

Studying novel sporulation genes in *Clostridium sporogenes*

Rizvi, Syeda Rubab Zehra

14267200



Thesis submitted to the Faculty of Life Sciences in fulfilment of the requirement of the degree of Doctor of Philosophy

JANUARY 30, 2020

UNIVERSITY OF NOTTINGHAM
University Park
NG7 2RD Nottingham

OVERVIEW OF CONTENTS

1. DECLARATION
2. ABSTRACT
3. ACKNOWLEDGEMENTS
4. TABLE OF CONTENTS
5. ABBREVIATIONS
6. CHAPTER ONE: INTRODUCTION
7. CHAPTER TWO: MATERIALS AND METHODS
8. CHAPTER THREE: SCREENING FOR SPORULATION GENES
9. CHAPTER FOUR: CHARACTERIZATION OF *yqfD*
10. CHAPTER FIVE: CHARACTERIZATION OF *divIVA* and *hypP*
11. CHAPTER SIX: CHARACTERIZATION OF *clpX*
12. CHAPTER SEVEN: CHARACTERIZATION OF *spoVK*
13. CHAPTER EIGHT: CHARACTERIZATION OF *spoIID*
14. CHAPTER NINE: GENERAL DISCUSSION AND CONCLUSION
15. REFERENCES
16. APPENDIX

DECLARATION

Regarding this thesis, I hereby declare that this is my own composition, based on my own research work; and all the supporting literature that has been used in this thesis has been appropriately referenced.

Syeda Rubab Zehra Rizvi

Nottingham, 30th January 2020

Studying novel sporulation genes in *Clostridium sporogenes*

Syeda Rubab Z. Rizvi, Nigel P. Minton, Alan Cockayne
Centre for Biomolecular Sciences, University of Nottingham, University Park, NG7 2RD, Nottingham, United Kingdom

ABSTRACT

The class Clostridia consists of bacteria that are ubiquitous in nature, but are of high significance in various aspects of life, namely health, disease, food spoilage, biofuel production, bioterrorism and cancer therapy. Most of the clostridial species produce resilient spores, which ensure their survival under harsh environments up to several years in the form of dormant refractile bodies; and once they find favourable growth conditions, they germinate into their metabolically active forms. As much as this phenomenon is astounding, the current knowledge about clostridial sporulation mechanisms is also very limited. This situation calls for exploiting the forward genetics approach, by using transposable elements to identify novel sporulation genes in a relatively-accessible, non-pathogenic strain of *Clostridium sporogenes*. In conditional transposon mutagenesis, a *mariner*-based transposon is allowed to jump into random genes on host chromosome by means of *Himar1* transposase enzymes, which eventually generates a library of transposon mutants that can be screened for the desired phenotype and ultimately, the genotype. Hence, transposon-insertion mutant libraries were constructed in *C. sporogenes*, from which 5,061 mutants were randomly screened for sporulation or germination defects, on the basis of little or complete loss of heat-resistant spores. Out of the 41 isolated sporulation- or germination-defective mutants, 28 unique genes were identified to have potential roles in sporulation or germination. According to annotations from the NCBI genome database, the identified genes encode for proteins involved in various stages of sporulation, cell division, DNA translocation, and some other biological processes. This indicates that the defects in sporulation or germination in the corresponding

C. sporogenes mutants were either caused through the direct or indirect disruption of their cascades. Among the identified sporulation or germination genes, CLSPO_RS15205 (*yqfD*), CLSPO_RS07560 (*divIVA*), CLSPO_RS16880 (*clpX*), CLSPO_RS17345 (*spoVK*) and CLSPO_RS00950 (*spoIID*) were selected for CRISPR/Cas9-based targeted genome editing and the resulting null mutants were further characterized. In brief, attempts were made to determine possible roles of the target genes in the survival, growth, haemolysis, proteolysis and cell and colony morphology of *C. sporogenes* mutant strains. This characterization of general biology and sporulation profile of the acquired mutants will not only aid in better understanding sporulation mechanisms in *C. sporogenes*, but also for its toxigenic surrogate Group A proteolytic *Clostridium botulinum*. The acquired data is aimed to help spores-based applications in cancer therapy and food safety research.

Keywords: clostridium; sporulation; transposon; genes; mutants; spores; CRISPR; cas9; yqfD; divIVA; clpX; ATPase; spoVK, spoIID; characterization

To my beloved family,
For their faith, love, prayers and support

ACKNOWLEDGEMENTS

There are several souls who are worth mentioning in my thesis; with whom I travelled on this quest to unearth the mysteries about science. First of all, I am very thankful to Prof Nigel P. Minton, who gave me this auspicious opportunity to work in his research group, as a research contributor on the CLOSPORE project. It is the benevolence of the Marie Skłodowska-Curie Actions (MSC), which generously funded this project and allowed me to develop my personal and career skills by being an Early Stage Researcher in the Innovative Training Network (ITN), and to which I am most gratified. The figure I am most grateful to, is Dr Alan Cockayne, my mentor and primary supervisor, who believed in me and encouraged me each time I was going through a rough period in the PhD-related processes. Dr Cockayne always helped me troubleshoot the experimental problems experienced in laboratory work, and I simply cannot thank him enough for his patience in enduring my overly apprehensive persona. I should mention my brief encounter with my bench supervisor Dr Aleksandra Ola Kubiak, who helped me in adapting with the new laboratory tools and techniques during my initial days at the Centre for Biomolecular Sciences (CBS).

I am also thankful to Dr Alexander Grosse-Honebrink and Dr Craig Woods, who helped me with Transposon mutagenesis; Dr Peter Rowe and Daphne Groothuis, who helped me with CRISPR/Cas9-based genome editing; François Seys, who helped me with Target-AID base editing; and Ms Denise McLean and Ms Michelle Kelly, who helped me with Electron Microscopy. I have learned so much from my other colleagues, Dr Maria Zygouropoulou, Dr Ines Canadas Blasco, and Dr Raquel Rodrigues, with whom I worked on *Clostridium sporogenes*. I have my sincere regards to my CBS laboratory technician Evangelos Karvonis, for his caring personality and

his efforts in making the growth media, autoclaving the consumables and managing the laboratory-generated waste. Working with the whole 100+ member-based Clostridia Research Group (CRG) was a very wonderful experience for me. With everyone's support I was able to achieve a lot.

I am also thankful to my brother Syed Abbas Rizvi, who influenced and trained me to use Inkscape v1.0 for making informational vector art. Lastly, I give my thanks to both my family, who appreciated my efforts in the purse of knowledge; and to the deity who created these opportunities for me to learn more about science.

Syeda Rubab Zehra Rizvi

TABLE OF CONTENTS

OVERVIEW OF CONTENTS ii

DECLARATION 3

ABSTRACT 4

ACKNOWLEDGEMENTS 6

TABLE OF CONTENTS 8

ABBREVIATIONS 16

1. INTRODUCTION 19

1.1. THE GENUS *Clostridium* 19

1.1.1. Importance of Clostridia in pathogenesis 20

1.1.2. Beneficial uses of *Clostridium* 21

1.2. SPORES AND SPORULATION 23

1.2.1. History and background 23

1.2.2. Spore formation 23

1.2.3. Spore structure 27

1.2.4. Spore resistance 29

1.2.5. Spore germination and outgrowth 30

1.3. *Clostridium sporogenes* – THE MODEL ORGANISM 32

1.3.1. General characteristics of *C. sporogenes* 33

1.3.2. Applications of *C. sporogenes* 33

1.3.3. A different side of *C. sporogenes* 38

1.3.4. Some lesser known facts of *C. sporogenes* 39

1.4. MOLECULAR TOOLS FOR STUDYING *Clostridium* 40

1.4.1. Transposon mutagenesis – A forward genetics approach 40

1.4.2. ClosTron gene knock-out system – A reverse genetics approach 44
1.4.3. CRISPR/Cas9-based genome editing – A reverse genetics approach 46
1.4.4. Target-AID base editing – A reverse genetics approach 48
1.5. THE *CLOSPORE* PROJECT 50
1.5.1. Some insights from literature on clostridial sporulation 50
1.5.2. Aims and objectives 51
2. MATERIALS AND METHODS 53
2.1. MATERIALS 53
2.1.1. Bacterial strains used. 53
2.1.2. Plasmids used. 55
2.1.3. Growth media used. 57
2.1.4. Antibiotics and supplements used. 57
2.2. BASIC MICROBIOLOGICAL METHODS 58
2.2.1. Preparing nutrient media plates 58
2.2.2. Growing bacterial cultures 58
2.2.3. Storage of bacterial stocks 59
2.2.4. Preparation of competent cells 59
2.3. GENERAL MOLECULAR METHODS 60
2.3.1. DNA Purification 60
2.3.2. Polymerase Chain Reaction 61
2.3.3. Agarose gel electrophoresis 62
2.3.4. DNA Digestion 62
2.3.5. DNA Ligation 63
2.3.6. Transformation and cloning 63
2.3.7. Conjugation 65

2.3.8. DNA Sequencing 66
2.4. BIOINFORMATICS AND DATA ANALYSIS 66
2.4.1. Genome editing 66
2.4.2. Sequencing data analysis 67
2.4.3. BLAST and InterPro 67
2.4.4. Graphics Analysis 67
2.5. MICROSCOPY 68
2.5.1. Light Microscopy 68
2.5.2. Transmission Electron Microscopy 68
2.6. CONDITIONAL TRANSPOSON MUTAGENESIS 70
2.6.1. Pre-requisites to transposon libraries 70
2.6.2. Generation of transposon libraries 70
2.6.3. Screening of transposon mutants 71
2.6.4. Identification of transposon insertion sites 72
2.7. CRISPR/CAS9-BASED GENOME EDITING 73
2.7.1. Designing and generating guide RNAs 73
2.7.2. Generation of homology arms 73
2.7.3. Genome editing in *C. sporogenes* 74
2.8. TARGET-AID GENOME EDITING 75
2.8.1. Designing and generating guide RNAs 75
2.8.2. Generating Target-AID mutants in *C. sporogenes* 76
2.9. CLOSTRON GENE KNOCK-OUT SYSTEM 76
2.9.1. Designing the ClosTron intron 76
2.9.2. Generation of *C. sporogenes* ClosTron mutants 77
2.10. GENE CHARACTERIZATION 77

2.10.1. Gene complementation 77
2.10.2. Standard spore assay 78
2.10.3. Growth assay 78
2.10.4. Haemolysis assay 78
2.10.5. Protease assay 79
2.10.6. Cell and colony morphology 79
3. SCREENING FOR NOVEL SPORULATION/ GERMINATION GENES 81
3.1. INTRODUCTION 81
3.1.1. Conditional transposon mutagenesis 82
3.1.2. Aim of this chapter 83
3.2. RESULTS 84
3.2.1. Screening transposon libraries 85
3.2.2. Identifying transposon insertion sites 87
3.2.3. Initial characterization of transposon mutants 91
3.3. DISCUSSION 92
3.3.1. Selection of transposon plasmid 92
3.3.2. Generation and screening of mutant libraries 93
3.3.3. Challenges of the process and problem-solving techniques 94
3.3.4. Mutants versus Wild type 97
3.3.5. Selection for mutant characterization 99
4. CHARACTERIZATION OF *yqfD* 102
4.1. INTRODUCTION 102
4.1.1. Aim of this chapter 105
4.2. RESULTS 106
4.2.1. CRISPR/Cas9-based genome editing in the *C. sporogenes yqfD* gene 106

4.2.2. Mutations in *yqfD* halts *C. sporogenes* sporulation 107
4.2.3. *yqfD* is not an essential gene in *C. sporogenes* 109
4.2.4. *C. sporogenes* β-haemolysis does not require *yqfD* 110
4.2.5. Proteolytic activity of *C. sporogenes* is independent of *yqfD* 110
4.2.6. Microscopy reveals absence of mature spores of Δ*yqfD* mutants 111
4.2.7. Δ*yqfD* mutant colonies become translucent on old TYG agar plates 115
4.2.8. Illumina sequencing reveals SNPs in the Δ*yqfD* mutant genome 115
4.3. DISCUSSION 116
4.3.1. The problems of cloning knock-out/in cassettes in *E. coli* 117
4.3.2. Inefficient CRISPR/Cas9 homology-directed repair in *C. sporogenes* 120
4.3.3. The success of Target-AID base editing 121
4.3.4. The unexpected phenotype of *yqfD* complementation 122
4.3.5. SNPs in the Δ*yqfD* mutant 123
4.3.6. Characterization of *yqfD* and its role in sporulation 124
4.3.7. Proposed function of *yqfD* in *C. sporogenes* 126
5. CHARACTERIZATION OF *divIVA* and *hypP* 128
5.1. INTRODUCTION 128
5.1.1. The *hypP* gene 129
5.1.2. The *divIVA* gene 129
5.1.3. Aim of this chapter 132
5.2. RESULTS 133
5.2.1. Some characterization of *hypP* 133
5.2.2. CRISPR/Cas9-based genome editing in the *C. sporogenes divIVA* gene 137
5.2.3. Mutation in *divIVA* alleviates *C. sporogenes* sporulation 138
5.2.4. *divIVA* is not an essential gene in *C. sporogenes* 139

5.2.5. *C. sporogenes* β-haemolysis does not require *divIVA* 140

5.2.6. Proteolytic activity of *C. sporogenes* is independent of *divIVA* 141

5.2.7. Microscopy reveals filamentous rods and minicells of Δ*divIVA* mutants 141

5.2.8. Δ*divIVA* mutant colonies have a unique morphology and become translucent on old TYG agar plates 144

5.2.9. Illumina sequencing reveals SNPs in the Δ*divIVA* mutant genome 144

5.3. DISCUSSION 145

5.3.1. Challenges in processes and problem-solving techniques 145

5.3.2. Other unexpected phenotypes 149

5.3.3. SNPs in the Δ*divIVA* mutant 150

5.3.4. Characterization of *divIVA* and its role in sporulation 151

5.3.5. Proposed function of *divIVA* in *C. sporogenes* 154

6. CHARACTERIZATION OF *clpX* 156

6.1. INTRODUCTION 156

6.1.1. Aim of this chapter 160

6.2. RESULTS 161

6.2.1. CRISPR/Cas9-based genome editing in the *clpX* gene 161

6.2.2. Mutation in *clpX* alleviates *C. sporogenes* sporulation 162

6.2.3. *clpX* is not an essential gene in *C. sporogenes* 164

6.2.4. *C. sporogenes* β-haemolysis does not require *clpX* 165

6.2.5. Proteolytic activity of *C. sporogenes* is independent of *clpX* 165

6.2.6. Microscopy reveals filamentous or swollen cells of Δ*clpX* mutants 166

6.2.7. Δ*clpX* mutant colonies become translucent on old TYG agar plates 169

6.2.8. Illumina sequencing reveals SNPs in the Δ*clpX* mutant genome 169

6.3. DISCUSSION 170

6.3.1. Challenges in cloning and problem-solving techniques 170
6.3.2. The unexpected mutant phenotype from the Target-AID system 171
6.3.3. Other unexpected mutant phenotypes 172
6.3.4. SNPs in the Δ*clpX* mutant 172
6.3.5. Characterization of *clpX* and its role in sporulation 173
6.3.6. Proposed function of *clpX* in *C. sporogenes* 176
7. CHARACTERIZATION OF *spoVK* 178
7.1. INTRODUCTION 178
7.1.1. Aim of this chapter 181
7.2. RESULTS 182
7.2.1. CRISPR/Cas9-based genome editing in the *spoVK* gene 182
7.2.2. Mutation in *spoVK* affects *C. sporogenes* sporulation 183
7.2.3. *spoVK* is not an essential gene in *C. sporogenes* 184
7.2.4. *C. sporogenes* β-haemolysis might require *spoVK* 185
7.2.5. Proteolytic activity of *C. sporogenes* might require *spoVK* 185
7.2.6. Microscopy reveals numerous enlarged spores of Δ*spoVK* mutants 186
7.2.7. Δ*spoVK* mutant colonies become translucent on old TYG agar plates 189
7.2.8. Illumina sequencing reveals SNPs in the Δ*spoVK* mutant genome 189
7.3. DISCUSSION 190
7.3.1. Challenges in cloning and problem-solving techniques 190
7.3.2. SNPs in the Δ*spoVK* mutant 192
7.3.3. Characterization of *spoVK* and its role in sporulation 193
7.3.4. Proposed function of *spoVK* in *C. sporogenes* 195
8. CHARACTERIZATION OF *spoIID* 197
8.1. INTRODUCTION 197

8.1.1. Aim of this chapter 201
8.2. RESULTS 201
8.2.1. CRISPR/Cas9-based genome editing in *spoIID* 201
8.2.2. Mutation in *spoIID* halts *C. sporogenes* sporulation 202
8.2.3. *spoIID* is not an essential gene in *C. sporogenes* 204
8.2.4. *C. sporogenes* β-haemolysis does not require *spoIID* 205
8.2.5. Proteolytic activity of *C. sporogenes* is independent of *spoIID* 206
8.2.6. Microscopy reveals Δ*spoIID* mutants lacking forespore and mature spores. 206
8.2.7. Δ*spoIID* mutant colonies become translucent on old TYG agar plates 209
8.2.8. Illumina sequencing reveals SNPs in the Δ*spoIID* mutant genome 209
8.3. DISCUSSION 210
8.3.1. Target-AID base editor and the off-targeting 210
8.3.2. SNPs in the Δ*spoIID* mutant 212
8.3.3. Characterization of *spoIID* and its role in sporulation 212
8.3.4. Proposed function of *spoIID* in *C. sporogenes* 214
9. GENERAL DISCUSSION AND CONCLUSION 216
REFERENCES 222
APPENDIX 246

ABBREVIATIONS

ABE	Acetone-butanol-ethanol
ACDP	Advisory Committee on Dangerous Pathogens
ACE	Allele-Coupled Exchange
BLAST	Basic Local alignment search Tool
Cas9	CRISPR-associated protein 9
CDEPT	Clostridial-directed enzyme prodrug therapy
CDI	*Clostridioides difficile* infection
cDNA	Complementary DNA
CDS	Coding sequence
CFU	Colony forming unit
Cm	Chloramphenicol
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
Cs	D-Cycloserine
::CT	ClosTron mutant
DNA	Deoxyribonucleic acid
DPA	Dipicolinic acid
Erm	Erythromycin
GC	Guanine Cytosine
HDR	Homology-directed repair
HG	Hazard Group
IEP	Intron-coded protein
IM	Inner membrane
IPTG	Isopropyl β- d-1-thiogalactopyranoside
ITR	Inverted Terminal Repeats
Kan	Kanamycin
LHA	Left homology arm
min	Time in minutes
NCBI	National Centre for Biotechnology Information
NEB	New Englands Biolabs
NHEJ	Non-homologous end joining

NT	Non-toxic
OD_{600}	Optical density at 600 nm wavelength
OM	Outer membrane
ORF	Open reading frame
PAM	Protospacer adjacent motif
PBS	Phosphate buffer solution
PCIA	Phenol:Chloroform:Isoamyl Alcohol
PCR	Polymerase Chain Reaction
RAM	Retrotransposition-Activated Marker
RBS	Ribosomal binding site
RHA	Right homology arm
RNA	Ribonucleic acid
RNP	Ribonuclear protein
SASPs	Small acid-soluble proteins
SDS	Sodium dodecyl sulphate
sgRNA	Small guide RNA
SNP	Single-nucleotide polymorphism
spo/ger$^-$	Sporulation/germination nil
TA	Thymine Adenine
::TA	Target-AID mutant
TAE	Tris-acetate-EDTA
Target-AID	CRISPR/Cas with activation-induced cytidine deaminase
TE	Transposable element
TEM	Transmission electron microscopy
Tm	Thiamphenicol
::Tn	Transposon mutant
tracrRNA	Trans-activating CRISPR RNA
TraDIS	Transposon Directed Insertion Sequencing
v/v	Volume by volume
w/v	Weight by volume
WGS	Whole-genome sequencing
WHO	World Health Organization
xg	Times gravity

CHAPTER ONE

1. INTRODUCTION

1. INTRODUCTION

1.1. THE GENUS *Clostridium*

Genus *Clostridium* includes the group of bacteria belonging to the phylum Firmicutes, the class Clostridia, the order Clostridiales and the family *Clostridiaceae*. Heller (1922) classified members of *Clostridiaceae* as anaerobic rods, with a cell envelop consisting of a thick Gram positive cell wall that surrounds a cell membrane, and an oval or spherical endospore, which can be centrally or terminally located within the cell. Furthermore, they can be proteolytic or non-proteolytic, saccharolytic or non-saccharolytic; with some that can produce H_2S, fix nitrogen and liquefy gelatine (Heller, 1922). This genus was originally described in detail, by Prazmowski (1880) for all strictly anaerobic, Gram positive, endospore-forming bacteria that lacked the ability to reduce sulphate (Andreesen et al., 1989, Minton and Clarke, 2013). However, with recent discoveries in science, this designation has been revised, and there are now phylogenetic studies that have revealed the occurrence of clostridial species that can be sporulating rods or non-sporulating cocci; with a single membrane and thick peptidoglycan layer (i.e., Gram positive) or with two membranes sandwiching a thin peptidoglycan layer (i.e., Gram negative); and with an anaerobic or an aerobic mode of respiration (Finegold et al., 2002, Minton and Clarke, 2013, Yutin and Galperin, 2013, Silhavy et al., 2010).

