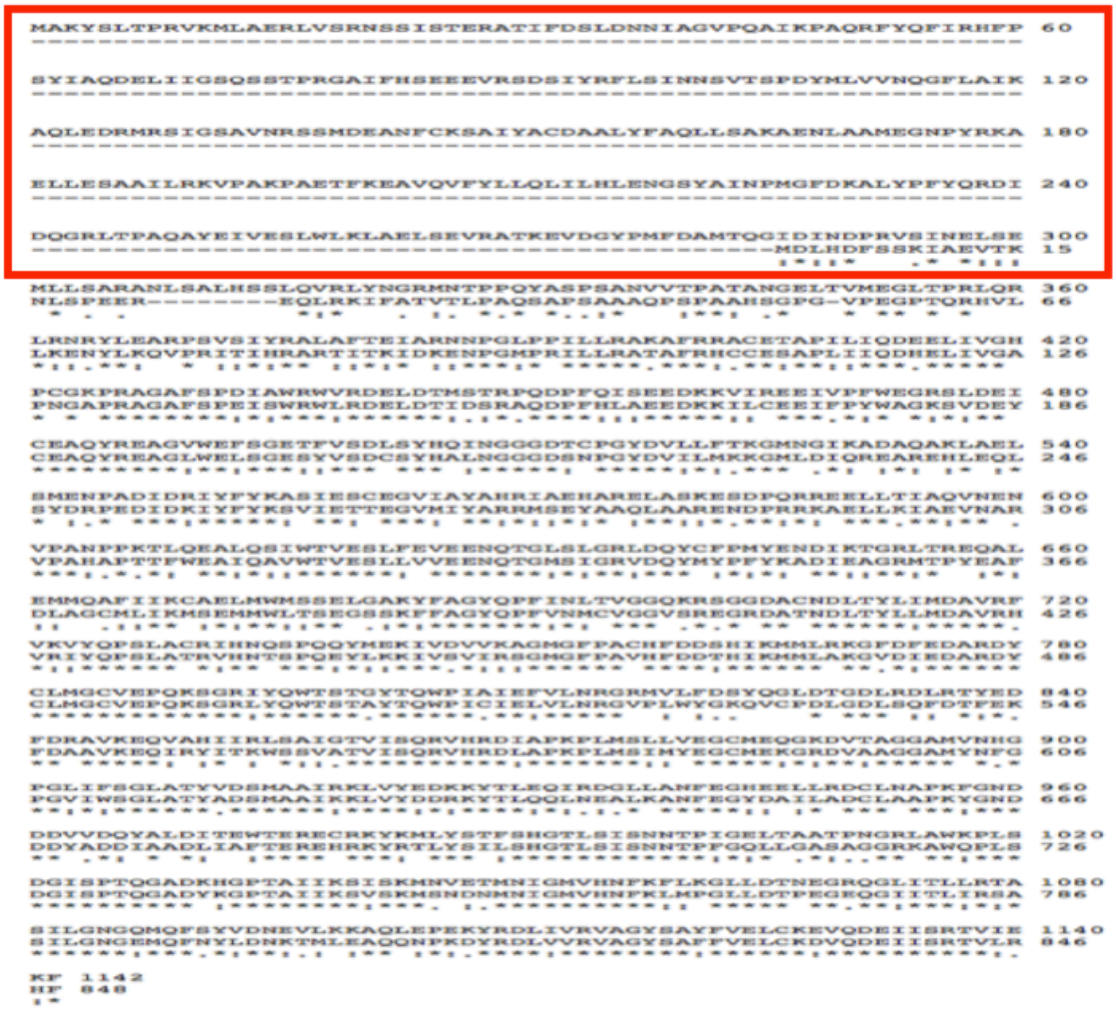


Figure 4.11 The CutC sequence alignment of *P. mirabilis* (upper) and *D. desulfuricans* (lower). The red box shows the extended N-terminus of 300 amino acids in *P. mirabilis*.

4.3 Discussion

RNA-Seq was employed in this chapter in order to identify the genes that were differentially expressed between culture grown in the presence or absence of choline, in the model bacterium, *P. mirabilis* DSMZ4479. Two independent statistical analyses have revealed a novel gene cluster involved in choline metabolism, including the newly discovered choline-TMA lyase encoded by *cutC/D* and several genes encoding microcompartment shell proteins.

I have therefore chosen CutC as the functional maker to investigate choline-to-TMA transformation in genome-sequenced bacteria. Phylogenetic analyses have shown that many isolates of *Gammaproteobacteria* (particularly *Enterobacteriaceae*), *Firmicutes* and *Deltaproteobacteria* have the genetic potential for choline degradation to TMA **(Fig 4.10)**. Interestingly, multiple sequence alignment of CutC homologues and subsequent phylogenetic analyses have revealed the presence of two clades of CutC. The type II.a CutC found in *Enterobacteriaceae*, including *P. mirabilis*, is longer than the type I or II.b CutC, such as found in the obligate anaerobe *D. desulfuricans*, with an additional ~300 amino acids at the N-terminus **(Fig 4.11)**. The function of the extra N-terminal of type II.a CutC is still unknown. From the genome analysis it suggests that type II.a CutC was only identified from facultative anaerobes. Thus, it is hypothesised that this additional N-terminus of CutC may play a role in stabilising the enzyme during transient oxygen exposure, which would allow facultative anaerobes to thrive on the intestinal or urinary tract epithelia. Therefore, further work has been carried out to investigate the extra N-terminal in type II.a CutC **(See Chapter 5)**.

The proposed microcompartments in choline utilization comprise five different proteins in the *cut* gene cluster: PMI2714, PMI2718, PMI2720, PMI2721 and PMI2722 **(Fig 4.8 & Table 4.6)**. Except for PMI2718, the other four microcompartment shell proteins (PMI2714, PMI2720, PMI2721, PMI2722) are homologues to the shell proteins PduJ/PduA in the 1,2-propanediol utilization (Pdu) microcompartments and EutM in the ethanolamine utilization (Eut) compartment (Jameson *et al*., 2015). RNA-Seq data shows that these microcompartment shell proteins were heavily induced by choline in *P. mirabilis* **(Fig 4.8, Table 4.3 & 4.6)**,

which suggests a link between choline degradation and microcompartment formation. Further investigation was required to confirm the microcompartment protein formation in the presence of choline. Growth experiments were carried out in liquid culture and in swarming cells on agar plates **(See Chapter 5)** to investigate the formation of microcompartments within the cells and their occurrence in response to choline addition in the culture medium. Moreover, I also investigated whether the formation of microcompartments is directly related to the enhancement of growth rate and swarming motility in *P. mirabilis* in the presence of choline **(See Chapter 5)**.

The initial statistic analysis carried out by Jameson *et al.* identified 20 differentially expressed genes identified by DESEQ analysis, with significant $\log_2$-fold change values **(Table 4.2)**. In the DESEQ analysis, *p*-values rather than FDR values was used to estimate the significance level of the differential gene expression. However, *p*-value is only applicable for a set of genome data when most of the genes are regulated or expressed. In the actual experiment, only the *cut* operon was over-expressed, *p*-value, in this case, will calculate the average of expression level by taking the whole genome expression into account, which lead to false positives. Therefore, FDR values should be applied in this case and thus I carried out the new analysis using Bedtool and the Degust program.

The Degust program uses EdgeR as the normalization and statistical calculation method. I carried out RNA-Seq in Degust using EdgeR package. **Fig 4.5** shows the gene expression and distribution under the two growth conditions: glucose and glucose plus choline at FDR cut off of 0.05. The down regulation level is shown as purple lines/dots while the up-regulation level is shown in red. 18 genes were consistently highly expressed in the presence of choline, at $\log_2$-fold change. **Fig 4.4** indicates there was good repeatability between the biological replicate samples. High consistency in the biological replicates reduces variation and increases the statistical power of the analysis, resulting in a more sensitive detection of genes that were differentially expressed between conditions. Although I have reproducible samples, the method EdgeR relies precisely on measuring biological variability to establish the statistical significance of differences in gene expression across conditions.

Among the 18 highly expressed genes identified by the Degust analysis, only one gene is indicated in blue dots with FDR value equal to 0.05, the 10 kDa chaperonin (PMI2544) **(Fig 4.6 & Table 4.3)**. My analyses also revealed that there is one gene encoding phage protein that is up-regulated by choline (PMI0908). Interesting, phage-like particles were observed in *P. mirabilis* during preparation of microcompartments **(See chapter 5)**.

Apart from those that are up-regulated in the presence of choline, three genes were down regulated, which encoded: a membrane protein, tRNA-Met and plasmid-related protein. The role of these proteins in choline metabolism remains to be established. tRNA-Met was found to be highly differentially expressed at FDR cut off of 0.05 **(Table 4.3)**. Genome sequencing of strain HI4320 revealed the presence of seven complete rRNA operons, one 16S-23S-5S-5S operon and six 16S-23S-5S operons (Pearson *et al.*, 2008) and tRNA-Met is usually located nearby these rRNA operons. However, it remains unclear why the bacterium downregulated tRNA-Met during its growth on choline.

The occurrence of an elongated N-terminus of CutC in *P. mirabilis* is intriguing, which may either have a role in the folding of CutC in this bacterium or help *P. mirabilis* to cope with transient oxygen exposure in this facultative anaerobe. Protein folding *in vivo* is vectorial, meaning that the N-terminus of a protein becomes solvent exposed in the cytoplasm before its more C-terminal regions. As a result, it is assumed that spatial and temporal constraints are produced on the nascent polypeptide chain in its folding trajectory so that folding dynamics of the N terminus of a protein will affect the correct formation of the entire polypeptide (Spencer and Barral, 2012). In this case, the extra N-terminus of CutC in facultative anaerobes may enhance the protein folding of CutC in order to minimize the exposure of nascent polypeptide to transient oxygen exposure during protein folding. In strict anaerobes, however, CutC is unlikely to encounter oxygen during CutC expression and it is likely that this extra N-terminus is therefore unnecessary in these bacteria. Therefore, mutagenesis of this extra N-terminus would be necessary to further confirm this hypothesis.

PMI2725, encoding the *tetR*-family transcriptional regulator, is located in the 5' of the *cutC* gene cluster **(Fig 4.8 & Table 4.5)**. I postulated that the *tetR* gene plays an

important role in choline degradation in *P. mirabilis*, functioning as a negative repressor. When choline is not present, it suppresses the transcription of the whole *cut* gene cluster. However, in the presence of choline, TetR will bind to choline, and thus the whole gene cluster will be switched on. Besides, the structure and folding of tRNA may change to assist the regulation of the *tetR* with or without choline. To prove the function of this regulator, knock-out mutagenesis and overexpression are required to confirm the role of *tetR* **(See Chapter 5)**.

Chapter 5

Anaerobic choline degradation in *Proteus mirabilis*

5.1 Introduction

Choline is a B-type vitamin and plays a fundamental role as an essential nutrient for human beings (Wang *et al.*, 2011). Our daily demand for choline (~70 %) largely comes from an extensive range of dietary intake such as red meat, milk, eggs, liver, fish and shell fish (Wang *et al.*, 2011). However, current research has demonstrated that choline degradation to TMA by gut microbiota will lead to the accumulation of TMAO in the human liver by the hepatic enzyme FMO_3 and can lead to further cardiovascular and non-alcoholic liver disease (Wang *et al*., 2011; Dumas *et al*., 2006).

Yet the genes responsible for encoding this choline utilization to TMA were unknown at the start of this PhD project. Using comparative transcriptomics analysis, I have identified a gene cluster that is highly induced by choline **(See Chapter 4)**. The aim of the work presented in this Chapter is therefore to identify the functional genes involved in the choline degradation to TMA using *P. mirabilis* as the model system through targeted mutagenesis. Using the marker exchange mutagenesis method established in **Chapter 3**, I investigated whether the *cutC* gene is indeed essential for choline metabolism and subsequent TMA production in *P. mirabilis*. In addition, I investigated the physiological role of choline degradation to TMA in this bacterium in both liquid culture and on solid agar plates. The data presented in this Chapter demonstrated that choline enhances *P. mirabilis* growth in broth culture, and plays a greater role in swarming where it significantly enhances swarming speed and reduces the necessity for consolidation. Furthermore, I observed the formation of microcompartments in this bacterium during its growth on choline and the role of the shell protein in choline degradation, as well as the role of CutC protein in the formation of microcompartment protein.

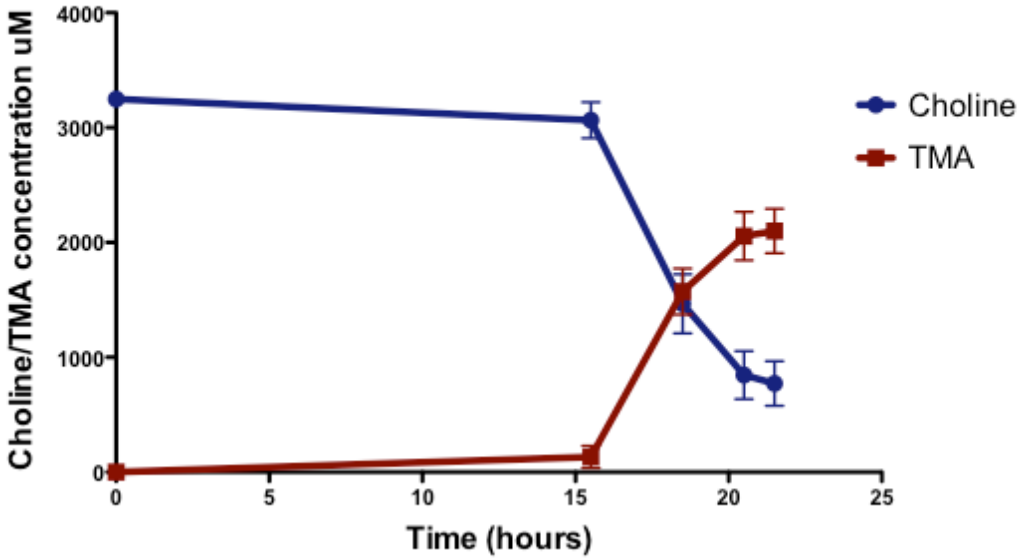
5.2 Results

5.2.1 Heterologous expression of the putative *cutCD* from *P. mirabilis* to *E. coli*

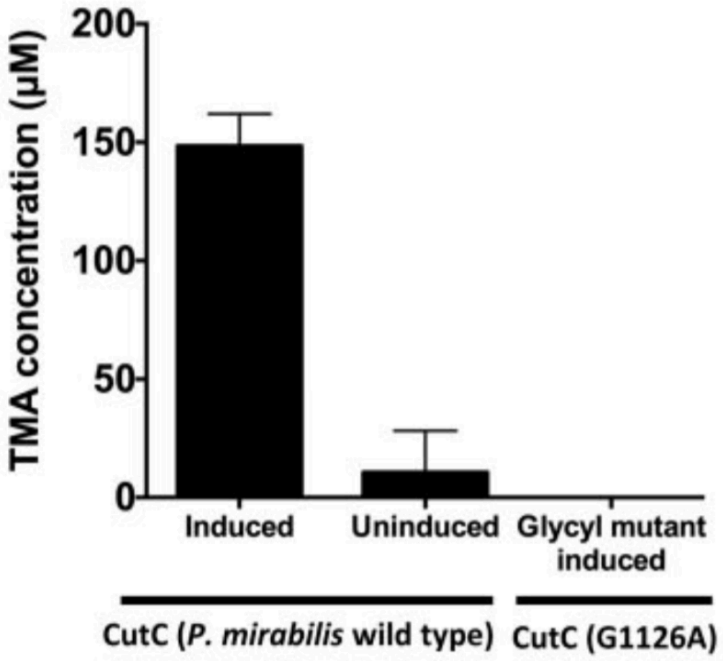
Prior to the start of my PhD project, the initial growth of *P. mirabilis* was carried out by Dr Jameson in our group and the anaerobic degradation of choline to TMA in *P. mirabilis* has been confirmed **(Fig 5.1A)**. The *cutC* gene cluster, including the *cutC* and *cutD* genes, was heavily induced by choline as revealed by comparative transcriptomics (RNA-Seq) **(See Chapter 4)**. Therefore, heterologous expression of the putative *cutCD* genes from *P. mirabilis* to *E.* coli was carried out to test whether *P. mirabilis* CutC and the accompanying activating enzyme CutD were sufficient for TMA formation from choline in a foreign host.

Several initial attempts had been carried out by Dr Jameson and during my project, we collaboratively verified the anaerobic choline-dependent TMA production in recombinant *E. coli* **(Fig 5.1B)**. The codon-optimised putative *cutCD* **(See Chapter 2, Materials and Methods)** was first cloned into *E. coli* under the regulation of the T7-*lac* promoter, and the inducible TMA production was observed anaerobically **(Fig 5.1B)**. Both the *cutC* and *cutD* genes were required to form TMA from choline since *E. coli* expressing either *cutC* or *cutD* alone was unable to degrade choline to TMA **(data not shown)**. To test the function of CutC in choline degradation, the crucial glycine residue that provides the glycyl radical was mutated to alanine by site-directed mutagenesis (CutC mutant: VRVA**A**YSA vs wide type: VRVA**G**YSA). TMA was not detectable in the culture of the site-directed mutant. Therefore, the conserved glycine residue (G1126 in CutC) for this group of radical enzymes is essential **(Fig 5.1B)**.

Together, the data presented above demonstrated that a functional choline-TMA lyase was confirmed to be encoded by the *cutC* homolog in *P. mirabilis*, despite the fact that the CutC of *P. mirabilis* is distinct from the characterized CutC of *D. desulfuricans*, according to the phylogenetic analyses presented in **Chapter 4**.



(A)

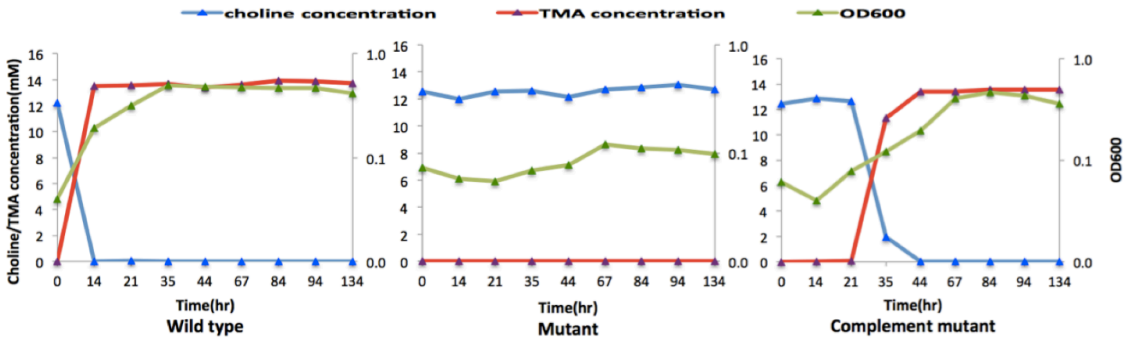


(B)

Figure 5.1 (A) Initial growth of *P. mirabilis* in the defined NH_4^+ medium supplemented with glucose plus choline. Choline and TMA concentrations were quantified by ion-exchange chromatography. **(B)** Quantification of TMA production in recombinant *E. coli* over-expressing both the *cutC*-*cutD* genes, first two bars: wild-type *CutC* with or without induction, third bar: a glycyl radical site CutC mutant (G1126A). Defined M9 medium for *E. coli* was used supplemented with 1 mM choline chloride. 2 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) was added and induction occurred at 5 hours. Three replicate cultures were carried out and error bars represent standard deviations (adapted from Jameson *et al.*, 2015, Creative Commons Attribution-3.0 Unported license, CC BY 3.0).

5.2.2 Marker exchange mutagenesis of the essential *cutC* in TMA formation from choline in *P. mirabilis*

In order to test if *cutC* is required for choline metabolism in *P. mirabilis*, a marker-exchange *cutC* mutant was generated using the method established in **Chapter 3**. The mutant was cultivated in a defined NH_4^+ medium **(See Chapter 2 Materials and Methods)** anaerobically to validate its role in choline metabolism by quantifying TMA production and choline consumption **(Fig 5.2)**. The medium was supplemented with choline as a sole carbon source. The wild type strain was able to metabolize choline to TMA and 10 mM of choline was completely degraded to TMA within 14 hours **(Fig 5.2A)**. As predicted, *P. mirabilis cutC::kan* mutant lost the ability to catalyze TMA formation from choline, TMA was no longer produced and choline remained unused in the culture supernatant **(Fig 5.2B)**. However, when the *cutC::kan* mutant was complemented with the native copy of the *cutCD*, conversion of choline to TMA was restored **(Fig 5.2C)**. Overall, the experiments confirmed that *cutC*D is required for TMA formation from choline.



(A) **(B)** **(C)**

Figure 5.2 Quantification of choline and TMA concentrations and measurement of optical density (at 600 nm) of *P. mirabilis* wild type **(A)**, the *cutC::kan* mutant **(B)** and the *cutC::kan* complemented mutant **(C)** respectively. The cells were grown anaerobically in a defined NH_4^+ medium with choline as the sole carbon source.

5.2.3 Identification of the physiological role of choline degradation in *P. mirabilis*

5.2.3.1 Choline metabolism promotes anaerobic growth of *P. mirabilis* in liquid broth

To identify the role of choline metabolism in *P. mirabilis*, anaerobic liquid growth experiments were first carried out in NH_4^+ defined medium supplemented with five different carbon sources: glucose, glycerol, choline, glucose plus choline or glycerol plus choline. The optical density at 600 nm (OD_{600}) was recorded as the measurement of growth rate over a time course. As can be seen from **Fig 5.3A**. *P. mirabilis* was able to grow on all of the five conditions utilising choline as a sole carbon/energy source. The bacterium grew fastest on glucose alone, followed by glycerol alone and then choline alone. The additional choline with either glucose or glycerol enhanced the growth rates and increased final cell yields. However, there was no significant difference in the final OD for glucose or glycerol with additional choline and the growth was the slowest when it was grown on choline alone.

5.2.3.2 Choline metabolism promotes anaerobic swarming of *P. mirabilis*

Anaerobic swarming assay of *P. mirabilis* was carried out on swarming agar plates with different carbon sources; glucose, glycerol, glucose plus choline and glycerol plus choline. *P. mirabilis* can swarm anaerobically on glucose or glycerol as a sole carbon source, and differences in swarming-associated colony expansion rates were observed **(Fig 5.3B)**. The fastest colony expansion was found on plates supplemented with glucose, whereas the slowest rate was found when it was grown on glycerol plates. The addition of choline significantly enhanced the colony expansion speed **(Fig 5.3B)**. Unlike a uniform ring pattern formed under aerobic swarming growth, anaerobic swarming of *P. mirabilis* formed a dendritic growth pattern instead, starting from irregular points around the initial inoculation colony **(Fig 5.3C, D, E&F)**. There was no well-defined consolidation zone demarking shorter swimmer cells between swarming intervals, compared to aerobic swarming patterns on rich medium (Williams and Schwarzhoff 1978). The addition of choline also affected the uniformity of the formed colony **(Fig 5.3C, D, E&F)**.

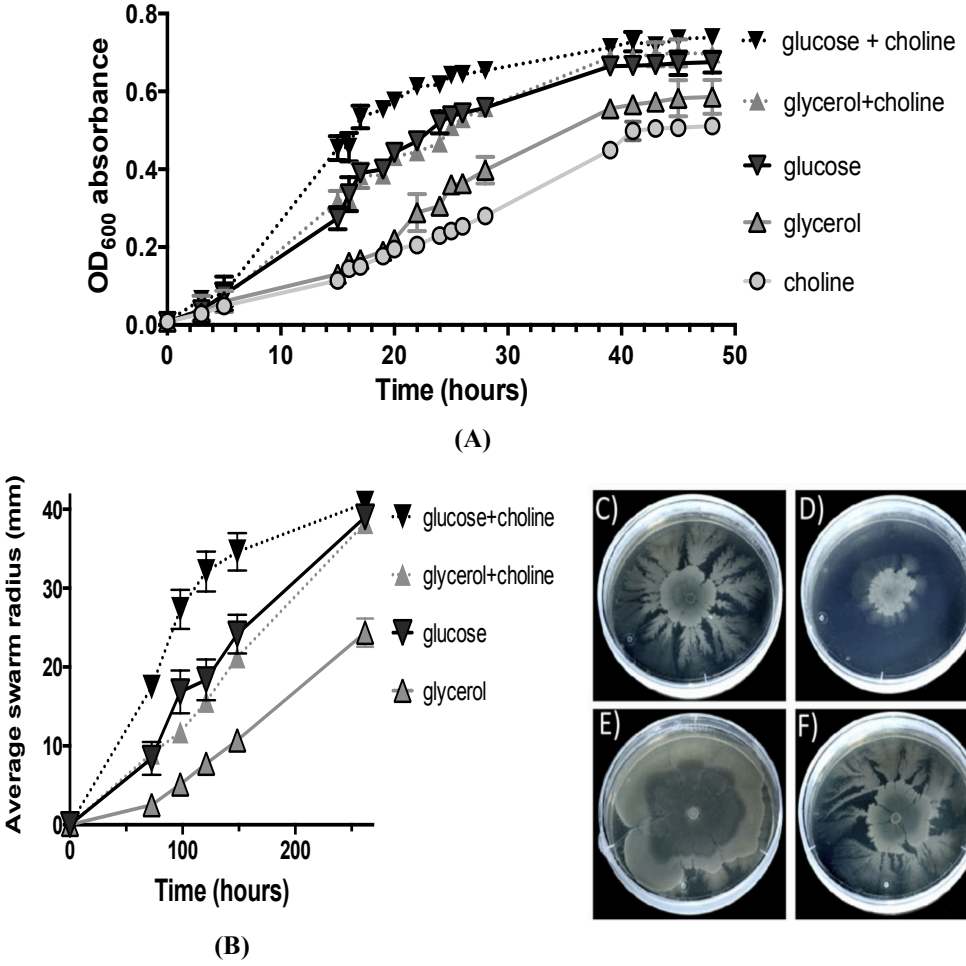


Figure 5.3 (A&B) Anaerobic liquid growth of *P. mirabilis* in the defined NH_4^+medium at 37 °C and anaerobic swarm radiuses of *P. mirabilis* on the swarming agar plates at 30 ºC in anaerobic cabinet on different carbon sources: glucose and choline (dashed ▼), glucose only (solid ▼), glycerol and choline (dashed ▲), glycerol only (solid ▲) and choline only (solid ●). Three replicates were used for each growth experiments and error bars indicate standard deviation **(C, D, E, F)** Anaerobic swarming pattern of *P. mirabilis* on agar plates supplemented with **(C)** glycerol and choline (10 days' growth); **(D)** glycerol (10 days' growth); **(E)** glucose and choline (6 days' growth); **(F)** glucose (10 days' growth) (pictures adapted from supplementary figures from Jameson *et al.*, 2015, Creative Commons Attribution-3.0 Unported license, CC BY 3.0).