Among the anaerobic bacteria, Clostridia are a significant focus of scientific studies due to their involvement in various aspects of life. These organisms are ubiquitous in nature: found in environments from soil to marine habitats, human and animal gut, as well as sewage; and they can easily find routes to contaminate foods and wounds (Hatheway, 1990, Kator and Rhodes, 2003). Pathogenic clostridial species have been exploited in the past to be used in bioterrorism (Arnon et al., 2001). Conversely, some

species have also been engineered to augment their usefulness in biotechnology (Minton et al., 2016, Atmadjaja et al., 2019), the cosmetic industry (Noland et al., 2016), and health and medicine (Van Mellaert et al., 2006, Kubiak and Minton, 2015, Staedtke et al., 2016, Zhang et al., 2016).

1.1.1. Importance of Clostridia in pathogenesis

Some of the Clostridia are notorious for causing mild to severe infection, following colonization, production of gas and lytic enzymes, and their ultimate weapons, toxins. Clostridial toxins are antigenic proteins that can lead to extreme toxicity in human and animals. Depending on the toxin-producing species, they can result in potentially lethal human diseases, like tetanus, botulism, gas gangrene, food poisoning and colitis (Wells and Wilkins, 1996). Most of the clinically isolated Clostridia are opportunistic and proliferate only in immune-compromised patients (Minton and Clarke, 2013). At least 14 clostridial species have been confirmed as serious pathogens, among which *Clostridioides difficile*, *Clostridium perfringens*, *Clostridium tetani*, and *Clostridium botulinum* are the most infamous toxigenic quartet (Hatheway, 1990). Note that *C. difficile* was initially characterized as a member of *Clostridium* during the orthodox grouping of a large number of Gram positive, sporulating, anaerobic rods in one genus, but later had to be reclassified as a separate genus due to high genetic diversity (Lawson et al., 2016). Some of the pathogenic Clostridia and features of the disease they cause are summarized in Table 1.1.

Table 1.1. List of some of the pathogenic Clostridia and diseases caused

SPECIES	PATHOGENESIS	REFERENCES
Clostridium botulinum	Toxins (BoNT A-G); botulism; nausea; diarrhea; slurred speech; blurred vision; constipation in infants; respiratory defect; paralysis; can be fatal	(Hughes et al., 1981, Shapiro et al., 1998, Lindstrom and Korkeala, 2006)
Clostridium perfringens	Toxins (alpha, beta, epsilon, iota, etc.); gas gangrene; diarrhea; enterotoxaemia; necrotizing colitis; can be fatal	(McDonel, 1980, Rood and Cole, 1991, Wells and Wilkins, 1996)
Clostridium tetani	Toxins (TeNT and tetanolysin); tetanus; paroxysmal muscular spasms; lockjaw; can be fatal	(Veronesi and Focaccia, 1981, Bruggemann et al., 2003)
Clostridioides difficile	Toxins (TcdA, TcdB and binary toxin CDT); *C. difficile* infection or CDI; mild to severe diarrhea; abdominal pain; pseudomembrane colitis; relapse is possible; can be fatal	(Borriello, 1998, Goorhuis et al., 2008, Rupnik et al., 2009)
Clostridium sordellii	Toxins (alpha, beta, hemolysin); nausea; diarrhea; septic shock; edema; endometritis; hemorrhage; can be fatal	(Nakamura et al., 1983, Hatheway, 1990, Sinave et al., 2002, Aldape et al., 2006)

1.1.2. Beneficial uses of *Clostridium*

Although predominantly known for its pathogenicity, the class Clostridia actually harbours a large number of non-pathogenic members that are also ubiquitous in nature. According to the 16S rRNA and rDNA-based studies on human faeces, members of *Clostridium* have been found to be a prominent part of the gut microflora (Wilson and Blitchington, 1996, Franks et al., 1998, Hold et al., 2002). It has been suggested that Clostridia play a vital role in the maintenance of gut homeostasis (Lopetuso et al., 2013) and the prevention of inflammatory bowel diseases in human (Frank et al., 2007). Moreover, the indigenous gut Clostridia have been found to induce accumulation of colonic $CD4^+$ regulatory T cells, contributing to resistance against colitis in a mouse model (Atarashi et al., 2011). The benign clostridial species have

also been the focus of studies during the last few decades for their beneficial roles in the industrial production of biochemicals (Kuehne and Minton, 2012, Cartman, 2011, Rajagopalan et al., 2013, Samul et al., 2013, Minton et al., 2016), as well as for their use in the development of novel therapeutic strategies for cancer (Van Mellaert et al., 2006, Kubiak and Minton, 2015). Some of the frequently studied aspects of clostridial species, regarding their usefulness are briefly mentioned in Table 1.2.

Table 1.2. List of some of the beneficial *Clostridia*

SPECIES	POTENTIAL APPLICATIONS	REFERENCES
Clostridium sporogenes	Cancer therapy	(Minton, 2003, Kubiak and Minton, 2015)
Clostridium novyi-NT	Cancer therapy	(Dang et al., 2001, Staedtke et al., 2016)
Clostridium acetobutylicum	Acetone-Butanol-Ethanol fermentation	(Weizmann and Rosenfeld, 1937, Papoutsakis, 2008, Sreekumar et al., 2015)
Clostridium beijerinckii	Acetone-Butanol-Ethanol fermentation	(Qureshi and Blaschek, 2000, Bellido et al., 2014)
Clostridium butyricum	Biobutanol production; probiotics	(Bergstrom and Foutch, 1983) (Isa et al., 2016)
Clostridium thermocellum	Biofuel	(Lynd et al., 2005)
Clostridium saccharoperbutylacetonicum	Biobutanol production	(Nakayama et al., 2011)

In addition to using various non-toxigenic strains of clostridial species, pathogenic Clostridia have also been engineered for their applications in the therapeutic and cosmetic industries. For instance, *C. botulinum* A toxin Botox has been reported to treat dystonia of eye and cervix, hyperhidrosis and muscle spasticity (Verheyden and Blitzer, 2002, Brin et al., 2009, Awan, 2017). Similarly, clinical trials have proven the use of botulinum toxin in treating headaches and migraines (Binder et al., 2002). Furthermore, Botox has been widely used in anti-ageing and wrinkle-reducing treatments (Small, 2014). Intriguingly, antitoxins of *C. sordellii* cytotoxic strains have

also been reported to neutralize the toxins of cytotoxic *C. difficile* strains and vice versa; however, the clinical efficacy of clostridial antitoxins has not been thoroughly examined (Nakamura et al., 1983, Sinave et al., 2002).

Besides toxins, clostridial spores have been an important focus in oncology research for their role in cancer therapy since the mid-19th century (Möse and Möse, 1964, Van Mellaert et al., 2006). Spores, their structure and their potential applications are described in the following sections.

1.2. SPORES AND SPORULATION

1.2.1. History and background

Studies on bacterial sporulation started about 140 years ago, when Ferdinand Cohn, a German botanist discovered the existence of certain bacteria, possibly *Bacillus* spp., that could survive boiling temperatures, unlike other groups of bacteria. He observed the presence of refractive bodies surrounding bacterial rods in a culture, and also proposed the theory that these bodies are produced and germinated in contrasting environments (Cohn, 1896). Around the same time, a German bacteriologist, Robert Koch managed to observe sporulation and germination of *Bacillus anthracis* under a microscope; and demonstrated this survival-to-multiplication process for the first time in history (Koch, 1888). Research on bacterial spores has flourished greatly since then and dedicated spore conferences have been regularly held in Europe and elsewhere.

1.2.2. Spore formation

Under unfavourable conditions, when the growth medium becomes deprived of nutrients or in the presence of physical, chemical or biological stresses, some bacteria have evolved to survive in the form of spores. Sporogenesis or sporulation, is the

process of endospore formation in bacteria. It is a complex process, and is best described in *Bacillus* spp. As identified in the earlier genomic studies, the main features of sporulation cascade are similar in all endospore-forming Clostridia and Bacilli (Henriques and Moran, 2007), owing to the conservation of certain important genes or gene homologues, such as *spo0A*, transcriptional factors etc., (Galperin et al., 2012). However, these genome-based studies and more recent advancements in molecular techniques have also revealed that not only the presence of several other important sporulation genes varies in different *Bacillus* and *Clostridium* spp., but also that their roles and expression profiles differ despite their conservation in the bacterial genome (Al-Hinai et al., 2015, Galperin et al., 2012).

1.2.2.1. Stages of sporulation

The process of spore formation can be divided into distinct stages as described in model organisms: *C. acetobutylicum*, *C. difficile* and *Bacillus subtilis.* As illustrated in Figure 1.1, sporulation initiates with the division of chromosomal DNA into two daughter nucleoid, hence allowing the cell to enter stage I. During the stage II of sporulation, the cell membrane invagintes to form an asymmetric septum at a cell pole, which further separates the two nuclei in the sporulating mother cell. Later, engulfment of this septum by the mother cell cytoplasm marks the stage III of sporulation. Sporulating cells in stage IV have layers of cell wall peptidoglycan and cortex surrounding the forespore protoplast. Protein coats are assembled around the forespore in stage V of sporulation. The maturation of the spore inside the mother cell defines stage VI. Eventually, the mother cell is lysed in the stage VII of sporulation, and the resistant spore is liberated into the environment (Al-Hinai et al., 2015, Henriques and Moran, 2007, Piggot and Coote, 1976, Ryter, 1965).

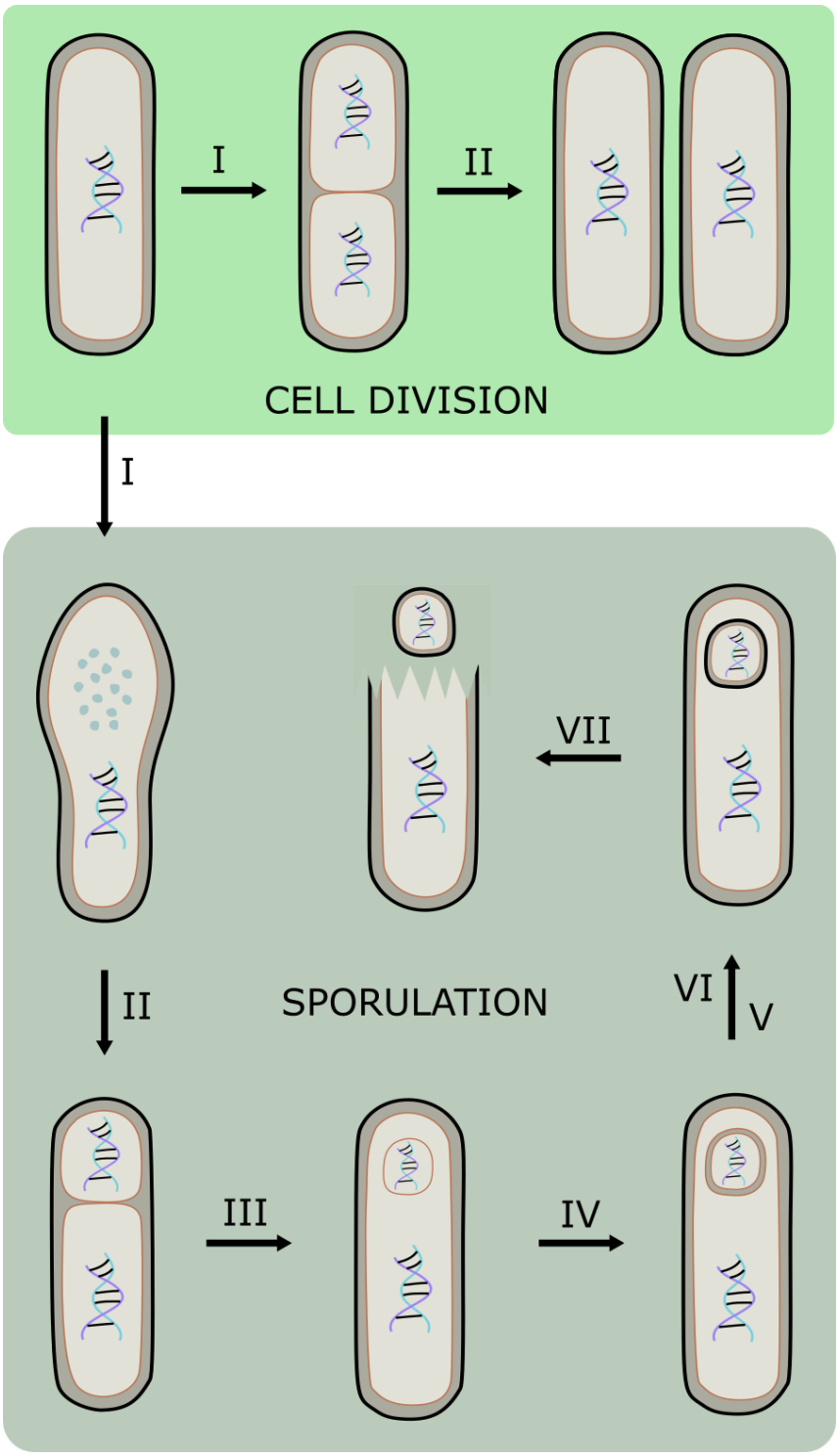


Figure 1.1. **Stages of spore formation in endospore-forming clostridial species**. Cell division takes place during favourable growth conditions (green background), in which binary fission forms two daughter cells (I-II). However, during unfavourable conditions (grey background), granules accumulate in the cells and the characteristic cigar-shape is observed (I). Then, asymmetric septum is formed (II), which gets engulfed by the mother cell (III). Eventually, spore cortex is synthesized (IV) and coat proteins get assembled (V). The spore matures (VI), and later gets released from the mother cell (VII). This diagram was inspired by (Al-Hinai et al., 2015).

In the form of spores, bacteria are in their survival phase. They do not require energy or protein generating machineries, since they will be metabolically inactive. However, they need to retain the necessary components, which can help restore them back to their metabolically active forms under appropriate growth conditions.

1.2.2.2. Requisites for successful sporulation

Firstly, the success of asymmetric cell division ensures that a complete copy of the parental genomic DNA is present in the forespore that contains all the genes responsible for cell structure, house-keeping, metabolism, DNA repair, etc. To prepare vegetative cells for sporulation, stimulus is perceived through a two-component regulatory system, in which sensory histidine kinases phosphorylate the master regulator of sporulation Spo0A (Stephenson and Lewis, 2005), in response to cellular and metabolic signals. The next important step is the transcriptional activation of other sporulation genes, by Spo0A, found in all *Bacillus* and clostridial endospore formers (Deakin et al., 2012, Galperin et al., 2012). In addition to *spo0A*, five sigma factor genes *sigH*, *sigF*, *sigE*, *sigG* and *sigK* have their own vital transcriptional roles at various stages of sporulation, and all of which are well conserved in *Clostridium* and *Bacillus* genera (Kirk et al., 2012, Galperin et al., 2012).

Moreover, the water content of the spore must be lower than 55% in order to confer resistant properties, and this is achieved by the mineral complex of dipicolinic acid or DPA and Ca^{2+} (Piggot and Coote, 1976, Setlow, 2014). During sporulation, DPA is synthesized in the mother cell and later taken up by the developing forespore (Daniel and Errington, 1993). The Ca^{2+}-DPA dimer binds to free water molecules in the spore core, which dehydrates the spore protoplast and confers spore heat resistance (Gerhardt, 1989, Paidhungat et al., 2000).

1.2.2.3. Other factors involved in sporulation

Studies have revealed that the presence and absence of certain compounds in growth media or environment can affect both sporulation, as well as resistant characteristics of a sporulating bacterium, and that these effects vary among different bacterial species. For instance, *C. perfringens* sporulates *in vitro* in the presence of special ingredients, such as starch, peptone, trypticase, magnesium sulphate and sodium thioglycolate, which are the main constituents of the Ellner and SEC media (Ellner, 1956, Duncan and Strong, 1968, Kim et al., 1967). Furthermore, heat-shock and sub-optimal growth conditions have been found to increase the spore yield in *C. sporogenes* (Yang et al., 2009). In another study, $CaCO_3$ and $CaCl_2$ supplementation not only increased the spore yield, but also elevated the heat resistance of *C. sporogenes* spores. In contrast, $FeCl_2$ and $MnCl_2$ had the opposite effects on yield and resistance (Mah et al., 2008). The supplementation of calcium, zinc and manganese in growth medium has been shown to increase the production of spores of *Bacillus* spp. and *C. botulinum*; however, the addition of iron was found to drastically decrease *C. botulinum* spore production (Charney et al., 1951, Kihm et al., 1990). Interestingly, a negative effect of iron deficiency on the heat resistance of *C. botulinum* spores has also been documented in a previous study (Sugiyama, 1951).

1.2.3. Spore structure

Spores of all endospore-forming bacteria possess a similar basic spore morphology, as presented in Figure 1.2. The most dominant part of a spore is the outer covering called the spore coat, which is a multi-layered crystalline structure composed of tyrosine- and cysteine-rich proteins (Driks, 1999, Plomp et al., 2007). This coat facilitates survival under extreme environmental conditions, such as in the presence of enzymes and bactericidal chemicals, while protecting the spore DNA present in the innermost core. Mutation in *cotE*, a gene responsible for the spore coat assembly in *Bacillus*, revealed an impaired spore resistance against chemicals (Setlow, 2014). In

addition, there are several spore layers in between the spore coat and the core (Warth, 1978, Atrih et al., 1996, Setlow, 2014). Surrounding the spore protoplast, is the phospholipid inner membrane or IM, which also serves as the site of germinant-receptor-binding for spore germination (Zheng et al., 2016, Hudson et al., 2001). IM is further protected by two peptidoglycan layers. The peptidoglycan layer next to the IM is the primordial cell wall, which is destined to be the cell wall of the future vegetative cell. The second peptidoglycan layer forms a thick cortex, which is the vital component for the maintenance of spore dormancy and spore mineralization (Henriques and Moran, 2007). Cortex is assembled on the spore outer membrane or OM, which eventually gets dissolved during spore germination in Gram-positive bacteria (Vollmer, 2011). Although spore OM lacks lipopolysaccharides present in the OM of Gram-negative bacteria, its presence in endospores might hint about the possible evolution history of Gram-negative Proteobacteria (Tocheva et al., 2011). The OM separates the peptidoglycan layers with the spore coat. An additional balloon-like exosporium serves as the outermost layer in several spore-forming *Bacillus* and *Clostridium* spp. (Takumi et al., 1979, Henriques and Moran, 2007, Brunt et al., 2015).

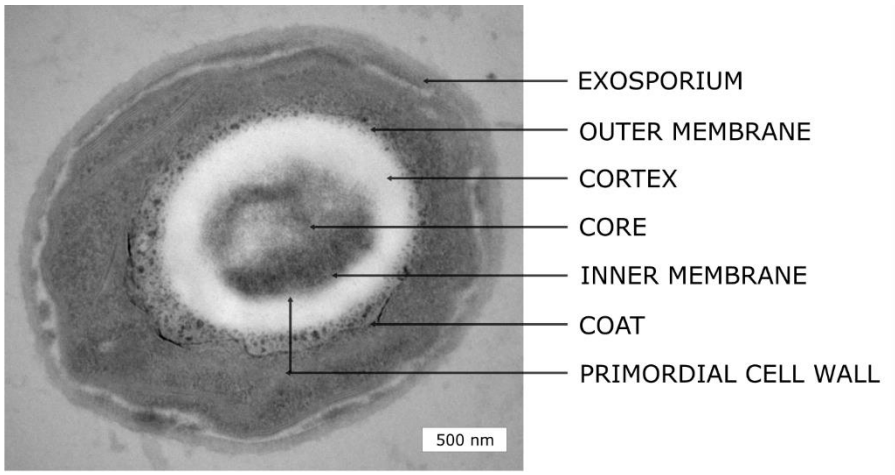


Figure 1.2. **Structure of a clostridial spore**. The inner core contains DNA, which is protected by inner membrane. Next, are two peptidoglycan layers: one that become cell wall of the future cell and the second thick layer is the spore cortex. An outer membrane separates the cortex with the assembled coat proteins. The outermost layer is exosporium, which is only found in some species. This is a micrograph of a *C. sporogenes* spore, obtained from TEM at 500 nm in this study.

1.2.4. Spore resistance

Spores are resistant to various physical stresses, such as desiccation, dry or wet heating, frosting, thawing, and ultraviolent radiation; and chemical agents, such as acids, bases, halides, aldehydes, alkylating and strong oxidizing compounds (Setlow, 2014), which are otherwise perilous to vegetative cells. Moreover, mature spores of *B. subtilis* have also been demonstrated to resist phagocytosis by bacterivorous nematodes and protozoa (Laaberki and Dworkin, 2008, Klobutcher et al., 2006). In their dehydrated state, they can survive hundreds of years of dormancy under unfavourable environmental conditions (Kennedy et al., 1994). Controversially, some spore-forming *Bacillus* spp. have also been reported to survive millions of years, being fossilized inside amber and salt crystal (Cano and Borucki, 1995, Vreeland et al., 2000). Despite such long periods of dormancy, they are capable of germinating into normal vegetative cells (Setlow, 1994); (Henriques and Moran, 2007). Such spore resistance properties are attributed to many factors.

The most obvious foundation of spore resilience is based on the rigid spore structure, comprising of spore coat proteins and exosporium that serve as the first lines of defence. The flexibility of exosporium and the presence of protective layer of spore coat proteins ensure that the DNA inside the spore core is not exposed to the harshness of the surrounding environment (Henriques and Moran, 2007, Plomp et al., 2007). Impermeability of the spore coat restricts the entry of lytic enzymes and bactericides into the core (Setlow, 2014). In contrast, spore coat-defective spores of *B. subtilis* Δ*cotE* mutants are sensitive to lysozyme (Zheng et al., 1988), as well as to predation by *Tetrahymena thermophile* (Klobutcher et al., 2006).

In addition to the outer protective layers, mineral composition of the spore play a vital role in spore resistance. Calcium has been shown to enhance the resistance of *Bacillus* spores against heat (Sinnela et al., 2019). About 10% of the total spore dry weight comprises of DPA, which combines with calcium to form a Ca^{2+}-DPA complex that is

responsible for dehydrating the spore protoplast in order to induce resistance properties (Mah et al., 2008). Furthermore, small acid-soluble spore proteins (SASPs), present in some spore-formers, contribute to resistance against heat and radiations as well as several mutagens. DNA repair enzymes in the spore core are also programmed to rescue any damaged DNA during spore germination and outgrowth (Setlow, 2014).

1.2.5. Spore germination and outgrowth

In the presence of sufficient nutrients and favourable environmental conditions, i.e., in the absence of stress stimuli (physical, chemical or biological), spores germinate from their dormant state into their metabolically active vegetative forms, without losing any of their genotypic or phenotypic assets (Burns et al., 2010).

1.2.5.1. Stages of germination

The germination process is best described in three distinctive stages, as illustrated in Figure 1.3. Stage I starts with the release of protons, cations and zinc ions from the core, which allows the release of Ca^{2+} from the Ca^{2+}-DPA complex. This forces the uptake of water by the spore undergoing germination. In the stage II of germination, hydrolysis of the peptidoglycan cortex occurs due to the action of cortex-degrading enzymes, like SleB and CwlJ (Li et al., 2013). This allows further uptake of water into the spore, causing the spore core to swell and lose resistance properties. However, the spore gets prepared to start essential metabolic activities. Spore outgrowth takes place in the stage III of germination, when the previously inactive spore takes the form of an active vegetative cell. Cell metabolic machineries generate energy and essential macromolecules; the inner peptidoglycan layer of the spore becomes the cell wall of the newly active bacterium; whereas, SASPs, cortex, coat and the vegetative cell's other unwanted macromolecules are completely dissolved (Setlow, 2003).

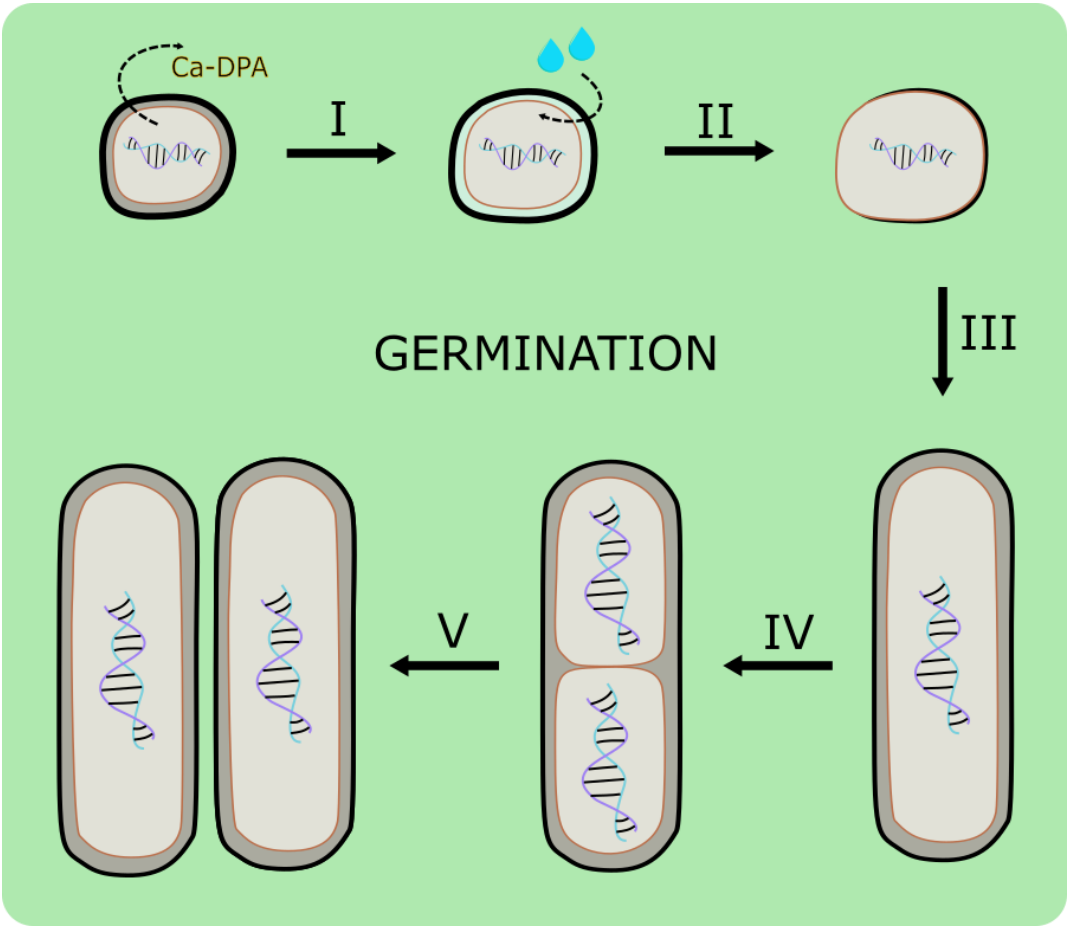


Figure 1.3. **Stages in a typical bacterial spore germination process**. When favourable growth conditions arise (green background), spores release Ca^{2+}-DPA and some water is taken up (I). Spore cortex gets hydrolysed, due to which spore water content increases and the spore swells up (II). Spore coat and other unwanted layers gets dissolved and spore outgrowth takes place III. The germinated cell divides by binary fission (IV) and produces two daughter cells (V). This diagram was inspired by (Setlow, 2003).

1.2.5.2. Requisites for germination

Germination is generally triggered by the binding of nutrient germinants to germinant receptors present on the spore IM, which induces a change in the permeability of spore layers, and prepares the spore for the main event (Moir and Smith, 1990, Paidhungat and Setlow, 2002). Nutrient germinants can be amino acids, nucleosides or saccharides that are present in the surrounding environment. L-alanine, potassium ions, asparagine, fructose, glucose, Ca^{2+}-DPA, bile salts, and even lysozymes and high pressure, have been found to serve as effective germinants or co-

germinants in endospore-forming bacteria (Setlow, 2003, Burns et al., 2010). The activity of germinants requires the presence of germinant receptors, which are generally encoded by genes, *gerA*, *gerB* and *gerK* in *Bacillus* spp.; and *gerA* in *C. botulinum* and *C. sporogenes*.

Interestingly, since growing cultures of bacteria release certain muropeptides into the environment, this can inform the dormant spores present in the same environment about the presence of favourable growth conditions that can support bacterial proliferation (Shah et al., 2008). Supplementing *B. subtilis* spores with cell-free supernatants from the metabolically active *B. subtilis* cultures have shown to efficiently trigger the germination cascade. Moreover, cell-free supernatants of other bacterial species can also induce germination of *Bacillus* spores (Shah et al., 2008). The same study identified the activity of PrkC, a Ser/Thr kinase, vital for this germination response. Muropeptides released from the growing cultures can bind to PkrC and induce germination of *B. subtilis* spores (Pompeo et al., 2016).

In addition to the genes for germinant recognition, some other vital germination genes, like *sleB* and *cwlJ* are present in *Bacillus* and clostridial species, which contribute to the essential steps in the germination process via cortex-lysis (Burns et al., 2010, Paidhungat and Setlow, 2000, Paidhungat et al., 2001).

1.3. *Clostridium sporogenes* – THE MODEL ORGANISM

For the purpose of studying clostridial sporulation, the in-house *C. sporogenes* NCIMB 10696 strain was used in this study. *C. sporogenes* has been classified as a Hazard Group HG1 organism by the UK Advisory Committee on Dangerous Pathogens (ACDP). It is a free-living organism that naturally dwells in soil (Smith, 1975); however, it was first isolated by Metchnikoff from faeces of both healthy individuals and patients with colitis (Princewill, 1978). *C. sporogenes* strains have also been found in marine (Finne and Matches, 1974) and freshwater lake (Molongoski and Klug, 1976) sediments.

1.3.1. General characteristics of *C. sporogenes*

C. sporogenes strains are motile, rod-shaped bacteria that possess the Gram-positive type cell wall. The cells strictly undergo an anaerobic mode of respiration and are, therefore, oxygen-labile. *C. sporogenes* is a mesophile and grows at an optimum temperature range of 30-40°C, at an ideal pH of 7 (Poehlein et al., 2015). The size of the *C. sporogenes* NCIMB 10696 genome is ~4.14 MB, which has 3,732 protein-encoding gene sequences and a total of 27.98% GC content (Kubiak et al., 2015).

C. sporogenes is a proteolytic organism and does not require carbohydrate supplements for growth in laboratory media (Smith and Hobbs III, 1974). It instead ferments certain amino acids by the Stickland reaction to produce similar compounds as with the carbohydrate fermentation (Barker, 1981), and has also been observed to deaminate amino acids even at a highly acidic pH (Flythe and Russell, 2004). The haemolytic nature of *C. sporogenes* has also been documented in previous literature (Gonzalez et al., 2010, Poehlein et al., 2015).

1.3.2. Applications of *C. sporogenes*

The favourable roles of *C. sporogenes* in various fields of research and therapy have been increasingly explored during the last few decades. Its non-pathogenic profile and efficient spore-forming trait makes the bacterium a good candidate for not only studying sporulation in general, but also for gaining a better understanding of the parameters that affect sporulation in species which are closely related to *C. sporogenes*.

1.3.2.1. Surrogate for *C. botulinum*

The group A proteolytic *C. botulinum* is of extreme importance in food industries, as the highly heat-resistant spores are capable of evading inadequate thermal processing methods, and resulting in food spoilage and food-borne diseases (Townsend et al., 1938, Ingram and Robinson, 1951, Butler et al., 2017). These *C. botulinum* strains

produce Type A neurotoxin, which has a minimum lethal dose of 0.00003 µg/kg, potentially making it the most potent toxin known (Schantz and Johnson, 1992). This property of botulinum toxin has also been exploited in the past for its potential in bioterrorism (Henderson, 1999). Therefore, in the absence of safe or toxin-negative strains of *C. botulinum*, working with toxigenic *C. botulinum* strains for research can be seriously hazardous. Moreover, handling the organism in a laboratory requires a permit from the public health officials in some states, in addition to the implementation of strict biosafety measures.

When compared with other Clostridia, *C. sporogenes* 16S rRNA has been found to be 99-100% identical to that of the proteolytic Group I *C. botulinum* strains (Bradbury et al., 2012, Sebaihia et al., 2007). In addition, some botulinum neurotoxin-producing strains of *C. sporogenes* have also been reported (Hill et al., 2007, Smith et al., 2018), possibly owing to the horizontal transfer of the toxin-producing genes. Nevertheless, *C. sporogenes* can serve as a perfect surrogate for toxigenic Group I *C. botulinum* in food preservation studies, such as, for the validation of food processing strategies (Luechapattanaporn et al., 2004, Taylor et al., 2013, Schill et al., 2016, El Jaam et al., 2017, Butler et al., 2017). *C. sporogenes* has already been used as a research model for studying heat resistance profile of proteolytic *C. botulinum* strains (Nakamura et al., 1977, Mah et al., 2009, Brown et al., 2012).

1.3.2.2. Drug delivery vector in cancer therapy

The life expectancy of human beings in the 21st century is wildly challenged by the fatal nature of cancer. In 2015, the World Health Organisation (WHO) estimated cancer as the leading cause of death in nearly 100 countries. According to the GLOBOCAN 2018 data, 9.6 million deaths worldwide were estimated to be as a result of cancer, with 18.1 million new cases in 2018 (Bray et al., 2018). It has been further estimated that by 2040, there is a chance of 27.5 million new cancer cases that are feared to claim 16.3 million deaths worldwide (Street, 2019), unless poor hygiene

conditions, unhealthy lifestyle, and conventional diagnostic and therapeutic procedures to combat cancer are not adequately improved.

Cancer is a disease hallmarked by the unrestricted proliferation of cells that evade the normal homeostatic regulation of their host, and in the form of tumour tissues, such cells are capable of invading other sites of the body (Hanahan and Weinberg, 2011). The inner core of solid tumour tissue is characterized by hypoxia, which arises from the imbalance between supply and consumption of blood oxygen (Vaupel and Harrison, 2004). The unique physiology of solid tumours makes radio- and chemotherapy challenging for specifically targeting tumour regions (Brown and Wilson, 2004, Baronzio et al., 2009). The concentrations of doses required for absolute elimination of tumour are neither therapeutically-efficient nor suitable for the viability of the remaining healthy host tissues. This challenge can be overcome by exploiting the non-pathogenic, obligatory anaerobic and conditionally-sporulating strains of Clostridia for their potential in cancer drug delivery in oncological therapy (Lemmon et al., 1997, Möse and Möse, 1964). *C. sporogenes* spores have been known as one of the best candidates for their application in the success of the Clostridial-Directed Enzyme Prodrug Therapy or CDEPT, due to their superior sporulation and colonization traits (Minton et al., 2001, Kubiak and Minton, 2015, Liu et al., 2002).

In CDEPT, *C. sporogenes* strains are engineered to express therapeutic proteins or prodrug-activating enzymes, such as nitroreductases. As portrayed in Figure 1.4, the aim is to administer these engineered spore suspensions into a cancer patient, after which, the spores specifically locate the tumour hypoxic cells and germinate into metabolically active vegetative cells that are capable of producing the prodrug-activating nitroreductases. These nitroreductases can then activate the inactive cancer prodrug that is also administered to the patient, which causes localized anti-tumour activity. In this co-treatment, both hypoxic and oxic areas of tumour cells are synergistically targeted (Staedtke et al., 2016). It is to note that depending on the type

and size of the tumour, copies of nitroreductase genes are engineered into the *C. sporogenes* genome, which can allow ambient activation of prodrugs, like 5-(aziridin-1-yl)-2,4-dinitrobenzamide, namely CB1954 (Heap et al., 2014, Kubiak and Minton, 2015).

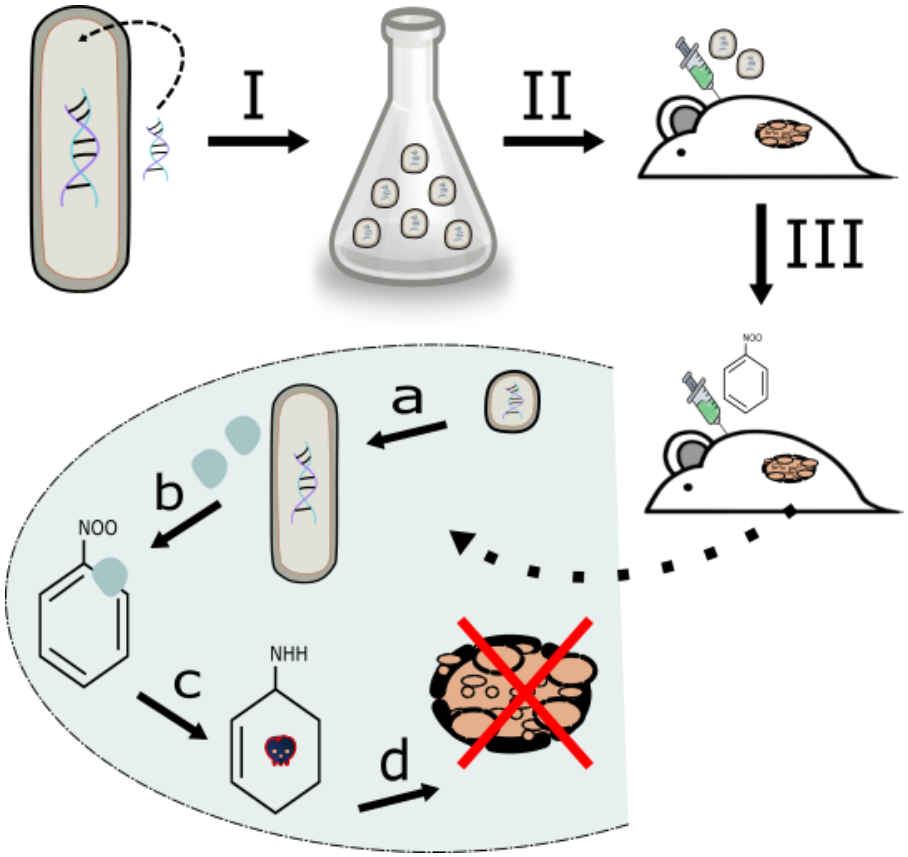


Figure 1.4. **Clostridial-directed enzyme-prodrug therapy/CDEPT**. Benign strains of Clostridia are engineered with nitroreductase genes and the strains are induced to sporulate (I). Clostridial spores are injected into tumour-infected mouse (II), after cancer prodrug is also injected (III). Spores specifically locate to hypoxic tumour tissues inside mouse and germinate (a). Nitroreductases are released from the cells, which activates prodrug to an active drug (b-c). Tumour is regressed by the direction action of cancer drug (d). This diagram was inspired from (Minton, 2003, Williams et al., 2015).

Intriguingly, *C. sporogenes* spores can also potentially serve as diagnostic tools for the presence of solid tumours (Minton et al., 2001). Germination of spores in human tissues induces antibody production in response to the active vegetative cells, which can be detected in the host's blood serum. Because spores are only capable of germinating in an oxygen-deficient environment, the presence of specific serum

antibodies can be indicative of the presence of potential hypoxic, tumour necrotic regions in the host's body.

Large quantities of' spores can be easily mass produced and conveniently injected intravenously. Spores are immunologically inert and cannot trigger the host's immune response, as tested in healthy mice (Minton et al., 2001). Additionally, they are incapable of colonization in healthy tissues due to the lack of anoxic environment; although spores have also been known to persist in healthy tissues for up to four weeks without germination (Lambin et al., 1998). When spores are administered to patients, the resilient spore coats protects them from enzymes, low pH and other host stresses. Spores tend to specifically locate the hypoxic or anoxic necrotic tumour cell regions and germinate subsequently. Inside the hypoxic tumour tissues, the proteolytic Clostridia are capable of secreting proteases that can digest tumours (Heap et al., 2014). Finally, since the non-pathogenic species, like *C. butyricum*, *C. sporogenes* and *C. novyi*-NT are not directly involved in the causation of disease or toxin production, they are considered unlikely to cause additional harm to the patient's body (Möse and Möse, 1964), unless perhaps in severely immune-compromised patients.

The development of a tightly regulated inducible sporulation process has been proposed as being beneficial for the application of *C. sporogenes* spores in real-life systems. In this inducible system, the engineered strains would only sporulate in the presence of an inducer; thereby, allowing the host body to get rid of the residual bacterial cells and spores after the successful necrosis of tumour cells. Hence, a better understanding of sporulation genes and their regulation would allow the development of such novel systems that can optimize spore production and also help in controlling the process in therapeutic applications.

1.3.3. A different side of *C. sporogenes*

Unlike some of the notoriously harmful members of Clostridia, *C. sporogenes* is more commonly known for its non-pathogenic nature. However, a rare case of *C. sporogenes* bacteraemia have also been reported in an immuno-competent patient (Abusnina et al., 2019). The severity of pathogenesis was likely due to the β-haemolytic (Gonzalez et al., 2010) and strongly proteolytic (Princewill, 1978) profile of the organism. *C. sporogenes* haemorrhagic toxicosis in rabbits and *in vitro* collagenase activity has also been previously described (Hara-Kudo et al., 1996, Hara-Kudo et al., 1997). Penicillin is generally the drug of choice for the treatment of human infections caused by *C. sporogenes*, *C. perfringens*, *C. sordellii* and *C. septicum* (Dornbusch et al., 1975, Schwartzman et al., 1977, Stevens et al., 1987, Roberts et al., 2006). However, some penicillin-resistant strains of these clostridial species have also been reported (Sarvari and Schoblocher, 2020).