5.2.3.3 Choline degradation to TMA enhances anaerobic swarming motility in *P. mirabilis*

Anaerobic swarming assays were carried out in order to determine the role of choline metabolism in this bacterium. A defined medium containing NH_4^+ was used **(See Chapter 2 Materials and Methods)**. Agar plates were inoculated anaerobically with 2 µl liquid culture of *P. mirabilis* of wild type, the *cutC::kan* mutant or the complemented mutant **(Fig 5.4)**. Because the *cutC::kan* mutant does not grow on choline as the carbon source, I supplemented the medium with glycerol.

Wild type *P. mirabilis* strain showed increased colony expansion rates in the presence of additional choline **(Fig 5.4A)**. The *cutC::kan* mutant was still able to grow on glycerol agar plates, however, no difference in colony expansion rates was observed when compared to the growth on plates with additional choline **(Fig 5.4B)**. The complemented mutant, however, showed increased colony expansion rates again on plates supplemented with additional choline **(Fig 5.4C)**. These experiments, therefore, confirmed the role of choline metabolism in enhancing swarming rates in *P. mirabilis*. In order to verify the defective growth of the *cutC::kan* mutant and the complemented mutant, their growth in liquid culture was also carried out. These strains were grown in the defined NH_4^+ medium supplemented with either glycerol, glycerol plus choline or choline alone **(Fig 5.5)**. The results showed no increased growth of the *cutC::kan* mutant on glycerol with additional choline **(Fig 5.5C)**, therefore confirming that *cutC* is essential for choline-dependent growth promotion in this bacterium.

To further explain why additional choline enhanced swarming rates in *P. mirabilis*, growth experiments were carried out to test two hypothesises:

1) Choline degradation provides cellular energy for swarming cells and/or differentiation. To test this hypothesis, the wild type of *P. mirabilis* was grown on NH_4^+ agar plates using 10 mM of glycerol supplemented with different concentrations of choline (1, 10, 50 mM) (**Fig 5.6A**), and swarming radius were determined and compared with increasing concentrations of glycerol (32, 50, 130 mM) as the energy source (**Fig 5.6B**). Interestingly, choline alone shows significant increased swarming rate than glycerol alone. High amounts of choline further

increased the swarming rate while increased amounts of glycerol shows no difference on swarming rates (**Fig 5.6A&B**). This suggests that choline did have the effect to enhance swarming motility and it is likely that choline serves as an energy source instead of carbon source for this bacterium.

2) TMA produced from choline degradation counteracts decrease of pH during glycerol fermentation in liquid culture (Letoffe *et al.*, 2014). In order to assess the impact of pH on swarming cells, phenol red was added to the agar plates as a pH indicator (Alteri *et al.*, 2012). The anaerobic swarming assay was thus carried out using 10 mM of glycerol supplemented with different concentrations of TMA (1, 10, 50 mM) (**Fig 5.6C**), and the swarming radius was compared with 10 mM of glycerol supplemented with different concentrations of choline (1, 10, 50 mM) (**Fig 5.6A**) and different concentrations of glycerol (32, 50, 130 mM) (**Fig 5.6B**). Additional TMA did not increase the swarming rates, suggesting that pH variation could not explain the enhanced rates of choline-promoted swarming (**Fig 5.6A, B&C**).

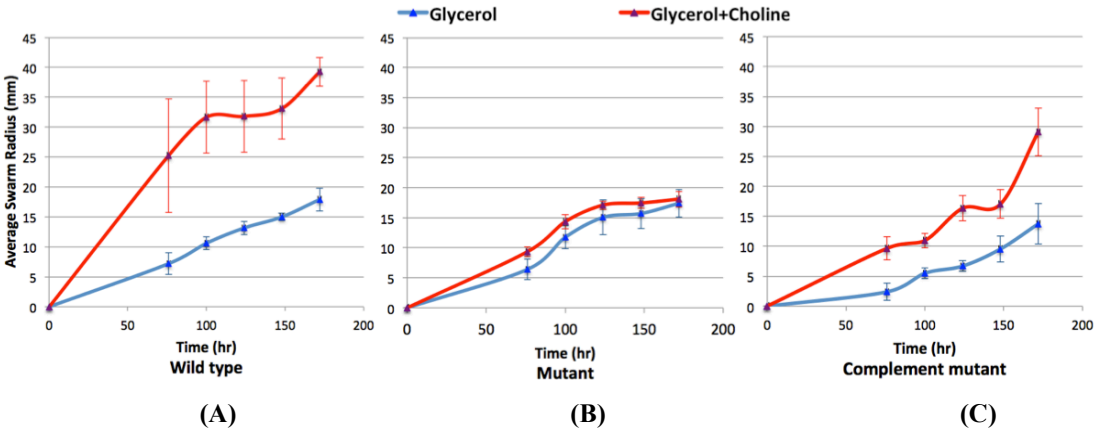


Figure 5.4 Anaerobic swarming radiuses of *P. mirabilis* grown on glycerol or glycerol plus choline for (**A)** wild-type (**B)** the *cutC::kan* mutant and (**C)** the complemented mutant with native *cutCD*. Three replicate plates were used and error bars show standard deviation. The agar plates were incubated at 30 ºC in an anaerobic cabinet.

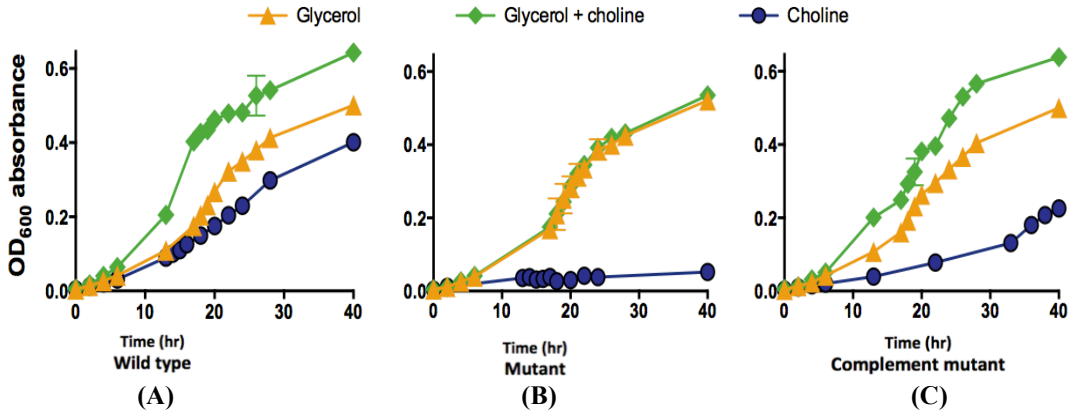


Figure 5.5 Anaerobic growth of *P. mirabilis* in defined NH_4^+ liquid medium supplemented with glycerol, glycerol plus choline or choline alone for **(A)** wild-type **(B)** the *cutC::kan* mutant and **(C)** the *cutC::kan* mutant complemented with native *cutCD* (adapted from supplementary figure from Jameson *et al.*, 2015, Creative Commons Attribution-3.0 Unported license, CC BY 3.0).

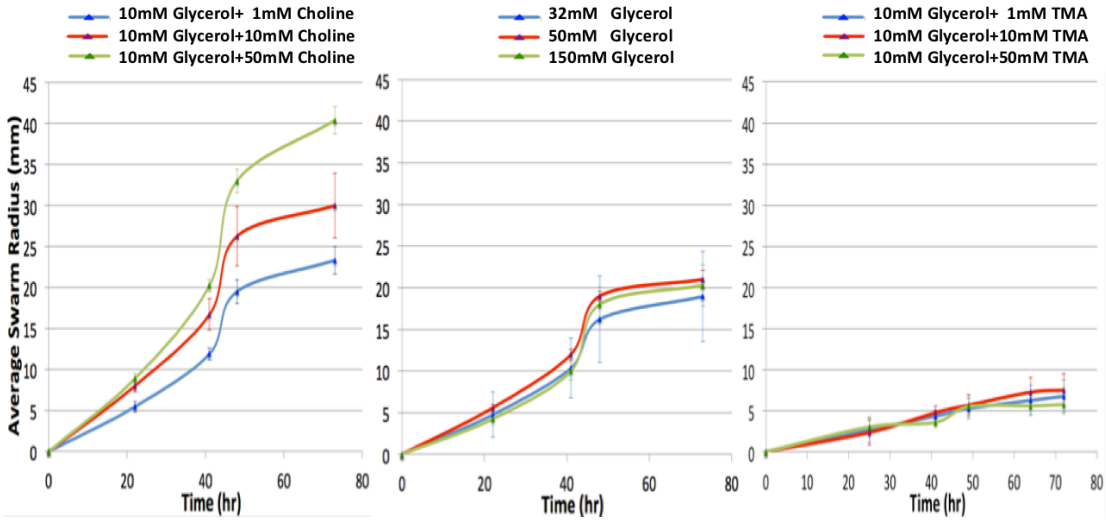


Figure 5.6 Anaerobic swarming for *P. mirabilis* wild type in 10 mM of glycerol supplemented with different concentrations of choline (1, 10, 50 mM) **(A)**, compared with different concentrations of glycerol (32, 50, 130 mM, the total amount of carbon source was same as glycerol plus choline one respectively) **(B)**, and also compared 10 mM of glycerol supplemented with different concentrations of TMA (1, 10, 50 mM) **(C)**. The agar plates were incubated at 30 ºC in an anaerobic cabinet.

5.2.4 Choline induces microcompartment protein formation in anaerobic liquid cultures or on swarming plates of *P. mirabilis*

Comparative transcriptomic analysis in **Chapter 4** has shown that the choline utilisation (*cut*) operon of *P. mirabilis* is heavily induced by choline. In addition to *cutC* and *cutD*, five genes annotated as microcompartment shell proteins: PMI2722, PMI2721, PMI2720, PMI2718 and PMI2714 are among the highly expressed genes that respond to choline addition **(See RNA-Seq data in Chapter 4)**.

Four of the five shell proteins (PMI2722, PMI2721, PMI2720, PMI2714) share high sequence similarity to the shell proteins PduJ/PduA involved in the formation of the 1,2-propanediol utilisation (Pdu) microcompartments and EutM involved in the formation of the ethanolamine utilisation (Eut) microcompartment (Crowley *et al.*, 2010; Kofoid *et al.*, 1999). PMI2718, however, is homologues to the EutN/CcmL family of shell proteins (Kofoid *et al.*, 1999). Previous studies have suggested that choline could be involved in the induction and formation of shell proteins in *cut* clusters of *Desulfovibrio desulfuricans* and *Desulfovibrio alaskensis* (Kuehl *et al.*, 2014, Martinez-del Campo *et al.*, 2015). Therefore, it is likely that choline could also induce the formation of microcompartments in *P. mirabilis*.

5.2.4.1 Microcompartment protein formation in anaerobic liquid cultures supplemented with/without choline

A liquid culture growth experiment was first carried out in order to test the hypothesis of choline-induced microcompartment formation. *P. mirabilis* wild-type was grown in the defined NH_4^+ medium anaerobically supplied with glucose, choline or glucose plus choline as sole carbon sources **(Fig 5.7)**. Cultures were collected at regular time intervals (at 2, 4, 6, 8, 13, 25, 48, 77 hrs) for optical density (OD) measurement and TMA quantification. The OD began to increase at 4-6 hrs and TMA was first detected at 4 hrs in cultures supplemented with choline. Thus, samples from 4 hrs onwards were collected for transmission electron microscopy (TEM) analysis.

TEM images presented in **Fig 5.7** showed that microcompartments were observed (indicated by red arrows) in the presence of choline from 6 hrs onwards. The

appearance of microcompartments coincided with the sharp increase of OD and the production of TMA in cultures supplied with choline. However, in glucose only cultures, no microcompartments were observed. The data therefore suggests that microcompartment formation in anaerobic cultures is induced by choline. For all three media, at 13 and 25 hrs, the cells had a large layer of extracellular polysaccharides (EPS) around, however, at 48 and 77 hrs the EPS layer was either thinner or disappeared.

5.2.4.2 Microcompartment protein formation on anaerobic swarming plates supplemented with/without choline

Similar to the TEM images in anaerobic liquid cultures, no microcompartment proteins were observed in cells cultivated on glycerol swarming plates **(Fig 5.8A)**, and larger amounts of regular BMCs were observed in cells cultivated with choline and glucose plus choline, indicated by red arrows **(Fig 5.8B&C)**. The data again suggests that microcompartment formation in cells cultivated on anaerobic swarming plates was induced by choline.

5.2.4.3 Microcompartment protein formation in the wild type, the *cutC::kan* mutant and the complemented mutant

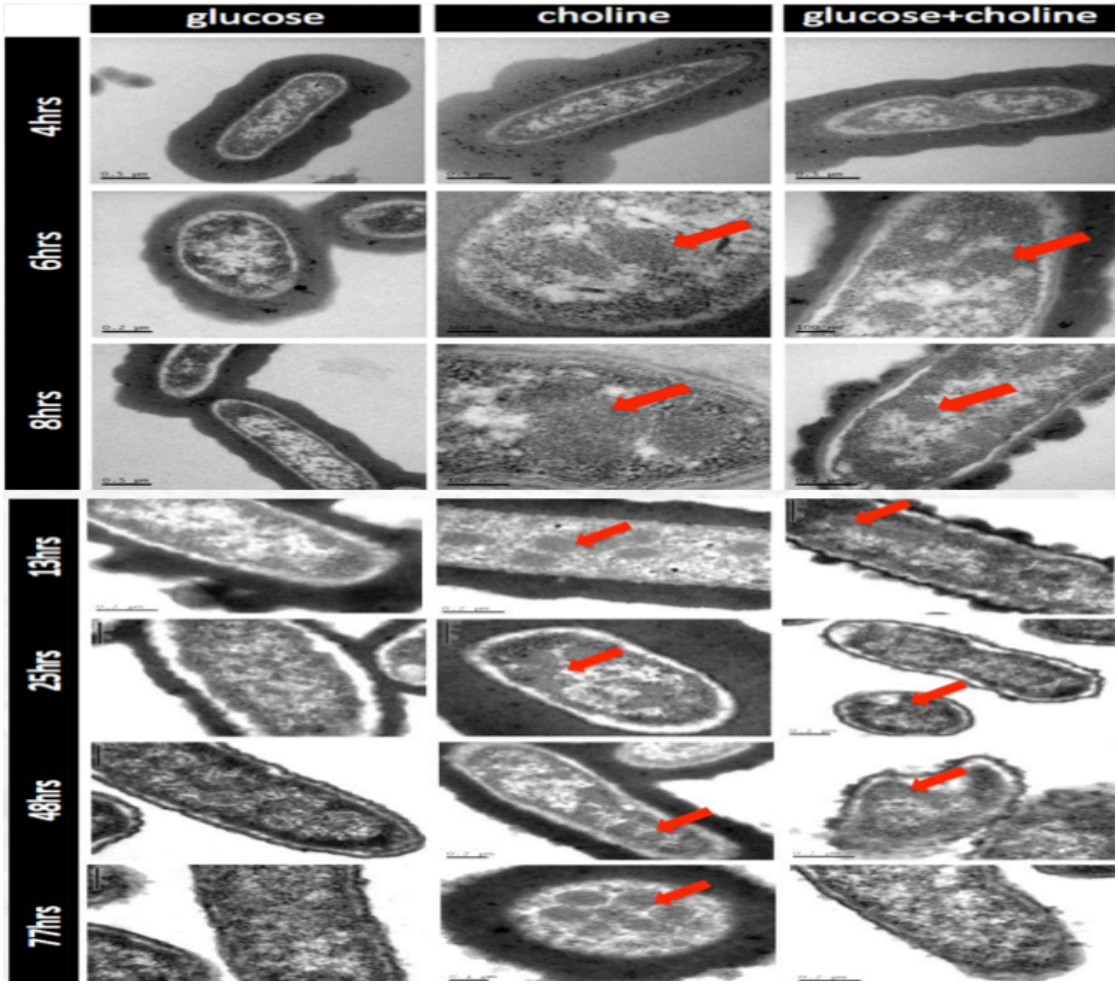
P. mirabilis wild type, the *cutC::kan* mutant and the complemented mutant were grown on glucose supplemented with choline anaerobically in the defined NH_4^+ medium for 77 hrs and the cells were harvested and fixed for TEM.

The wild type BMCs appear electron dense, around 100 nm in diameter and are regularly shaped **(Fig 5.9A)**. Interestingly, abnormal shapes of BMC were found for the *cutC::kan* mutant and the complemented mutant **(Fig 5.9B&C)**. Red arrows indicate the different structures of BMCs observed from the three strains. For the *P. mirabilis cutC::kan* mutant cultures, microcompartment formation was severely disrupted. The *cutC::kan* mutant microcompartments have angular structures, forming sheets and then the sheets rolling up with multiple layers per structure. Next to these structures are quite often what seems to be protein aggregates **(Fig 5.9B)**. It was not possible to assess the microcompartment structures of the mutant grown on choline only because the cells did not grow. Aberrant multi-layered structures and protein aggregations were observed when the mutant was grown on glucose and

choline **(Fig 5.9B)**. For the complemented mutant, when grown on choline, the normal wild-type microcompartment phenotype was recovered after complementing with the native *cutCD.* Regular shaped microcompartments were observed in the complimented mutant but the *cutC::kan* mutant microcompartments with angular stuctures still existed **(Fig 5.9C)**. When grown on glucose plus choline, a combination of malformed multi-layered structures, protein aggregates and wild-type microcompartments were detected by TEM **(Fig 5.9A&C)**.

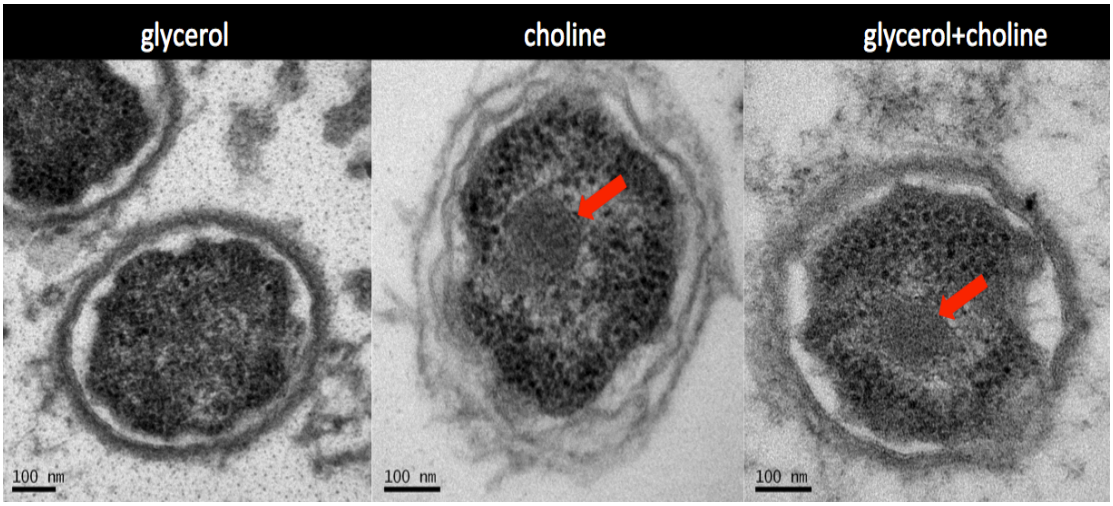
Since choline has also been found to enhance anaerobic swarming of *P. mirabilis*, anaerobic swarming plates were prepared and swarmer cells were processed for TEM. Interestingly, large numbers of microcompartments were found in elongated, hyper-flagellated *P. mirabilis* swarmer cells in the presence of choline.

Together, the simultaneous appearance of microcompartments in the *cutC::kan* complemented mutants, accompanied with TMA production and cell growth, indicated that the microcompartment structures are associated with choline metabolism. The distinct differences in the formation of microcompartments for the *cutC* mutants and the complimented mutant suggest that CutC plays a crucial role in the formation of the microcompartments.



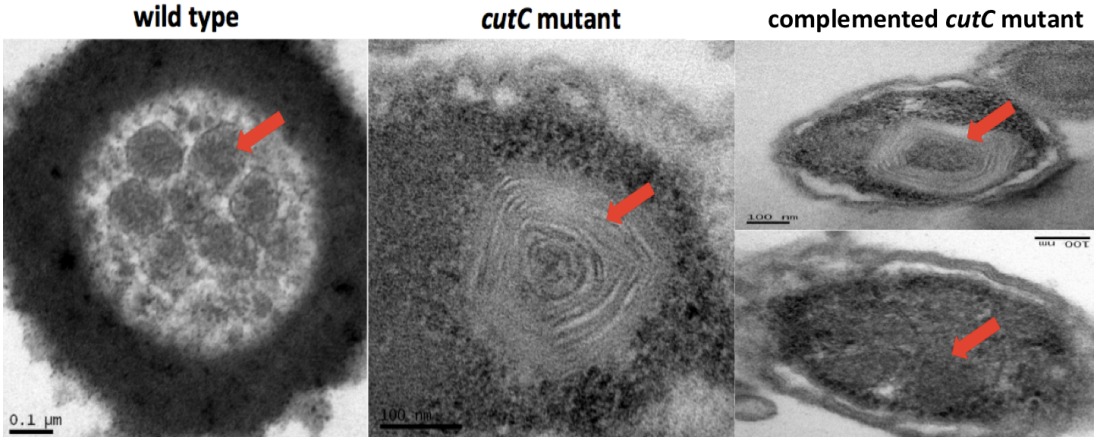
(A) **(B)** **(C)**

Figure 5.7 Wild type *P. mirabilis* was grown anaerobically on glucose **(A)**, choline **(B)** or glucose plus choline **(C)** in the NH_4^+ defined medium. At 5 different time points (4 hrs, 6 hrs, 8 hrs, 13 hrs, 25 hrs, 48 hrs, 77 hrs), the cells were fixed respectively for TEM images to investigate the formation of microcompartments. Red arrows indicated the microcompartment formed in choline and glucose plus choline mediums from 6 hrs onwards (TEM pictures obtained in collaboration with Dr Stefanie Frank from University of Kent).



(A) **(B)** **(C)**

Figure 5.8 Anaerobic swarming of *P. mirabilis* wild type on glycerol **(A)**, choline **(B)**, glycerol plus choline **(C)** on NH_4^+ agar plates. The agar plates were incubated at 30 ºC in an anaerobic cabinet. The agar slices were cut from the plates including the outer-ring to the inner ring, and stained with 0.1 % alcian blue in 0.1 % acetic acid before fixing for TEM. Red arrows indicate the microcompartments formed on swarming plates supplemented with choline and glycerol plus choline (TEM pictures obtained in collaboration with Dr Stefanie Frank from University of Kent).



(A) **(B)** **(C)**

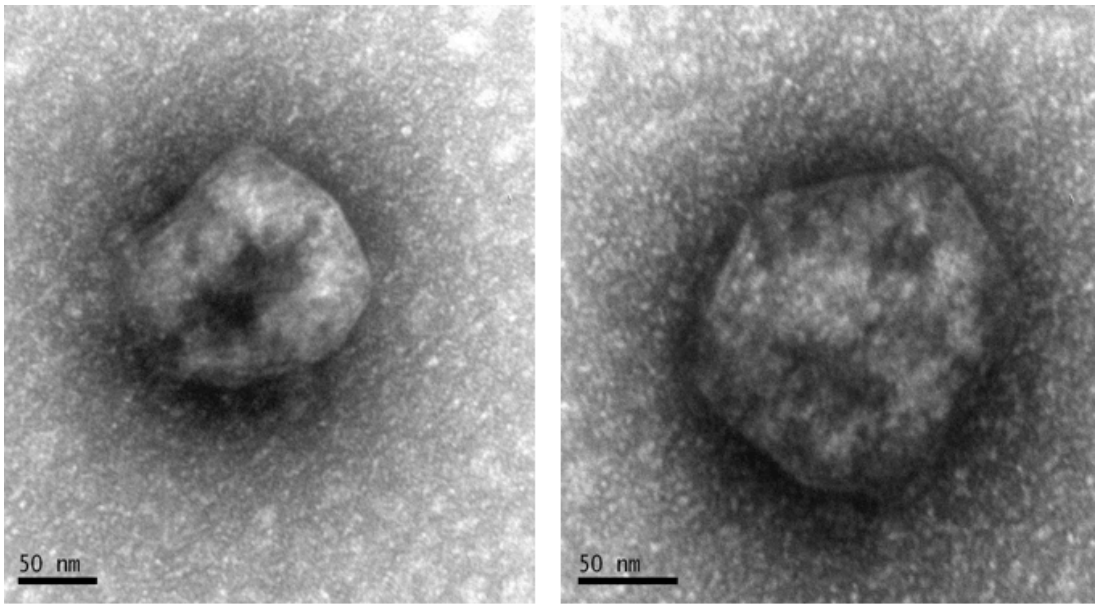
Figure 5.9 TEM for wild type **(A)**, the *cutC::kan* mutant **(B)** and the *cutC::kan* complemented mutant **(C)** BMCs. Anaerobic cells were grown on glucose supplemented with choline for 77 hrs and then harvested and fixed for TEM. Red arrows show the different structures of BMCs for the three phenotypes (TEM pictures obtained in collaboration with Dr Stefanie Frank from University of Kent).

5.2.4.4 Microcompartment protein purification

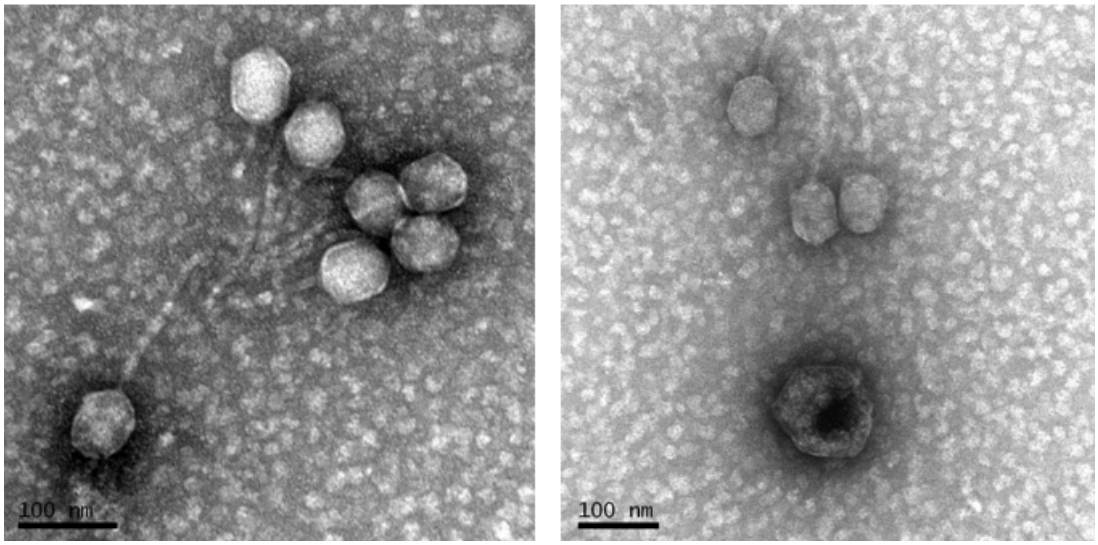
To better understand the structure of the shell protein, attempts were made to purify the microcompartments in liquid culture growth on choline as the sole carbon source. All purification work was carried out at the University of Kent in collaboration with Dr Stefanie Frank.

2 litres of liquid culture was harvested and ~ 2 g of cells were gained for purification using the protocol from Sinha *et al.* (2012) **(See Chapter 2, Materials and Methods)** and later TEM imaging.

Fig 5.10A shows the semi-purified microcompartment protein in liquid culture grown on choline. The size of the organelle is estimated to be around 100 nm. Interestingly, there are large amounts of phage-like particles present in the culture, ~ 50 nm in size under TEM with a long phage tail **(Fig 5.10B)**. After several attempts, I was not successful in obtaining the pure microcompartments from choline-grown cultures. Due to time constraints, this experiment was abandoned and further optimization is therefore required in order to obtain purified microcompartments.



(A)



(B)

Figure 5.10 TEM image for the purified microcompartment in liquid culture grow on choline **(A)** purified microcompartment protein, 100nm under TEM **(B)** purified phage alongside the microcompartment, ~ 50 nm in size under TEM (TEM pictures obtained in collaboration with Dr Stefanie Frank from University of Kent).

5.2.4.5 Microcompartment protein in recombinant *E. coli*

The role of microcompartments is hypothesized either to help retain the volatile intermediates such as aldehyde and carbon dioxide or to protect the cell from the potential toxic effects of the intermediates (e.g. acetaldehyde) (Hampel *et al.*, 2014). Since one of the products of choline metabolism is acetaldehyde, I presumed that the microcompartment proteins are essential in preventing the release of this toxic compound into the cell. However, the exact composition of the BMC formed in *P. mirabilis* during choline metabolism is still unknown. The operon of propanediol utilisation genes (Pdu) encodes eight shell proteins and it has been proved that each single gene is required for the formation of the BMC (Parson *et al.*, 2010). My BlastP results showed that the five predicted BMCs of *P. mirabilis* are composed of PduA/J or PduN like proteins (length: PduA 93 aa, PduJ 91 aa, PduN 91 aa) **(See Table 5.1 & Fig 5.11)**. Therefore, to investigate if the five shell proteins are sufficient for BMC formation, they were cloned into an expression vector (pET28a) and overexpressed in *E. coli* **(Fig 5.12)**.

Table 5.1 The predicted function for the five shell proteins and the BlastP results

No.	Locus tag	Predicted function	Amino acid Length	BlastP	
1	PMI2722	Microcompartment protein	92 aa	CFRPduA MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDA 60 CFRPduJ -MNNALGLVETKGLVGAIEAADAMVKSANVQLVGYEKIGSGLITVMVRGDVGAVKAAVDA 59 PMI2722 -MGDALGLIETKGLVACIEAADAMCKAANVELIGYENVGSGLVTAMVKGDVGAVKAAVDS 59 CFRPduA GAAAARNVGEVKAVHVIPRPHTDVEKILPKGIS 93 CFRPduJ GSAAASAVGEVKSCHVIPRPHSDVEAILPKSA- 91 PMI2722 GVESAQRVGEVVTSLVIARPHNDINKIVIKHKA 92	PduA/J like
2	PMI2721	Microcompartment protein	94 aa	CFRPduA MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDA 60 CFRPduJ -MNNALGLVETKGLVGAIEAADAMVKSANVQLVGYEKIGSGLITVMVRGDVGAVKAAVDA 59 PMI2721 -MGDALGLIETQGLVACIEAADAMCKAANVELIGYENVGSGLVTAMVKGDVGAVKAAVDS 59 CFRPduA GAAAARNVGEVKAVHVIPRPHTDVEKILPKGIS-- 93 CFRPduJ GSAAASAVGEVKSCHVIPRPHSDVEAILPKSA--- 91 PMI2721 GLEAAQRVGTVVTSLVIARPHNDIQKIVAQYKVTE 94	PduA/J like
3	PMI2720	Microcompartment protein	92 aa	CFRPduA MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDA 60 CFRPduJ -MNNALGLVETKGLVGAIEAADAMVKSANVQLVGYEKIGSGLITVMVRGDVGAVKAAVDA 59 PMI2720 -MKEALGLVETKGLVACIEAADAMCKAANVELIGYENVGSGLVTAMVKGDVGAVNAAVES 59 CFRPduA GAAAARNVGEVKAVHVIPRPHTDVEKILPKGIS 93 CFRPduJ GSAAASAVGEVKSCHVIPRPHSDVEAILPKSA- 91 PMI2720 GVEAAKRIGTVVTSRVIARPHNDIEKIAEQHKA 92	PduA/J like

4	PMI2718	Microcompartment protein	87 aa	PMI2718 MILAKVIGHVVATQKSPELKGSNLLMIATL--DDEL--NPLKNKTYVAVDSVGAGINDVV 56 CFRPduN MHLARVTGVVVSTQKSPSLVGKKLLLVRRVSADGELPASPVSGDE-VAVDSVGAGTGELV 59 PMI2718 LA-----EEYFALNKDRYKAMSVVAIVEKVYRDSKE 87 CFRPduN LLSSGSSARHVFSGPNEAIDLAIVGIVDTLSR---- 91	PduN like
5	PMI2714	Microcompartment protein	The first 93 aa	PduA MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDA 60 PduJ -MNNALGLVETKGLVGAIEAADAMVKSANVQLVGYEKIGSGLITVMVRGDVGAVKAAVDA 59 PMI2714_AA93 --MNSLGVIETRGLTAAIQAADAACKAASVEIIGYRKIGSGLVSVCFQGEISAVKTAVEH 58 PduA GAAAARNVGEVKAVHVIPRPHTDVEKILP--KGIS 93 PduJ GSAAASAVGEVKSCHVIPRPHSDVEAILP--KSA- 91 PMI2714_AA93 GVDVVSQKELVIGSLVIARPEPSVITKLLTIKGKK 93	PduA/J like

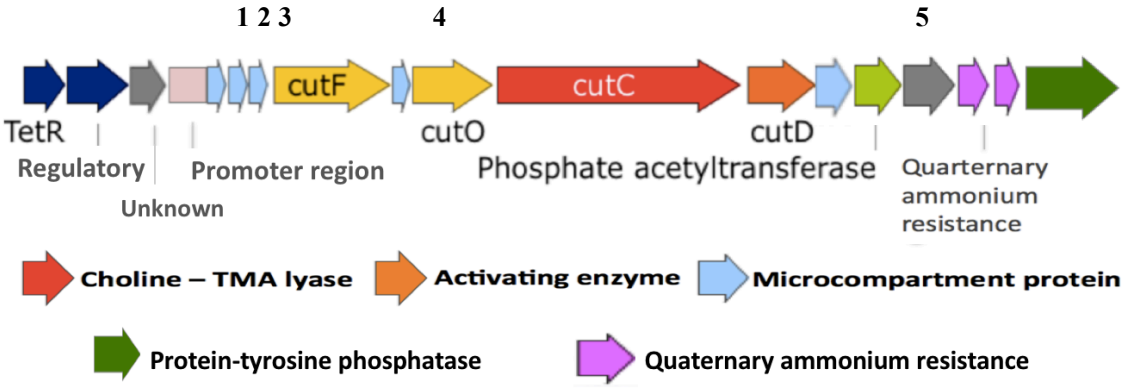


Figure 5.11 The *cut* gene cluster involved in choline degradation in the genome of *P. mirabilis* and the five BMC shell proteins are highlighted in light blue.

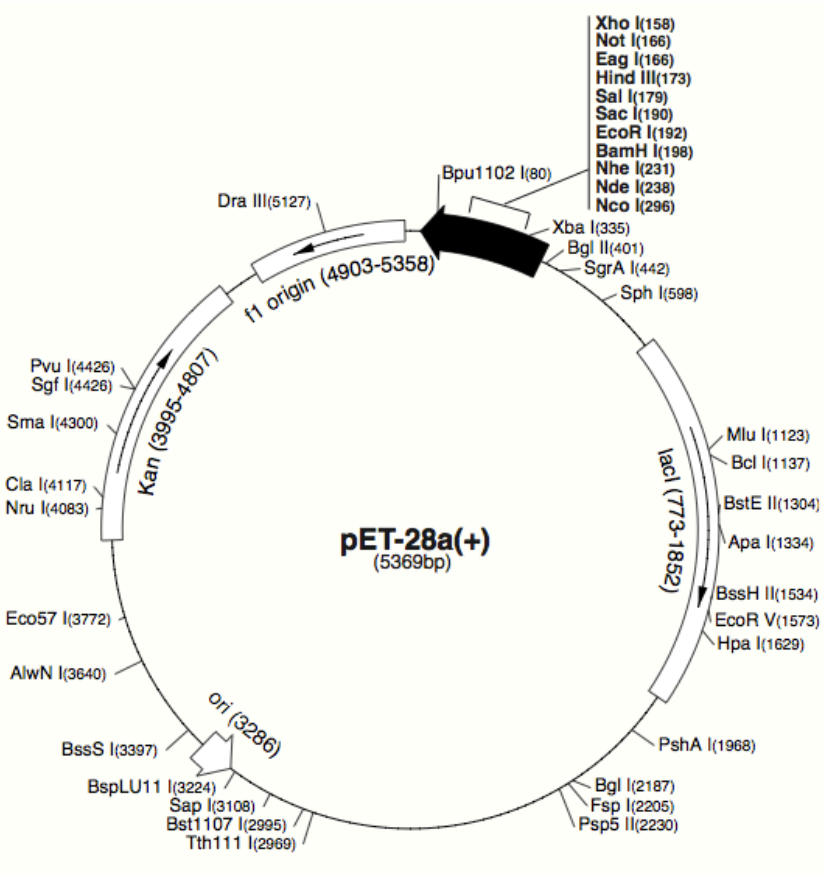


Figure 5.12 The map of plasmid pET28a

Firstly, the five genes encoding microcompartment proteins were amplified as PMI2722-20, PMI2718 and PMI2714 by PCR using specific primers **(Chapter 2)** and high-fidelity Tag DNA polymerase **(Fig 5.13A)**. The primers were engineered with NdeI and NheI sites for PMI2722-20, and NcoI and NheI sites for both PMI2718 and PMI2714. The three PCR fragments were cloned individually into pGEMT. The three fragments were then ligated step by step as described in **Fig 5.13C**, resulting in the cloning of all five BMC genes in one expression vector. Note that the NheI site is compatible to that of XbaI (see details below for the cloning strategy).

The three PCR products were purified respectively, by running through 1 % (w/v) agarose gel and the band of the expected size were purified and further ligated to the cloning vector pGEM-T. The ligation DNA was chemically-transformed into *E. coli* JM109 competent cells for blue-white screening and colony PCRs were performed**.** Plasmids with the correct insert were further purified and digested with relevant restriction enzymes, which resulted in the three fragments: 877 bp (PMI2722-20), 264 bp (PMI2718) and 507 bp (PMI2714), suggesting the three BMC genes were all successfully cloned into pGEM-T, and the sequences were further confirmed by DNA sequencing **(Fig 5.13B)**.

The three fragments were cut out from pGEM-T vector using relevant restriction enzymes, i.e. NdeI and NheI for PMI2722-20, NcoI and NheI for PMI2718 and PMI2714, and then cloned into the vector pET28a respectively, to create pET28a (PMI2722-20), pET28a (PMI2718) and pET28a (PMI2714). PMI2718 was then cut out from pET28a (PMI2718) using XbaI and HindIII, and ligated with pET28a (PMI2722-20) cut by NheI and HidIII, because NheI on this plasmid is compatible to XbaI, to create pET28a (PMI2722-20-18). Using the same principle, PMI2714 was cut out from pET28a (PMI2714) using XbaI and HindIII, and ligated with pET28a (PMI2722-20-18) which is already digested with NheI and HidIII, to create the expression vector pET28a (PMI2722-14) with all the five BMC genes **(Fig 5.13C)**.

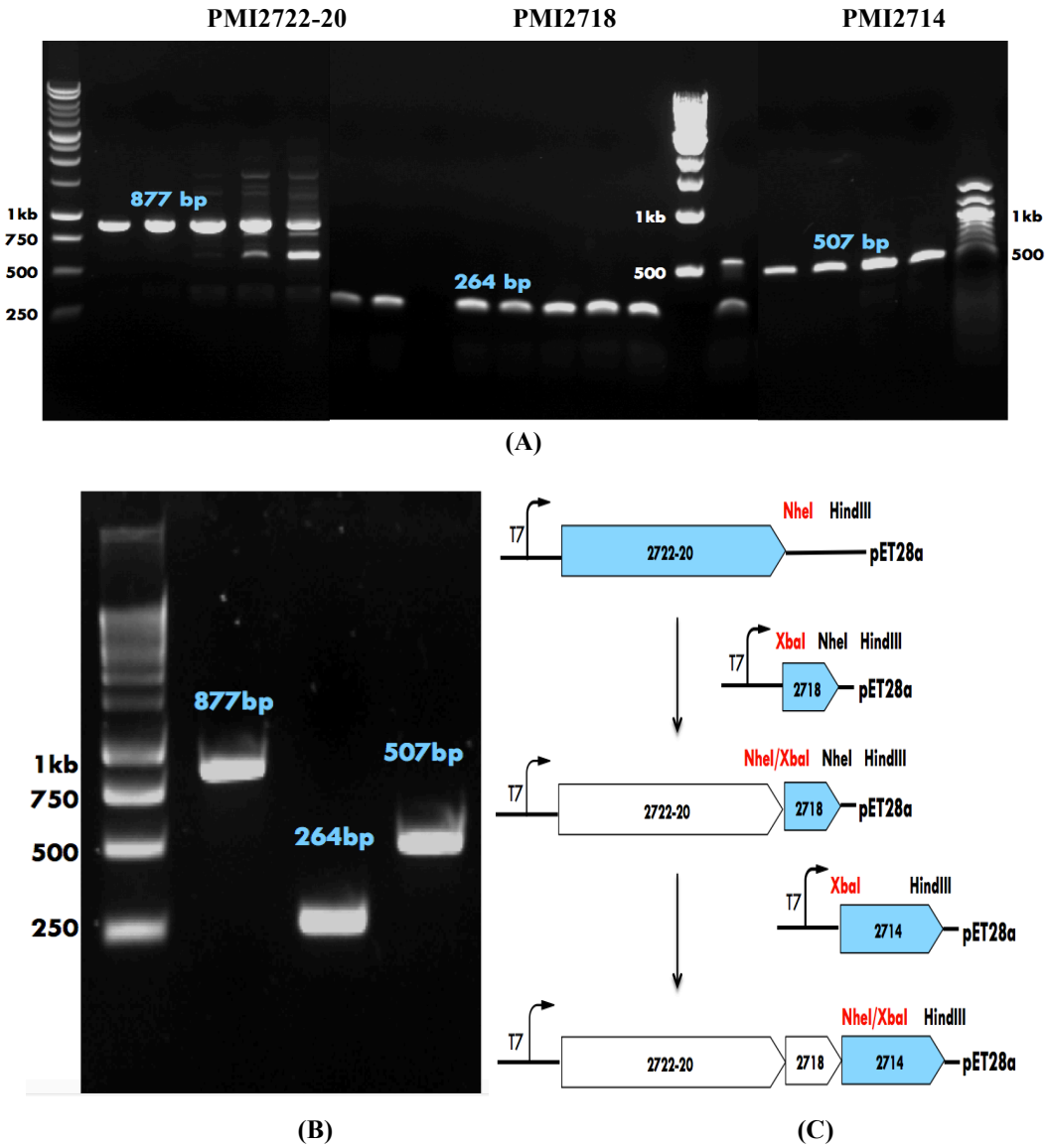


Figure 5.13 The gel pictures of the three amplified microcompartment proteins **(A)** the three PCR purified proteins **(B)** and the cloning strategy of the three fragments into pET28a **(C)**.

The constructed pET28a (PMI2722-14) was then transformed into the expression *E. coli* host BLR (DE3) pLysS. pET28a (PMI2722-20) was also transformed into *E. coli* to generate a comparison control. The transformants were cultivated anaerobically in the defined M9 medium and were induced with 400 μM IPTG when OD_{600} reached 0.4-0.6 **(See Chapter 2 Materials and Methods)**. The cells were harvested and cellular proteins were prepared by passing through a French press to assess the overexpression of the five microcompartment proteins by SDS-PAGE gel **(Fig 5.14)**. Further TEM imaging was carried out to visualize the assemble of shell proteins **(Fig 5.15)**.

Fig 5.14 are the SDS-PAGE gel pictures of the overexpression of PMI2722-20 (three shell proteins) and PMI2722-14 (five shell proteins). Lane 1 is the induced total cell protein for PMI2722-20 (70 μg). Lane 2-5 are induced and un-induced supernatant samples for PMI2722-20, with different amount of protein: 120 μg (lane 2-3) and 70 μg (lane 4-5). Lane 10 is the induced total cell protein for PMI2722-14 (90 μg). Lane 6-9 are induced and un-induced supernatant samples for PMI2722-14, with different amount of protein: 60 μg (lane 6-7) and 90 μg (lane 8-9). The total cell protein always contained a higher amount of proteins, thus has brighter band and more expression levels than the supernatant **(Fig 5.14 lane 1 vs 2-5, lane 10 vs 6-9)**. Samples with higher amount of protein show brighter bands with more expression levels than samples with lower amount of protein **(Fig 5.14 lane 2 vs 4, lane 8 vs 6)**. Several unspecific bands were found ~ 6 kDa (bottom red arrows) for all the induced samples of PMI2722-20 and PMI2722-14, while the un-induced samples do not have any of these bands **(Fig 5.14 lane 2 vs 3, lane 4 vs 5, lane 6 vs 7 and lane 8 vs 9)**.

The size of the five shell proteins is 9.4 kDa (PMI2722), 9.6 kDa (PMI2721), 9.5 kDa (PMI2720), 9.6 kDa (PMI2718) and 17.7 kDa (PMI2714) respectively, therefore, as can been see from **Fig 5.14**, one or two or all of the three shell proteins, PMI2722, PMI2721 and PMI2720, are significantly expressed with the right sizes in the IPTG-induced samples. However, there were no expressed bands with the correct size for PMI2722-14 induced samples. Induced samples of PMI2722-20 and PMI2722-14 all showed an overexpressed band around 130 kDa, as highlighted by the red arrows **(Fig 5.14 lane 1, 2, 4, 6, 8 and 10)**. One possibility of the source of the protein is from the proteome of the host *E. coli* overexpression itself.

The TEM images of the recombinant shell proteins in *E. coli* were shown in **Fig 5.15**, with the three constructions: PMI2722-21-20 (three shell proteins) **(Fig 5.15A)**, PMI2722-21-20-18 (four shell proteins) **(Fig 5.15B)** and PMI2722-21-20-18-14 (all the five shell proteins) **(Fig 5.5C)**. Large aggregated proteins were found in the constructed PMI2722-21-20 **(Fig 5.15A)** and PMI2722-21-20-18 **(Fig 5.15B)**, but hardly any protein aggregates in the constructed PMI2722-21-20-18-14 with all the five shell proteins **(Fig 5.15C)**. There are no proper structures in any of the samples.

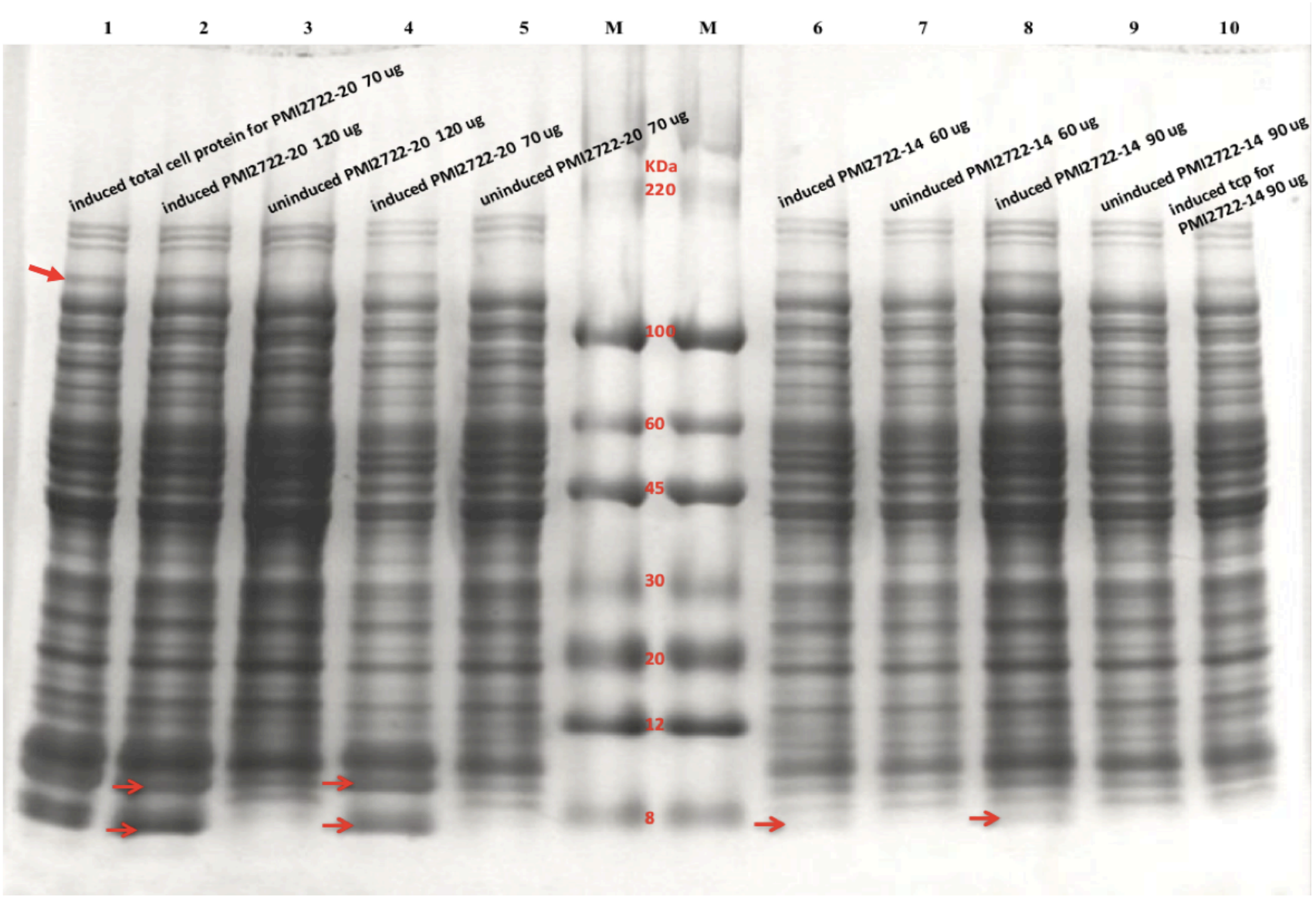

1
2
3
4
5
M
M
6
7
8
9
10
induced total cell protein for PMI2722-20 70 ug
induced PMI2722-20 120 ug
uninduced PMI2722-20 120 ug
induced PMI2722-20 70 ug
uninduced PMI2722-20 70 ug
KDa
220
100
60
45
30
20
12
8
induced PMI2722-14 60 ug
uninduced PMI2722-14 60 ug
induced PMI2722-14 90 ug
uninduced PMI2722-14 90 ug
induced tcp for PMI2722-14 90 ug

Figure 5.14 SDS-PAGE gel shows the protein overexpression results of the total cell protein and supernatant for pET28a (PMI2722-20) (Lane 1-5) and pET28a (PMI2722-14) (Lane 6-10). Lane 1: induced total cell protein for PMI2722-20 (70 μg); lane 2: induced supernatant for PMI2722-20 (120 μg); lane 3: un-induced PMI2722-20 (120 μg); lane 4: induced supernatant for PMI2722-20 (70 μg); lane 5: un-induced PMI2722-20 (70 μg); maker; lane 6: induced supernatant for PMI2722-14 (60 μg); lane 7: un-induced PMI2722-14 (60 μg); lane 8: induced supernatant for PMI2722-14 (90 μg); lane 9: un-induced PMI2722-14 (90 μg); lane 10: induced total cell protein for PMI2722-14 (90 μg).