Strains of *C. sporogenes* and *C. botulinum* have been previously found to possess the clostridiolysin S gene cluster in their genome, in which ClosA-I genes are responsible for *in vitro* β-haemolysis (Gonzalez et al., 2010). This cluster consists of a structural propeptide, some proteases for the post-translational modification of the propeptide and ABC transporters for transporting the mature haemolysin outside the cell. The ClosA-I gene cluster is not only comparable to the SagA-I streptolysin S gene cluster of group A *Streptococcus*, but the haemolytic activity of group A streptococcal Δ*sagA*, Δ*sagC* and Δ*sagD* mutants has also been found to be complemented by *closA*, *closC* and *closD* genes on a plasmid that was transferred to the respective hosts. *C. botulinum* haemolysin has also been previously purified and characterized in another study (Haque et al., 1992).

Despite its haemolytic and proteolytic nature, *C. sporogenes* has been repeatedly identified as a safe organism to work with in laboratory-based research.

1.3.4. Some lesser known facts of *C. sporogenes*

There have also been studies that have unearthed other interesting characteristics of this bacterium. For instance, it has been reported that non-acetone producing strains of *C. sporogenes* have been capable of producing butanol and ethanol when grown not only on refined sources like sorbitol and instant dry yeast, but also on rice straw and crude glycerol sources (Kaushal et al., 2017, Kaushal et al., 2019). In a similar study, *C. sporogenes* was also observed to produce butanol, volatile fatty acids and hydrogen from the rice straw hydrolysate (Gottumukkala et al., 2013). With their ease of availability worldwide, rice straws are considered as an attractive feedstock for biofuel production (Satlewal et al., 2018). These findings suggest *C. sporogenes* can serve as a potential producer of biofuels from rice straws, when followed by the genetically engineered enhancements.

Astoundingly, in a series of studies for confirming the safety of the first manned flight in space, *C. sporogenes* cells have been sent into space (Taylor, 1974). Upon investigation of the inflight viability, the cells were observed to remain viable throughout the flight; however, some instability was observed owing to the damage caused by high solar radiations during the flight.

Methioninase-producing *C. sporogenes* strains have also been characterized for their tumour-inhibitory activity by the degradation of L-methionine, an amino acid critically required for the growth and progression of cancer cells. The effectiveness of methioninase supplementation on tumour suppression was found to be superior to using methionine-free diets in the Walker 256 tumour model (Kreis and Hession, 1973, Hoffman et al., 2019).

1.4. MOLECULAR TOOLS FOR STUDYING *Clostridium*

Many Clostridia possess valuable properties in industrial and health sectors, as already briefly mentioned in the earlier sections of this thesis. With their worth in fulfilling the increasingly high demands for renewable energy resources, these naturally occurring industrially useful microorganisms require adequate metabolic engineering tools for the enhancement of product-related properties (Lynd et al., 2005). Likewise, in order to pursue the non-pathogenic clostridial species as drug delivery vehicles for cancer therapy, clostridial strains have been previously engineered to express prodrug-activating enzymes; and more recently an inducible sporulation system for CDEPT has also been developed (Heap et al., 2009, Minton, 2003, Canadas et al., 2019).

While there has been some fundamental information available regarding the sporulation process of *C. sporogenes*, there are many details still missing relating to the functions of several genes involved in sporulation and germination. Therefore, genomic studies are required to identify such novel sporulation and germination genes, including their regulons as well as the interactive mechanisms by which these genes assemble various vital biochemical events to manifest the sporulation process.

1.4.1. Transposon mutagenesis – A forward genetics approach

Transposable elements (TEs) or transposons are fragments of DNA, capable of changing their positions within a genome, which makes them user-friendly candidates for studying genetics in the majority of living species (Ivics and Izsvak, 2010). Transposons are widely found in nature and they have been discovered in several eukaryotes and prokaryotes. For example, TEs comprise of 12% of the genome of some fish species, 37% of the mouse genome, 45% of the human genome, and in extreme cases, the genome of some plant and fly species even comprise of up to 80% TEs (Munoz-Lopez and Garcia-Perez, 2010).

The concept of transposons as "jumping genes" dates back to 1950, when Barbara McClintock discovered TEs in maize and learned that some genes could jump from their native locus to a different locus on the same chromosome. Depending on the site of transposon insertion, this transposition alters transcription and gene expression, as a consequence (Biemont, 2010).

Molecular studies on different types of transposons have demonstrated the success of *in vivo* and *in vitro mariner* TE transposition in several organisms; and thereby, confirmed them as powerful tools for genome manipulation (Picardeau, 2010). *Mariner* transposon was first discovered as an unusual insertion in the gene encoding the white eye colour of *Drosophila mauritiana*, which was able to change the eye colour phenotype of the mutant flies from white to white-peach (Jacobson et al., 1986). These transposons are now known to be most abundantly found TEs in nature; being present in flies, frogs, plants, fungi, animals, arthropods, nematodes, fish species and even human. They belong to the Tc1/*mariner* superfamily of transposons, members of which are identical in structure and transposition mechanism (Plasterk et al., 1999). Most of these elements are silenced in their hosts due to mutations, population fitness and evolutionary events (Agren and Clark, 2018). However, some have been re-activated synthetically, such as the *Sleeping Beauty* transposon from the ancient transposons in fish genomes (Ivics et al., 1997).

A *mariner* transposon contains inverted terminal repeats (ITRs) that flank a transposase gene or a selective antibiotic marker gene (Cartman and Minton, 2010). The transposition of a transposon is executed by the help of transposase enzymes, which contain DNA-binding and catalytic domains (Plasterk et al., 1999). The *mariner* transposition is a cut and paste mechanism. As shown in Figure 1.5, the transposase binds to both ITRs and excise them from their original locus and form a transposon-transposases complex, whereas the original site is repaired by the host. The transposase randomly targets any TA site in the host chromosome, cleaves the DNA,

inserts the flanking ITRs and joins the DNA excision site (Munoz-Lopez and Garcia-Perez, 2010). Depending on the location of transposon insertion, this can lead to the impairment or loss of gene function.

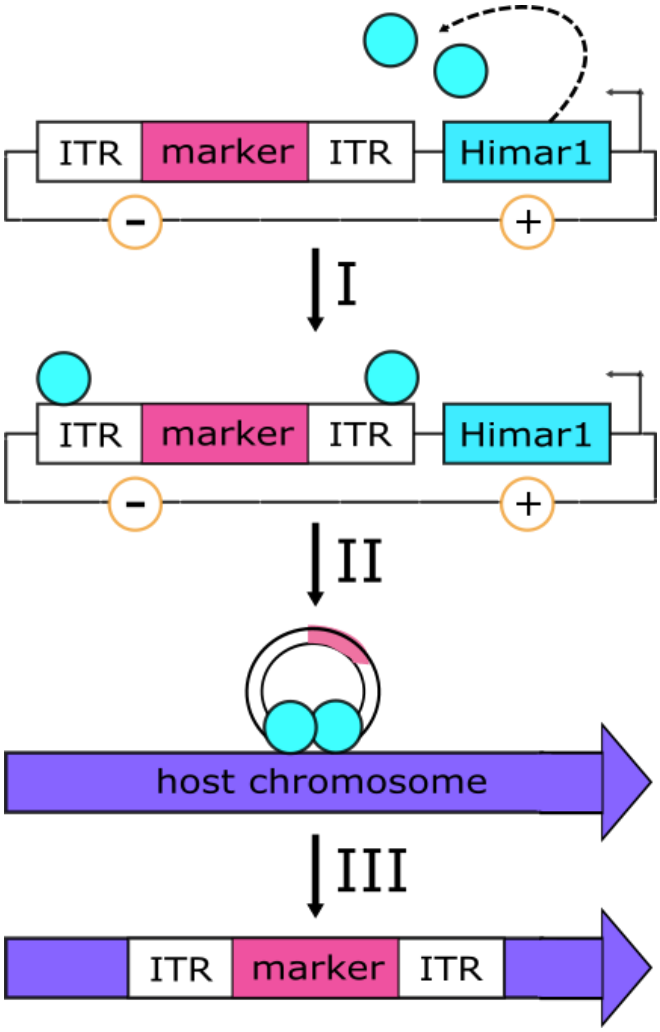


Figure 1.5. **Transposition of Tc1/*mariner* superfamily transposons**. Transposase enzymes are expressed by the *Himar1* gene, which bind to each of the two DNA-binding sites of transposon ITRs (I). Transposon is cut from the original locus and the two transposases make a complex (II). This complex randomly locates a TA site in the host genome and inserts the transposon cassette into the genome (III). This diagram was inspired by (Munoz-Lopez and Garcia-Perez, 2010).

Transposon mutagenesis has been demonstrated in several research studies on both Gram-positive and Gram-negative bacteria, in attempts to identify essential genes and to study other novel genes functions (Reznikoff and Winterberg, 2008). In general, these studies involve the construction of random transposon mutant libraries, which

are then screened for the desired mutant phenotype using various strategies. One of the screening approaches is also described in the later chapters of this thesis.

Making use of the mutagenesis property of transposable elements, thousands of previously uncharacterized bacterial genes have been studied through their mutant phenotypes (Price et al., 2018). In one such genetic study, a transposon mutant library was generated in *Mycobacterium tuberculosis* by infecting the bacterium with an engineered mycobacteriophage, and essential genes in the *M. tuberculosis* mutants were identified on the basis of the pristinamycin-I inducible system (Forti et al., 2011). In another study, Transposon Mediated Differential Hybridization or TMDH was carried out on a mutant library of *Staphylococcus aureus*, which identified 351 genes that are important for growth and survival of the host bacterium (Chaudhuri et al., 2009). Similarly, Transposon site hybridization or TraSH was used to identify essential genes in mycobacteria (Sassetti et al., 2001). Based on a DNA sequencing approach, several virulence genes were identified in *S. aureus* using the Signature-tagged mutagenesis or STM technique (Mei et al., 1997). By following a different strategy for gene identification, a transposon mutant library was phenotypically analysed to discover previously unidentified genes involved in the twitching motility and prototrophic growth of *Pseudomonas aeruginosa* (Jacobs et al., 2003). In a similar study, transposon mutants defective in gliding motility were identified by simply plating the *C. perfringens* transposon mutant library on BHI agar plates and screening for colonies with smooth edges (Liu et al., 2013). Transposon mutagenesis has also been used to screen genes involved in hepatocellular carcinoma in a mouse model (Keng et al., 2009).

1.4.2. ClosTron gene knock-out system – A reverse genetics approach

The Alan Lambowitz laboratory demonstrated the use of mobile group II introns in genetic engineering for exploiting biological systems in bacteria. It involves the insertion of mobile genetic elements into the host genome, which generates mutants (Karberg et al., 2001, Kuehne and Minton, 2012). These introns are catalytic RNAs that are capable of excising from RNA transcripts and inserting into the target genes. This mechanism is controlled by the ribonuclear protein (RNP) complex, which consists of intron-encoded protein or IEP, i.e., LtrA protein, and an RNA transcript. The mobile group II introns are naturally found in the DNA of eubacteria (Michel and Ferat, 1995), among which *Lactococcus lactis* L1.LtrB intron has been well-studied for its efficient mobility in both Gram-positive and Gram-negative bacteria (Mills et al., 1996, Karberg et al., 2001). The target specificity of the group II intron can be controlled by modifying a few base pairs in the intron RNA sequence. These target sites are predicted by the Perutka algorithm, which calculates the best scores for the on-target insertion in the desired gene (Karberg et al., 2001, Kuehne and Minton, 2012).

ClosTron is a genetic modification tool, designed using the Lambowitz's group II intron, for knocking-out genes in Clostridia (Heap et al., 2007). The process begins with the transformation of host clostridial species with the pMTL007-EC2 plasmid that carries a retrotransposition-activated marker or RAM, along with the mobile group II intron. RAM harbours the *ermB* gene for erythromycin resistance, which serves as a marker for indicating a successful intron insertion. To control the expression of *ermB*, a group I intron phage *td* is inserted in the very gene, in the orientation that keeps the RAM in an inactive state (Kuehne and Minton, 2012).

As concisely illustrated in Figure 1.6, the transcription of the DNA strand around the group II intron is followed by the binding of LtrA protein with the transcript, forming the RNP complex. This forces the group I *td* intron to correct its orientation and splice out. In order to insert intron RNA at the target site, LtrA recognizes and binds to the

targeted gene sequence in the host chromosome. The reverse transcriptase activity of LtrA protein synthesizes the complementary DNA or cDNA, while the host nucleases degrade the RNA insert. The opposite DNA strand is synthesized by DNA polymerases and all the gaps are sealed by host ligases. The functionality of *ermB* gene is observed by the appearance of erythromycin-resistant colonies of the desired mutants on selective agar plates (Kuehne and Minton, 2012). To determine the success of RAM insertion in the target locus, PCR and DNA sequencing are done using relevant primers, and the insertion is further confirmed by a Southern blot assay (Heap et al., 2010).

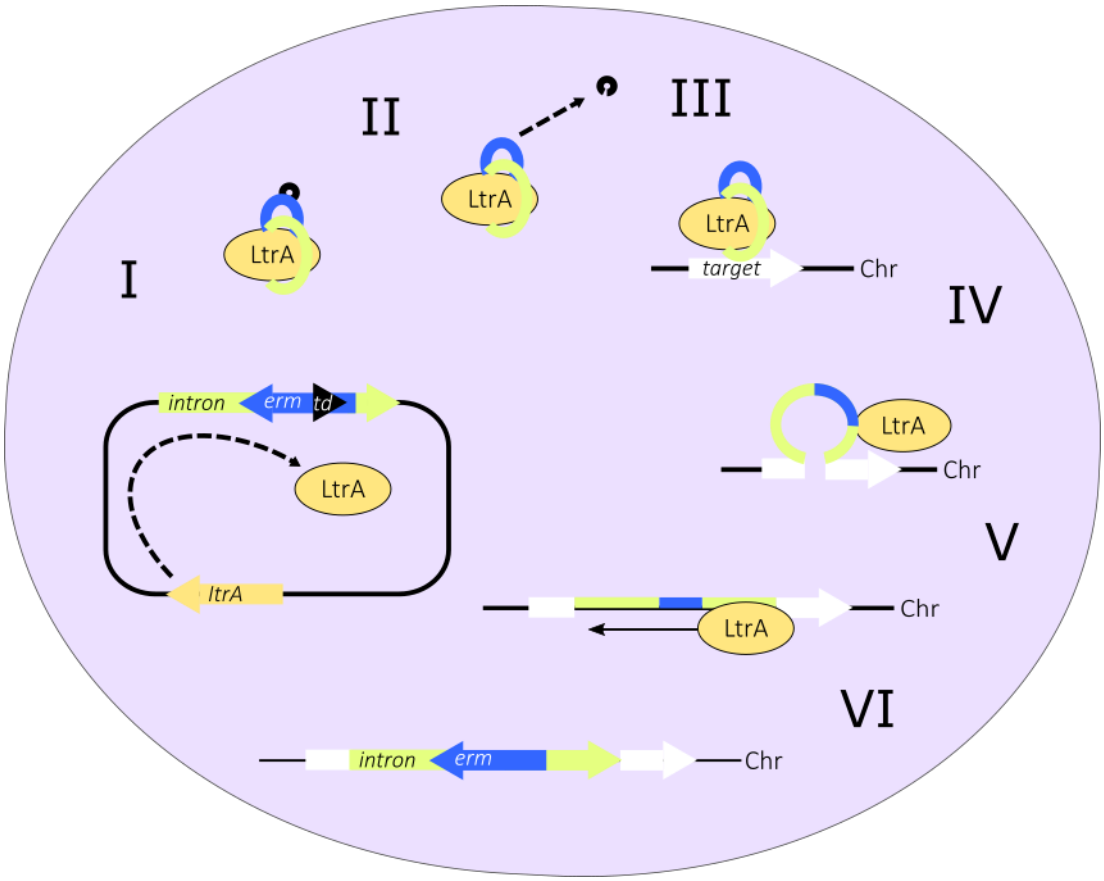


Figure 1.6. **ClosTron-mediated mutant generation in Clostridia**. Mobile II intron and LtrA are allowed to express in the host chromosome and the intron forms a ring around the LtrA protein (I). Mobile I intron gets excised from the complex and the orientation of RAM gets corrected (II). Mobile II intron-LtrA complex recognises the target sequence (III) and inserts the intron into the gene (IV). LtrA synthesise the complementary DNA (V), and ClosTron mutant is generated, in which the *ermB* RAM is functional. This diagram was inspired by (Kuehne and Minton, 2012).

The ClosTron mutagenesis system has been previously used to study the importance of the *spo0A* gene in the sporulation process of several clostridial species (Heap et al., 2007, Heap et al., 2010). Multiple gene mutations per genome have also been achieved by using a RAM-less version of ClosTron plasmid in a different study on orphan histidine kinases of *C. acetobutylicum* (Steiner et al., 2011).

1.4.3. CRISPR/Cas9-based genome editing – A reverse genetics approach

Inspired from nature's defence mechanism, in which bacteria and archaea protect their genomes from bacteriophage (Sorek et al., 2013), the clustered regularly-interspaced short palindromic repeats Type II system a.k.a. CRISPR/Cas9 has been proved to be an effective tool for targeted genome editing in prokaryotes (Wang et al., 2016). The CRISPR/Cas9 mutagenesis functions via the activity of *Streptococcus pyogenes* Cas9 nuclease that is guided to the target DNA by small guide RNAs or sgRNAs (Huang et al., 2016). The chimeric guide RNA consists of CRISPR RNA or crRNA and trans-activating crRNA or tracrRNA (Wang et al., 2015). The target specificity is ensured by the 20 bp seed in crRNA, which guides the Cas9 to the target DNA sequence by the Watson-Crick base pairing (Hsu et al., 2014, Cao et al., 2016).

The Cas9 gene, the guide RNA expression module (i.e., 20 bp seed + guide RNA scaffold) and the homology cassette modules are incorporated on a plasmid platform and transferred to host cells (Huang et al., 2016). As concisely illustrated in Figure 1.7, the Cas9 nuclease forms a complex with the guide RNA, which leads the complex to the target DNA sequence of the host chromosome (Hsu et al., 2014, Cao et al., 2016). In the presence of a protospacer adjacent motif or PAM sequence (i.e., NGG), the Cas9 nuclease forms double-stranded DNA breaks or DSB in the target chromosome (Sander and Joung, 2014). The homology arms cassette serves as a DNA template for repairing the host DNA by recombination events. In an alternative to this homology-directed repair or HDR mechanism, the DNA repair system can also work with non-

homologous end-joining or NHEJ, in which the host's machinery fills the gap and join the breaks by incorporating random nucleotide bases (Ran et al., 2013, Ding et al., 2016). In this way, gene deletions and insertions are easily made, which allow the generation of desired mutants to observe the mutant phenotype, as well as gene complemented strains for restoring the gene and understanding the role of gene in the host organism.

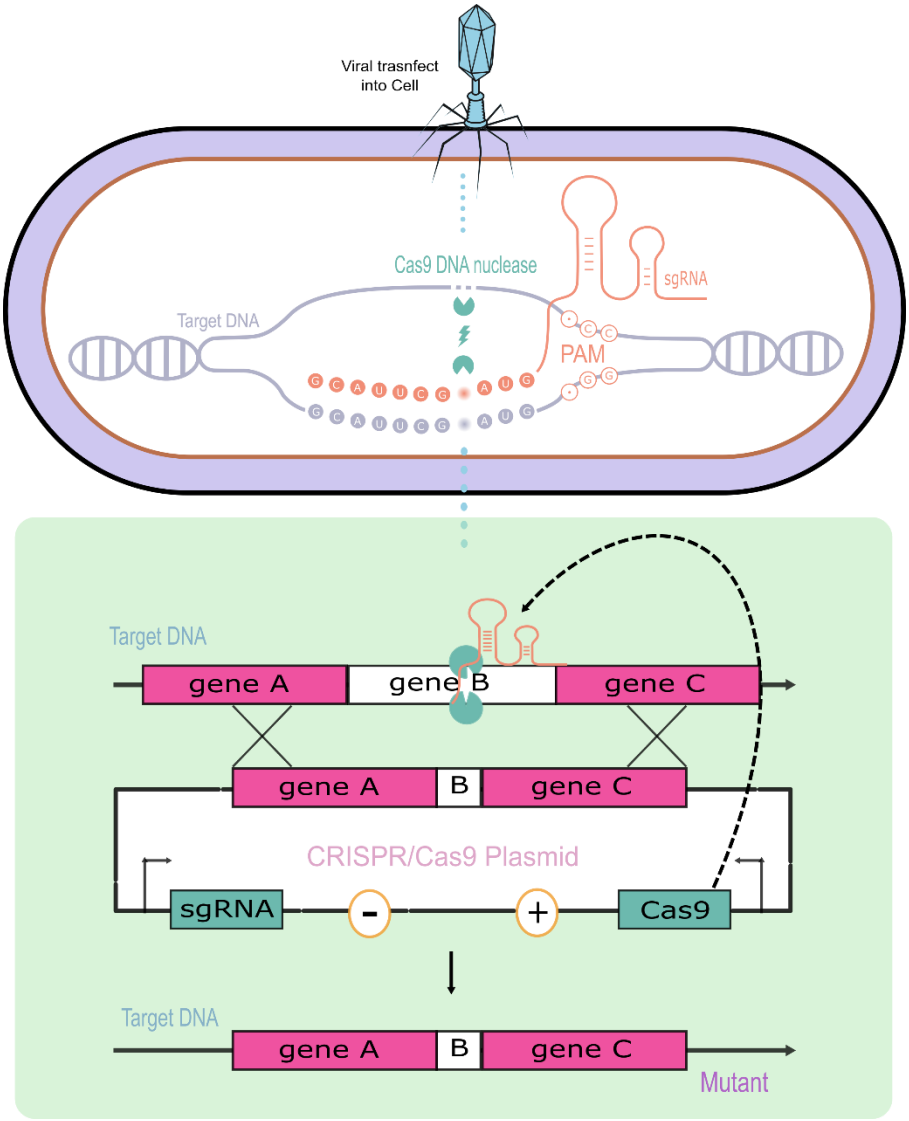


Figure 1.7. **Highlights of a CRISPR/Cas9-based targeted genome editing**. CRISPR RNA is generated in bacteria in response of a foreign DNA. This small RNA guides the Cas9 nuclease to the target DNA sequence for making double-stranded breaks. Genome engineering utilises this phenomenon to generate mutants or engineered strain. The example shown here in the green background explains how mutant gene B is generated by homology-directed repair in CRISPR/Cas9 system. This diagram was inspired by (Lee, 2015).

Genome editing tools, like CRISPR/Cas9 also open up new possibilities for developing better yields of products, useful synthetics, industrial strains, resistant crops, and novel therapeutic strategies (Hsu et al., 2014).

CRISPR/Cas9 has been previously used to genetically modify and improve the industrially-important strains of *Clostridium beijerinckii* (Wang et al., 2015), *Clostridium cellulolyticum* (Xu et al., 2015), *Clostridium autoethanogenum* (Nagaraju et al., 2016), *Clostridium ljungdahlii* (Huang et al., 2016) and *Clostridium tyrobutyricum* (Zhang et al., 2018).

Despite the revolutionary practicality of CRISPR/Cas9-based systems for genetic modifications, there are certain obstacles in the process, and one of the biggest problem is the Cas9 toxicity in both donor and host strains (Cui and Bikard, 2016). Recently, a theophylline-dependent riboswitch-inducible CRISPR/Cas9 tool, RiboCas has been developed to combat the toxicity and efficacy problems of CRISPR/Cas9 mutagenesis in Clostridia (Canadas et al., 2019), which is also exemplified in the later chapters of this thesis. Furthermore, the widely used CRISPR/Cas9 technique depends on generating DSB in the host DNA and repairing it by using the provided homology arms template for recombination. However, these DSB are dangerous (Rich et al., 2000) and the efficiency of HDR can be very low (Zu et al., 2013). Strategies that induce point mutations without the use of DSB and HDR can help to overcome such problems in the mutant generation process (Tanaka et al., 2018).

1.4.4. Target-AID base editing – A reverse genetics approach

Genome engineering tools that depend on homology-directed recombination become a great challenge if the DNA sequence in the homology arms has toxic effects in the intermediate cloning strain, like *E. coli* (Ribis et al., 2018). While gene knock-in strategies majorly depend on integrating an entire gene CDS using HDR, gene mutations can actually be done independent of using a template for recombination.

One such base editor is the CRISPR/Cas activation-induced cytidine deaminase or Target-AID system, in which the CRISPR/Cas9 components works in synergy with AID to specifically target host gene and introduce a premature STOP codon in the start of the gene through a C→T point mutation (Nishida et al., 2016, Tanaka et al., 2018).

Efficient guide RNA recognizes the target gene, while the engineered Cas9 nikase makes single-strand breaks or SSB in the target gene (Brezgin et al., 2019, Eid et al., 2018). Within its catalytic window, AID deaminates and converts a cytosine to uracil on the target strand, for which the host DNA repair machinery replaces guanine on the complementary strand to an adenine. Since in a DNA molecule, thiamine is the uracil-equivalent base, the previously converted uracil will become thiamine on a new DNA strand after replication (Nabel et al., 2012, Komor et al., 2018, Eid et al., 2018, Brezgin et al., 2019). In short, Target-AID induces the C→U→T conversion in the host genome to generate mutants.

Although Target-AID's small catalytic window frame avoids the off-targeting problem better than the standard CRISPR/Cas9 tool, there are instances of non-specific targeting. For example, it has been known in the literature that that the -18 position upstream of PAM is more likely to be targeted by AID for base-editing than the -16 position in a yeast host (Nishida et al., 2016), and this phenomenon has also been exemplified in Chapter eight of this thesis.

1.5. THE *CLOSPORE* PROJECT

The current study is a part of a grander project *CLOSPORE*, which was coordinated to derive intensive knowledge about clostridial sporulation and germination; and for exploiting these processes in the development of innovative tools and therapies that can potentially benefit the food, industrial and health sectors.

1.5.1. Some insights from literature on clostridial sporulation

Endospore formation in Clostridia is triggered by either starvation or acidic pH as a result of organic acid accumulation in the environment (Dürre, 2005). Similar to *Bacillus*, the *spo0A* gene is well conserved in the genome of all spore-forming clostridial species, and the deletion of which, causes the loss of sporulation phenotype and consequently, inability of the mutant species to persist in the environment for pathogenesis (Deakin et al., 2012). On the other hand, the clostridial Spo0A is directly activated by histidine kinases, in contrast to the phosphorelay activation system found in *Bacillus* spp. (Al-Hinai et al., 2015, Steiner et al., 2011). Although the sporulation-specific sigma factors σ^F, σ^E, σ^G and σ^K are also well conserved and essential for the sporulation cascade in all Clostridia, their expression and regulation pattern varies from species to species (Galperin et al., 2012, Al-Hinai et al., 2015, Fimlaid and Shen, 2015). The activities of sigma factors are compartment-specific; with σ^E and σ^K detected in the sporulating mother cell, while σ^F and σ^G being accumulated in the forespore of *C. difficile* strains (Pereira et al., 2013). Gene expression of *sigF*, *sigE* and *sigG* has been previously detected in 11 hour old cultures of proteolytic *C. botulinum* wild type strain, where *sigK* expression was observed to be highest in both 5 hour and 24 hour old *C. botulinum* cultures (Kirk et al., 2014).

The spore exosporium layer has also been reported and characterized in some of the Clostridia, including *C. sordellii*, *C. botulinum* and *C. sporogenes* (Hodgkiss et al., 1967, Takumi et al., 1979, Stewart, 2015, Brunt et al., 2015, Janganan et al., 2016).

Profound data to understand stages of sporulation has been mostly obtained in the past by constructing different classes of sporulation mutants in endospore-formers. For example, phenotypes of asporogenous mutants (that lose their ability to produce functional spores) and oligosporogenous mutants (that have decreased sporulation frequency than their wild type counterparts) have been well-studied in *B. subtilis* (Coote, 1972). Hence, majority of the information that is currently known about the sporulation process of endospore-forming bacteria is based on studies done on various *Bacillus* spp., presumably due to their great accessibility and widespread use in both industry and research laboratories. In contrast, many well-known clostridial species, such as *C. difficile, C. botulinum*, etc., are fastidious and require special growth supplements along with the constant maintenance of anaerobic conditions (Burke and Lamont, 2014, Annan et al., 2019). Accessibility of genetic tools for their use in the clostridial genome manipulation is also limited due to host's restriction modification systems (Minton et al., 2016). Additionally, stricter safety rules and risk assessment protocols are followed in order to work with the toxigenic strains of Clostridia.

In such circumstances, *C. sporogenes*, with its comparatively minimum nutrient requirements, relative ease of genetic manipulation and non-pathogenic nature, serves as a good model for studying clostridial sporulation and germination processes.

1.5.2. Aims and objectives

In order to better understand the sporulation process in *C. sporogenes*, specific aims were proposed in this study, which are as follows:

a. To construct and screen *C. sporogenes* transposon libraries for sporulation-defective mutants

b. To generate *C. sporogenes* mutants in selected genes using CRISPR/Cas9 genome editing

c. To study these mutants to characterize their function in *C. sporogenes*.

CHAPTER TWO

2. MATERIALS AND METHODS

2. MATERIALS AND METHODS

2.1. MATERIALS

All reagents and consumables were supplied from Thermo Fisher Scientific, Sigma Aldrich, Greiner Bio-One, Promega or New England Biolabs (NEB), unless stated otherwise.

2.1.1. Bacterial strains used.

Table 2.1. List of the bacterial strains used in this study

STRAINS	PROPERTIES	REFERENCE
Escherichia coli T0P10	F^- strain for cloning	Invitrogen
Escherichia coli 5-alpha F' Iq	F^- ($lacI^q$) strain for cloning toxic genes	NEB
Escherichia coli CA434	Conjugation donor	(Williams et al., 1990)
Escherichia coli sExpress	Dcm $methylation^-$ conjugation donor	(Woods et al., 2019)
Clostridium sporogenes NCIMB 10696	Wild type strain	(Kubiak et al., 2015)
Clostridium sporogenes *spo0A*::CT	*spo0A*-249s::CT	(Heap et al., 2007)
Clostridium sporogenes TcdR (CRG3817)	Wild type strain + integrated TcdR CDS	(Zhang et al., 2015)
Clostridium sporogenes #42*	*hypP*-419s::Tn; TcdR	This study
Clostridium sporogenes ProhypP	#42 + pMTL82251:*hypP*	This study
Clostridium sporogenes ProdivIVA	#42 + pMTL82251:*divIVA*	This study
Clostridium sporogenes #64*	*yqfD*-485s::Tn; TcdR	This study
Clostridium sporogenes Δ*yqfD*_old	CLSPO_RS15205 (*yqfD*) null mutant; in #64 background)	This study
Clostridium sporogenes Δ*yqfD*_catP	CLSPO_RS15205 (*yqfD*) null mutant; with catPsgRNA3 bookmark**	This study
Clostridium sporogenes YqfDcomp	#64 + integrated *yqfD* CDS	This study

Clostridium sporogenes Δ*yqfD*	CLSPO_RS15205 (*yqfD*) null mutant; with botsgRNA1 bookmark**	This study
Clostridium sporogenes YqfDCO	Δ*yqfD* + integrated *yqfD* CDS; with alternate codons watermark***	This study
Clostridium sporogenes *yqfD*::TA	*yqfD*-391a::TA (for G→A mutation)	This study
Clostridium sporogenes *yqfD*::TA:ProyqfD	*yqfD*::TA + pMTL87121:*yqfD*	This study
Clostridium sporogenes Δ*hypP*	CLSPO_RS07555 null mutant	This study
Clostridium sporogenes Δ*divIVA*_old	CLSPO_RS07560 null mutant; with SNPs construct	This study
Clostridium sporogenes Δ*divIVA*	CLSPO_RS07560 null mutant	This study
Clostridium sporogenes DivCO	Δ*divIVA* + integrated *divIVA* CDS; with alternate codons watermark***	This study
Clostridium sporogenes *divIVA*::CT	*divIVA*-273s::CT	This study
Clostridium sporogenes Δ*clpX*	CLSPO_RS16880 null mutant; with botsgRNA1 bookmark**	This study
Clostridium sporogenes ClpXcomp	Δ*clpX* + integrated *clpX* CDS	This study
Clostridium sporogenes *clpX*::TA	*clpX*-70s::TA (C→T mutation)	This study
Clostridium sporogenes Δ*spoVK*	CLSPO_RS17345 null mutant; with botsgRNA1 bookmark**	This study
Clostridium sporogenes SpoVKcomp	Δ*spoVK* + integrated *spoVK* CDS; with SNPs construct	This study
Clostridium sporogenes SpoVKCOs	Δ*spoVK* + integrated *spoVK* CDS; with alternate stop codon watermark***	This study
Clostridium sporogenes ProSpoVK	Δ*spoVK* + pMTL87121:*spoVK*	This study
Clostridium sporogenes Δ*spoIID*	CLSPO_RS00950 null mutant; with botsgRNA1 bookmark**	This study
Clostridium sporogenes SpoIIDCOs	Δ*spoIID* + integrated *spoIID* CDS; with alternate stop codon watermark***	This study
Clostridium sporogenes *spoIID*::TA	*clpX*-437s::TA (C→T mutation)	This study

* see the complete list of *C. sporogenes* transposon mutants in the Appendix Table A2

**mutant strains were engineered with a guide RNA bookmark for future complementation using CRISPR/Cas9 tool; bookmark consisted of an efficient 20 bp guide RNA sequence, PAM sequence and an extra base to keep the gene in frame

***complemented strains were engineered to have alternate codon-based watermark in order to identify them among possible contaminants

2.1.2. Plasmids used.

Table 2.2. List of plasmid vectors used in this study

PLASMIDS	PROPERTIES	REFERENCE
AVAILABLE PLASMIDS		
pMTL-YZ14	Conditional transposon mutagenesis plasmid	(Zhang et al., 2015)
pMTL-CW20	pMTL-YZ14; *codA* cassette	(Craig Woods, unpublished)
pMTL83151-CRISPR/Cas9	pMTL83151; Cas9; SIBS sgRNA	(Ingle et al., 2019)
pMTL83121-pRECas1	pMTL83121; Rb3; Cas9; SIBS sgRNA	(Canadas et al., 2019)
pvFS36-TargetAID	pMTL83151; Rb3; Cas9; SIBS sgRNA; AID	(François Seys, unpublished
pMTL82251	Modular vector	(Ng et al., 2013)
pMTL87151-p19	pMTL87151; low copy Gram positive replicon p19	(James Millard, unpublished
PLASMIDS CONSTRUCTED		
pMTL007::Csp-*divIVA*-273s	ClosTron plasmid for *divIVA*::CT	This study
pMTL82251:*divIVA*	pMTL82251 + *divIVA* CDS + RBS + promoter	
pMTL82251:*hypP*	pMTL82251 + *hypP* CDS + RBS + promoter	
pCRISPR-yqfD	pMTL83251; catPsgRNA3	
pCRISPR-yqfD1	pMTL83221; catPsgRNA3	
pCRISPR-yqfD2	pMTL83221; yqfDsgRNA1	
pCRISPR-YqfDcomp	pMTL83221; catPsgRNA3; *yqfD* CDS	
pCRISPR-hypP	pMTL83221; hypPsgRNA1	
pCRISPR-divIVA	pMTL83221; divsgRNA1; 2 kb arm	
pCRISPR-divIVA1	pMTL83221; divsgRNA1; 1.4 kb arm	
pMTL83221-pRECas1	pMTL83121-pRECas1, but with *ermB* plasmid marker	
pRECas1-yqfD	pMTL83121-pRECas1; *yqfD*-knock-out cassette	
pRECas1-YqfDCO	pMTL83121-pRECas1; *yqfD*-knock-in cassette (with watermark)	
pRECas1-Erm-divIVA	pMTL83221-pRECas1	
pRECas1-divIVA	pMTL83121-pRECas1; *divIVA*-knock-out cassette	
pRECas1-DivCO	pMTL83121-pRECas1; *divIVA*-knock-in cassette	
pRECas1-clpX	pMTL83121-pRECas1; *clpX*-knock-out cassette	

pRECas1-ClpXcomp	pMTL83121-pRECas1; clpX-knock-in cassette	This study
pRECas1-Erm-spoVK	pMTL83221-pRECas1	
pRECas1-Erm-SpoVKcomp	pMTL83221-pRECas1	
pRECas1-spoVK	pMTL83121-pRECas1; *spoVK*-knock-out cassette	
pRECas1-SpoVKCOs	pMTL83121-pRECas1; *spoVK*-knock-out cassette (with watermark)	
pRECas1-spoIID	pMTL83121-pRECas1; spoIID-knock-out cassette	
pRECas1-SpoIIDCOs	pMTL83121-pRECas1; spoIID-knock-in cassette (with watermark)	
pTargetAID-yqfDsgA	pvFS36-TargetAID; yqfDsgA guide RNA	
pTargetAID-clpXsgC	pvFS36-TargetAID; clpXsgC guide RNA	
pTargetAID-spoIIDsgA	pvFS36-TargetAID; spoIIDsgA guide RNA	
pMTL87121-p19	pMTL87151-p19, but with low copy Gram negative replicon p15a	
pMTL87121:*yqfD*	pMTL87121-p19 + *yqfD* CDS + RBS + promoter	
pMTL87121:*divIVA*	pMTL87121-p19 + *divIVA* CDS + RBS + promoter	
pMTL87121:*hypP*	pMTL87121-p19 + *hypP* CDS + RBS + promoter	
pMTL87121:*spoVK*	pMTL87121-p19 + *spoVK* CDS + RBS + promoter	

2.1.3. Growth media used.

Table 2.3. List of nutrient media used for growing bacterial strains

MEDIA	RECIPE* (w/v)	PURPOSE	SOURCE
Tryptone Yeast Extract (TYG)	3% Bacto-tryptone; 2% Yeast extract; 1% Sodium thioglycollate	Growth of *C. sporogenes* strains	Oxoid
Luria-Bertani (LB)	1% Bacto-tryptone; 0.5% Yeast extract; 1% Sodium chloride	Growth of *E. coli* strains	Oxoid
S.O.C.	2% Bacto-tryptone; 0.5% Yeast extract; 10mM NaCl; 2.5mM KCl; 10mM MgCl2; 10mM MgSO4; 20mM Glucose	Recovery of *E. coli* strains after transformation	Invitrogen

* agar medium was prepared by addition of 1.5% Agar No.1 (Oxoid) prior autoclaving

2.1.4. Antibiotics and supplements used.

Table 2.4. List of supplements used in this study

ANTIBIOTICS	CONCENTRATION	PURPOSE	SOURCE
D-Cycloserine (Cm)	250 µg/ml (agar/broth)	Selective against *E. coli*	Apollo Scientific
Erythromycin (Erm)	50 µg /ml (agar) 30 µg /ml (broth)	Plasmid selection for *E. coli* and *C. sporogenes*	Sigma Aldrich
Thiamphenicol (Tm)	15 µg/ml (agar)	Plasmid or transposon mutant selection for *C. sporogenes*	MP Biomedicals
Chloramphenicol (Cm)	25 µg /ml (agar/broth)	Plasmid selection for *E. coli*	Alfa Aesar
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	1 mM (agar)	Plasmid loss induction in *C. sporogenes*	NZYTech
Theophylline (Theo)	5 mM (agar)	Cas9 induction in *C. sporogenes*	Sigma Aldrich

Oligonucleotides used in this study are listed down in Appendix Table A1.

2.2. BASIC MICROBIOLOGICAL METHODS

2.2.1. Preparing nutrient media plates

The pre-autoclaved nutrient agar media in DURAN® bottles (200 ml or 400 ml) were melted in a microwave, cooled to 50-55°C in a water bath, and then carefully poured into sterile disposable petri plates. Agar was allowed to completely solidify and the plates were dried at 37°C for 20-45 min. Plates were stored at 4°C in a plastic bag, until further use.

Fisherbrand 55 mm Polystyrene petri dishes were used for routine culture (growing bacteria, cloning, etc.), Thermo Scientific™ Sterilin™ 100 mm square petri dishes were used for patch-plating, and Greiner Bio-One 120 mm square petri dishes were used for spore assays.

2.2.2. Growing bacterial cultures

All *C. sporogenes* strains were grown in TYG medium, with or without selection, at 37°C in an MG1000 anaerobic workstation (Don Whitley, UK), with the anaerobic gas mixture (v/v) of Nitrogen (80%), Hydrogen (10%) and Carbon dioxide (10%). All the media, buffers and supplements were pre-reduced in the anaerobic cabinet, for at least 3 hours, prior to be used for growing clostridial strains.

LB medium supplemented with the relevant antibiotics as required was used to grow all the *E. coli* strains, while the heat-shocked or electroporated *E. coli* cells were recovered in S.O.C. The broth cultures were grown in a 30°C shaker at 200 rpm, while the inoculated agar plates were either placed in a 30°C incubator (Thermo Fisher Scientific, UK) or incubated at room temperature.

2.2.3. Storage of bacterial stocks

Bacterial stocks prepared from plates or broth cultures were stored at 80°C in TYG +10% v/v glycerol (for *C. sporogenes*) or LB + 10% v/v glycerol (for *E. coli*), in 2 ml cryogenic vials.

2.2.4. Preparation of competent cells

For routine cloning, competent cells of *E. coli* TOP10, *E. coli* CA434 and *E. coli* sExpress were prepared in house. However, for effective cloning of potentially-toxic DNA sequences, the *E. coli* TOP10 chemically-competent cells were later replaced by the ready-to-use NEB® 5-alpha F'Iq Competent *E. coli* cells.

2.2.4.1. Chemically competent *E. coli*

To prepare chemically competent *E. coli* TOP10 cells, 1.5 ml of overnight LB broth culture was inoculated into a 500 ml conical flask containing 200 ml pre-warmed LB broth, and incubated with shaking at 200 rpm, 37°C until an optical density of 0.4-0.6 OD_{600} was achieved. The culture was aliquoted (3 × 33 ml) into falcon tubes and incubated on ice for 10-30 min. The tubes were then centrifuged at 3,100 × *g*; 4°C for 10 min in the Allgera X-30 Benchtop Centrifuge (Beckman Coulter, UK). The pellets were suspended in 10 ml ice-cold 0.1 M $CaCl_2$ solution, incubated on ice for 15 min, and re-centrifuged. The pellets were re-suspended in 4 ml ice-cold 0.1 M $CaCl_2$ solution with 15% (v/v) glycerol. 120 µl of the cell suspension was aliquoted into 1.5 ml Eppendorf tubes and stored at - 80°C until use.