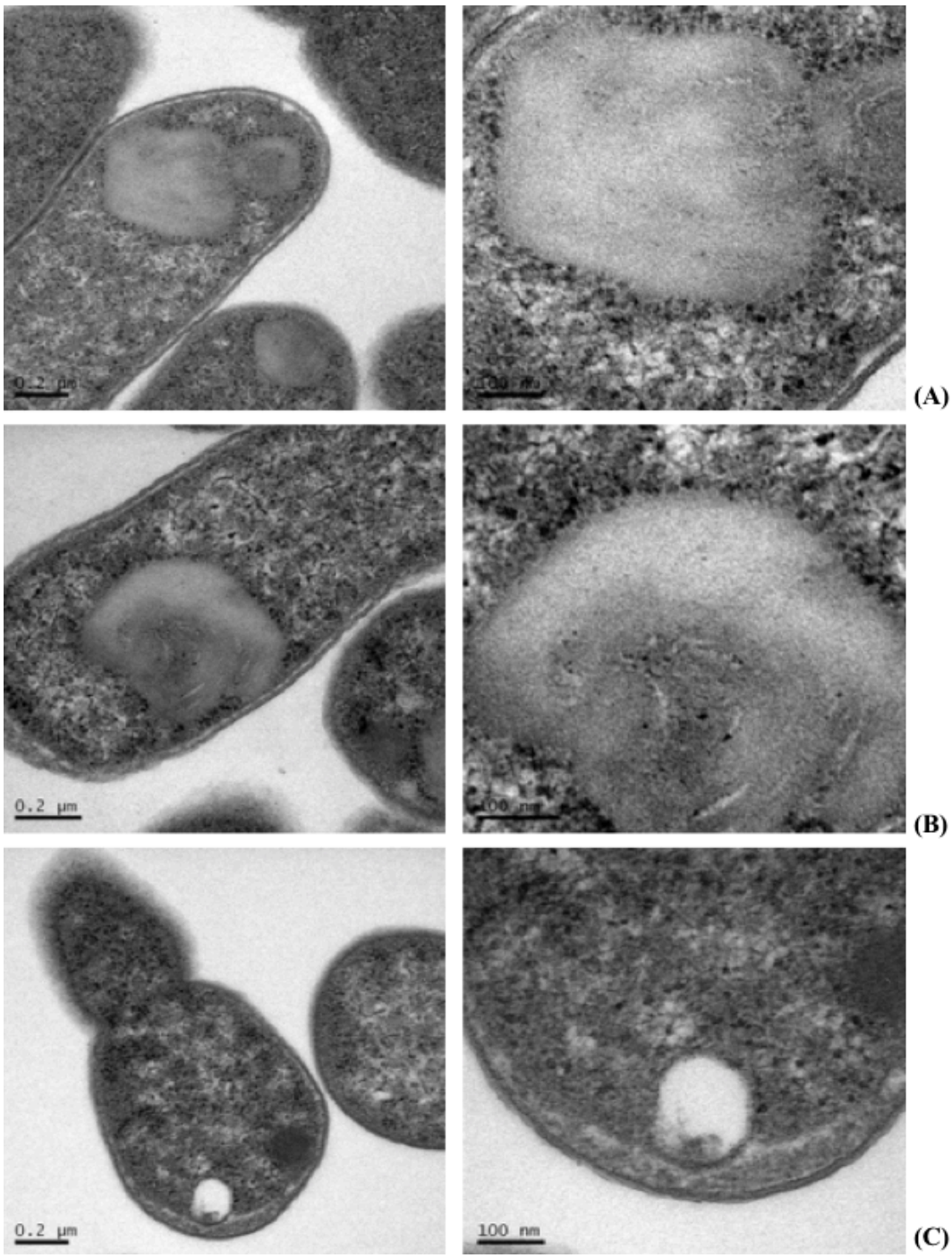
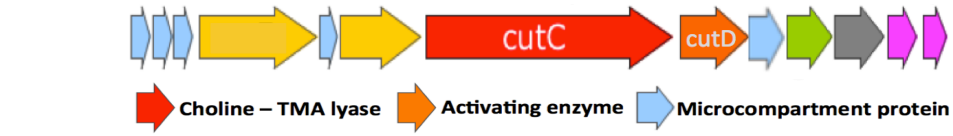
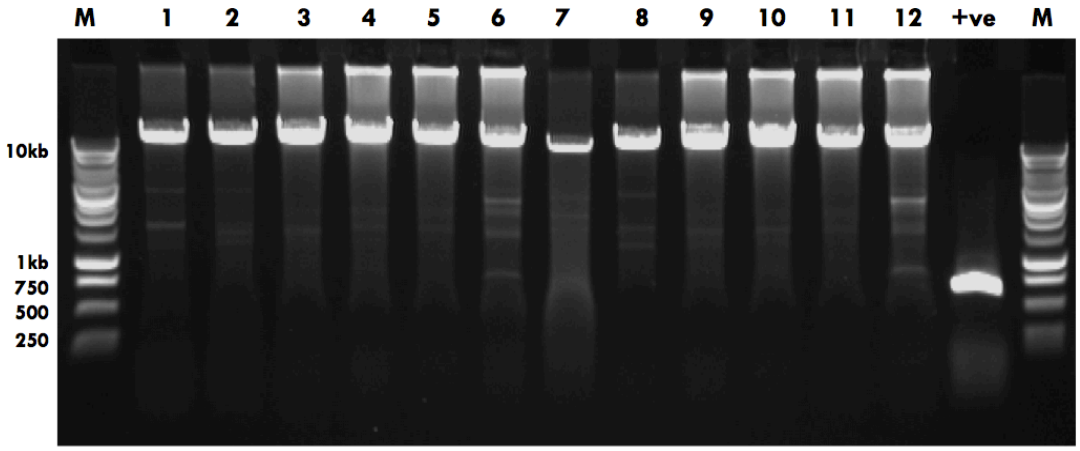


Figure 5.15 TEM images of the recombinant microcompartments in *E. coli*. The recombinant shell proteins for PMI2722-21-20 **(A)**, PMI2722-21-20-18 **(B)** and PMI2722-21-20-18-14 **(C)** (TEM pictures obtained in collaboration with Dr Stefanie Frank from University of Kent)

Furthermore, the difference in the formation of microcompartments in the wild type, the *cutC::kan* mutant and the complemented mutant suggests that CutC may be required for the microcompartment assembly **(Fig 5.16)**. Therefore, to test this hypothesis, the entire *cutC* operon, 11.325kb from the first shell protein (PMI2722) to the last gene on this operon (PMI2710), was amplified using high-fidelity Taq DNA polymerase **(Fig 5.16A)**. Although the whole *cutC* operon was successfully amplified **(Fig 5.16B)**, a correct overexpression plasmid vector in *E. coli* was not obtained during the course of the project.



(A)



(B)

Figure 5.16 (A) The whole *cutC* operon from the first shell protein (PMI2722) to the last gene on this operon (PMI2710)**. (B)** The gel pictures of the amplified whole *cutC* operon PMI2722-10. PCR was carried out at a range of annealing temperatures from 72 °C - 55 °C, purified genomic DNA was used as template, 2 % DMSO was added into the PCR reagents to enhance the genomic DNA.

5.2.5 Investigate the role of the N-terminal domain in CutC of *P. mirabilis*

The presence of two clusters of CutC in genome sequenced bacteria is supported by high bootstrap values in the phylogenetic analysis **(See Chapter 4)**. CutC of *P. mirabilis* has an extra N-terminus of about 300 amino acids compared to that of *D. desulfuricans* **(See Chapter 4, Fig 4.4&4.5)**. The CutC cluster represented by *D. desulfuricans* (848 amino acids in length) is found mainly in obligate anaerobes whereas the cluster represented by *P. mirabilis* (1142 amino acids in length) is found in facultative anaerobes. It is therefore likely that these extra amino acids in the N-terminus of CutC in the *P. mirabilis* cluster may be essential in facultative anaerobic choline degradation**.** Therefore, to test this hypothesis, a N-terminus truncated CutC of *P. mirabilis* was produced in order to compare the activity with its native form.

The full-length *cutC*, together with *cutD* of *P. mirabilis* was already cloned (by Dr Jameson before I started) into the dual expression vector pCOLADuet (Novagen) to obtain the vector pCOLADuet-*cutC/D* **(Fig 5.17)**. The N-terminus truncated *cutC* (*cutC'*, 2546 bp) was amplified by PCR using specific primers and high-fidelity Taq DNA polymerase **(Chapter 2)**. The primers were engineered with NcoI and PstI restriction sites on each end **(Fig 5.17)**. The truncated *cutC'* was then cloned into the MCS 1 of pCOLADuet-*cutD* (digested with NcoI and PstI to remove the full length *cutC*) which was transformed into the expression host of *E. coli* BLR (DE3) pLysS **(Fig 5.17)**.

The PCR product was purified by running through a 1 % (w/v) agarose gel and the band of the expected size was purified and further ligated to the cloning vector pGEM-T. The ligation DNA was chemically transformed into *E. coli* JM109 competent cells for blue-white screening and colony PCR was performed**.** Plasmids with the correct insert were further purified and digested with EcoRI, which resulted in a small fragment ~2.5kb (*cutC'*) and a large fragment of ~3 kb (pGEM-T backbone) suggesting *cutC'* was successful cloned into pGEM-T. The *cutC'* sequence was further confirmed by DNA sequencing **(Fig 5.18)**.

The *cutC'* fragment was then cut out from pGEM-T-*cutC'* with NcoI and PstI and inserted into the vector pCOLADuet-*cutC/D* to create pCOLADuet-*cutC'*/*cutD*, which was then transformed into the expression strain *E. coli* BLR (DE3) pLysS.

In addition, using the same procedure, a second version of the truncated *cutC'* was also produced by Joshua Lee, a final-year undergraduate project student under my supervision, which was inserted to the expression host to generate pCOLADuet-*cutC2'/D*. The wild-type *cutC* is 3432 bp, and the N-terminus truncated *cutC* is 2577 bp (*cutC'*) and 2546 bp (*cutC2'*), respectively.

The transformants were cultivated anaerobically in the M9 defined medium and choline and TMA concentrations were measured. The cells were harvested and cellular proteins were prepared by passing through a French press to assess the overexpression of CutC'. Data presented in **Fig 5.19** shows expression of wild-type CutC as highlighted by the red arrows (around 128.4 kDa). No significant induction of the CutC' truncated mutants (96.4 kDa and 93.7 kDa respectively) was observed in the IPTG-induced samples.

Choline consumption and TMA production data obtained by ion exchange chromatography shows TMA production occurred for wild type but not for the truncated mutants. However, since the protein were not overexpressed for truncated *cutC*, I cannot conclude whether or not they are able to metabolise choline or the role of the N-terminus in anaerobic choline degradation**.** Further experiments are required to optimize the overexpression of CutC' in recombinant *E. coli*, for example by varying the concentrations of IPTG.

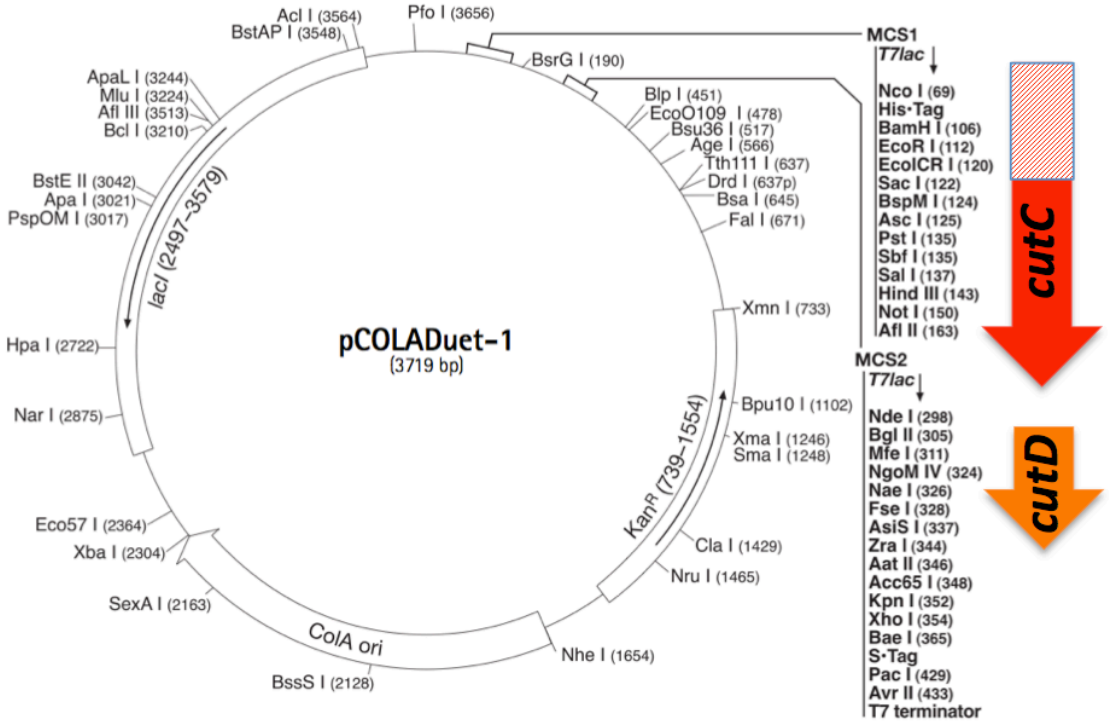


Figure 5.17 An N-terminus truncated *cutC* cloned into the multiple cloning site MCS 1 using the NcoI & PstI restriction sites to investigate the role of its N-terminus on enzyme activity. *cutD* has already been cloned into MCS 2 under the sites NdeI &KpnI.

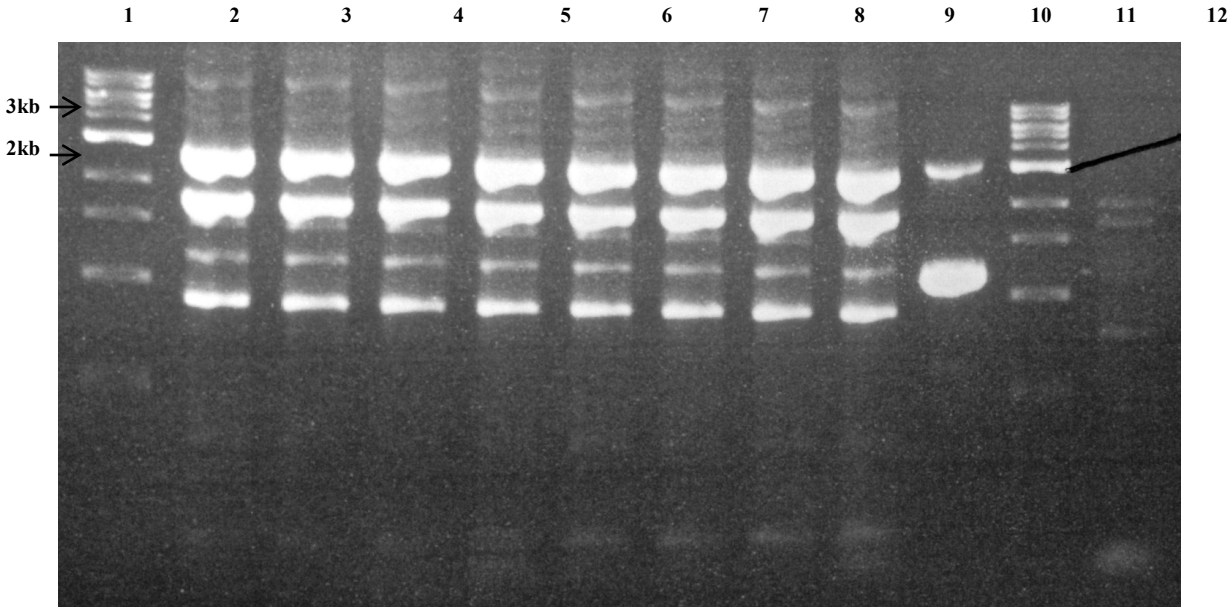


Figure 5.18 The PCR products using primers to generate the N-terminus truncated *cutC* (*cutC'*). 1kb DNA adder (lanes 1 and 11); PCR products (lanes 2-9); positive control using T7 promoter and T7 terminator primers (lane 10) and negative control (lane 12). The top bright band at ~2.5 kb represents the expected size of the truncated *cutC'*. Other bands < 2 k bp represent non-specific amplification by PCR.

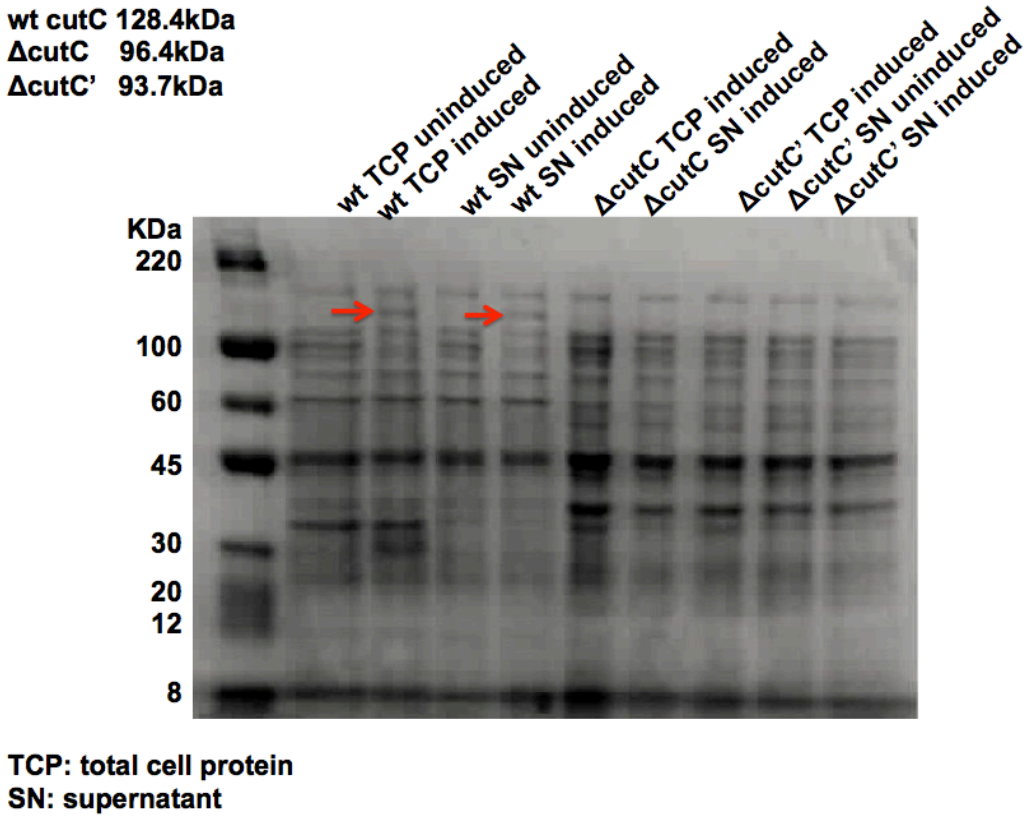


Figure 5.19 SDS-PAGE gel shows the protein overexpression results of the total cell protein and supernatant for wt (un-inducted and induced respectively), the induced total cell protein and supernatant for CutC truncated mutants.

5.2.6 Investigation of the role of a putative choline-responsive *tetR*-type transcriptional regulator in choline-TMA lyase expression

Genome analysis of the choline-TMA lyase gene cluster **(Chapter 4, Fig 4.2&Table 4.2)** revealed the presence of a *tetR* type transcriptional regulator in *P. mirabilis*, which is conserved in other *cutCD*-containing facultative anaerobes **(Chapter 4, Fig 4.2&Table 4.2)**. The *tetR* transcripional regulator and the hypothetic promoter region of the *cut* gene cluster are shown in **Fig 5.11**, highlighted in dark blue and pink resprectively. The TetR group of transcriptional regulators usually work as negative regulators, i.e. when a ligand is absent, the regulator will bind to the promoter causing an obstruction of RNA polymerase binding to the promoter. This results in switching off downstream gene expression. This change of the expression of TetR regulons can be quantified by transcriptional fusion reporter assay, e.g. *lac*Z assay. The promoter region of the *cut* gene cluster (P_{tetR}) is cloned into the plamid pBIO1878 (*lac*Z) (Lidbury *et al.*, 2014) **(Fig 5.20)** to construct the transcriptional fusion plasmid pBIO1878 (P_{tetR}-*lac*Z), which is then electroporated into *P. mirabilis*. The cells were grown anaerobically in the presence of choline, compared with cultures grown in glucose and other structurally related compounds. Cells were then collected and the β-galactosidase assays were carried out (Hampel *et al.*, 2014). The results presented in **Fig 5.21** demonstrated a two-fold induction of the promoter activity in the presence of choline, indicating that the *cut* cluster is indeed switched on by choline.

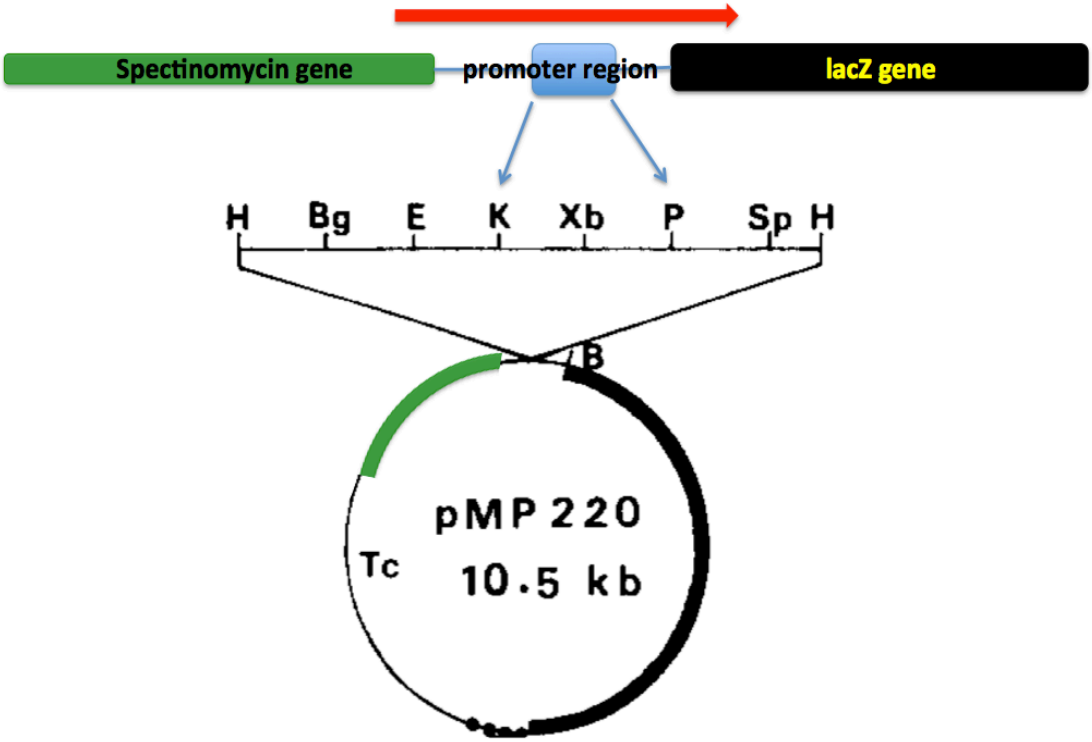
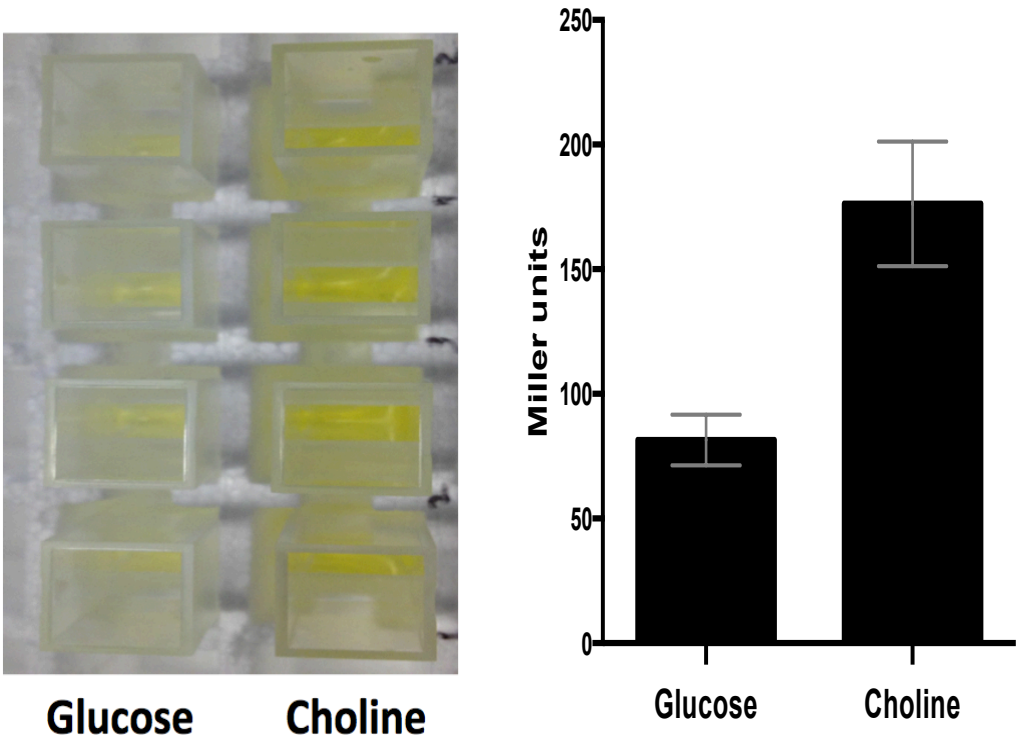


Figure 5.20 The map of the plasmid pBIO1878 (*lac*Z) with unique restriction sites, and the cloning scheme to construct the pBIO1878 (P_{tetR}-*lac*Z).



(A) **(B)**

Figure 5.21 The β-galactosidase assay was carried out by using β-Gal levels as a readout for the promoter activity for *P. mirabilis* [pBIO1878 (P_{tetR}-*lacZ*)] growing in the NH_4^+ defined medium supplemented with glucose and choline. **(A)** shows colour change of choline samples compared with glucose; **(B)** shows 2-fold increase of β-galactosidase activity in choline samples compared with glucose.

5.3 Discussion

5.3.1 The role of CutC and the role of choline degradation in TMA formation in *P. mirabilis*

The human gut microbiome is a very complex but important community for human health, being involved in many physiological activities such as metabolism of xenobiotics, nutrient acquisition, development of mucosal immunity, inflammatory responses and the control of pathogens (Lawley and Walker 2013; Petersson *et al.*, 2011; Reinhardt *et al.*, 2012; Swan *et al.*, 2009). There are more than 1000 bacterial species present in human gut, but the most abundant species in a healthy gut belongs to the *Bacteroidetes* and *Firmicutes* phyla, whilst *Proteobacteria* only persist at low levels (Arumugam *et al.*, 2011; Yatsunenko *et al.*, 2012; El Kaoutari *et al.*, 2013; Winter *et al.*, 2013b).

Recent studies have demonstrated the fundamental role of gut microbiota metabolism of quaternary amines (such as choline and carnitine) to TMA (Craciun and Balskus 2012; Zhu *et al.*, 2014). Winter *et al.* (2013a) have found that *Gammaproteobacteria* could benefit from the host immune and inflammatory response to enhance anaerobic respiration. When inflammatory responses occur, the rich carbon sources in the gut contents are limited for these microbes by the host via diarrhoea, however, choline provided by the host enterocytes and gut mucosa is highly available. The resultant TMA is then able to react with superoxide species provided by the host and converted to TMAO, which enhances anaerobic respiration by these *Gammaproteobacteria* (Balagam and Richardson 2008). This mechanism of taking advantage of the host immune response at least partially accounts for the success of *Gammaproteobacteria* in this competitive environment.

Recent studies have demonstrated that choline degradation to TMA by gut microbiota can trigger the accumulation of TMAO in the circulation, which can lead to cardiovascular diseases and non-alcohol liver diseases (Wang *et al*., 2011; Dumas *et al*., 2006). The facultative anaerobe *P. mirabilis* was used as a model organism to study the molecular mechanism of choline degradation to TMA in this PhD project. I have combined phylogenetic, metabolic and genetic analyses to confirm the

presence and assign the function of a choline utilisation (*cut*) cluster in a group of human *Gammaproteobacteria*. The *cut* cluster was found highly expressed when *P. mirabilis* was growing in a defined NH_4^+ medium supplemented with choline **(See Chapter 4)**. *P. mirabilis* can use choline as a sole carbon source, but is unable to utilize the nitrogen and thus excretes TMA. Choline metabolism in *P. mirabilis* enhanced the anaerobic growth rates in liquid culture and anaerobic colony expansion rates on swarming agar plates.

Anaerobic growth experiments showed that *P. mirabilis* can very rapidly uptake and degrade high concentrations of choline to TMA. The *cut* promoter is up-regulated two-fold in the presence of choline in liquid cultures **(Fig 5.21)** and transcription of the *cut* gene cluster (PMI2710-22) is strongly up-regulated in the presence of choline **(See Chapter 4).** Similar observations have been made in *D. desulfuricans* (Martínez-del Campo *et al.*, 2015).

P. mirabilis was able to swarm both aerobically and anaerobically on a solid surface. The growth phase and source of the *P. mirabilis* inoculum was crucial to swarming-associated colony expansion rates. Inoculum from a fresh agar plate with established anaerobic swarm cells resulted in the fastest colony expansion rate, whereas inoculum from liquid cultures caused slower swarming **(data not shown)**. Sodium fumarate was used as electron acceptor in the defined NH_4^+ medium for anaerobic growth in liquid culture or on swarming plates. This is the first time that true anaerobic swarming of this bacterium has been studied in detail **(Fig 5.3)** as previously studies used sodium azide as an indicator to mimic anaerobic respiration in an aerobic culture (Alteri *et al.*, 2012; Wilkerson and Niederhoffer 1995).

Anaerobic swarming is important in the largely anaerobic human gut and urinary tract environment, offering a competitive advantage for *P. mirabilis* to cause infection (Armbruster and Mobley, 2012). In this study, choline was observed to promote anaerobic swarming by increasing the intervals between consolidation phases and colony expansion rates **(Fig 5.3)**. However, the mechanism by which choline promotes swarming in *P. mirabilis* is still unclear. It is known that several compounds known as environmental cues, for example, glutamine, can increase swarming activity although it has no effect on cell growth (Armbruster *et al.*, 2013). Increased concentrations of choline enhanced swarming expansion rates, whereas

increased glycerol concentrations showed no impact on swarming rates (**Fig 5.6A&B**). The data indicate that choline is very likely functioning as an energy source instead of a carbon source for this bacterium. Indeed, in the downstream of choline metabolism, acetyl-CoA is produced, which is an important energy intermediate during the ATC cycle **(Fig 5.5)**. The *cutC::kan* mutant had no enhancement of swarming when grown on choline **(Figure 5.4)**, suggesting that choline-derived chemosensing in promoting swarming is unlikely (Armbruster *et al.*, 2013).

5.3.2 The role of microcompartment proteins in choline degradation

Martínez-del Campo *et al.* (2015) have identified two different types of *cutC* genes and the phylogenetic analysis has suggested that different *cutCs* are associated with different microcompartment genes. The glycyl radical enzymes-containing microcompartment 1 (GRM1) contains both facultative and obligate anaerobes (Type I and Type II.b in this study), with around 2.5 kb of *cutC*, while the glycyl radical microcompartment 2 (GRM2) contains only facultative anaerobes (Type II.a in this study), mainly *Gammaproteobacteria*, with an extra 300 amino acids in the N-terminus of CutC (Martínez-del Campo *et al.*, 2015). As also can be seen from the phylogenetic tree in **Chapter 4, Fig 4.10,** all identified *cut* gene clusters contain homologues of microcompartment proteins, suggesting a key role of microcompartments in choline degradation (Martínez-del Campo *et al.*, 2015). Indeed, a mutant of a microcompartment gene in *Desulfovibrio alaskensis* was unable to degrade choline (Kuehl *et al.*, 2014).

The *P. mirabilis cut* cluster belongs to the GRM2, which is widely found in opportunistic human pathogens (Axen *et al.*, 2014). There were five genes in the *P. mirabilis cut* cluster predicted to encode for microcompartment shell proteins (PMI2722, PMI2721, PMI2720, PMI2718 and PMI2714). It has been proved by extensive TEM image analyses that microcompartments did form in *P. mirabilis* liquid cultures or swarming agar plates supplemented with choline anaerobically **(Fig 5.6&5.7)**.

During growth in anaerobic liquid culture, a large layer of extracellular polysaccharides (EPS) were found around the cells at 13 and 25 hrs **(Fig 5.6)**. EPS has been reported to play a role in biofilm formation and protection of pathogenic bacteria. Interestingly, at 48 and 77 hrs the EPS layer was breaking up, becoming thinner or disappears, which coincided with the increasing cell density **(Fig 5.6)**. It is possible that when the cell numbers increase, carbon source becomes a limiting factor and EPS was broken down by the bacterium for the production of energy. I have reported, for the first time, that swarmer cells can produce microcompartments. Because the formation of microcompartments is very energy-consuming, it thus has been assumed that swarmer cells need to carefully balance the energy expenditure for cell growth, maintaining cellular metabolism and motility (Armitage 1981;

Falkinham and Hoffman, 1984) **(Fig 5.7)**.

One interesting observation in this project is that *cutC* is very likely to be linked with the assemble of microcompartment protein in *P. mirabilis* **(Fig 5.8)**. The wild type microcompartments appear electron dense, around 100 nm in diameter and are irregularly shaped. However angular structured microcompartments were produced by the *cutC::kan* mutant. The complemented mutant produced not only normal shaped microcompartments, but also many angular structures (the same as those observed in the *cutC::kan* mutant). The data therefore suggests that choline degradation is likely housed inside microcompartments, presumably up until the formation of acetyl-CoA. Further catabolism of acetyl-CoA possibly takes place outside the microcompartment and similar observations have been made in 1,2-propanediol utilizing microcompartments (Parsons *et al*., 2010).

To better understand the structure of microcompartment protein, attempts were made to purify the *P. mirabilis* BMCs **(Fig 5.9A)**. However, this was not successful and future work should focus on modifying the protocol from Sinha *et al.* (2012) for *P. mirabilis* using cells cultivated on choline as the sole carbon source. Interestingly, some unexpected phage like particles were found in the isolated BMCs **(Fig 5.9B)**.

It was not straightforward for any growth, purification and TEM work on *P. mirabilis*, since it is a human pathogen. Attempts have been made to overexpress microcompartment proteins into *E. coli* to facilitate further investigation and manipulation of BMCs. **Fig 5.14** shows the assembly of the five microcompartment proteins in *E. coli*, PMI2722-21-20**,** PMI2722-21-20-18 and the assembly of PMI2722-21-20-18-14, however, there are no proper structures in any of the samples. It is likely that the presence of a N-terminal His-tag in PMI2722 interfered with protein assembly. It is therefore necessary to produce a new construct by removing the N-terminus His-tag and assess the impact of His-tag on microcompartment assembly.

In addition, the whole *cutC* operon, PMI2722-10 (11.325kb) has been successfully amplified **(Fig 5.15)** and further overexpression will be required to assess the assembly of microcompartments in *E. coli*. In addition, the entire *cut* operon (15.238kb in total) including the two regulator genes, *tetR* regulator (PMI2725) and

the protein-tyrosine phosphatase (PMI2709), should be cloned into *E. coli* to investigate their roles in the formation of BMC. **(Fig 5.22)**. Once the shell proteins of *E. coli* recombinants are created, further purification is then required to investigate its structure under TEM.

5.3.3 The role of the N-terminal domain in CutC of *P. mirabilis*

My data presented in **Fig 4.10** showed that *cut* genes are predominately found in *Gammaproteobacteria* (particularly *Enterobacteriaceae*), *Firmicutes* and *Deltaproteobacteria* **(Chapter 4)**. Two distinct *cutC* genes were identified: the first kind of *cutC* gene includes the type I and type II.b *cutC* gene **(as classified in Fig 4.10, Chapter 4)**, both are approximately 2.5 kb in length and are mainly found in *Firmicutes* and *Deltaproteobacteria*. The other kind of *cutC* gene includes the type II.a *cutC* **(as classified in Fig 4.10, Chapter 4)**, about 3.4 kb in length with a 300 amino acids N-terminus. Type II.a is mainly found in *Gammaproteobacteria* including *P. mirabilis.*

There are a few hypotheses regarding the function of the additional N-terminus in CutC. One possibility is that this extra N-terminal domain may be involved in stabilising the enzyme during transient oxygen exposure, because the type II.a CutC was only found in facultative anaerobes. This additional N-terminal domain may therefore help these facultative anaerobes better adapt to living close to intestinal or urinary tract epithelia. Another possibility is that this N-terminus domain may play a role in the assembly of microcompartments. Abnormal microcompartment structures were observed in the cells of both the *cutC::kan* mutant and the *cutC::kan* complemented mutant **(Fig 5.8)**, which resemble the swiss-roll like structures that are formed when only three Pdu shell proteins (PduA, PduB and PduJ) are recombinantly produced in *E. coli* (Parsons *et al.*, 2010). Pang *et al.* (2014) have reported that PduA is a major component of the Pdu shell and may play a crucial role in self-assembly the sheets and higher ordered structures in the over-expressed *E. coli* host. Therefore, it is possible that layered swiss-roll like structures may be formed by multiple PduA-like Cut proteins, in the absence of CutC. This suggests that CutC may be essential for correct assembly of the microcompartments. Clearly the role of CutC in microcompartment formation warrants further investigation.

The two truncated CutC' of *P. mirabilis* were used to investigate if the elongated N-terminus is essential in choline degradation. However, the SDS-PAGE gel showed no obvious expression for truncated *cutC*, thus further experiments are required to optimize the overexpression of CutC' in recombinant *E. coli*. A number of strategies

can be used to improve expression of CutC', for example by varying the concentrations of IPTG or lowing the induction temperature.

5.3.4 The role of the *tetR* transcriptional regulator in choline-TMA lyase expression

The β-galactosidase assay showed a 2-fold increase of the promoter activity in choline samples compared with glucose in *P. mirabilis* [pBIO1878 (P_{tetR}-*lac*Z)], suggesting that choline acts as an inducer for the *cutC* operon. This can be further validated by RT-PCR experiment targeting *cutC*. RNA will be extracted from cells cultivated on choline or glucose and transcription of *cutC* will be quantified by RT-PCR. Furthermore, overexpressing *tetR* in recombinant *E. coli* and protein purification can be carried out. Gel shift assays will be required to determine the promoter sequences which interact with the regulator. Furthermore, *tetR* gene knock-out mutant can be made using the TargeTron system as described previously in **Chapter 3**. Once a mutant is obtained, its growth on choline can be further investigated.

In conclusion, the choline utilization mechanism in the model organism *P. mirabilis* represents an efficient example of a human-niche adaptation in *Enterobacteriaceae* and further *in vivo* studies to confirm this advantageous adaptation is certainly warranted. Metagenomics analysis of the gut microbiome has provided an important insight into the microbial communities in relation to gut health and disease, but understanding microbial metabolism helps to explain the ecology of the microbiota in our gut. Here I have shown how choline is utilised as a carbon and energy source by *P. mirabilis* to facilitate growth in liquid broth and swarming on solid surface. I have also identified a novel microcompartment involved in choline metabolism in this bacterium. These observations represent a mechanism to help better understand the interplay between the host and the gut microbiota and may contribute to explain the success of these bacteria during gut dysbiosis in future *in vivo* studies.

Chapter 6

The impact of gut microbial TMA formation on *Caenorhabditis elegans*

6.1 Introduction

Previous research carried out in the group by Zhu *et al.* (2014) has identified the metabolic pathway and the functional genes involved in carnitine degradation to TMA, *cntA* and *cntB*, which are responsible for converting carnitine to TMA and malate in the human pathogen *Acinetobacter baumannii*. Zhu *et al.* identified the essential role of *cntAB* for this bacterium to grow on carnitine by showing the disability of the knockout mutants for carnitine utilisation, and the abolished TMA production. Carnitine degradation to TMA by CntAB results in the formation of malate and it has been shown previously that malate can extend the life span of *Caenorhabditis elegans* (Edwards *et al.*, 2013).

To gain a better understanding of the role of carnitine degradation in host-microbe interaction, *C. elegans* was used as the model organism. *A. baumannii* wild-type (wt), *cntA* and *cntB* mutants were grown on a defined medium (Zhu *et al.*, 2014) on carnitine as the sole carbon source to induce the expression of CntAB. *C. elegans* was then cultivated on these *A. baumannii* strains to determine the impact of carnitine transformation to TMA on the life span of the host, by counting live/death of every single nematode for every single day during its life span. To further determine the impact of bacterial strains on the life span of the worm, a non-pathogenic *Escherichia coli* SE11 strain that is known to metabolize carnitine to TMA (Zhu *et al*., 2014), was also used.

6.2 Results

6.2.1 The impact of TMA formation by *A. baumannii* on the life span of *C. elegans*

In order to investigate the impact of carnitine degradation and TMA formation on *C. elegans*, *cntA* and *cntB* mutants were compared against the wild type *A. baumannii*, strain (ATCC 19606) to determine the impact of carnitine-to-TMA formation on the worm's life span. Cultivation of *A. baumannii* strains and *C. elegans* life span assays are detailed in **Chapter 2 Materials and Methods**.

Ten age-synchronized adult nematodes per plate were seeded for *A. baumannii* wild type, *cntA*, and *cntB* mutant strains, respectively, with ten plates (*i.e.* 100 worms in total for each strain). The *E. coli* OP50 strain was used for cultivation and general maintenance of *C. elegans* strain N2 before synchronization. *A. baumannii* wild type, *cntA* and *cntB* mutants were grown in the defined M9 medium supplemented with carnitine or succinate at a final concentration of 20 mM. 100 µl of M9-washed bacterial culture was spotted on NH_4^+ free M9 medium supplemented with carnitine or succinate. The plates were maintained at 20 °C every day and the numbers of nematodes (alive, dead or missing) and those that were dried on the wall of the petri dishes were all recorded. The percentage of surviving nematodes was plotted over time and log-rank tests were performed using Kaplan-Meier survival analysis in Sigmaplot (version 11.0).

Preliminary practices were first carried out following the life span protocol for *C. elegans* strain N2 seeded on *A. baumannii* wild type, or the *cntA* and *cntB* mutants. Due to the poor practical skill at the beginning of this experiment, a large number of censored worms were counted (more than 50% of the total nematodes), most of which were caused by accidently killing from picking. Therefore, more practices were carried out until sufficient nematodes for the actual experiment were obtained. However, the actual life span experiment, shown in **Fig 6.1 & Table 6.1**, suggests that there was no statistical difference between wt, *cntA* and *cntB*, and the survival rates are almost the same.

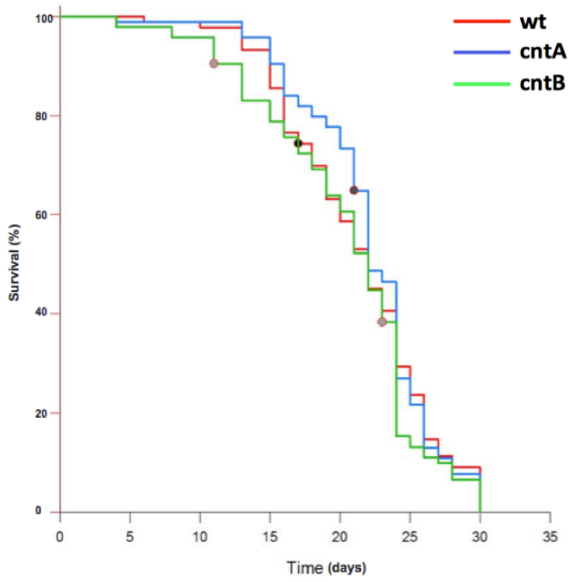


Figure 6.1 Lifespan analyses showing the survival rates of *C. elegans* N2 strain over time at 20 °C. 100 nematodes were used for each strain: *A. baumannii* wild type (red line), *cntA* mutant (blue line) and *cntB* mutant (green line).

Table 6.1 Survival analysis (log rank test/multiple comparison) of *C. elegans* N2 strain fed on *A. baumannii* wild type, *cntA* or *cntB* mutant

Strain	Total	Event	Censored	% of censored	Median time	Unadjusted *p* value	Significance?
wt	100	89	11	11	22	Log rank *p* value	No
cntA	100	93	7	7	22	0.313	
cntB	100	93	7	7	22		

6.2.2 The role of TMA formation in *E. coli* SE11 on the life span of *C. elegans*

6.2.2.1 Feeding of *E. coli* SE11 by *C. elegans* strain N2

To investigate the role of carnitine degradation on the life span of *C. elegans* N2 strain, an *E. coli* SE11 strain was investigated which is also capable of producing TMA from carnitine (Zhu *et al.*, 2014). Since it has been well documented that *A. baumannii* is pathogenic against *C. elegans*, it may have a strong pathogenic effect itself on *C. elegans* rather that the effect of TMA formation from carnitine degradation on the worm (Jayamani, 2015). In contrast to *A. baumannii*, the *E. coli* SE11 strain is nonpathogenic and its genome sequence revealed no known virulence factors (Oshima *et al.*, 2008). SE11 was cultivated on different carbon sources, 20 mM of carnitine, malate, succinate, malate plus TMA, or succinate with TMA, respectively. The experiments were carried out at two temperatures, i.e. 25 °C and 20 °C.

Data presented in **Fig 6.2 and Table 6.2** show that nematodes seeded on *E. coli* SE11 grown on malate and carnitine lived longer than those grown on succinate. The survival analysis suggests that, at 25 °C, there is no significant difference between carnitine and malate (log rank p=0.910), but there is significant difference between succinate and carnitine (log rank p<0.001), and between succinate and malate (log rank p<0.001). In addition, TMA with either succinate or malate **(Fig 6.2 D&E**) did not affect the longevity of *C. elegans* (log rank p=0.221/p=0.094).

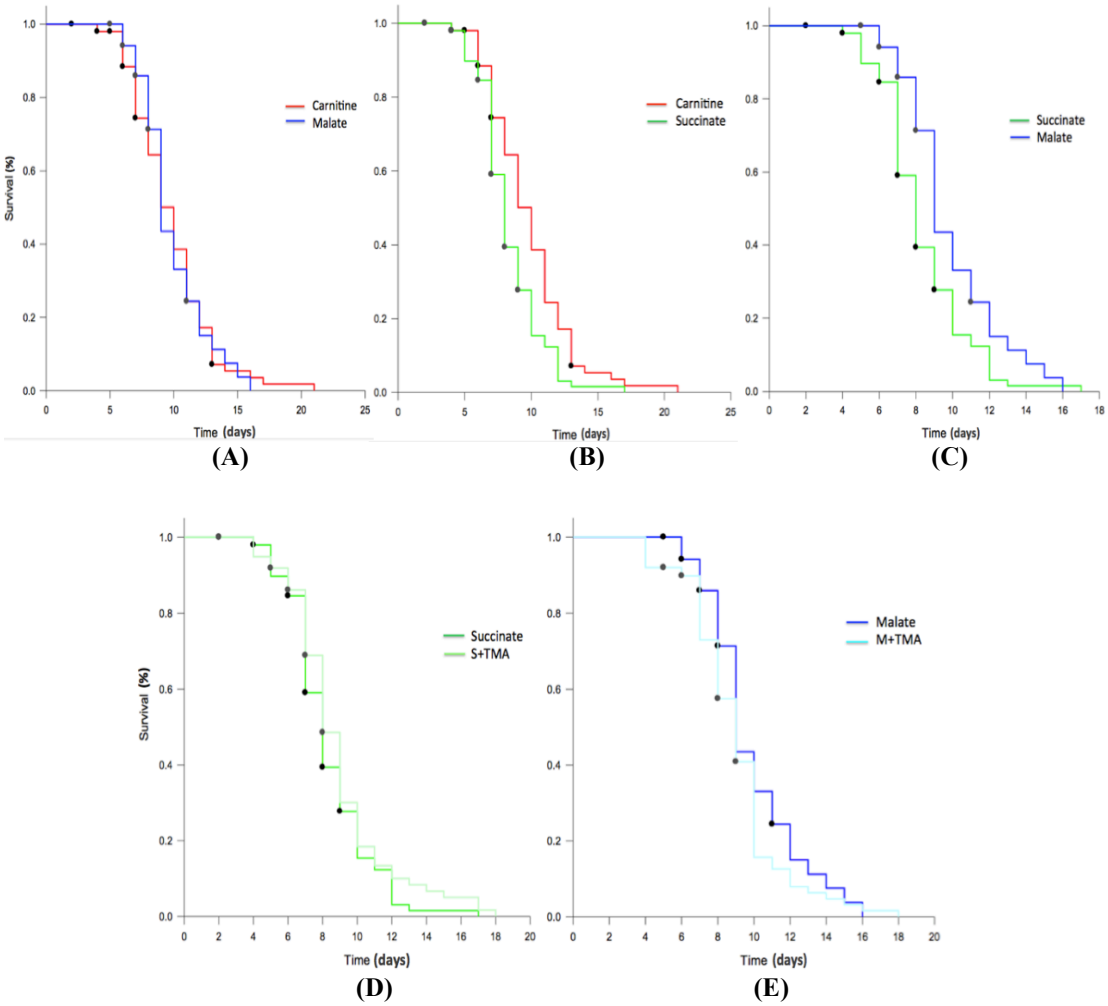


Figure 6.2 Lifespan analysis of *C. elegans* strain N2 at 25 °C. The nematode was fed on *E. coli* SE11 grown in the presence of different metabolites in the carnitine degradation pathway: 20 mM final concentration of carnitine, malate, succinate, malate plus TMA, or succinate with TMA, respectively. **(A)** Carnitine vs malate. **(B)** Carnitine vs succinate. **(C)** Malate vs succinate. **(D)** Succinate alone vs succinate plus TMA. **(E)** Malate alone vs malate plus TMA.

Table 6.2 Survival analysis (log rank test) of *C. elegans* N2 fed on *E. coli* SE11 grown on various carbon sources at 25 °C

Substrate	Total	Event	Censored	% of censored	Median time	Comparisons	Log rank *p* value	Significance?
Carnitine	100	71	29	29	10	carnitine vs malate	p=0.910	No
Malate	100	60	40	40	9	carnitine vs succinate	p=0.001	Yes
Succinate	100	78	22	22	8	succinate vs malate	p<0.001	Yes
Succinate + TMA	100	67	33	33	8	succinate vs succinate + TMA	p=0.221	No
Malate + TMA	100	70	30	30	9	malate vs malate + TMA	p=0.094	No

Because the nematode cultivated at 25 °C had a high rate of censored individuals **(Table 6.2)**, indicating a stress response at this temperature, the experiment was repeated at 20 °C. Furthermore, a higher number of nematodes (120) for each group was used in order to improve statistical power. In agreement with the previous results, data presented in **Fig 6.3** also shows a longer survival time on malate and carnitine than those seeded on succinate. There is a significant difference in life span between succinate and malate (log rank p=0.001), and a slight difference between succinate and carnitine (log rank p=0.015<0.05), but no difference between carnitine and malate (log rank p=0.253). No effect of additional TMA on the worm's longevity was found (p=0.232/p=0.177) **(Fig 6.3 & Table 6.3)**.

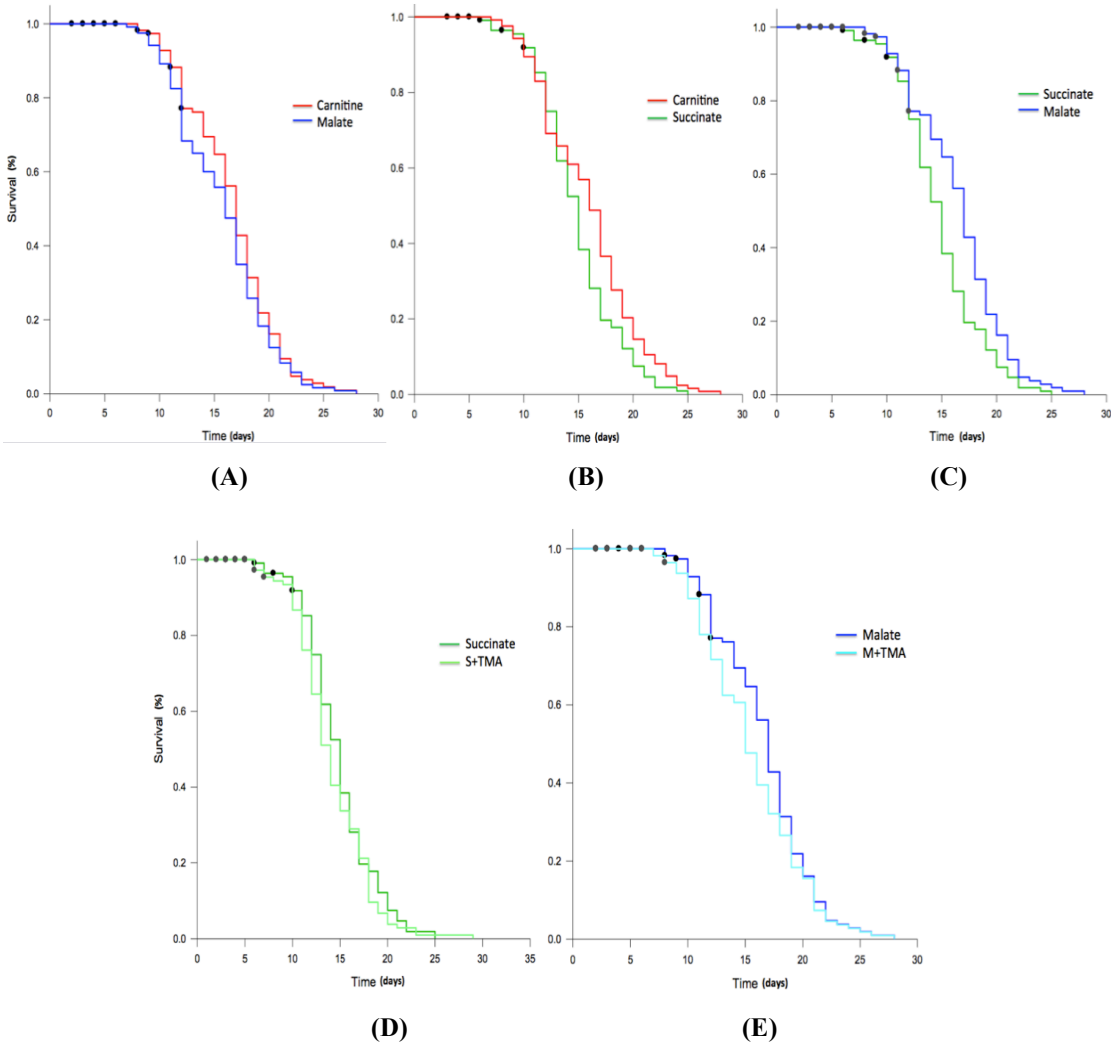


Figure 6.3 Life span analysis of *C. elegans* N2 strains cultivated at 20 °C. The nematode was fed on *E. coli* SE11 grown in the presence of different metabolites in the carnitine degradation pathway, 20 mM concentrations of carnitine, malate, succinate, malate plus TMA, or succinate with TMA. **(A)** Carnitine vs malate degradation on lifespan. **(B)** Carnitine vs succinate on lifespan. **(C)** Malate vs succinate treatment on lifespan. **(D)** Succinate vs succinate plus TMA. **(E)** Malate vs malate plus TMA.

Table 6.3 Survival analysis (log rank test) of *C. elegans* N2 fed on *E. coli* SE11 grown on various carbon sources at 20 °C

Substrate	Total	Event	Censored	% of censored	Median time	Comparisons	Log rank *p* value	Significance?
Carnitine	120	120	0	0	16	carnitine vs malate	*p*=0.253	No
Malate	120	106	14	12	17	carnitine vs succinate	*p*=0.015	Yes
Succinate	120	107	13	11	15	succinate vs malate	*p*=0.001	Yes
Succinate + TMA	120	104	16	13	14	succinate vs succinate + TMA	*p*=0.232	No
Malate + TMA	120	109	11	9	15	malate vs malate + TMA	*p*=0.177	No

6.2.2.2 Feeding of *E. coli* SE11 by *C. elegans* strain CF512

The survival analysis on *C. elegans* N2 strain shows a large proportion of censored cases during the life span **(Tables 6.2)**. Reducing the growth temperature for the worm from 25 °C to 20 °C reduced the percentage of censored worms, however, the experiments lasted much longer (4 weeks at 20 °C vs 2 weeks at 25 °C). In order to shorten the time for each life span experiment, a *C. elegans* mutant CF512 strain was tested. CF512 strain is a temperature-sensitive strain derived from N2 wild type, which does not lay eggs and remains sterile at 25 °C. The same procedure of lifespan analysis was thus carried out using this strain as model organism (Hsin, 2007).