2.2.4.2. Electrocompetent *E. coli*

The culture for preparing electrocompetent *E. coli* CA434 and *E. coli* sExpress cells was prepared, as described for chemically competent cells in Section 2.2.4.1. However, the 200 ml culture was distributed into 5x 40 ml in falcon tubes, and centrifuged at 3,100 × *g*; 4°C for 10 min. The pellets were washed 2-3 times with pre-chilled distilled

water and eventually, combined into one tube. Following the re-suspension of the washed pellet in 1000 µl 1 M MOPS solution, 60 µl of the cell suspension was aliquoted into Eppendorf tubes and stored at - 80°C until use.

2.3. GENERAL MOLECULAR METHODS

2.3.1. DNA Purification

DNA was purified using either commercially available kits or by the Phenol/Chloroform method.

2.3.1.1. Commercial kit method

GenElute™ Bacterial Genomic DNA Kit (Sigma Aldrich) was used to extract the genomic DNA from broth cultures, while the DNA inserts were purified from PCR reactions using the Monarch® PCR & DNA Cleanup Kit (NEB). Monarch® DNA Gel Extraction Kit (NEB) was used to extract DNA from agarose gels. Monarch® Plasmid Miniprep Kit (NEB) was used to extract plasmid copies from *E. coli* broth cultures. All kits were used according to the manufacturer's protocol. For eluting DNA, nuclease-free water was used instead of the provided elution buffers.

2.3.1.2. Phenol/Chloroform method

For whole genome sequencing, Phenol:Chloroform:Isoamyl Alcohol or PCIA 25:24:1 (Sigma Aldrich) was used to extract genomic DNA. In this method, 2 ml of 4 hours old broth culture was transferred into a 2 ml tube and centrifuged at ~3,100 × *g* for 2 min in a 5427 R Centrifuge (Eppendorf, UK). The pellet was suspended in 180 µl freshly-prepared Lysozyme solution (10 mg/ml in phosphate buffered saline or PBS) and incubated in a 37°C water bath for 30 min, with occasional shaking. The lysate was treated with 20 µl RNase for 15 min at room temperature. Later, the proteins were denatured by adding 25 µl Proteinase K, 85 µl nuclease-free water and 110 µl 10% (w/v)

Sodium dodecyl sulphate or SDS to the lysate and incubating at 65°C for 30 min in a Digital Dry Bath dual position (Flowgen, UK).

In a fume hood, the cell lysate was mixed with 400 µl PCIA, and transferred in a phase-lock tube, which was centrifuged at 20,000 × *g* for 3 min. The top phase was transferred into an Eppendorf tube and re-treated with RNase for a further 15 min. 400 µl PCIA was mixed with the contents of treated cell lysate; transferred into a fresh phase-lock tube; and centrifuged twice in a similar manner. The top phase was eventually transferred into an Eppendorf tube containing 40 µl 3 M Sodium acetate with 800 µl ice-cold ethanol. Following the 30 min incubation at - 80°C, the tube was centrifuged at 20,000 × *g* for 15 min and the supernatant was carefully removed. The DNA pellet was washed in 70% (v/v) ethanol; re-pelleted and incubated at 37°C for 10 min to evaporate any traces of ethanol. Finally, the DNA was eluted in 50 µl nuclease-free water and stored at -20°C until use.

2.3.1.3. Quantification of DNA

DNA concentrations (ng/µl) were measured using the Nano Drop Lite Spectrophotometer (Thermo Fisher Scientific, UK), with nuclease-free water as blank, or by using the Qubit dsDNA BR Assay Kit and Fluorometer (Invitrogen).

2.3.2. Polymerase Chain Reaction

To amplify DNA, the DNA template, primers and polymerase master mix were mixed in 0.2 ml PCR tubes, according to the manufacturer's protocol. KOD Hot Start Master Mix (Novagen) was used for the amplification of chromosomal inserts and oligomerization of guide RNAs; whereas, DreamTaq Green PCR Master Mix (Thermo Scientific) was used for colony PCR. Colony PCR for both *E. coli* and *C. sporogenes* cultures was done by suspending a small portion of a colony in 10 µl nuclease-free water and using 1 µl of suspension as DNA template. All the PCR programs were run on the Eppendorf Mastercycler® nexus thermal cycler.

For SOEing PCR, two DNA templates and relevant primers were added to the PCR reaction mix and amplified by a thermal cycler.

All the primers used in this study are listed in the Appendix Table A1.

2.3.3. Agarose gel electrophoresis

DNA was identified and separated on the basis of its size using the Biometra Compact XS/S/M gel casting system. 100-180 ml of 1% (w/v) melted and cooled (50-55°C) agarose gel, prepared with 1 × Tris-Acetate-EDTA (TAE) buffer, was poured into the gel tray, and mixed with 10-18 µl SYBR Safe DNA gel stain (Thermo Fisher Scientific, UK). Gel combs were inserted and the gel was allowed to solidify. The gel casting tank was filled with 1 × TAE buffer, and the gel tray was submerged in the tank. Combs was then removed and DNA was loaded onto the gel wells. The apparatus was run at 100-120V; 400W for 30-45 min, according to the size of gel. DNA bands were observed by imaging the gel with the Molecular Imager Gel Doc XR+ System (Bio-Rad, UK) against the 2-log ladder (NEB) that was loaded on the gel along with the DNA samples.

2.3.4. DNA Digestion

For digesting DNA, oligonucleotide inserts and plasmids were mixed with the restriction enzymes and CutSmart buffer (NEB) in a 25 µl digestion mix and incubated in a 37°C water bath, according to the manufacturer's protocol. 1 µg DNA inserts were digested for 1.5 hours and then purified; whereas, the modular plasmids were incubated for 3-4 hours, dephosphorylated with Alkaline Phosphatase (Promega) for 30 min, and ran on gel or directly purified using a commercial kit.

Restriction endonuclease enzymes were used according to the cloning sites present on the target DNA. Plasmid pMTL83121-pRECas1 was digested with SalI-HF (NEB) and AscI (NEB), the relevant guide RNA was digested with SalI-HF and AatII (NEB) and the sewed-homology arms amplicon was digested with AatII and AscI. Plasmid

pTargetAID and the relevant guide RNA were digested with SalI-HF and AsiSI (NEB). Plasmid pMTL87121 and the relevant inserts were digested with SalI-HF and BamHI (NEB).

To maximise digestion efficacy, CRISPR/Cas9-based or low copy number plasmids (~15), such as pMTL83121-pRECas1, pTargetAID and pMTL87121, were digested overnight, dephosphorylated for 45 min and purified. Digestion and dephosphorylation was then repeated and the digested plasmid was separated on an agarose gel. The correct size DNA band was extracted from the gel.

2.3.5. DNA Ligation

The digested and purified DNA inserts were ligated using the Promega LigaFast™ Rapid DNA Ligation System, in a final volume of 10 µl ligation mix at room temperature. The amount of each insert to ligate was calculated by the following formula:

$$Volume\ of\ insert = \frac{100\ ng\ plasmid * size\ of\ insert\ (kb)}{size\ of\ plasmid\ (kb)} * 3$$

The fragmented DNA for self-ligation was incubated for 30 min; whereas two-way and three-way ligations were incubated for 1 or 2 hours, respectively.

For CRISPR/Cas9 cloning, relevant guide RNA and sewed-homology arms amplicon were ligated into pMTL83121-pRECas1. Cloning for Target-AID was done by ligating relevant guide RNA into pTargetAID. Whereas, for the complementation on plasmid-based assays, the relevant gene CDS, RBS and promoter sequences were ligated into pMTL87121.

2.3.6. Transformation and cloning

The ligated plasmid mix was used to transform chemically competent *E. coli* TOP10 or the NEB® 5-alpha F'Iq Competent *E. coli* cells. Plasmids were extracted from the

verified *E. coli* transformants and used for the electroporation of electrocompetent *E. coli* CA434 or sExpress cells.

The following protocols were used to transformation *E. coli* strains:

2.3.6.1. Chemical transformation

The entire contents of ligation reaction tube were transferred into 50 µl (thawed) chemically competent *E. coli* cells and incubated on ice for 30 min. The cell-plasmid mixture was heat-shocked in a 42°C water bath for 30 sec, and then instantly cold-shocked in a wet ice bath for 5 min. Cells were immediately recovered by the addition of 350 µl S.O.C. medium, and were kept in the 30°C shaking incubator for 2 hours. All the recovered cell suspension was spread on a selective LB agar plate, which was incubated at 30°C for 48-72 hours, until colonies were big enough for screening. For potentially toxic plasmid constructs, the inoculated agar plates were first incubated at 30°C for 24 hours and then kept at room temperature for a further 72-96 hours, until colonies were visible.

2.3.6.2. Electroporation

50 µl of (thawed) electrocompetent *E. coli* cells were mixed with 2-5 µl pure plasmid prep and transferred into a pre-chilled 0.2 cm-gap electroporation cuvette (Bio-Rad, UK). The mixture was pulsed using the Ec1 program in the MicroPulser Electroporator (Bio-Rad, UK), and cells were immediately recovered in 400 µl S.O.C. medium by incubating with constant shaking at 30°C for 1 hour. 200 µl of the recovered cell suspension was spread on a selective LB agar plate, which was incubated at 30°C for 24-48 hours. The cells transformed with potentially toxic plasmid constructs were incubated at room temperature for 48- 72 hours.

2.3.6.3. Cloning gene knock-out/-in constructs in *E. coli*

Plasmid constructs were made for cloning gene knock-out and knock-in constructs in *E. coli*. Plasmids and other DNA inserts were extracted or generated, respectively, and later digested, purified and ligated, as described in Sections 2.3.1 to 2.3.5. The ligation mix was used to transform 5-alpha F'Iq Competent *E. coli* cells, as described in Section 2.3.6.1. The obtained transformants were screened by performing colony PCR, using primers that amplify from the plasmid backbone into the ligated cassette, and running an aliquot of the PCR reactions on an agarose gel with the 2-log DNA ladder, as described in Sections 2.3.2 and 2.3.3. The remaining reaction mix of the those transformant colonies that generated the correct size amplicon was purified and sent for sequencing with the relevant sequencing primers, as described later in Section 2.3.8. The plasmid construct was extracted from the verified transformants and used for transforming electrocompetent *E. coli* sEXpress cells, as described in Section 2.3.6.2. The construct was confirmed again by colony PCR, using the relevant primers, which amplify the entire ligated cassette on the plasmid. The PCR amplicons were purified and sent for Sanger sequencing.

2.3.7. Conjugation

For transferring plasmid constructs from *E. coli* to *C. sporogenes* strains, 5 ml overnight cultures were prepared in their relevant broth media, mainly *E. coli* grown with selective antibiotic to maintain the plasmid and *C. sporogenes* grown without any antibiotic. 1 ml *E. coli* culture was centrifuged at ~3100xg in a 5427 R centrifuge (Eppendorf, UK) for 1 min and washed with 1 ml PBS. The pellets were taken into the anaerobic cabinet and carefully re-suspended with 200 µl of the respective *C. sporogenes* broth culture. The cell mixture was then spotted onto a pre-reduced TYG plate as 5-6 spots and incubated for 7 hours. The growth was then scraped from the plate using a 10 µl disposable loop and suspended in 500-1000 µl pre-reduced PBS, depending on the growth. Subsequently, 80 µl, 160 µl and the remaining suspension

was spread on pre-reduced selective TYG plates to select for *C. sporogenes* transconjugants, but against *E. coli* cells. The selective plates were then incubated for 24-72 hours, until the appearance of transconjugants. Several colonies were randomly selected, streaked onto fresh selective agar plates and further processed according to the respective genome editing protocol.

2.3.8. DNA Sequencing

2.3.8.1. Region-specific sequencing

All the PCR products were purified using the commercially available kits, and along with their respective primers, and were sent for DNA sequencing via the Source BioScience Sanger Sequencing service, the University of Nottingham's Deep Seq - Next Generation Sequencing facility, or the Eurofins Genomics Sanger Sequencing service.

2.3.8.2. Whole-genome sequencing

The genomic DNA was extracted by the PCIA method and sent for whole genome sequencing using University of Birmingham's MicrobesNG Illumina Sequencing service or University of Nottingham's Deep Seq Illumina MiSeq service.

2.4. BIOINFORMATICS AND DATA ANALYSIS

2.4.1. Genome editing

The annotated genome of *C. sporogenes* NCIMB 10696 was imported from the NCBI databases using the reference number NZ_CP009225.1. To view and edit nucleotide sequences, SnapGene v.3.2 was used. All the primers, knock-in/-out cassettes and plasmid constructs were virtually designed using SnapGene.

2.4.2. Sequencing data analysis

2.4.2.1. Region-specific data

The data obtained from Sanger sequencing of the PCR products was analysed in SnapGene. A virtual construct of the desired genome mutagenesis was constructed, which was then aligned against the obtained files from sequencing to determine the presence of the desired nucleotide sequence.

2.4.2.2. Whole-genome sequencing data

The data obtained from Illumina sequencing was analysed using the QIAGEN's CLC Genomics Workbench v.11. The trimmed reads were mapped to the reference genome from the NCBI genome database and SNPs in the genome were determined using the Fixed Ploidy Variant Detection function. The resulting SNPs table was then imported in an Excel spreadsheet.

2.4.3. BLAST and InterPro

For determining the presence of homologues of the *C. sporogenes* genes and proteins in other species of *Clostridium* as well as *Bacillus* spp., the NCBI BLAST suites (Altschul et al., 1990): blastn and blastp (https://blast.ncbi.nlm.nih.gov/Blast.cgi) were used, respectively.

InterPro tool (Mitchell et al., 2019) was used to identify conserved domains in the target protein sequences and find relevant literature. This scanning tool can be easily accessed online (http://www.ebi.ac.uk/interpro/).

2.4.4. Graphics Analysis

The data obtained from spore assays was first recorded in Excel spreadsheets and CFU/ml was calculated for each sample. To construct graphical figures, these values

were copied to the GraphPad Prism v.8.2 program, and were analysed using the relevant graph types.

2.5. MICROSCOPY

2.5.1. Light Microscopy

Nikon Eclipse Ci microscope was used for visualizing cells and spores, with and without spore staining and the microscopic field images were taken using the NIS-Elements F v.4.60 program.

Slides were prepared by centrifuging 1 ml of 120 hour old broth cultures (~3,100 × *g*; 1.5 min) and suspending the obtained pellets with 50-250 µl nuclease-free water, depending on the cell density. 5 µl of the cell suspension was loaded onto a glass slide and the drop was either covered by a coverslip to visualize using the 40 × phase contrast lens or was air-died and covered with immersion oil (Sigma Aldrich) to be visualized using a 100 × magnification lens on bright field.

For spore staining, 5 µl of the cell suspension was loaded on a glass slide, air-dried and heat-fixed using a Bunsen burner. The smear containing the cells and endospores were stained by the Schaeffer and Fulton Spore Stain Kit (Sigma Aldrich), according to the manufacturer's protocol. The slides were air-dried and visualized with the 100 × magnification lens on the bright field.

2.5.2. Transmission Electron Microscopy

All the starved control and mutant cultures were chemically fixed, stained, washed, dehydrated and embedded in resin, which were carefully sectioned and stained for visualization by Transmission Electron Microscopy or TEM. Briefly, 72 hour old TYG broth cultures of *C. sporogenes* wild type and sporulation mutant strains were centrifuged at 3,100 × *g* for 10 min and the supernatant was discarded. All the

subsequent TEM processing work was carefully carried out inside a fume hood. The pellets were fixed with freshly prepared solution of 3% (v/v) Glutaraldehyde in 0.1 M Cacodylate buffer for 3 hours and washed twice with 0.1 M Cacodylate buffer, which was also used as a medium to preserve the fixed cells until further processing.

Later on, Cacodylate buffer was discarded from the bacterial suspension and the pellets were stained with 1% (v/v) Osmium tetroxide for 1 hour and then washed twice with distilled water. Cultures were then treated with 2% (v/v) Tannic acid for 30 min and then rinsed with 1% (v/v) sodium sulphate. In order to avoid excessive loss of culture from subsequent processing steps, the pellets were submerged in 200 µl of 2% (w/v) low gelling agarose gel, which was allowed to solidify and then sliced into small pieces. The pellet gel slices were gradually dehydrated twice in 50%, 70%, 90% and 100% (v/v) ethanol for 10 min, while on a shaker. Ethanol was carefully removed and the gels were incubated twice in Propylene oxide for 15 min.

TAAB low viscosity resin (24 g) was freshly mixed with the hardener (26 g) and was first added to the pellet gels in 3:1 Propylene oxide:resin mixture and left to incubate for 1 hour at room temperature. The pellet gel were then transferred to 1:1 Propylene oxide:resin mix, in which they were incubated overnight with the lids open. Pellet gels were later removed and added to the undiluted resin in the presence of the accelerator and incubated under pressure twice for 2 hours. Eventually, pellet gels were placed in fresh resin, which was polymerized in an embedding oven at 70°C for 48 hours.

By the courtesy of Ms Denise McLean, ultra-thin (90 nm) sectioning of the resin-embedded samples was done using a diamond knife, while visualizing with the Leica EM UC6 Ultramicrotome. On a 200 carbon coated copper grid, samples were stained with lead citrate. Grids were finally visualized by the FEI Tecnai T12 BioTWIN transmission electron microscope, running at 100 kV, in the Nanoscale and Microscale Research Centre, Nottingham. Images were captured with the MegaView SIS camera, also by the courtesy of Ms Michelle Kelly.

2.6. CONDITIONAL TRANSPOSON MUTAGENESIS

2.6.1. Pre-requisites to transposon libraries

Initially, the transposon plasmid (pMTL-YZ14 or pMTL-CW20) was extracted from the available *E. coli* strains, and used to transform *E. coli* CA434 competent cells, as described in Section 2.3.6.2. The stability of the transposon cassette in the plasmid was determined by digesting it with EcoRV restriction enzyme (NEB) and confirming the size of the obtained fragments.

2.6.2. Generation of transposon libraries

2.6.2.1. Obtaining *C. sporogenes* transconjugants

Libraries of transposon mutants were generated by transferring the transposon plasmid from the *E. coli* CA434 strain to the *C. sporogenes*-TcdR strain via a 7 hours conjugation process, as described in Section 2.3.7. The growth was spread on TYG plates, supplemented with D-cycloserine and erythromycin (TYG/Cs/Erm) and the resulting transconjugants were re-streaked on fresh TYG/Cs/Erm plates to eliminate any *E. coli* contaminant.

2.6.2.2. Transposon induction

The transconjugants were scraped off from the TYG/Cs/Erm plates, suspended in 1 ml PBS and serially diluted to 1×10^{-6}. 100 µl of the appropriate dilutions were spread on TYG plates, containing 1 mM IPTG and thiamphenicol (TYG/IPTG/Tm), and the plates were incubated for 24 hours.

2.6.2.3. Plasmid loss

After the generation of each transposon library, 48-144 colonies were randomly picked from the TYG/IPTG/Tm plates and patch-plated on 100 mm square petri plates

of TYG supplemented with D-cycloserine and thiamphenicol (TTC), as well as the TYG/Cs/Erm plates. Colonies that were resistant to thiamphenicol, but sensitive to erythromycin were declared as "transposon mutants".

These TTC plates were treated as "Master plates" from which the transposon mutants were sourced and used for screening for sporulation-defective phenotypes. All master plates were carefully sealed and stored in the anaerobic cabinet until further use.

2.6.3. Screening of transposon mutants

Initially, the transposon mutants with sporulation deficiency were directly identified by spore assay. However, for increasing the screening efficiency, the approach of using a transparency assay paired with the spore assay was adopted later on.

2.6.3.1. Spore assay for the direct screening of asporogenous and oligosporogenous mutants

Transposon mutants were grown in 500 µl TYG/Tm broth in 48 well plates in duplicates, such that there were are two exact replicas of the culture plates. Control strains comprising *C. sporogenes* wild type and *C. sporogenes spo0A*::CT were also grown alongside the transposon mutants. In order to induce sporulation, the broth cultures were starved in the anaerobic cabinet for a week. After that, both sets of plates were briefly brought out of the anaerobic cabinet. One set of plates was heat-treated at 80°C for 20 min in a Techne HB-1D Hybridizer and the second stored at room temperature. Then, both sets were taken back into the anaerobic cabinet to proceed with the spore assay.

Using an INTEGRA Viaflo 2 multi-channel pipette, 5 µl of neat untreated cultures were spotted in triplicates on 120 mm square TTC plates. Similarly, 5 µl of heat-treated cultures were spotted in triplicates on the square TTC plates, as neat or 1×10^{-3} diluted. The control strains, *C. sporogenes* wild type and *C. sporogenes*

spo0A::CT, were plated on TYG/Cs and TYG/Cs/Erm plates, respectively. All plates were incubated for 24 hours and the transposon mutants that grew on the untreated set, but not from the heat-treated set (neat or diluted) were declared as sporulation-defective mutants.