Data presented in **Fig 6.4 and Table 6.4** suggest that there is a significant difference (p<0.05) between succinate and malate (log rank p=0.033). Additional TMA with succinate/malate did not affect the longevity of *C. elegans* CF512 strain (log rank p=0.922/p=0.256). There is no significant difference between carnitine and malate (log rank p=0.985), however, there is no difference between carnitine and succinate, either (log rank p=0.056).

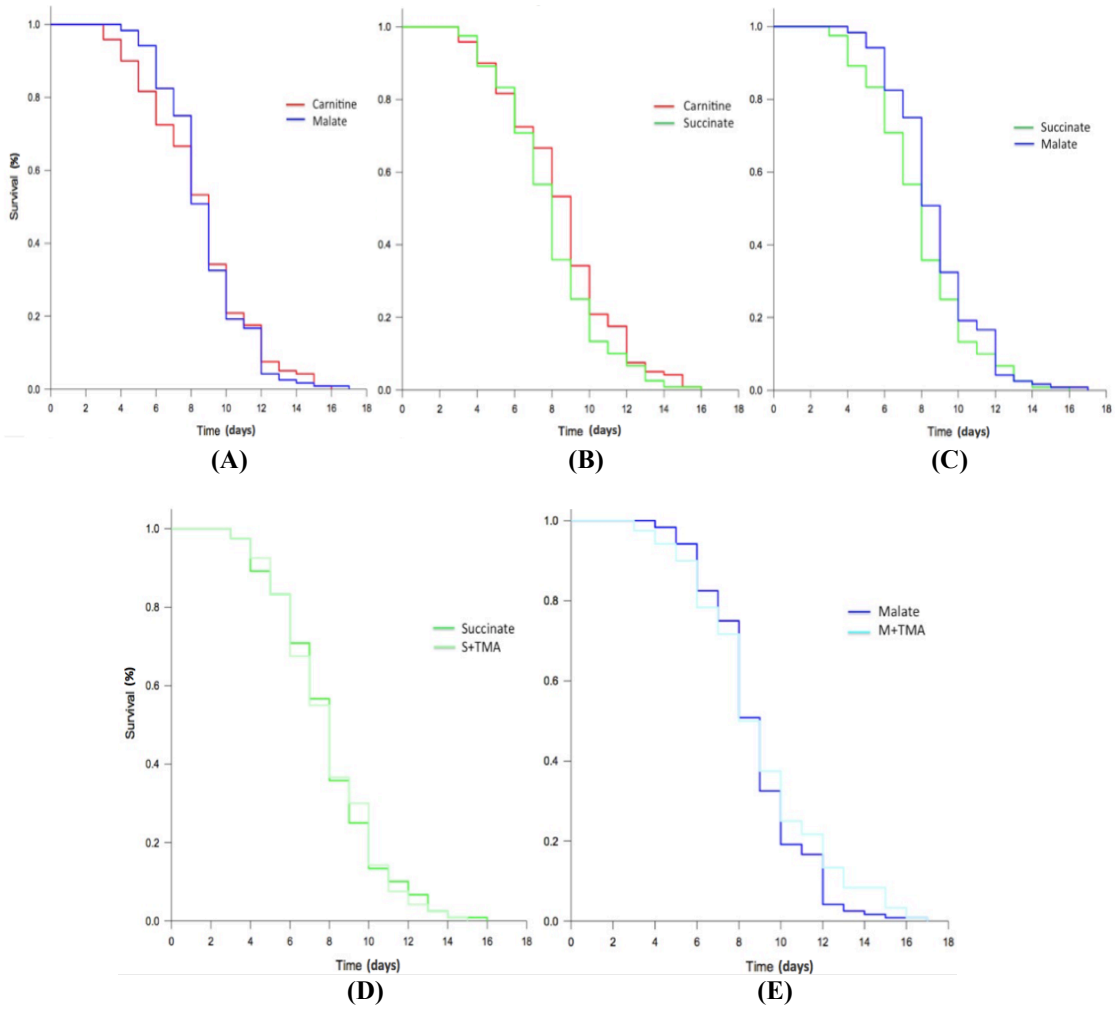


Figure 6.4 Lifespan analysis of *C. elegans* CF512 strains at 25 °C fed on E. coli SE11 cultivated on different carbon sources, 20 mM concentrations of carnitine, malate, succinate, malate plus TMA, or succinate with TMA. **(A)** Carnitine vs malate. **(B)** Carnitine vs succinate. **(C)** Malate vs succinate. **(D)** Succinate vs succinate plus additional TMA. **(E)** Malate vs malate supplemented with TMA.

Table 6.4 Survival analysis (log rank test) of *C. elegans* CF512 at 25 °C fed on *E. coli* SE11 grown on difference carbon sources

Substrate	Total	Event	Censored	% of censored	Median time	Comparisons	Log rank *p* value	Significance?
Carnitine	120	120	0	0	9	carnitine vs malate	*p*=0.985	No
Malate	120	120	0	0	9	carnitine vs succinate	*p*=0.056	No
Succinate	120	120	0	0	8	succinate vs malate	*p*=0.033	Yes
Succinate + TMA	120	120	0	0	8	succinate vs succinate + TMA	*p*=0.922	No
Malate + TMA	120	120	0	0	9	malate vs malate + TMA	*p*=0.256	No

6.3 Discussion

6.3.1 Compare *C. elegans* N2 life span during carnitine degradation in *A. baumannii* wild type, *cntA* and *cntB* mutants

At the beginning of the preliminary practices of *C. elegans* strain N2 life span experiments on the three strains of *A. baumannii*, poor reproductivity was observed in each practice and indicated contrasting results of the effect of carnitine degradation on *C. elegans* longevity **(data not shown)**. The reason for the inconsistency is probably due to the lack of technical proficient skills of performing such experiments at that moment. This can be seen by the high rates of censored cases (the percentage of censored cases was almost 50%), in which accidental death of worm was caused by improper transfer of worms by picking. Therefore, more practices were carried out before the actual experiment was performed. With the improved experimental skills, the censored rate became reasonably low for the actual experiment. However, in the experiment, the survival rates for wt, *cntA* and *cntB* were almost the same, i.e. 30 days for wt, *cntA* and *cntB*, although the log rank *p* value is improved and no pairwise comparison is required. The above data suggests there is no statistically significant difference ($p<0.05$) in life span of *C. elegans* fed on wt or the mutants of *A. baumannii* (**Fig 6.1 & Table 6.1)**.

6.3.2 Compare *C. elegans* N2 life span fed on *E. coli* SE11 grown on carnitine and other carbon sources

For *C. elegans* N2 strain, the life span experiment was first carried out at 25 °C, which promoted a fast growth rate for the worm, however, the nematodes may have experienced stresses under this temperature as evidenced by the higher percentage of censored cases. Therefore, the experiment was repeated at a lower temperature (20 °C) for *C. elegans*. At both incubation temperatures, the survival analyses confirmed that there is a significant difference (p<0.05) between succinate vs carnitine, succinate vs malate, but no significant difference is found between carnitine and malate. Interestingly, it is confirmed that TMA did not affect the longevity of *C. elegans* **(Fig 6.2 & Table 6.2, Fig 6.3 & Table 6.3)**.

At 25 °C, multiple group comparison **(Fig 6.5 & Table 6.5)** showed the same significance between carnitine and succinate as in the single group comparison **(Fig 6.2 & Table 6.2)**. At 20 °C, when single pairs of groups **(Fig 6.3 & Table 6.3)** were compared, the survival analysis confirmed that there is a difference of the life span of the worm between succinate vs carnitine. However, there was no difference when multiple comparisons were carried out **(Fig 6.6 & Table 6.6)**.

This is because p values used for single-group comparison vs multiple group comparison are different in the Kaplan-Meier survival analysis. When single comparison was carried out, p<0.001 was regarded as significant, in which if the null hypothesis is true, there is a 1% chance of ending up with a 'statistically significant' result simply by chance. However, when multiple-group comparison was carried out, this 1% value applies to the entire *family* of comparisons. The 1% is defined as the *family-wise* error rate, an adjusted p value. This means that if all the null hypotheses were true, there is a 1% chance that one or more of the differences will be 'statistically significant' simply due to random variation, leaving a 99% chance that all the comparisons will be 'not significant' (Krzywinski and Altman, 2014). Under this circumstance, multiple comparison will not be reliable, for example, comparison between different conditions. The analysis will calculate the whole group including conditions, replicates etc. and take the average as adjusted p value. Therefore, the single comparison is usually carried out initially to see whether

there is a significant level, before the multiple groups of different conditions are compared. (Krzywinski and Altman, 2014).

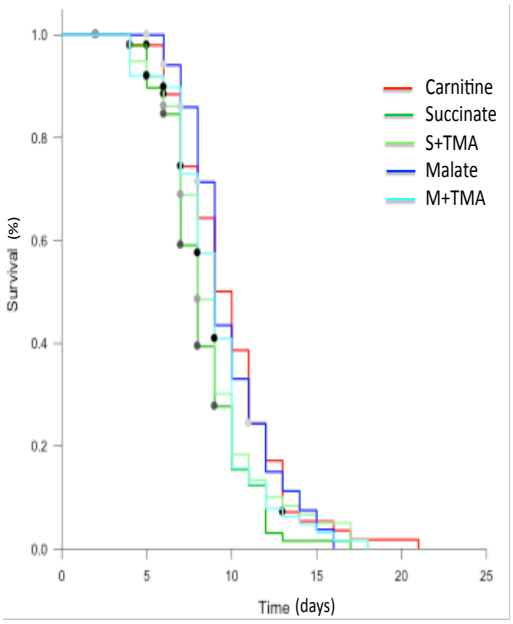


Figure 6.5 Survival curves of *C. elegans* N2 strains at 25 °C growth on different substrates in carnitine degradation in *E. coli* SE11: carnitine, malate, succinate, malate plus TMA, and succinate with TMA.

Table 6.5 Survival analysis (multiple group comparisons) of *C. elegans* N2 in different substrates in *E. coli* SE11 at 25 °C

Comparisons	Statistic	Unadjusted P value	Significance?
Succinate vs malate	12.675	0.000371	Yes
Carnitine vs succinate	10.507	0.00119	Yes
Malate vs malate +TMA	2.807	0.0938	No
Succinate vs succinate +TMA	1.497	0.221	No
Carnitine vs malate	0.0129	0.910	No

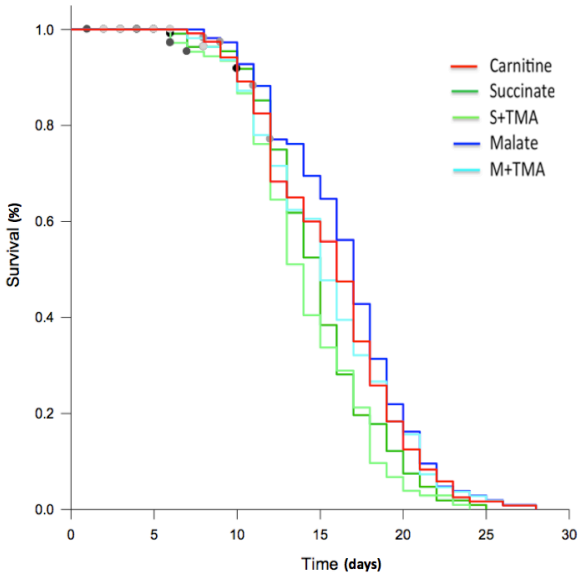


Figure 6.6 Survival curves of *C. elegans* N2 strains at 20 °C growth on different substrates in carnitine degradation in *E. coli* SE11: carnitine, malate, succinate, malate plus TMA, and succinate with TMA.

Table 6.6 Survival analysis (multiple group comparisons) of *C. elegans* N2 in different substrates in *E. coli* SE11 at 20 °C

Comparisons	Statistic	Unadjusted P value	Significance?
Succinate vs malate	10.560	0.00116	Yes
Carnitine vs succinate	4.042	0.0444	No
Succinate vs succinate +TMA	1.823	0.177	No
Malate vs malate +TMA	1.822	0.177	No
Carnitine vs malate	1.308	0.253	No

6.3.3 Comparison of *C. elegans* CF512 life span during carnitine degradation in *E. coli* SE11 wild type

C. elegans CF512 strain was used to reduce the high rate of censored nematodes and improve the efficiency of the survival analyses. There is a significant difference between succinate and malate, but no difference of carnitine vs succinate and carnitine vs malate was found. Additional TMA with succinate/malate did not affect the longevity of *C. elegans* CF512 strain **(Fig 6.4 & Table 6.4)**, but the multiple comparison showed no significant differences in the three groups **(Fig 6.7 & Table 6.7)**. One possible reason is that CF512 is a temperature-sensitive mutant stain of N2 (mutations in *rrf-3(b26)* II and *fem-1(hc17)* IV genes, which is involved in promoting activity of repeated genes and required for masculinization of germline and somatic tissues, respectively) and will stay sterile without producing progeny at this temperature, however, the mutation may have caused unknown impact on its life span during carnitine degradation.

Overall, the data presented in this Chapter suggest that the bio-degradation of carnitine by *E. coli* SE11 may affect *C. elegans* life span. This difference cannot be attributed to the formation of TMA from carnitine degradation because additional TMA did not affect the longevity of the worm. Therefore, in agreement with previous reports (Edwards *et al.*, 2013), it is likely that the intermediate, malate which is formed from carnitine degradation by the bacterium, may have extended the life span of *C. elegans*. However, due to the time limit of the PhD project, biological replicates have not been carried out for any of the three experiments: *C. elegans* N2 life span during carnitine degradation in *A. baumannii* wild type, *cntA* and *cntB* mutant, N2 life span during carnitine degradation in SE11, and CF512 life span during carnitine degradation in SE11. Therefore, to further ascertain the conclusion, 2 or more biological replicates of life span experiments will need to be carried out independently. Besides, once the *cnt*A and *cntB* of mutants of SE11 are obtained, further tests will also be required to compare the wt vs *cnt*A or *cntB* mutant of SE11 with at least 3 biological replicates for each life span experiment.

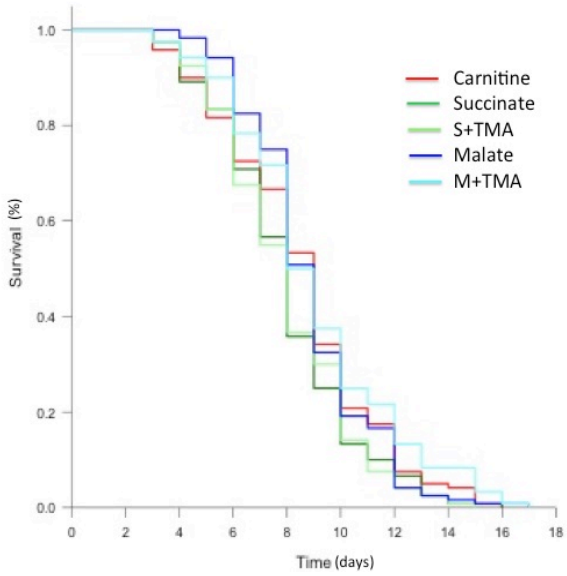


Figure 6.7 Survival curves of *C. elegans* CF512 strains at 25 °C growth on different substrates in carnitine degradation in *E. coli* SE11: carnitine, malate, succinate, malate plus TMA, and succinate with TMA.

Table 6.7 Survival analysis (multiple group comparisons) of *C. elegans* CF512 in different substrates in *E. coli* SE11 at 25 °C

Comparisons	Statistic	Unadjusted P value	Significance?
Succinate vs malate	4.565	0.0326	No
Carnitine vs succinate	3.656	0.0559	No
Malate vs malate +TMA	1.290	0.256	No
Succinate vs succinate +TMA	0.00964	0.922	No
Carnitine vs malate	0.000352	0.985	No

Chapter 7

Summary and future perspectives

7.1 Chapter 3 - Developing a method for genetic manipulation of *Proteus mirabilis*

This chapter aimed to establish the genetic mutagenesis method for targeted gene knockout in *P. mirabilis*.

The initial test of four *P. mirabilis* strains were first carried out to confirm their ability to convert choline to TMA. MIC tests were then carried out for the four stains of *P. mirabilis*, which showed no difference in antibiotic resistance profiles and the strain DSMZ4479 was therefore used as the model strain in this study. A few attempts were made to establish the plasmid transfer and targeted gene knock-out mutagenesis methods. pKNG101 was initially chosen as a suicide vector in order to facilitate homologous recombination; however, due to the incompatibility of pR6K-type plasmid present in the strain DSMZ4479, it was impossible for conjugating pKNG101 into this bacterium. The TargeTron gene knockout kit from Sigma Aldrich was then adopted and the *cutC* gene was successfully deleted from *P. mirabilis* DSMZ4479. The *cutC::kan* mutant was further complemented by cloning and expression of the native *cutC/D* gene from DSMZ4479, with growth tests subsequently carried out to confirm their phenotypes. However, a major limitation of the TargeTron method is the low efficiency of electroporation-mediated plasmid transfer, which is estimated to be ~ 4×10^5 CFU/μg. Further optimization is required in order to improve the efficiency. Furthermore, it remains unknown if similar pR6K-type plasmids are present in other clinic *P. mirabilis* strains, such as P14 and P19, and it remains to be established whether pKNG101-based conjugation method will work in those isolates.

7.2 Chapter 4 - Comparative transcriptomic analysis of *Proteus mirabilis* by RNA-Seq

Chapter 4 took the advantage of comparative transcriptomics (RNA-Seq) to investigate the key genes involved in choline degradation to TMA in *P. mirabilis*. From the RNA-Seq data analysis, 18 highly-expressed genes were identifed in the presence of additional choline, including the choline-TMA lyase gene (*cutC*), activating enzyme (*cutD*) and genes encoding for putative microcompartments. The subsequent phylogenetic and genome analyses revealed the functional choline-TMA lyase and the *cutC* gene cluster in *P. mirabilis* and the *cutC* gene cluster was compared between *P. mirabilis* and *D. desulfuricans* **(Aim 1)**.

The phylogenetic tree presented in **Chapter 4, Fig 4.10** has clearly divided species according to the different types of *cutC* genes. There are two distinct groups of CutC. Type I and Type II.b contained both facultative and obligate anaerobes with CutC around 2.5 kb in length whereas the Type II.a group contained only facultative anaerobes (mainly *Gammaproteobacteria*) with an extra 300 amino acids long N-terminus domain in CutC. As previously described by Martínez-del Campo *et al.*, the *cut* gene clusters contain homologues of microcompartment proteins, therefore, the role of microcompartments in choline degradation was further examined.

Phylogenetically, CutC belongs to the so-called glycyl radical protein family, which also includes B12-independent glycerol dehydrogenase and pyruvate formate lyase. Recently, an interesting study from Romano *et al.* (2015) screened choline utilization and TMA formation using 79 bacterial strains isolated from human intestine, many of which contained the predicted choline-TMA lyase according to genome analyses. The result indicated that two species, *Providencia alcalifaciens* and *Providencia rustiganii*, did not generate TMA although their genomes do encode a seemingly entire choline-TMA lyase pathway. On the other hand, one strain, *E. tarda* strain 23685, does not contain *cut* genes in its genome but produced TMA from choline (Romano *et al.*, 2015). These results suggest cautions need to be taken when using *cutC* as the functional gene marker for predicting choline utilization and it is likely that CutC homologs shown in

the phylogenetic tree may have other functions rather than choline utilization. It is possible that other choline-to-TMA transformation pathways may also exist (e.g. in *E. tarda*), which warrants further investigation.

Both genome analysis and the constructed phylogenetic tree suggest the presence of an extra N-terminus of CutC in some gut microbiota. The type II.a CutC, containing mainly sequences of *Enterobacteriaceae* such as *P. mirabilis*, has an additional ~300 amino acids towards the N-terminus **(Fig 4.11)**. Interestingly, the type II.a CutC was only identified from facultative anaerobes, therefore the role of this extra N-terminus of CutC is postulated to have an association with stabilizing the enzyme during transient oxygen exposure **(Chapter 4)**. It may also be involved to facilitate CutC folding during translation. It is unfortunate that an attempt to generate a truncated CutC of *P. mirabilis* was not successful due to the lack of expression of the truncated protein **(Chapter 5)**. Therefore, further work on mutagenesis and overexpression of the N-terminus of CutC will be necessary to confirm these hypotheses.

The *tetR*-family transcriptional regulator was also highly expressed. It is hypothesized that the *tetR* gene plays a role as a negative repressor. In the presence of choline, TetR will bind to choline instead of the promoter, and thus switch on the transcription of the whole gene cluster. Besides, the structure and folding of tRNA may change to assist the regulation of the *tetR* with or without choline. To prove the function of this regulator, knock-out mutagenesis and overexpression of this protein followed by gel-shift assays are therefore required.

7.3 Chapter 5 - Anaerobic choline degradation in *Proteus mirabilis*

The work presented in this Chapter clearly demonstrated the functional genes involved in the choline metabolism to TMA using *P. mirabilis* as the model system through targeted mutagenesis. Using the marker exchange mutagenesis method established in **Chapter 3**, I investigated whether the *cutC* gene is indeed essential for choline degradation and subsequent TMA production in *P. mirabilis*. In addition, I investigated the physiological role of choline degradation to TMA in this bacterium in both liquid culture and on solid agar plates. The data presented in this Chapter indicates that choline enhanced growth of *P. mirabilis* in broth culture, and played a greater role in swarming where it significantly enhanced swarming speed and reduced the need for consolidation **(Aim 2)**. Furthermore, I observed the formation of microcompartments in this bacterium during its growth on choline and the role of the shell protein in choline degradation **(Aim 3)**. Surprisingly, it was observed that the *cutC* gene was also required for the formation of microcompartments.

My data showed that choline metabolism enhanced the anaerobic growth rates in liquid culture. The addition of choline not only increased swarming rates, but also enhanced intervals between colony expansion distance and consolidation phases under anaerobic conditions **(Fig 5.3)**. *P. mirabilis* can use choline as a sole carbon source and is able to uptake and degrade high concentrations (~ 20 mM) of choline to TMA rapidly. Furthermore, I have confirmed that the *cut* promoter was up-regulated two-fold with the presence of choline in liquid cultures using the beta-galactosidase assay **(Fig 5.21)**. This was also confirmed by the RNA-Seq analysis, showing that the transcription of the *cut* gene cluster (PMI2710-22) was strongly up-regulated in the presence of choline **(Chapter 4)**. Similarly, induction of *cut* gene cluster by choline has also been observed in *D. desulfuricans* by Martínez-del Campo *et al.* (2015).

However, there are still a few outstanding questions that merit further investigation.

1. How do microcompartments assemble in *P. mirabilis*? Overexpression of microcompartment proteins into *E. coli* has been carried out to facilitate further

investigation and microcompartment purification. However, no proper structures have been gained in any of the samples (**Fig 5.14)**. The likely interference from the N-terminal His-tag on PMI2722 with protein assembly is a plausible explanation. It is therefore wise to re-clone these genes into *E. coli* without a His-tag and possibly include the natural ribosome binding regions for all the 5 genes encoding the microcompartment to achieve a desirable ratio of each polypeptide for correct assembly of the microcompartment. It is also possible to overexpress the whole *cut* operon in *E. coli*, together with the putative regulator genes *tetR* (PMI2725) and/or the protein-tyrosine phosphatase (PMI2709) **(Fig 5.22)**.

2. The role of the extra N-terminus in CutC remains elusive. The SDS-PAGE gel of the truncated CutC' of *P. mirabilis* did not show obvious expression for the two versions of the truncated CutC. The concentrations of IPTG and induction temperature may affect the expression efficiency in recombinant *E. coli*. Therefore, further experiments are required to optimize the overexpression of CutC' in recombinant *E. coli*. Once the overexpression is successful, growth experiments of the recombinant *E. coli* in the presence of choline will be required to investigate if the elongated N-terminus is indeed essential in choline degradation.

3. Validation of the role of the *tetR* regulator in choline metabolism. The β-galactosidase enzyme assay suggests that choline acted as an inducer for the *cut* operon and a 2-fold increase of the promoter activity was observed in choline samples compared with glucose in *P. mirabilis* [pBIO1878 (P_{tetR}-*lacZ*)]. A *tetR* gene knock-out mutant is required to firmly establish the role of this regulator in regulating choline metabolism (as discussed in **Chapter 3)**. Once a mutant is obtained, its growth on choline can be further investigated. Overexpression of *tetR* in recombinant *E. coli* and protein purification will be further carried out to confirm the interaction of TetR with a range of ligands including choline by gel shift assays.

4. Experiments carried out in this chapter have clearly demonstrated a growth advantage from choline metabolism in *P. mirabilis*, however, further *in vivo* studies to confirm this advantageous adaptation in interaction with the host is certainly warranted. A mouse *in vivo* model can be set up for dietary interventions. Specifically, mice could

be treated with choline and non-choline food (e.g. carbohydrate), and various health parameters will be assessed including glucose homeostasis, fatty lipids profile and TMAO accumulation. The concentrations of glucose and TMAO in the serum could be determined by liquid chromatography to assess the link between choline metabolism and diabetes and cardiovascular disease. Lipids in the gut metabolites could be measured by LC-Mass spectrometry.

5. Meta-transcriptomics analysis can be used to investigate the microbial community under different conditions. Unlike metagenomics, which indicates which microbes are present in the microbial community, meta-transcriptomics allows study of the diversity of actively transcribed genes, gene expression abundance and gene differential expression in a specific microbial environment (Bashiardes *et al.*, 2016). By using meta-transcriptomics, sample cultures from human or animals can be prepared and the complex microbial community across all species can be identified. Furthermore, cultures could be grown under choline and non-choline fed conditions, and the differentially expressed genes and the diversity of functional pathways can be explored and the abundance of *cutC* across many taxa can be quantified using this technology.

7.4 Chapter 6 - The impact of gut microbial TMA formation on *Caenorhabditis elegans*

C. elegans was used as the model organism in this chapter to investigate the role of carnitine degradation in host-microbe interaction. The worms were fed on *Acinetobacter baumannii* wild-type (wt), *cntA* and *cntB* mutants, which were grown on carnitine to determine the impact of carnitine transformation to TMA on the life span of the host. *Escherichia coli* SE11 strain was further used to determine the impact of bacterial strains on the life span of the worm metabolizing carnitine to TMA **(Aim 4)**. Carnitine degradation in *A. baumannii* wt, *cntA* and *cntB* showed no difference on the life span of *C. elegans* N2 strain **(Fig 6.1 & Table 6.1)**. For carnitine degradation in *E. coli* SE11, both multiple group comparison **(Fig 6.5 & Table 6.5)** and single group comparison showed clear difference between carnitine and succinate on *C. elegans* N2 strain at 25 °C. At 20 °C, single group comparison **(Fig 6.3 & Table 6.3)** confirmed that there is a difference of the life span of the worm between succinate vs carnitine treatments, however, such difference was not obvious when multiple comparisons were carried out **(Fig 6.6 & Table 6.6)**. When the *C. elegans* CF512 strain was used, it was found that there is a significant difference between succinate and malate treatments, but no difference for carnitine vs succinate or carnitine vs malate **(Fig 6.4 & Table 6.4)**, however, multiple comparison showed no significant differences in the three groups **(Fig 6.7 & Table 6.7)**.

The Kaplan-Meier survival analysis uses the log rank test for significance assessment. The log rank test calculates chi-squared values for each time event for each group and sums the results to give an ultimate chi-squared value. As can be seen from **Table 6.1**, throughout the analyses, unadjusted *p* value was used. Multiple testing corrections have been applied here to reduce the number of false positive results (Rich *et al.*, 2010).

Different results were obtained when using single group comparison and multiple group comparison **(Fig 6.3 & Table 6.3 vs Fig 6.6 & Table 6.6, Fig 6.4 & Table 6.4 vs Fig 6.7 & Table 6.7)**. This is because *p* values used for single-group comparison vs multiple group comparison are different in the Kaplan-Meier survival analysis

(Krzywinski and Altman, 2014). When single comparison was carried out, $p<0.001$ was regarded as significant. If the overall p value is below the desired significance level, it regards the result as significant. Specifically, if the null hypothesis is true, there is a 1 % chance of ending up with a 'statistically significant' result simply by chance. However, when multiple-group comparison was carried out, this 1 % value applies to the entire *family* of comparisons. The 1 % is defined as the *family-wise* error rate, an adjusted p value. This means that if all the null hypotheses were true, there is a 1 % chance that one or more of the differences will be 'statistically significant' simply due to random variation, leaving a 99 % chance that all the comparisons will be 'not significant' (Krzywinski and Altman, 2014).

Single comparison in **Fig 6.4 and Table 6.4** suggest that there is a significant difference ($p<0.05$) between succinate and malate (log rank $p=0.033$). The conventional cut-off is $p<0.05$ (5 %). It did give some confidence for significance, however, the p value (3.3 %) was close to the cut-off value (5 %). When small differences appear in comparison, the simplest approach is to repeat the experiment, for example succinate-malate comparison, to see if similar results will be obtained. If repeatable, the difference is likely to be real. If not, the difference obtained in the first comparison is very likely to be a false positive. Under this circumstance, the multiple comparison will not be reliable since the analysis will calculate the whole group including conditions, replicates etc. and take the average as adjusted p value. Therefore, single comparison is usually preferred in the first instance to see whether there is a significant level under multiple conditions, however, these experiments need to be repeated in order to confirm the observations (Krzywinski and Altman, 2014).

Two *C. elegans* strains (e.g. CF512 strain and N2 strain) have been used in this study and differences in life span have been observed on the same treatment. The mutation in the activity of repeated genes and masculinization of germline in CF512 strain may have caused unknown impact on its life span during carnitine degradation. Additional TMA with either succinate or malate did not affect the longevity of *C. elegans* CF512 strain **(Fig 6.4 & Table 6.4)**, suggesting that TMA is not harmful for the worm.

As discussed above, the bio-degradation of carnitine by *E. coli* SE11 may have affected

the life span of *C. elegans*. Since the additional TMA did not affect the longevity of the worms, it is likely that the intermediate, malate, which is formed from carnitine degradation by the bacterium, may have extended the life span of *C. elegans*. *cnt*A and *cntB* mutants of SE11 are thus required in order to test this hypothesis using the experimental procedures established in this thesis.

In summary, the data presented in this thesis demonstrates the ability of *Enterobacteriaceae* in human gut to utilize choline and carnitine as a source of nutrients. A number of key genes and enzymes essential for the catabolism of choline, such as choline-TMA lyase and microcompartment proteins, have been identified by RNA-Seq and the functions of these key genes have been confirmed by mutagenesis and culture growth in the presence/absence of choline. A number of genes revealed by RNA-Seq analysis, which are highly expressed in the presence of choline, still remain uninvestigated, providing an avenue for further research. In addition, malate, the intermediate of carnitine degradation to TMA, may play a role in host-microbe interaction that expends the life span of *C. elegans*, and this hypothesis warrants further investigation. The results presented in this thesis help better understand the interplay between the host and the gut microbiota in heath and disease.

References

Aballay, A., Yorgey, P., and Ausubel, F.M. (2000) *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of *Caenorhabditis elegans*. *Curr Biol* **10:** 1539-1542.

Ahringer, J. (2006) Reverse genetics. *WormBook: the online review of C.elegans* biology, pp. 1-11.

Aktas, M., Jost, K.A., Fritz, C., and Narberhaus, F. (2011) Choline uptake in *Agrobacterium tumefaciens* by the high-Affinity ChoXWV transporter. *J Bacteriol* **193:** 5119-5129.

Alteri, C.J., Himpsl, S.D., Engstrom, M.D., and Mobley, H.L. (2012) Anaerobic respiration using a complete oxidative TCA cycle drives multicellular swarming in *Proteus mirabilis*. *MBio* **3:** pii: e00365-12.

Altun, Z.F. and Hall, D.H. (2017) Handbook of *C. elegans* Anatomy. In *WormAtlas*: http://www.wormatlas.org/hermaphrodite/hermaphroditehomepage.htm.

Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol* **11:** R106.

Andresen, P.A., Kaasen, I., Styrvold, O.B., Boulnois, G., and Strom, A.R. (1988) Molecular cloning, physical mapping and expression of the bet genes governing the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. *J Gen Microbiol* **134**: 1737-1746.

Armbruster, C.E., and Mobley, H.L.T. (2012) Merging mythology and morphology: the multifaceted lifestyle of *Proteus mirabilis*. *Nat Rev Microbiol* **10:** 743-754.

Armbruster, C.E., Hodges, S.A., and Mobley, H.L. (2013) Initiation of swarming motility by *Proteus mirabilis* occurs in response to specific cues present in urine and requires excess L-glutamine. *J Bacteriol* **195:** 1305-1319.

Armitage, J.P. (1981) Changes in metabolic activity of *Proteus mirabili*s during swarming. *J Gen Microbiol* **125:** 445-450.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., *et al*. (2011) Enterotypes of the human gut microbiome. *Nature* **473:** 174-180.

Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014) A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. *PLoS Comput Biol* **10:** e1003898.

Baker, F.D., Papiska, H.R., and Campbell, L.L. (1962) Choline fermentation by *Desulfovibrio desulfuricans*. *J Bacteriol* **84:** 973-978.

Balagam, B., and Richardson, D.E. (2008) The mechanism of carbon dioxide catalysis in the hydrogen peroxide *N*-oxidation of amines. *Inorg Chem* **47:** 1173-1178.

Baumgart, M., Dogan, B., Rishniw, M., Weitzman, G., Bosworth, B., Yantiss, R., *et al*. (2007) Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of *Clostridiales* in Crohn's disease involving the ileum. *ISME J* **1:** 403-418.

Bennasar, A., Luna, G.D., Cabrer, B., and Lalucat, J. (2000) Rapid identification of *Salmonella typhimurium*, *S. enteritidis* and *S. virchow* isolates by Polymerase Chain Reaction based fingerprinting methods. *Internatl Microbiol* **3:** 31-38.

Bennett, B.J., de Aguiar Vallim, T.Q., Wang, Z., Shih, D.M., Meng, Y., Gregory, J. *et al.* (2013) Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metab* **17:** 49-60.

Bentley, R., and Meganathan, R. (1982) Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiol Rev* **46:** 241-280.

Bidulescu, A., Chambless, L.E., Siega-Riz, A.M., Zeisel, S.H., and Heiss, G. (2007) Usual choline and betaine dietary intake and incident coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) study. *BMC Cardiovasc Disord* **7:**20.

Billi, A.C., Fischer, S.E., and Kim, J.K. (2014) Endogenous RNAi pathways in *C. elegans*. *WormBook: the online review of C. elegans biology,* pp. 1-49.

Blaser, M., Bork, P., Faser, C., Knight, R., and Wang, J. (2013) The microbiome explored: recent insights and future challenges. *Nat Rev Microbiol* **11:** 213-217.

Blaut, M., and Clavel, T. (2007) Metabolic diversity of the intestinal microbiota: implications for health and disease. *J Nutr* **137:** 751-755.

Boch, J., Kempf, B., and Bremer, E. (1994) Osmoregulation in *Bacillus subtilis*: synthesis of the osmoprotectant glycine betaine from exogenously provided choline. *J Bacteriol* **176**: 5364-5371.

Bonkat, G., Braissant, O., Widmer, A.F., Frei, R., Rieken, M., Wyler, S., *et al.* (2012) Rapid detection of urinary tract pathogens using microcalorimetry: principle, technique and first results. *BJU Int* **110:** 892-897.

Bradbeer, C. (1965) The clostridial fermentations of choline and ethanolamine. I. Preparation and properties of cell-free extracts. *J Biol Chem* **240:** 4669-4674.

Brenner, S. (1988) "Foreword" in *The Nematode Caenorhabditis elegans*, eds. W.B. Wood and Community of *C. elegans* Researcher. New York: Cold Spring Laboratory Cold Spring Harbor, pp. IX-XIII.

Brenner, S. (1973) The genetic of behavior. *British Medical Bulletin* **29:** 269-271.

Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics* **77:** 71-94.

Bull, M.J., and Plummer, N.T. (2014) Part 1: The human gut microbiome in health and disease. *Integr Med* **13:** 17-22.

Burall, L.S., Harro, J.M., Li, X., Lockatell, C.V., Himpsl, S.D., Hebel, J.R., *et al*. (2004) *Proteus mirabilis* genes that contribute to pathogenesis of urinary tract infection: identification of 25 signature-tagged mutants attenuated at least 100-fold. *Infect Immun* **72:** 2922-2938.

Byerly, L., Cassada, R.C. and Russell, R.L. (1976) The life cycle of the nematode *Caenorhabditis elegans*. I. Wild-type growth and reproduction. *Dev boil* **51:** 23-33.

Canny, G.O., and McCormick, B.A. (2008) Bacteria in the intestine, helpful residents or enemies from within? *Infect Immun* **76:** 3360-3373.

C. elegans Sequencing Consortium. (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282:** 2012-2018.

Chalmers, R.A., Bain, M.D., Michelakakis, H., Zschocke, J., and Iles, R.A. (2006) Diagnosis and management of trimethylaminuria (FMO3 deficiency) in children. *J Inherit Metab Dis* **29:** 162-172.

Chao, C.K., and Zeisel, S.H. (1990) Formation of trimethylamine from dietary choline by *Streptococcus sanguis* I, which colonizes the mouth. *J Nutr Biochem* **1:** 89-97.

Chen, P., Andersson, D.I., and Roth, J.R. (1994) The control region of the *pdu/cob* regulon in *Salmonella typhimurium*. *J Bacteriol* **176:** 5474-5482.

Chen, Y., Patel, N.A., Crombie, A., Scrivens, J.H., and Murrell, J.C. (2011) Bacterial Flavin-containing monooxygenase is trimethylamine monooxygenase. *PNAS* **108**: 17791-17796.

Chester, D.N., Goldman, J.D., Ahuja, J.K., and Moshfegh, A.J. (2007-2008) Dietary intakes of choline. What we eat in American, NHANES. *US Department of Agriculture*.

Chowdhury, C., Sinha, S., Chun, S., Yeates, T.O., and Bobik, T.A. (2014) Diverse bacterial microcompartment organelles. *Microbiol Mol Biol Rev* **78:** 438-468.

Conlon, M.A., and Bird, A.R. (2015) The impact of diet and lifestyle on gut microbiota and human health. *Nutrients* **7:** 17-44.

Cox, J.C. and Knight, R. (1981) Trimethylamine N-oxide (TMAO) reductase activity in chlorate-resistant or respiration-deficient mutants of *Escherichia coli*. *FEMS Microbiol Lett* **12:** 249-252.

Craciun, S., and Balskus, E.P. (2012) Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *PNAS* **109:** 21307-21312.

Craig, S.A. (2004) Betaine in human nutrition. *Am J Clin Nutr* **80:** 539-549.

Crosatti, M. (2014) Use of non-mammalian models to assess the virulence of

Pseudomonas aeruginosa. PhD thesis. University of Leicester, UK.

Crowley, C.S., Cascio, D., Sawaya, M.R., Kopstein, J.S., Bobik, T.A., and Yeates, T.O. (2010) Structural insight into the mechanisms of transport across the *Salmonella enterica* Pdu microcompartment shell. *J Biol Chem* **285:** 37838-37846.

Cruden, D.L. and Galask, R.P. (1988) Reduction of trimethylamine oxide to trimethylamine by *Mobiluncus* strains isolated from patients with bacterial vaginosis. *Microb Ecol Health Dis* **1:** 95-100.

Darby, C. (2005) Interactions with microbial pathogens. *WormBook: the online review of C.elegans biology,* pp. 1-15.

D'Argenio, V., and Salvatore, F. (2015) The role of the hut microbiome in the healthy adult status. *Clinica Chinica Acta* **452:** 97-102.

Dark, M.J. (2013) Whole-genome sequencing in bacteriology: state of the art. *Infect Drug Resist* **6:** 115-123.

Davanloo, P., Rosenberg, A.H., Dunn, J.J., and Studier, F.W. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *PNAS* **81:** 2035-2039.

Davis, B. D., and Mingioli, S. (1950) Mutants of *E. coli* requiring methionine of vitamin B_{12}. *J. Bacteriol* **60:** 17-28.

Denby, K.J., Rolfe, M.D., Crick, E., Sanguinetti, G., Poole, R.K., and Green, J. (2015) Adaptation of anaerobic cultures of Escherichia coli K-12 in response to environmental trimethylamine-N-oxide. *Environ Microbiol* **17:** 2477-2491.

Ditullio, D., Anderson, D., Chen, C.S., and Sih, C.J. (1994) L-carnitine via enzyme-catalyzed oxidative kinetic resolution. *Bioorg Med Chem* **2:** 415-420.

Dumas, M.E., Barton, R.H., Toye, A., Cloarec, O., Blancher, C., Rothwell, A., *et al*. (2006) Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *PNAS* **103:** 12511-12516.

Dyer, F.E., and Wood, A.J. (1947) Action of *Enterobacteriaceae* on choline and related compounds. *J Fish Res Board Can* **7:** 17-21.

Easter, M.C., Gibson, D.M., and Ward, F.B. (1982) A conductance method for the assay and study of bacterial trimethylamine oxide reduction. *J Appl Bacteriol* **52:** 357-365.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., *et al*. (2005) Diversity of the human intestinal microbial flora. *Science* **308:** 1635-1638.

Eddy, B.P. (1953) Bacterial degradation of choline. *Nature* **171:** 573-574.

Edwards, C.B., Copes, N., Brito, A.G., Canfield, J., and Bradshaw, P.C. (2013) Malate and fumarate extend lifespan in *Caenorhabditis elegans*. *PLoS One* **8:** e58345.

El Kaoutari, A., Armougom, F., Gordon, J.I., Raoult, D., and Henrissat, B. (2013) The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* **11:** 497-504.

Ewbank, J.J. (2006) Signaling in the immune response. *WormBook: the online review of C. elegans biology,* pp. 1-12.

Falkinham, J., and Hoffman, P. (1984) Unique developmental characteristics of the swarm and short cells of *Proteus vulgaris* and *Proteus mirabilis*. *J Bacteriol* **158:** 1037-1040.

Fay, D.S. (2013) Classical genetic methods. *WormBook: the online review of C. elegans biology,* pp. 1-58.

Feller, A.G., and Rudman, D. (1988) Role of carnitine in human nutrition. *J Nutr* **118:** 541-547.

Fennema, D., Phillips, I. R., and Shephard, E. A. (2016) Trimethylamine and trimethylamine N-Oxide, a flavin-containing monooxygenase 3 (FMO3)-mediated host-microbiome metabolic axis implicated in health and disease. *Drug Metab Dispos* **44**: 1839-1850.

Fiebig, K., and Gottschalk, G. (1983) Methanogenesis from choline by a coculture of *Desulfovibrio* sp. and *Methanosarcina barkeri*. *Appl Environ Microbiol* **45:** 161-168.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391:** 806-811.

Fischer, L.M., da Costa, K.A., and Kwock, L. (2007) Sex and menopausal status influence human dietary requirements for the nutrient choline. *Am J Clin Nutr* **85:**1275-1285.

Food and Nutrition Board, Institute of Medicine, Washington, DC. (1998) Dietary reference intakes: thiamin, riboflavin, niacin, vitamin B_6, vitamin B_{12}, pantothenic acid, biotin, and choline. *National Academy of Sciences* p. 390-422.

Fraaije, M.W., Kamerbeek, N.M., van Berkel, W.J.H., and Janssen, D.B. (2002) Identification of a Baeyer–Villiger monooxygenase sequence motif. *FEBS Letters* **518**: 43-47.

Frankel, W., Zhang, L., Singh, A., Klurfeld, D., Don, S., Sakata, T., Modlin, I., *et al.* (1994) Mediation of the trophic effects of short-chain fatty acids on the rat jejunum and colon. *Gastroenterology* **106:** 375-380.

Gaci, N., Borrel, G., Tottey, W., O'Toole, P.W., and Brugere, J.F. (2014) Archaea and the human gut: new beginning of an old story. *World J Gastroenterol* **20:** 16062-16078.

Garrett, W.S., Gallini, C.A., Yatsunenko, T., Michaud, M., DuBois, A., *et al*. (2010) *Enterobacteriaceae* act in concert with the gut microbiota to induce spontaneous and

maternally transmitted colitis. *Cell Host Microbe* **8:** 292-300.

Garsin, D.A., Villanueva, J.M., Begun, J., Kim, D.H., Sifri, C.D., *et al.* (2003) Long-lived *C. elegans daf-2* mutants are resistant to bacterial pathogens. *Science* **300:** 1921.

Gerhard, G.T., and Duell, P.B. (1999) Homocysteine and atherosclerosis. *Curr Opin Lipidol* **10:** 417-428.

Gibellini, F., and Smith, T.K. (2010) The Kennedy pathway-De novo synthesis of phosphatidylethanolamine and phosphatidylcholine. *IUBMB Life* **62:**414-428.

Gosiewski, T., Jurkiewicz-Badacz, D., Sroka, A., Brzychczy-Wloch, M., *et al.* (2014) A novel, nested, multiplex, real-time PCR for detection of bacteria and fungi in blood. *BMC Microbiol* **14:** 144.

Graham, J.E., and Wilkinson, B.J. (1992) *Staphylococcus aureus* osmoregulation: roles for choline, glycine betaine, proline, and taurine. *J Bacteriol* **174**: 2711-2716.

Griffin, J. L., Wang, X., and Stanley, E. (2015) Does our gut microbiome predict cardiovascular risk? A review of the evidence from metabolomics. *Circ Cardiovasc Genet* **8**: 187-191.

Güell, M., van Noort, V., Yus, E., Chen, W.H., Leigh-Bell, J., *et al.* (2009) Transcriptome complexity in a genome-reduced bacterium. *Science* **326:**1268-71.

Guerrerio, A.L., Colvin, R.M., and Schwartz, A.K (2012) Choline intake in a large cohort of patients with nonalcoholic fatty liver disease. *Am J Clin Nutr* **95:** 892-900.

Hampel, K.J., LaBauve, A.E., Meadows, J.A., Fitzsimmons, L.F., Nock, A.M., *et al*. (2014) Characterization of the GbdR Regulon in *Pseudomonas aeruginosa*. *J Microbiol* **196:**7-15.

Han, Y., Gao, S., Muegge, K., Zhang, W., and Zhou, B. (2015) Advanced Applications of RNA Sequencing and Challenges. *Bioinform Biol Insights* **9:** 29-46.

Hauser, G. (1885) On putrefactive bacteria and their relationship to sepsis.

Havemann, G.D., Sampson, E.M., and Bobik, T.A. (2002) PduA is a shell protein of polyhedral organelles involved in coenzyme $B_{(12)}$-dependent degradation of 1, 2-propanediol in *Salmonella enterica* serovar Typhimurium LT2. *J Bacteriol* **184:** 1253-1261.

Hayward, H.R., and Stadtman, T.C. (1959) Anaerobic degradation of choline. I. Fermentation of choline by an anaerobic, cytochrome-producing bacterium, *Vibrio cholinicusn* sp. *J Bacteriol* **78:** 557-561.

Hayward, H.R., and Stadtman, T.C. (1960) Anaerobic degradation of choline. II. Preparation and properties of cell-free extracts of *Vibrio cholinicus*. *J Biol Chem* **235:** 538-543.

Hekimi, S., Lakowski, B., Barnes, T.M., and Ewbank, J.J. (1998) Molecular genetics of life span in *C. elegans*: how much does it teach us? *Trends Genet* **14:** 14-20.

Hendler, S.S., and Rorvik, D.R. (2008) PDR for Nutritional Supplements. 2nd edition. *Thomson Reuters*.

Hirsch, M.J., Growdon, J.H., and Wurtman, R.J. (1978) Relations between dietary choline or lecithin intake, serum choline levels, and various metabolic indices. *Metabolism* **27:** 953-960.

Hodgkin, J., Kuwabara, P.E., and Corneliussen, B. (2000) A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Current biology* **10:** 1615-1618.

Hormann, K., and Andreesen, J.R. (1989) Reductive cleavage of sarcosine and betaine by *Eubacterium acidaminophilum* via enzyme systems different from glycine reductase. *Arch Microbiol* **153:** 50-59.

Hsin, H. (2007) *CF512*. Homepage of CGC, Online available: http://www.cgc.cbs.umn.edu/strain.php?id=7607

Iancu, C.V., Ding, H.J., Morris, D.M., Dias, D.P., Gonzales, A.D., Martino, A., *et al.* (2007) The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography. *J. Mol. Biol* **372:** 764-773.

Ikawa, M., and Taylor, R.F. (1973) Choline and related substances in Algae. In *Marine Pharmacognosy: Action of Marine Biotoxins at the cellular level*. Martin, D., and Padilla, G.M. (eds). New York: Academic Press INC., pp. 203-236.

Irazoqui, J.E., and Ausubel, F.M. (2010) 99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: *Caenorhabditis elegans* as a model to study tissues involved in host immunity and microbial pathogenesis. *Clin Exp Immunol* **160:** 48-57.

Irazoqui, J.E., Ng, A., Xavier, R.J., and Ausubel, F.M. (2008) Role for beta-catenin and HOX transcription factors in *Caenorhabditis elegans* and mammalian host epithelial-pathogen interactions. *PNAS* **105:** 17469-17474.

Irazoqui, J.E., Urbach, J.M. and Ausubel, F.M. (2010) Evolution of host innate defense: insights from *Caenorhabditis elegans* and primitive invertebrates. *Nat Rev Immunol* **10:** 47-58.

Ishimoto, M., and Shimokawa, O. (1978) Reduction of trimethylamine N-oxide by *Escherichia coli* as anaerobic respiration. *Z Allg Mikrobiol* **18:** 173-181.

Jacob, R.A., Jenden, D.J., Allman-Farinelli, M.A., and Swendseid, M.E. (1999) Folate nutriture alters choline status of women and men fed low choline diets. *J Nutr* **129:** 712-717.

Jameson, E., Fu, T., Brown, I.R., Paszkiewicz, K., Purdy, K.J., Frank, S., and Chen, Y. (2015) Anaerobic choline metabolism in microcompartments promotes growth and swarming of *Proteus mirabilis*. *Environ Microbiol* **18:** 2886-2898.

Jayamani, E., Rajamuthiah, R., Larkins-Ford, J., Fuchs, B.B., Conery, A.L., *et al.* (2015) Insect-derived cecropins display activity against *Acinetobacter baumannii* in a whole-animal high-throughput *Caenorhabditis elegans* model. *Antimicrob Agents Chemother* **59:**1728-1737.

Jebamercy, G., and Balamurugan, K. (2012) Effects of sequential infections of *Caenorhabditis elegans* with *Staphylococcus aureus* and *Proteus mirabilis*. *Microbiol Immunol* **56:** 825-835.

Jünger, M., Vautz, W., Kuhns, M., Hofmann, L., Ulbricht, S., *et al.* (2012) Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. *App Microbiol Biotechnol* **93:** 2603-2614.

Kalnins, G., Kuka, J., Grinberga, S., Makrecka-Kuka, M., Liepinsh, E., *et al.* (2015) Structure and function of CutC choline lyase from human microbiota bacterium *Klebsiella pneumonia*. *J Biol Chem* **290:** 21732-21740.

Kaniga, K., Delor, I., and Cornelis, G.R. (1991) A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria: inactivation of the *bla*A gene of *Yersinia enterocolitica*. *Gene* **109:** 137-141.

Kawai, K., Fujita, M., and Nakao, M. (1974) Lipid components of two different regions of an intestinal epithelial cell membrane of mouse. *Biochim Biophys Acta* **369:** 222-233.

Kelly, B., and Appleman, M.D. (1961) Degradation of ergothioneine by cell-free extracts of *Alcaligenes faecalis*. *J Bacteriol* **81:** 715-720.

Kennedy, E.P., and Weiss, S.B. (1956) The function of cytidine coenzymes in the biosynthesis of phospholipids. *J Bio Chem* **222:** 193-214.

Kim, D.H., Feinbaum, R., Alloing, G., Emerson, F.E., Garsin, D.A., *et al.* (2002) A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* **297:** 623-626.

Kim, Y.I., Miller, J.W., and da Costa, K.A. (1994) Severe folate deficiency causes secondary depletion of choline and phosphocholine in rat liver. *J Nutr* **124:** 2197-2203.

King, G.M. (1984) Metabolism of trimethylamine, choline, and glycine betaine by sulfate-reducing and methanogenic bacteria in marine sediments. *Appl Environ Microbiol* **48**: 719-725.

Kleber, H.P., Seim, H., Aurich, H., and Strack, E. (1978) Interrelationships between carnitine metabolism and fatty acid assimilation in *Pseudomonas putida*. *Arch Microbiol* **116:** 213-220.

Kleber, H.P., Seim, H., Aurich, H., and Strack, E. (1977) Utilization of trimethylammonium compounds *by Acinetobacter calcoaceticus*. *Arch Microbiol* **112:** 201-206.

Koc, H., Mar, M.H., Ranasinghe, A., Swenberg, J.A., and Zeisel, S.H. (2002) Quantitation of choline and its metabolites in tissues and foods by liquid chromatography/electrospray ionization-isotope dilution mass spectrometry. *Anal Chem* **74:** 4734-4740.

Koeth, R.A., Wang, Z., Levison, B.S., Buffa, J.A., Org, E., Sheehy, B.T., *et al.* (2013) Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, pro- motes atherosclerosis. *Nat Med* **19:** 576–585.

Kofoid, E., Rappleye, C., Stojiljkovic, I., and Roth, J. (1999) The 17-Gene ethanolamine (*eut*) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. *J Bacteriol* **181:** 5317-5329.

Konstantinova, S.V., Tell, G.S., Vollset, S.E., Nygard, O., Bleie, O., and Ueland, P.M. (2008) Divergent associations of plasma choline and betaine with components of metabolic syndrome in middle age and elderly men and women. *J Nutr* **138:** 914-920.

Kornfeld, K. (1997) Vulval development in *Caenorhabditis elegans*. *Trends Genet* **13:** 55-61.

Krogius-Kurikka, L., Lyra, A., Malinen, E., Aarnikunnas, J., Tuimala, J., *et al*. (2009) Microbial community analysis reveals high level phylogenetic alterations in the overall gastrointestinal microbiota of diarrhoea predominant irritable bowel syndrome sufferers. *BMC Gastroenterol* **9:** 95.