2.6.3.2. Transparency assay for the rapid identification of putative sporulation mutants

Based on an initial observation that the colonies of *C. sporogenes* sporulation-defective mutants start to turn translucent within 48-72 hours on TYG agar plates, but gave opaque colony morphology in fresh cultures upon sub-culturing; the transposon mutant colonies were left on the TYG/IPTG/Tm plates for 72 hours after induction. Colonies that turned translucent within 48-72 hours, but not earlier, were selected for the spore assay screenings, as described in Section 2.6.3.1.

2.6.4. Identification of transposon insertion sites

Genomic DNA was extracted from the selected mutants and quantified. 500 ng/µl of genomic DNA was digested with HindIII-HF (NEB) in a 10 µl mix, and after an hour's incubation, the enzyme was heat-inactivated at 80°C for 15 min. 100 ng of digested genomic DNA was self-ligated during 30 min incubation in a 10 µl mix. For Inverse PCR, 12.5-50 ng DNA template was used along with the catP-INV-F1 and catP-INV-R1 primers. The PCR product was run on a gel to check the number of amplicons obtained. DNA was purified from the gels and send for sequencing with the catP-INV-R2 primer. The transposon-chromosome region in the sequenced data was aligned and matched with the *C. sporogenes* wild type genome in order to identify the targeted gene.

2.7. CRISPR/CAS9-BASED GENOME EDITING

2.7.1. Designing and generating guide RNAs

For determining the best guide RNA sequences for targeting the desired genes, Benchling's CRISPR design tool (https://benchling.com/crispr) was used. The three 20 bp sequences that had the best on-target and off-target scores, were selected for the targeted genome editing, and corresponding primer sequences were designed and ordered.

Guide RNAs were generated by the oligomerization of the relevant primers with the sgRNA_AatII_R primer in a Touch down PCR reaction and then purified.

2.7.2. Generation of homology arms

For constructing the knock-out cassette for each gene target, primers were designed to amplify the region 0.54 kbp upstream of the target gene (i.e., for left homology arm or LHA) and the region 0.54 kbp downstream of the target gene (i.e., for right homology arm or RHA). The primer sequences included the SalI and AatII restriction sites, the required extra bases and, the overlaps required for SOEing PCR, as well as a bookmark for future targeted genome editing. This bookmark consisted of 20 bp seed sequence of an efficient guide RNA, a PAM sequence or NGG for Cas9 activity and an extra base to maintain the nucleotide frame.

The gene knock-in cassettes, for targeting the bookmark-based mutants, were designed in a similar manner; however, the bookmark was replaced by an alternate codon-based watermark to omit any possibility of false positive results owing to contamination with wild type *C. sporogenes*. This watermark was designed in two ways. In one approach the positions of codons encoding the same amino acid were swapped with each other; e.g., two consecutive glycine codon sequences, GGTGGA

were changed to GGAGGT. In a different approach, the original STOP codon (e.g. TAA) was replaced with a different STOP codon (e.g. TAG).

LHA and RHA for both gene knock-out and knock-in were individually generated using the relevant LH_F/R and RH_F/R primers (Appendix Table A1), purified and sewed together by SOEing PCR, separated on an agar gel, and finally purified, as described in Sections 2.3.1 to 2.3.4.

2.7.3. Genome editing in *C. sporogenes*

All the gene knock-out and knock-in plasmid constructs were cloned in relevant *E. coli* strains, as described in Section 2.3.6.3, and later processed according to their desired goal.

2.7.3.1. Generation of *C. sporogenes* mutants

The transformed *E. coli* sExpress strains with the verified gene knock-out cassette were conjugated with *C. sporogenes* wild type strain, as described in Section 2.3.7. The thiamphenicol-resistant transconjugants were streaked on TYG agar plates supplemented with theophylline and thiamphenicol, which were incubated for 24 hours. The induced mutants were confirmed via colony PCR, using the relevant oLH_F and oRH_R primers that amplify from outside the homology arms region on the host chromosome, and running an aliquot on a gel to check the size of the obtained DNA product. To promote plasmid loss, the mutants were inoculated into 5 ml TYG broth and incubated for 48 hours. Cultures were then serially diluted and plated on TYG/Cs plates, from which the colonies were patch plated onto fresh TYG/Cs and TTC plates. The thiamphenicol-sensitive colonies were re-confirmed by colony PCR and sequencing and were stored at -80°C.

2.7.3.2. Complementing *C. sporogenes* mutants

The transformed *E. coli* sExpress strains with the verified gene knock-in cassette were conjugated with their corresponding *C. sporogenes* mutant strains, as described in Section 2.3.7. Similar to the mutant generation protocol described in Section 2.7.5.1, the Cas9 activity in the thiamphenicol-resistant transconjugants was induced with theophylline and recombination confirmed by colony PCR. Later, the cells were cultured to promote plasmid loss, which was confirmed by patch-plating. The complemented strains were re-confirmed by colony PCR and Sanger sequencing and the bacterial stocks were stored at -80°C.

2.8. TARGET-AID GENOME EDITING

2.8.1. Designing and generating guide RNAs

Efficient RNA spacer sequences were determined using the Benchling's CRISPR tool, as briefly mentioned in Section 2.7.1. Afterwards, an in-house CRISPR-guide RNA design tool (François Seys, unpublished) was used to find the best 20 bp spacer sequence that can guide the Cas9-AID complex to the target locus in the *C. sporogenes* genome. Preliminarily developed on Microsoft Excel, this in-house tool analyses Benchling-identified 20 bp guide RNA sequences for the presence of a Cytosine or C base that is capable of introducing a STOP codon at the start of the target gene. The resulting efficient 20 bp guide RNA sequences were also manually verified by analysing them on the genome database with respect to the closest PAM sequence. Ultimately, corresponding primer sequences were designed and ordered.

The guide RNAs were generated by the oligomerization of the relevant primers with the sgRNA_PD_Rv primer in a Touch down PCR reaction and then purified.

2.8.2. Generating Target-AID mutants in *C. sporogenes*

All the pTargetAID constructs were cloned in relevant *E. coli* strains, as described in Section 2.3.6.3. Relevant *E. coli* sExpress pTargetAID strains were conjugated with *C. sporogenes* wild type strain and the 7 hours conjugation mix was directly plated on the selective TYG plates containing theophylline, i.e., TTC/Theo plates. The transconjugants were re-streaked on fresh TTC/Theo plates. Colony PCR was performed using the relevant flanking primers and the purified PCR products were sent for sequencing. The transconjugants with the desired mutation were selected for plasmid loss. The loss of plasmid was confirmed by patch-plating the colonies and the resulting thiamphenicol-sensitive mutants were stored at -80°C.

2.9. CLOSTRON GENE KNOCK-OUT SYSTEM

2.9.1. Designing the ClosTron intron

The intron for mutant generation was designed using the Perutka algorithm's intron designing tool on the ClosTron website (http://www.ClosTron.com/). Based on the target gene sequence, the tool predicted the best intron sequences for efficient genome targeting. The intron sequence with the best score was selected and the ClosTron plasmid, harbouring the intron on the pMTL007-EC2 backbone, was ordered using the DNA 2.0 services (https://www.atum.bio). The ClosTron plasmids were delivered on filter membranes, and the plasmid DNA was eluted in 50 µl of RNase-free water.

The ClosTron plasmid was electroporated into the *E. coli* TOP10 and CA434, as described in Section 2.3.6.2.

2.9.2. Generation of *C. sporogenes* ClosTron mutants

The ClosTron plasmid was eventually transferred to *C. sporogenes* wild type strain via conjugation, as described in Section 2.3.7. The growth of the thiamphenicol-resistant transconjugants was suspended in 1 ml PBS, serially diluted and was spread on TYG/Cs/Erm plates. The erythromycin-resistant colonies were screened by colony PCR using three sets of primers: the gene flanking primers F/R, the RAM_F/R primers, and EBS Universal primer coupled with one relevant flanking primer; and an aliquot of the PCR reactions were run on a gel. The colonies with the correct sized amplicons for each primer set were re-streaked to lose the ClosTron plasmid and this plasmid loss was confirmed by patch-plating on TYG/Cs/Erm and TTC plates. The erythromycin-resistant, thiamphenicol-sensitive mutants were stored at -80°C.

2.10. GENE CHARACTERIZATION

A number of assays were done to study the generated mutants.

2.10.1. Gene complementation

The mutants generated by CRISPR/Cas9 genome editing were complemented by inserting the deleted gene back into the same locus on the host chromosome, as described in Section 2.7.5. Due to challenges with chromosomal insertions, the transposon- or CRISPR/Cas9-based mutants were additionally complemented by placing the wild type gene, potential promoter sequence and RBS on a modular plasmid: pMTL82251 or pMTL87121-p19 (low copy number plasmid, ~19), respectively. The complementation plasmid was constantly maintained in the strains using plasmid-selective antibiotics in the medium.

The restoration of gene functionality was determined by comparing the phenotype of the complemented strains with that of both wild type and the corresponding mutant strains.

2.10.2. Standard spore assay

In this assay, 4 hour old *C. sporogenes* culture was used to inoculate 10 ml fresh TYG broths (1/100 v/v) and the cells were grown for 120 hours. 200 µl culture was withdrawn at each 0, 24, 48, 72, 96 and 120 hours, from which 100 µl from each sample was heat-treated in a 96 well plate using the Eppendorf Mastercycler® nexus thermal cycler. Both untreated and heat-treated cultures were serially diluted in PBS and 5 µl of each culture was spotted in triplicates on 120 mm square TYG plates.

After 24 hours incubation, the number of colonies were recorded and CFU/ml was calculated using the following formula:

$$CFU/ml = \frac{Average\ of\ colonies * dilution\ factor * 1000}{inoculum}$$

2.10.3. Growth assay

The growth curve of *C. sporogenes* strains was determined by inoculating the overnight broth cultures in 10 ml TYG broths at a starting optical density of 0.01 OD_{600} for all the cultures. OD_{600} at 0, 2, 4, 6, 8, 10, 12 and 24 hours was measured in 96 well plates using the BMG LABTECH's CLARIOstar® *Plus* Multi-mode Microplate Reader. The data was exported to Microsoft Excel spreadsheets to perform the required calculations and the growth curve was generated in GraphPad.

2.10.4. Haemolysis assay

The haemolytic activity of *C. sporogenes* strains was tested on TYG agar plates prepared with 5% (v/v) horse blood (Thermo Fisher Scientific). Overnight TYG broth cultures were adjusted to 0.5 OD_{600} and 2 µl from each culture was carefully spotted in triplicates onto the blood agar plates. The plates were incubated for 24-48 hours and the presence or absence of clearance zones of β-haemolysis was recorded for each strain.

2.10.5. Protease assay

Skimmed milk was used to assess the proteolytic activity of the *C. sporogenes* strains. Briefly, 36% (w/v) freshly prepared solution of skimmed milk was used to prepare 4% (v/v) milk TYG agar plates. Similar to the haemolysis assay mentioned in Section 2.10.4, overnight broth cultures were adjusted to 0.5 OD_{600} and 2 µl from each culture was carefully spotted in triplicates on milk agar plates. After 24-48 hours incubation, the presence and size of zones of clearance were observed and compared with the relevant control strains.

2.10.6. Cell and colony morphology

C. sporogenes strains were grown on TYG agar plates to generate single colonies and the differences in the colony morphology of the mutated strains were noted against that of the *C. sporogenes* wild type and the complemented strains.

Similarly, the cells and the spores of all the strains were observed under the microscope, with and without spore staining, as described in Section 2.5.1.

CHAPTER THREE

3. SCREENING FOR NOVEL SPORULATION GENES

3. SCREENING FOR NOVEL SPORULATION/ GERMINATION GENES

3.1. INTRODUCTION

Studies which adapt the forward genetics approach are helpful for genotypic identification, as they are based on screening desired mutant phenotypes without prior process of hypothesis-making regarding a particular gene function. Transposons have been widely studied for their valuable role in the forward genetics studies for the identification of novel genes (Cartman and Minton, 2010). *Mariner* transposable elements, in particular, have been characterized as universally the best transposons for genomic studies in both prokaryotic and eukaryotic organisms (Zhang et al., 2015); (Trubitsyna et al., 2017). This transposon family has been extensively exploited for the development of genetic tools that optimise the *in vivo* mutant generation and gene identification strategies in both Gram positive and Gram negative bacteria (Rubin et al., 1999); (Le Breton et al., 2006); (Chaudhuri et al., 2009); (Liu et al., 2013); (Maglennon et al., 2013).

Genetic tools that utilize transposable elements have also been previously employed to study sporulation in endospore-forming bacteria. In a previous study, *mariner* transposon was used to generate mutant libraries in *B. subtilis*, which were then screened for their sporulation profile on the basis of growth and colony morphology on DSM agar plates (Le Breton et al., 2006). In another study, the transposon mutant library of *C. difficile* was screened by a method called Transposon Directed Insertion Sequencing or TraDIS, which revealed the identity of all the genes essential for *C. difficile* growth and sporulation (Dembek et al., 2015).

3.1.1. Conditional transposon mutagenesis

The conditional transposon mutagenesis tool is the improved version of the previously designed genetic tool (Cartman and Minton, 2010), which was initially designed to induce *in vivo* mutagenesis in *C. difficile*. This enhanced transposon system is aimed to conditionally induce the transposition of the *mariner Himar1* element in most of the clostridial host species for the generation of mutant libraries, and it has already been validated to work in *C. sporogenes* and *C. acetobutylicum* (Zhang et al., 2015). The first important feature of the conditional system is based on the regulation of the *Himar1 C9* transposase gene under the tight control of the P_{tcdB} promoter that was previously derived from *C. difficile*. Because P_{tcdB} is known to only work in clostridial strains expressing the transcriptional factor TcdR, it is aimed to prevent any non-specific transposition event in the *E. coli* donor host. The second feature consists of the lac-inducible promoter cassette, in which the *lacI* gene represses the expression of the artificial P_{fac} promoter (Heap et al., 2007), which has been implemented to control the expression of the transposon selection marker gene *catP*. Moreover, the placement of fdx transcriptional terminator (from *Clostridium pasteurianum*) in between P_{fac} and pCB102 clostridial replicon prevents the loss of plasmid under normal conditions. Addition of IPTG in the growth medium severs the *lacI* repression on P_{fac}, which results in the acquisition of thiamphenicol-resistant transposon mutants as well as the induction of plasmid loss to avoid multiple transposon insertions within the same genome. The exact mechanism of this plasmid loss is not known. However, it can be assumed that the de-repression of P_{fac} consequences in high expression of pCB102 replicon, which causes sufficient metabolic burden on the cells that makes them lose the plasmid (Zhang et al., 2015).

Figure 3.1 shows the map of one of the conditional transposon mutagenesis plasmid used in this study. The original conditional transposon plasmid (Zhang et al., 2015) was modified to integrate a negative selective marker *codA* gene under the control of

P_{thl} promoter (Craig Woods, unpublished), to ensure the stability of the *mariner* transposable element in the pMTL-80000 modular vector series plasmid.

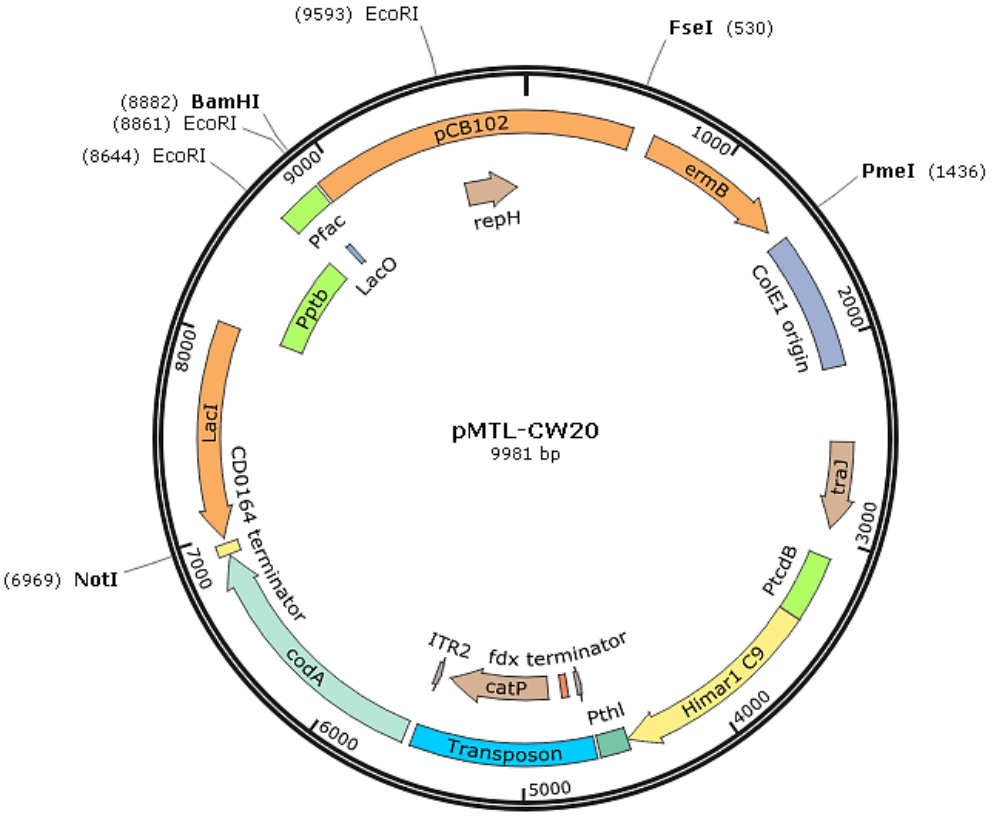


Figure 3.1. Plasmid map of the conditional transposon mutagenesis plasmid pMTL-CW20 (Craig Woods, unpublished). Key features: P_{tcdB} promoter from *C. difficile*; *Himar1 C9* transposase from *Haematobia irritans*; P_{thl} from *C. acetobutylicum*; flanking ITRs; fdx transcriptional terminator from *C. pasteurianum*; *catP* transposon marker; *codA* from *E. coli*; CD0164 transcriptional terminator from *C. difficile*; *lacI* from *E. coli*; P_{ptb} promoter from *C. beijerinckii*; artificial P_{fac} promoter; pCB102 Gram positive replicon; *ermB* plasmid marker; ColE1 Gram negative replicon; and *traJ* conjugation gene.

3.1.2. Aim of this chapter

This chapter is aimed at utilizing the conditional transposon mutagenesis tool for the construction of transposon mutant libraries in the *C. sporogenes*-TcdR strain and screening individual transposon mutants for any defects in sporulation or germination, in an attempt to discover novel genes involved in these processes.

3.2. RESULTS

In order to discover novel sporulation genes in *C.sporogenes*, transposon plasmids pMTL-YZ14 and pMTL-CW20 were used to introduce the *Himar1 C9*-based transposon into the *C. sporogenes* TcdR strain (Zhang et al., 2015). The resulting transposon mutant libraries were then screened to identify novel putative sporulation or germination genes, as summarized in Figure 3.2.

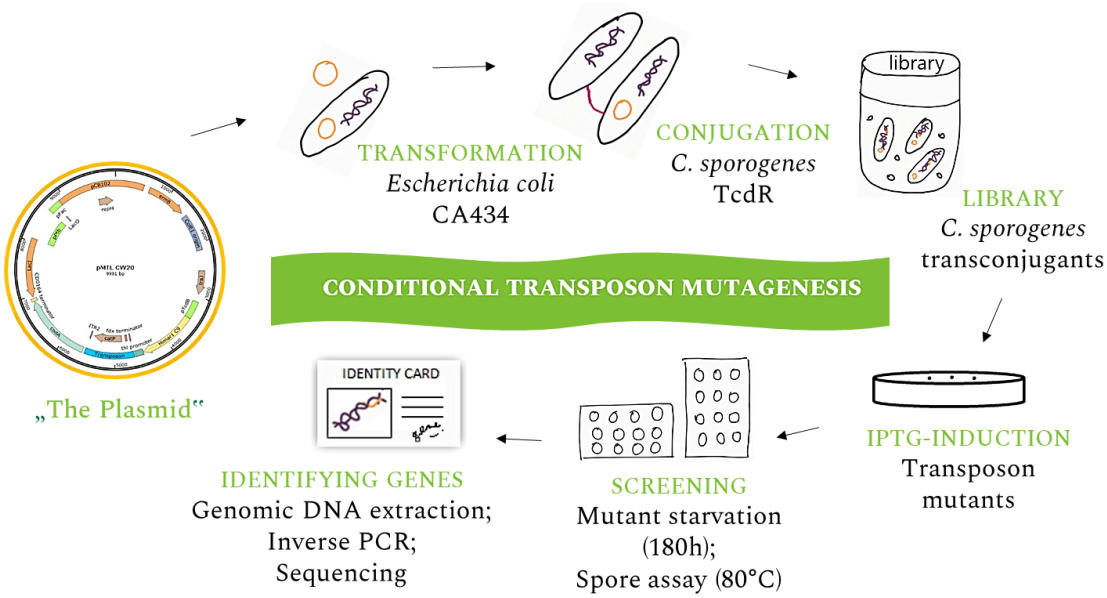


Figure 3.2. **Generating and screening transposon mutant library for sporulation or germination defects**. Transposon plasmid was cloned in *E. coli* strains and eventually transferred to *C. sporogenes* host by conjugation. The transconjugants were induced with IPTG and transposon mutants lacking plasmids were heat-treated to screen for sporulation defects. Transposon-interrupted genes were identified by inverse PCR and DNA sequencing the genome-transposon junction.