Krueger, S.K., and Williams, D.E. (2005) Mammalian flavin-containing monooxygenases: structure/ function, genetic polymorphisms and role in drug metabolism. *Pharmacol Therapeut* **106**: 357-387.

Krzywinski, M., and Altman, N. (2014) Comparing samples-part II: when a large number of tests are performed, *P* values must be interpreted differently. *Nat Methods* **11:** 355-356.

Kuehl, J.V., Price, M.N., Ray, J., Wetmore, K.M., Esquivel, Z., Kazakov, A.E., *et al*. (2014) Functional genomics with a comprehensive library of transposon mutants for the sulfate-reducing bacterium *Desulfovibrio alaskensis* G20. *MBio* **5:** e01041-14.

Kuka, J., Liepinsh, E., Makrecka-Kuka, M., Liepins, J., Cirule, H., *et al.* (2014) Suppression of intestinal microbiota-dependent production of proatherogenic trimethylamine N-oxide by shifting L-carnitine microbial degradation. *Life Sci* **117:** 84-92.

Kurz, C.L., and Ewbank, J.J. (2007) Infection in a dish: high-throughput analyses of bacterial pathogenesis. *Curr Opinn Microbiol* **10:** 10-16.

Kwan, H.S., and Barrett, E.L. (1983) Purification and properties of trimethylamine oxide reductase from *Salmonella typhimurium*. *J Bacteriol* **155:** 1455-1458.

Lamark, T., Kaasen, I., Eshoo, M., Falkenberg, P., McDougall, J., and Strøm, A. (1991) DNA sequence and analysis of the *bet* genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. *Mol Microbiol* **5:** 1049-1064.

Landfald, B., and Strøm, A.R. (1986) Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. *J Bacteriol* **165**: 849-855.

Lawley, T.D., and Walker, A.W. (2013) Intestinal colonization resistance. *Immunology* **138:** 1-11.

Leach, N.V., Dronca, E., and Vesa, S.C. (2014) Serum homocysteine levels, oxidative stress and cardiovascular risk in non-alcoholic steatohepatitis. *Eur J Intern Med* **25:** 762-767.

Lee, J.V., Gibson, D.M., and Shewan, J.M. (1977) A numerical taxonomic study of some Pseudomonas-like marine bacteria. *J Gen Microbiol* **98:** 439-451.

Lehtiö, L., and Goldman, A. (2004) The pyruvate formate lyase family: sequences, structures and activation. *Protein Eng Des Sel* **17:** 545-552.

Letoffe, S., Audrain, B., Bernier, S.P., Delepierre, M., and Ghigo, J.M. (2014) Aerial exposure to the bacterial volatile compound trimethylamine modifies antibiotic resistance of physically separated bacteria by raising culture medium pH. *Mbio* **5:** 1-12.

Liaw, S.J., Lai, H.C., Ho, S.W., Luh, K.T., and Wang W.B. (2001) Characterisation of p-nitrophenylglycerol-resistant *Proteus mirabilis* super-swarming mutants. *J Med Microbiol* **50:** 1039-1048.

Liaw, S.J., Lai, H.C., Ho, S.W., Luh, K.T., and Wang W.B. (2003) Role of RsmA in the regulation of swarming motility and virulence factor expression in *Proteus mirabilis*. *J of medical microbiology* **52:** 19-28.

Lidbury, I. (2015) Microbial methylated amine metabolism in marine surface waters. PhD thesis. University of Warwick, UK.

Lidbury, I., Murrell, C., and Chen, Y. (2014) Trimethylamine N-oxide metabolism by abundant marine heterotrophic bacteria. *PNAS* **111:** 2710-2715.

Linares, D., Ross, P., and Stanton, C. (2016) Beneficial microbes: the pharmacy in the gut. *Bioengineered* **7:** 11-20.

Lin, J.K., and Hurng, D.C. (1989) Potentiation of ferrous sulphate and ascorbate on the microbial transformation of endogenous trimethylamine N-oxide to trimethylamine and dimethylamine in squid extracts. *Food Chem Toxicol* **27:** 613-618.

Li, Z., and Vance, D.E. (2008) Phosphatidylcholine and choline homeostasis. *J Lipid*

Res. **49:** 1187-1194.

Ludwig, W., Strunk, O., Westran, R., Richter, L., Meier, H., Yadhukumar, *et al*. (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32:** 1363-1371.

Lundberg, P., Dudman, N.P., Kuchel, P.W., and Wilcken, D.E. (1995) 1H NMR determination of urinary betaine in patients with premature vascular disease and mild homocysteinemia. *Clin Chem* **41:** 275-283.

Marchesi, J., and Shanahan, F. (2007) The normal intestinal microbiota. *Curr Opin Infect Dis* **20:** 508-513.

Marin, I., and Baker, B.S. (1998) The evolutionary dynamics of sex determination. *Science* **281:** 1990-1994.

Martínez-del Campo, A., Bodea, S., Hamer, H.A., Haiser, H.J., Turnbaugh, P.J., *et al.* (2015) Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. *mBio* **6:** e00042-15.

Max, C.J., and Lidstrom, M.E. (2002) Broad-host-range *cre-lox* system for antibiotic marker recycling in gram-negative bacteria. *Biothchniques* **33:** 1062-1067.

McGhee, J.D. (2007) The *C. elegans* intestine. *WormBook: the online review of C. elegans biology,* pp. 1-36.

Meadows, J.A., and Wargo, M.J. (2015) Carnitine in bacterial physiology and metabolism. *Microbiology* **161:** 1161–1174.

Metzstein, M.M., Stanfield, G.M., and Horvitz, H.R. (1998) Genetics of programmed cell death in *C. elegans*: past, present and future. *Trends Genet* **14:** 410-416.

Mitchell, S.C., and Smith, R.L. (2001) Trimethylaminuria: The fish malodor syndrome. *Drug Metab Dispos* **29**: 517-521.

Mitchell, S.C., Zhang, A.Q., and Smith, R.L. (2002) Chemical and biological liberation of trimethylamine from foods. *J Food Compos Anal* **15:** 277-282.

Mitschke, J., Georg, J., Scholz, I., *et al*. (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *PNAS* **108:** 2124-2129.

Miura-Fraboni, J., Kleber, H.P., and Englard, S. (1982) Assimilation of g-butyrobetaine, and D-and L-carnitine by resting cell suspensions of Acinetobacter calcoaceticus and Pseudomonas putida. *Arch Microbiol* **133:** 217-221.

Mobley, H.L.T., and Chippendale, G.R. (1990) Hemagglutinin, urease, and hemolysin production by *Proteus mirabilis* from clinical sources. *J. Infect. Dis* **161:**525-530.

Möller, B., Hippe, H., and Gottschalk, G. (1986) Degradation of various amine compounds by mesophilic clostridia. *Arch Microbiol* **145:** 85-90.

Morgenstein, R.M., and Rather P.N. (2012) Role of Umo proteins and the Rsc phosphorelay in the swarming motility of the wide type and an O-antigen (waaL) mutant of *Proteus mirabilis*. *J Bacteriol* **194:**669-676.

Naumann, E., Hippe, H., and Gottschalk, G. (1983) Betaine: New oxidant in the Stickland reaction and methanogenesis from betaine and l-alanine by a *Clostridium sporogenes-Methanosarcina barkeri* coculture. *Appl Environ Microbiol* **45:** 474-483.

Nicholson, J.K., Holmes, E., and Wilson, I.D. (2005) Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* **3:** 431-438.

Noga, A.A., and Vance, D.E. (2003) A gender-specific role for phosphatidylethanolamine N-methyltransferase-derived phosphatidylcholine in the regulation of plasma high density and very low density lipoproteins in mice. *J Biol Chem* **278:** 21851-21859.

Noga, A.A., Zhao, Y., and Vance, D.E. (2002) An unexpected requirement for phosphatidylethanolamine N-methyltransferase in the secretion of very low density lipoproteins. *J Biol Chem* **277:** 42358-42365.

Nookaew, I., Papini, M., Pornputtapong, N., Scalcinati, G., Fagerberg, L., *et al.* (2012) A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **(20):** 10084-10097.

Normann, E., Fahlen, A., Engstrand, L., and Lilja, H.E. (2013) Intestinal microbial profiles in extremely preterm infants with and without necrotizing enterocolitis. *Acta Paediatr* **102:** 129-136.

Obeid, R., Awwad, H.M., Rabagny, Y., Graeber, S., Herrmann, W., and Geisel, J. (2016) Plasma trimethylamine N-oxide concentration is associated with choline, phospholipids, and methyl metabolism. *Am J Clin Nutr* **103:** 703-711.

Oshima, K., Toh, H., Ogura, Y., Sasamoto, H., Morita, H., Park, S.H., Ooka, T., *et al.* (2008) Complete genome sequence and comparative analysis of the wild-type commensal *Escherichia coli* strain SE11 isolated from a healthy adult. *DNA Res* **15:** 375-86.

Pang, A., Frank, S., Brown, I., Warren, M.J., and Pickersgill, R.W. (2014) Structural insights into higher order assembly and function of the bacterial microcompartment protein PduA. *J Biol Chem* **289:** 22377-22384.

Pang, P.K.T., Griffith, R.W., and Atz, J.W. (1977) Osmoregulation in elasmobranchs. *Am Zool* **17:** 365–377.

Parsons, J.B., Frank, S., Bhella, D., Liang, M., Prentice, M.B., Mulvihill, D.P., *et al.* (2010) Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Mol Cell* **38:** 305-315.

Pearson, M.M., Sebaihia, M., Churcher, C., Quail, M.A., Seshasayee, A.S., *et al*. (2008) Complete Genome Sequence of Uropathogenic *Proteus mirabilis*, a Master of both Adherence and Motility. *J Bacteriol* **11:**4027-4037.

Pellanda, H. (2013) Betaine homocysteine methyltransferase (BHMT)-dependent remethylation pathway in human healthy and tumoral liver. *Clin Chem Lab Med* **51:** 617-621.

Penrod, J.T., and Roth, J.R. (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in *Salmonella enterica*. *J Bacteriol* **188:** 2865-2874.

Petersson, J., Schreiber, O., Hansson, G.C., Gendler, S.J., Velcich, A., *et al*. (2011) Importance and regulation of the colonic mucus barrier in a mouse model of colitis. *Am J Physiol Gastrointest Liver Physiol* **300:** G327-G333.

Plotkin, J.B., and Kudla, G. (2011) Synonymous but not the same: The causes and consequences of codon bias. *Nat Rev Gene*t **12:** 32-42.

Postgate, J., and Kent, H. (1985) Expression of *Klebsiella pneumoniae* nif genes in *Proteus mirabilis*. *Arch Microbiol* **142:** 289-294.

Price-Carter, M., Tingey, J., Bobik, T.A., and Roth, J.R. (2001) The alternative electron acceptor tetrathionate supports B_{12}-dependent anaerobic growth of *Salmonella enterica* serovar Typhimurium on ethanolamine or 1, 2-propanediol. *J Bacteriol* **183:** 2463-2475.

Pujol, N., Link, E.M., Liu, L.X., Kurz, C.L., Alloing, G., Tan, M.W., Ray, K.P., *et al.* (2001) A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Curr Biol* **11:** 809-821.

Quinlan, A.R. (2015) BEDTools: the Swiss-army tool for genome feature analysis. *Curr Protoc Bioinformatics* **47:** 11.12.1-11.12.34.

Rather, P.N. (2005) Swarmer cell differentiation in *Proteus mirabilis*. *Environ Microbiol*. **7:**1065-73.

Ravcheev, D.A., Godzik, A., Osterman, A.L., and Rodionov, D.A. (2013) Polysaccharides utilization in human gut bacterium *Bacteroides thetaiotaomicron*: comparative genomics reconstruction of metabolic and regulatory networks. *BMC Genomics* **14:** 873.

Rebouche, C.J., and Chenard, C.A. (1991) Metabolic fate of dietary carnitine in human adults: identification and quantification of urinary and fecal metabolites. *J Nutr* **121:** 539-546.

Reinhardt, C., Bergentall, M., Greiner, T.U., Schaffner, F., Östergren-Lundén, G., *et al*. (2012) Tissue factor and PAR1 promote microbiota-induced intestinal vascular remodeling. *Nature* **483:** 627-631.

Riddle, D.L., Blumenthal, T., Meyer, B.J., and Priess, J.R. (1997) *C. elegans II.* USA: Cold Spring Harbor Laboratory Press, Plainview (NY).

Robinson, J., Gibbons, N.E., and Thatcher, F.S. (1952) A mechanism of halophilism in *Micrococcus halodenitrificans*. *J Bacteriol* **64:** 69-77.

Romano, K.A., Vivas, E.I., Amador-Noguez, D., and Rey, F.E. (2015) Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. *MBio* **6:** e02481.

Sambrook, J., Russell, D., and Irwin, N. (2001) *Molecular cloning: a laboratory manual*. New York, Cold Spring Harbor Laboratory Press.

Sandhu, S.S., and Chase, T., Jr (1986) Aerobic degradation of choline by *Proteus mirabilis*: enzymatic requirement and pathway. *Can J Microbiol* **32:** 743-750.

Schäfer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G., and Puhler, A. (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum. Gene* **145:** 69-73.

Schaffer, J.N., and Pearson, M. (2016) *Proteus mirabilis* and Urinary Tract Infections. *Microbiol Spectr* **3:** doi:10.1128.

Schmid, M.F., Paredes, A.M., Khant, H.A., Soyer, F., Aldrich, H.C., Chiu, W., *et al.* (2006) Structure of *Halothiobacillus neapolitanus* carboxysomes by cryo-electron tomography. *J Mol Biol* **364:** 526-535.

Schoberth, S., and Gottschalk, G. (1969) Consideration on the energy metabolism of *Clostridium kluyveri*. *Arch Microbiol* **65:** 318-328.

Secades, J.J. (2010) Citicoline: pharmacological and clinical review. *Rev Neurol* **2:** S1-S62.

Seim, H., Löster, H., Claus, R., Kleber, H.P., and Strack, E. (1982a) Formation of γ-Butyrobetaine and trimethylamine from quaternary ammonium compounds structure-related to L-carnitine and choline by *Proteus vulgaris*. *FEMS Microbiol Lett* **13:** 201-205.

Seim, H., Löster, H., Claus, R., Kleber, H.P., and Strack, E. (1982b) Splitting of the C-N bond in carnitine by an enzyme (trimethylamine forming) from membranes of *Acinetobacter calcoaceticus*. *FEMS Microbiol Lett* **15:** 165-167.

Sekirov, I., Russell, S.L., Antunes, L.C., and Finlay, B.B. (2010) Gut microbiota in health and disease. *Physioll Rev* **90:** 859-904.

Sellars, M.J., Hall, S.J., and Kelly, D.J. (2002) Growth of *Campylobacter jejuni* supported by respiration of fumarate, nitrate, nitrite, trimethylamine-N-oxide, or dimethyl sulfoxide requires oxygen. *J Bacteriol* **184:** 4187-4196.

Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., *et al*. (2010b) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464:** 250-5.

Shin, Y.C., Bischof, G.F., Lauer, W.A., and Desrosiers, R.C. (2015) Importance of codon usage for the temporal regulation of viral gene expression. *PNAS* **112:** 14030-14035.

Shreiner, A.B., Kao, J.Y, and Young, V.B. (2015) The gut microbiome in health and in disease. *Curr Opin Gastroenterol* **31:** 69-75.

Sifri, C.D., Begun, J., and Ausubel, F.M. (2005) The worm has turned-microbial virulence modeled in *Caenorhabditis elegans*. *Trends Microbiol,* **13:** 119-127.

Simenhoff, M.L., Saukkonen, J.J., Burke, J.F., Wesson, L.G., and Schaedler, R.W. (1976) Amine metabolism and the small bowel in uraemia. *Lancet* **2:** 818-821.

Sinha, S., Cheng, S., Fan, C., and Bobik, T.A. (2012) The PduM protein is a structural component of the microcompartments involved in coenzyme B_{12}-dependent 1,2-Propanediol degradation by *Salmonella enterica*. *J Bacteriol* **194:** 1912-1918.

Smith, L.T., Pocard, J.A., Bernard, T., and Le Rudulier, D. (1988) Osmotic control of glycine betaine biosynthesis and degradation in *Rhizobium meliloti*. *J Bacteriol* **170:** 3142-3149.

Spencer, P.S., and Barral, J.M. (2012) Genetic code redundancy and its influence on the encoded polypeptides. *Comput Struct Biotechnol J* **1:** 1-8.

Stead, L.M., Brosnan, J.T., Brosnan, M.E., Vance, D.E., and Jacobs, R.L. (2006) Is it time to re-evaluate methyl balance in humans? *Am J Clin Nutr* **83:**5-10.

Stenberg, E., Styrvold, O.B., and Strøm, A.R. (1982) Trimethylamine oxide respiration in *Proteus* sp. strain NTHC153: electron transfer-dependent phosphorylation and L-serine transport. *J Bacteriol* **149:** 22-28.

Stiernagle, T. (2006) Maintenance of *C. elegans*. *WormBook: the online review of C. elegans biology*, pp. 1-11.

Strøm, A.R., Olafsen, J.A., and Larsen, H. (1979) Trimethylamine oxide: a terminal electron acceptor in anaerobic respiration of bacteria. *J Gen Microbiol* **112:** 315-320.

Styrvold, O.B., Falkenberg, P., Landfald, B., Eshoo, M.W., Bjørnsen, T., *et al.* (1986) Selection, mapping, and characterization of osmoregulatory mutants of *Escherichia coli* blocked in the choline-glycine betaine pathway. *J Bacteriol* **165**: 856-863.

Swann, J., Wang, Y., Abecia, L., Costabile, A., Tuohy, K., Gibson, G., *et al*. (2009) Gut microbiome modulates the toxicity of hydrazine: a metabonomic study. *Mol Biosyst* **5:** 351-355.

Szabo, C., Ischiropoulos, H., and Radi, R. (2007) Peroxynitrite: biochemistry, pathophysiology and development of therapeutics. *Nat Rev Drug Discov* **6:** 662-680.

Takagi, M., Tsuchiya, T., and Ishimoto, M. (1981) Proton translocation coupled to trimethylamine N-oxide reduction in anaerobically grown *Escherichia coli*. *J Bacteriol*

148: 762-768.

Tanaka, S., Kerfeld, C.A., Sawaya, M.R., Cai, F., Heinhorst, S., Cannon, G.C., *et al.* (2008) Atomic-level models of the bacterial carboxysome shell. *Science* **319:** 1083-1086.

Tang, W.H.W., and Hazen, S.L. (2014) The contributory role of gut microbiota in cardiovascular disease. *J Clin Invest* **124:** 4204-4211.

Tang, W.H., Wang, Z., Levison, B.S., Koeth, R.A., Britt, E.B., Fu, X., Wu, Y., *et al.* (2013) Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med* **368:** 1575-1584.

Tan, M.W., Mahajan-Miklos, S., and Ausubel, F.M. (1999) Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *PNAS* **96:** 715-720.

Tenor, J.L., and Aballay, A. (2008) A conserved Toll-like receptor is required for *Caenorhabditis elegans* innate immunity. *EMBO Rep* **9:** 103-109.

Thiennimitr, P., Winter, S.E., Winter, M.G., Xavier, M.N., Tolstikov, V., *et al*. (2011) Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. *PNAS* **108:** 17480-17485.

Todd, J.D., Kirkwood, M., Newton-Payne, S., and Johnston, A.W.B. (2012) DddW, a third DMSP lyase in a model *Roseobacter* marine bacterium, *Ruegeria pomeroyi* DSS-3. *ISME J* **6:** 223-226.

Troemel, E.R., Chu, S.W., Reinke, V., Lee, S.S., Ausubel, F.M., and Kim, D.H. (2006) p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet* **2:** e183.

Ueland, P.M. (2011) Choline and betaine in health and disease. *J Inherit Metab Dis* **34:**3-15.

Vance, D.E. (2014) Phospholipid methylation in mammals: from biochemistry to physiological function. *Biochimica et Biophysica Acta* **1838:** 1477-1487.

Visalli, M.A., Murphy, E., Projan, S.J., and Bradford, P.A. (2003) AcrAB multidrug efflux pump is associated with reduced levels of susceptibility to tigecycline (GAR-936) in *Proteus mirabilis*. *Antimicrob Agents Chemother* **47:** 665-669.

Vollmer, W., and Tomasz, A. (2001) Identification of the teichoic acid phosphorylcholine esterase in *Streptococcus pneumoniae*. *Mol Microbiol* **39**: 1610-1622.

Wang, Z., Klipfell, E., Bennett, B.J., Koeth, R., Levison, B.S., Dugar, B., *et al*. (2011) Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472:**57-63.

Wargo, M.J., Ho, T.C., Gross, M.J., Whittaker, L.A., and Hogan, D.A. (2009) GbdR regulates *Pseudomonas aeruginosa plcH* and *pchP* transcription in response to choline catabolites. *Infect Immun* **77**: 1103-1111.

Wargo, M.J., Szwergold, B.S., and Hogan, D.A. (2008) Identification of two gene clusters and a transcriptional regulator required for *Pseudomonas aeruginosa* glycine betaine catabolism. *J Bacteriol* **190**: 2690-2699.

Wessel, M., Klüsener, S., Gödeke, J., Fritz, C., Hacker, S., and Narberhaus, F. (2006) Virulence of *Agrobacterium tumefaciens* requires phosphatidylcholine in the bacterial membrane. *Mol Microbiol* **62**: 906-915.

Wilkerson, M.L., and Niederhoffer, E.C. (1995) Swarming characteristics of *Proteus mirabilis* under anaerobic and aerobic conditions. *Anaerobe* **1:** 345-350.

Williams, D.E., Hale, S.E., Muerhoff, A.S., and Masters, B.S. (1985) Rabbit lung flavin-containing monooxygenase: Purification, characterization, and induction during pregnancy. *Mol Pharmacol* **28**: 381-390.

Williams, F.D., and Schwarzhoff, R.H. (1978) Nature of the swarming phenomenon in *Proteus*. *Annu Rev Microbiol* **32:** 101-138.

Winter, S.E., Lopez, C.A., and Bäumler, A.J. (2013a) The dynamics of gut-associated microbial communities during inflammation. *EMBO Rep* **14:** 319-327.

Winter, S.E., Thiennimitr, P., Winter, M.G., Butler, B.P., Huseby, D.L., *et al*. (2010) Gut inflammation provides a respiratory electron acceptor for *Salmonella*. *Nature* **467:** 426-429.

Winter, S.E., Winter, M.G., Xavier, M.N., Thiennimitr, P., Poon, V., *et al*. (2013b) Host-derived nitrate boosts growth of *E. coli* in the inflamed gut. *Science* **339:** 708-711.

Wolff, J.B. (1962) Ergothionase from *Escherichia coli*. *J Biol Chem* **237:** 874-881.

Wolkow, C.A., Kimura, K.D., Lee, M.S., and Ruvkun, G. (2000) Regulation of *C. elegans* life-span by insulin-like signaling in the nervous system. *Science* **290:** 147-150.

Wostmann, B. (1981) The germfree animal in nutritional studies. *Ann Rev Nutr* **1:** 257-279.

Wurtzel, O., Yoder-Himes, D.R., Han, K., Dandekar, A.A., Edelheit, S., *et al.* (2012) The single-nucleotide resolution transcriptome of *Pseudomonas aeruginosa* grown in body temperature. *PLoS Pathog* **8:** e1002945.

Yanasugondha, D., and Appleman, M.D. (1957) Degradation of ergothioneine by *Alcaligenes faecalis*. *J Bacteriol* **74:** 381-385.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., *et al*. (2012) Human gut microbiome viewed across age and geography. *Nature* **486:** 222-227.

Yeates, T.O., Crowley, C.S., and Tanaka, S. (2010) Bacterial microcompartment organelles: protein shell structure and evolution. *Annu Rev Biophys* **39:** 185-205.

Yu, F., Pan, J., Ding, B., Yang, L., Zhang, X., and Wang, L. (2010) Prevalence of plasmid-mediated 16S rRNA methylase genes among *Proteus mirabilis* isolates from a Chinese hospital. *Afr J Microbiol Res* **4:** 2790-2794.

Yu, D., Shu, X.O., and Xiang, Y.B. (2014) Higher dietary choline intake is associated with lower risk of nonalcoholic fatty liver in normal weight Chinese women. *J Nutr* **144:** 2034-2040.

Zeisel, S.H. (1990) Choline deficiency. *J Nutr Biochem* **1:** 332-349.

Zeisel, S.H., and Corbin, K.D. (2012) Choline. *Present Knowledge in Nutrition. 10th ed: John Wiley & Sons, Inc* 405-418.

Zeisel, S.H., and da Costa, K.A. (2009) Choline: an essential nutrient for public health. *Nut Rev* **67:** 615-623.

Zeisel, S.H., da Costa, K.A., Youssef, M., and Hensey, S. (1989) Conversion of dietary choline to trimethylamine and dimethylamine in rats: dose-response relationship. *J Nutr* **119:** 800-804.

Zeisel, S.H., Mar, M.H., Howe, J.C., and Holden, J.M. (2003) Concentrations of choline-containing compounds and betaine in common foods. *J Nutr* **133:**1302-1307.

Zeisel, S.H., Wishnok, J.S., and Blusztajn, J.K. (1983) Formation of methylamines from ingested choline and lecithin. *J Pharmacol Exp Ther* **225:** 320-324.

Zeng, W., and Mortazavi, A. (2012) Technical considerations for functional sequencing assays. *Nature Immunology* **13:** 802-807.

Zerbst-Boroffka, I., Kamaltynow, R.M., Harjes, S., Kinne-Saffran, E., and Gross, J. (2005) TMAO and other organic osmolytes in the muscles of amphipods (Crustacea) from shallow and deep water of Lake Baikal. *Comp Biochem Physiol A Mol Integr Physiol* **142:** 58-64.

Zhang, A.Q., Mitchell, S.C., and Smith, R.L. (1999) Dietary precursors of trimethylamine in man: a pilot study. *Food Chem Toxicol* **37:** 515-520.

Zhang, Y., Lu, H. and Bargmann, C.I. (2005) Pathogenic bacteria induce aversive olfactory learning in *Caenorhabditis elegans*. *Nature* **438:** 179-184.

Zhao, J., Cheah, S., Roberts, K.D., Nation, R.L., Thompson, P.E., Velkov, T., *et al*. (2016) Transcriptomic analysis of the activity of a novel polymyxin against *Staphylococcus aureus*. *mSphere* **1:** e00119-16.

Zhou, J., and Austin, R.C. (2009) Contributions of hyperhomocysteinemia to atherosclerosis: Causal relationship and potential mechanisms. *Biofactors*. **35:** 120-129.

Zhu, Y., Jameson, E., Crosatti, M., Schäfer, H., Rajakumar, K., Bugg, T.D.H., and Chen, Y. (2014) Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota. *PNAS* **111:** 4268-4273.

Zindel, U., Freudenberg, W., Rieth, M., Andreesen, J.R., Schnell, J., and Widdel, F. (1988) *Eubacterium acidaminophilum* sp. nov., a versatile amino acid-degrading anaerobe producing or utilizing H_2 or formate. *Arch Microbiol* **150:** 254-266.