Mutant libraries v1.0, v2.0 and v3.0 were generated using pMTL-YZ14. However, due to the inadvertent mobility of transposon cassette in *E. coli* CA434, the subsequent libraries, v4.0 and v5.0 were generated using the pMTL-CW20 plasmid. The pMTL-CW20 harbours the *codA* cassette, which stabilized the transposon in *E. coli* donor strain, in a manner not currently known.

3.2.1. Screening transposon libraries

For screening multiple individual mutants on each attempt, multiwell plates were used. Transposon mutant colonies were randomly picked and screened by spore assay. Depending on the load of work, mutants were screened in batches of 46, 94 or 142. In each attempt, two control strains: *C. sporogenes* wild type and *spo0A*::CT were also screened with the transposon mutants. Library v1.0 was screened using 96 well plates; whereas, the succeeding libraries were screened using 48 well plates in order to minimize the evaporation and cross-contamination problem that occurred with the use of high mutant load in the 96 well plates.

After combining the data from all screening attempts, a total of 5,061 individual *C. sporogenes* transposon mutants were screened for sporulation or germination defects. Compared to the wild type strain, 41 of the 5,061 mutants produced less heat-resistant CFUs or showed no growth after the heat treatment, but were able to give similar CFUs without any heat-shock. 25 of these 41 mutants were observed to have the spo/ger$^-$ phenotype, i.e., they did not produce any heat-resistant CFU, similar to the *C. sporogenes spo0A*::CT strain. Whereas, the remaining 16 mutants showed a reduced degree of sporulation or germination, compared to the *C. sporogenes* wild type strain (Figure 3.3). A detailed list of transposon mutants and their sporulation profile is mentioned in Appendix Table A2.

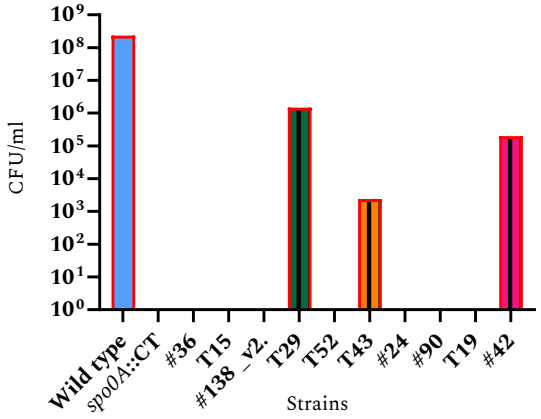


Figure 3.3. **Heat-resistant CFUs of some of the identified sporulation- or germination-deficient *C. sporogenes* transposon mutants**. Transposon libraries v1.0-5.0 were screened by spore assay, along with the sporulation-efficient *C. sporogenes* wild type strain and spo/ger⁻ *spo0A*::CT mutant. Strains #36, T15, #138_v2, T52, #24, #90, T19 were spo/ger⁻, i.e., no CFU was obtained after heat-shock. Strains, T29, T43 and #42 sporulated; however, with lesser frequency than the wild type. This data represents one set of observation at 180 hour culture incubation, owing to heavy workload of screening libraries. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4×10^6.

The first random 1,040 transposon mutants were screened individually by spore assay, which was labour-intensive; however, it identified 12 sporulation- or germination-deficient mutants. Interestingly, it was found that 11 of these mutants became translucent within 2-5 days of incubation on agar plates (Figure 3.4), but produced opaque colonies again upon sub-culturing (data not shown).

In order to decrease the workload and increase the screening efficacy, the mutant screening approach was revised on the basis of translucent phenotype. A further 4,021 transposon mutants were first screened through the transparency assay, as mentioned in Section 2.6.3.2, and the resulting 54 translucent colonies were later screened by spore assay, as described in Section 2.6.3.1. 29 out of these 54 mutants were observed to have a sporulation or germination deficiency.

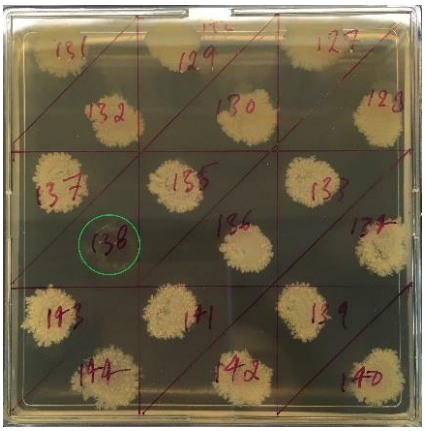


Figure 3.4. **The translucent colony phenotype of a sporulation-deficient mutant**. The spo/ger$^-$ transposon mutant #138_v5 appears translucent among the opaque growths of sporulation-efficient transposon mutants on a 180 hour old TTC Master plate, as indicated by a green circle.

3.2.2. Identifying transposon insertion sites

In order to identify transposon insertion sites in the sporulation-deficient mutants, the protocol from Cartman and Minton 2010 was adopted; however, much lower concentration of genomic DNA template was used for inverse PCR. Briefly, the extracted genomic DNA of sporulation or germination mutants was first fragmented by digesting with the HindIII restriction enzyme, and the DNA fragments were later self-ligated. The DNA sequence at the site of transposon insertion was amplified by inverse PCR using the catP_INV_F1 and catP_INV_R1 primers, as described in Section 2.3.2. In the first transposon library v1.0, multiple amplicons were detected on agarose gel for four of the transposon mutants, as shown in Figure 3.5. This observation may have been a consequence of multiple transposition events or cross-contamination among the transposon mutants. Nevertheless, this problem was solved in the succeeding libraries by reducing the sub-culturing steps of libraries prior plasmid loss, as well as minimizing any chance of cross-contamination during the individual mutant screening. Additionally, sub-culturing helped in purifying two of those four mutants, which eventually gave a single inverse PCR amplicon. Whereas, the other two mutants consistently showed multiple amplicons.

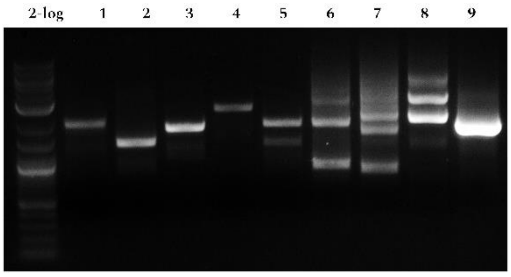


Figure 3.5. **Agarose gel of Inverse PCR amplicons of sporulation-deficient transposon mutants**. Mutants from transposon libraries v1.0, v1.1 and v1.2 showed either single band (lanes: 1-9), or multiple bands (lanes 5-8), beside the 2-log DNA ladder.

In order to identify the transposon insertion sites, DNA amplicons were extracted from agarose gels and sent for Sanger sequencing using the catP_INV_R2 primer. To determine the exact location of transposon insertion, transposon ITR sequences were detected in the sequenced data file and the nucleotide bases immediately next to the ITR1 sequence were matched with those in the *C. sporogenes* wild type genome database. In general, transposon insertion sites were detected in two types of regions in the *C. sporogenes* genome: within a gene and in the (potential) promoter or other intergenic region. Most of the identified affected genes are currently annotated in the NCBI databases as sporulation genes; however, genes involved in other bacterial cellular functions were also observed, which indicates their possible contribution to the sporulation- or germination-deficient phenotype (Table 3.1).

Table 3.1. Transposon insertion sites identified in examples of sporulation- or germination-deficient *C. sporogenes* transposon mutants mentioned in Figure 3.2

TRANSPOSON MUTANTS*	ANNOTATION OF AFFECTED GENE	ANNOTATED CELLULAR FUNCTION
#36	CLSPO_RS00710 (peptidase M23)	Sporulation/Germination
T15	CLSPO_RS13155 (*sigE*)	Sporulation
#138 _v2.0	Intergenic region, possibly CLSPO_RS18585 (*yabP*)'s promoter sequence	Sporulation
T29	CLSPO_RS13170 (*ftsA*)	Cell division
T52	CLSPO_RS12505 (*ftsK*)	DNA translocation
T43	CLSPO_RS16680 (*clpX*)	Protein coding
#24	CLSPO_RS12440 (serine recombinase)	DNA recombination
#90	CLSPO_RS11725 (oxidoreductase)	Metabolism
T19	CLSPO_RS17345 (ATPase)	Unknown
#42	CLSPO_RS07555 (hypothetical protein)	Unknown

* see the complete list of sporulation- or germination-deficient transposon mutants in Appendix Table A2

It was also observed that many transposon mutants had insertions in the same gene, but at different sites. Thus, only 29 unique genes were identified from the 41 sporulation- or germination-deficient mutants. Based on these transposon insertion sites, the potential sporulation or germination genes identified in this study were found to be distributed across the *C. sporogenes* genome (Figure 3.6).

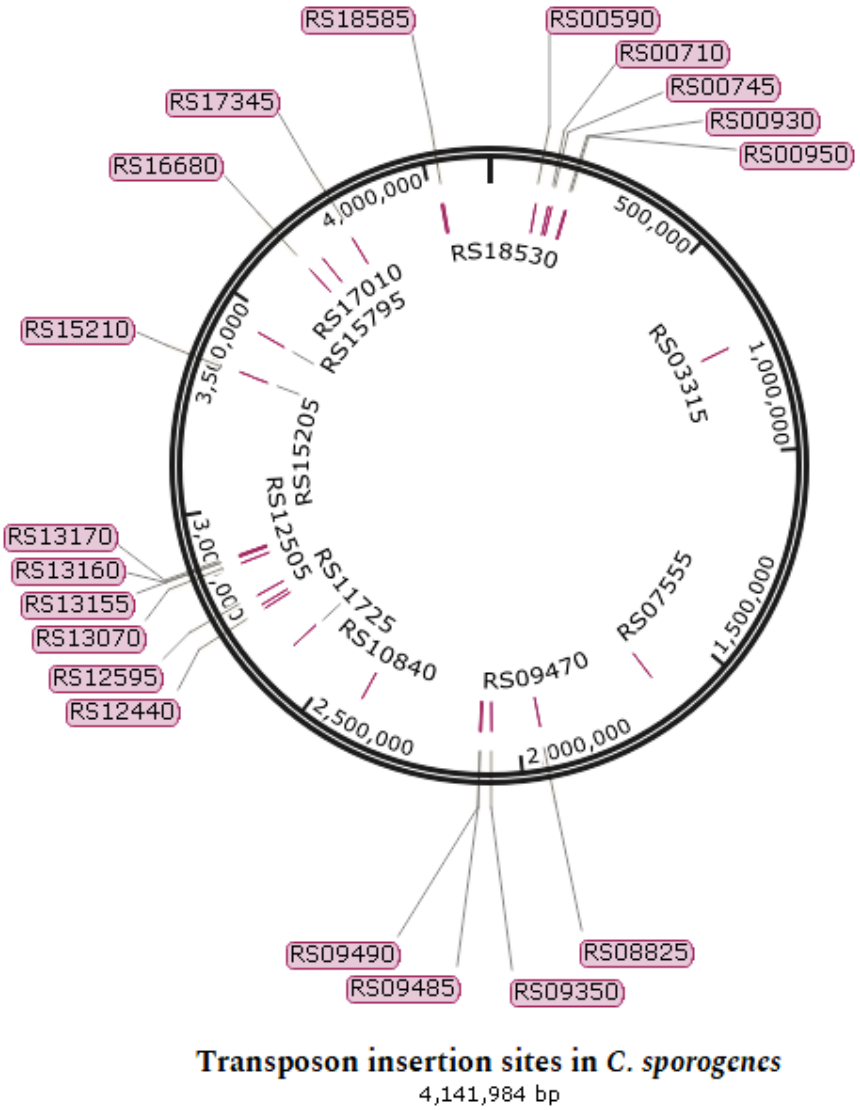


Figure 3.6. **Identification of some of the potential sporulation or germination genes across the *C. sporogenes* NCIMB 10696 genome by transposon mutagenesis**. Affected genes are represented by locus tags (CLSPO_RSxxxxx). This diagram was generated by SnapGene v3.2.

3.2.3. Initial characterization of transposon mutants

Some of the sporulation- or germination-deficient transposon mutants showed other interesting phenotypic differences to the *C. sporogenes* wild type strain. For example, mutant #42 that has transposon insertion in an unknown gene, not only showed a unique colony morphology on TYG agar plates (data not shown), but also a distinctive cell morphology under the microscope (Figure 3.7). Also, the spo/ger$^-$ mutant #64 didn't show any mature spore surrounding the bacterial rods; however, forespore was observed inside some of the mother cells.

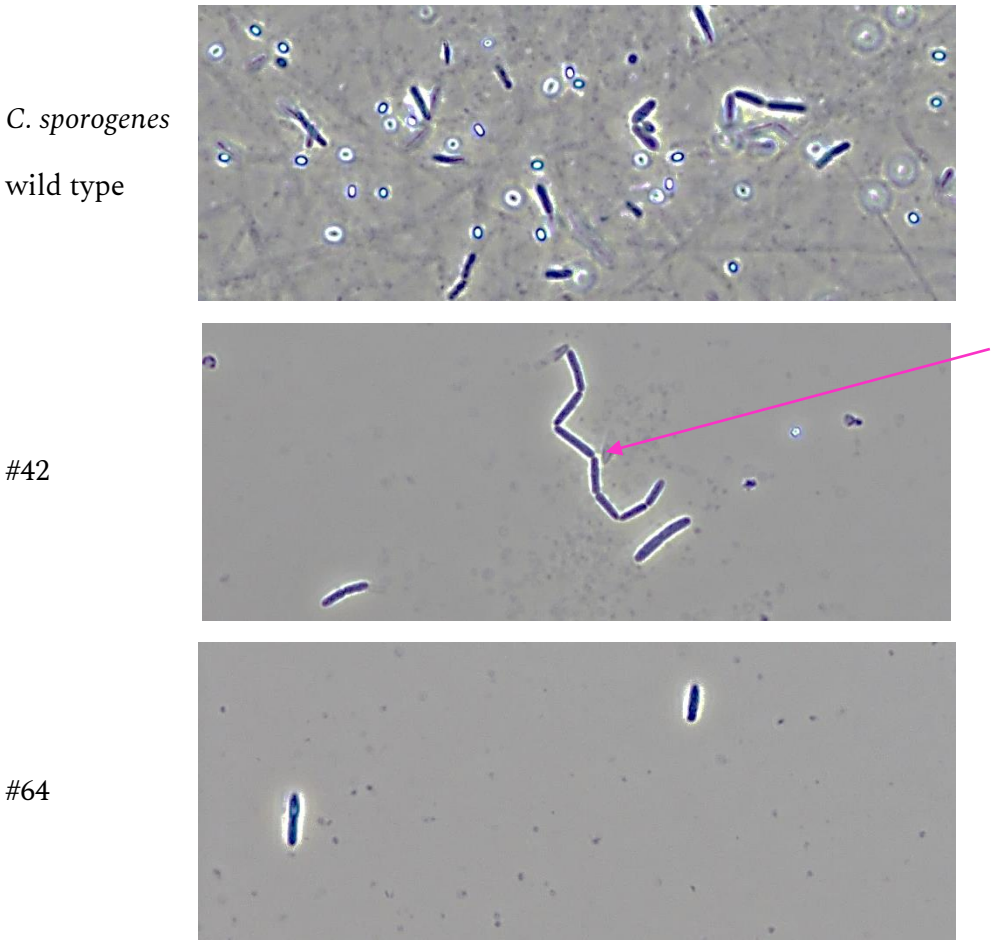


Figure 3.7. **Phase contrast micrographs of *C. sporogenes* transposon mutants**. Sporulation-deficient mutants #42 and #64 show no mature spores, unlike *C. sporogenes* wild type. Moreover, mutant #42 showed peculiar branching phenotype, as indicated by an arrow. The images represent 180 hour old TYG broth cultures of *C. sporogenes* strains at 1000 × magnification.

3.3. DISCUSSION

In this study, transposon mutant libraries in *C. sporogenes* NCIMB 10696 were generated by conditional transposon mutagenesis, and then screened to identify novel sporulation or germination genes. As mentioned earlier, the transcriptional factor TcdR from *C. difficile* was used to control the expression of the transposase *Himar1 C9* gene in the host strain. The *tcdR*-engineered *C. sporogenes* strain sporulates like the wild type strain (Zhang 2015). Therefore, CRG3817 strain, which has the *tcdR* CDS and RBS introduced at its *pyrE* locus, was used to generate transposon libraries.

3.3.1. Selection of transposon plasmid

The pMTL-YZ14 transposon plasmid has been previously verified to work in *C. sporogenes* (Zhang et al., 2015). Naturally, this was initially selected for the generation of mutant libraries in this study. However, the transposon-transposase cassette was highly unstable in *E.coli* CA434 strain, despite being under the tight control of clostridial promoters. In order to avoid the necessity of freshly transforming *E. coli* CA434 competent cells with pMTL-YZ14 before every conjugation attempt, it was decided to use plasmid pMTL-CW20 (Figure 3.1) instead.

Based on pMTL-YZ14, the pMTL-CW20 has the same vector backbone, but also contains the *codA* gene as a negative selection marker, which can be activated upon the rearrangement of the transposon cassette on pMTL-CW20 plasmid in *E. coli*. The *codA* gene encodes for cytosine deaminase, which deaminates 5-fluorocytosine into 5-fluorouracil in *E. coli* to cause conditional cell death (Koprek et al., 1999). In this manner, *E. coli* cells with unstable transposon cassettes are naturally eliminated. Moreover digesting the stable plasmid with EcoRV restriction enzyme results in three fragments of 1,242 bp, 2,303 bp and 6,436 bp (data not shown). This simple stability assay was helpful in the diagnosis of an unstable transposon cassette prior to starting large-scale experiments.

3.3.2. Generation and screening of mutant libraries

Five mutant libraries, v1.0, v2.0, v3.0, v4.0 and v5.0, were generated, which were not only screened directly after preparation, but were also preserved in a pre-induced state and later re-screened. Making fresh libraries not only ensured a higher mutant variety, but also allowed the screening of potentially cold-sensitive mutant strains that would not otherwise survive exposure to extreme freezing temperature.

In the first attempt of screening, mutants from library v1.0 were inoculated in two sets of 96 well plates, containing 200 µl TYG broth, and incubated for two weeks to allow the transposon mutant cells to completely sporulate. Unfortunately, this resulted in the evaporation of broth cultures and the addition of PBS buffer to the evaporated culture plates was required prior to plating. Therefore, 96 well plates were replaced by 48 well plates in the subsequent screening attempts. Furthermore, the transposon mutants were grown in 500 µl selective broths and incubated for one week inside a loosely lidded plastic box containing a little water. Use of a box not only prevented massive evaporation from the wells, but also came in handy for securely transferring the broth culture plates, in and out of the anaerobic cabinet.

In order to find a good screening approach for plating the control and heat-treated bacterial cultures, two techniques were tested. In one approach, the cultures were spotted on agar and later observed for the appearance of colonies. In the second approach, the cultures were inoculated in fresh broths, inoculated overnight and assessed for turbidity or no growth. The former method provided a clearer indication of the degree of sporulation or germination defect, and hence, it was adopted for all the screenings.

The *mariner*-based transposon system targets TA sites in the host genome, which makes it ideal to be used in the low GC *Clostridium* spp. (Cartman and Minton, 2010). Despite being random in the locus preference of transposon insertion, it was found that *mariner* transposon does have some hotspots in the *C. sporogenes* genome, which

are more likely be targeted than other loci. For e.g., several mutants were identified to have unique transposon insertions in *spoIIE* gene and the genes in the *spoIIIA* operon, as observed by spore assay of random individual mutants (data not shown). Preference of certain genome hotspots have been previously documented regarding the use of conjugative transposon Tn916 in *C. difficile* mutagenesis (Hussain et al., 2005).

3.3.3. Challenges of the process and problem-solving techniques

3.3.3.1. Inverse PCR protocol

The first troublesome task was the identification of transposon insertion sites by inverse PCR of the fragmented mutant genome. Previously used protocols (Zhang et al., 2015), (Cartman and Minton, 2010) and in-house ones didn't generate useful PCR products, as indicated on a gel (data not shown). Differences among the commercially used reagents, concentrations of DNA at each molecular technique and duration and temperature might have contributed to the challenges in the reproducibility of the previous protocols. Therefore, the protocol was customized accordingly. Using lower concentrations of DNA template (12.5 or 25 ng), depending on different genomic DNA preparations, greatly solved the problem of getting smears or no amplicon on the agarose gels.

3.3.3.2. Multiple amplicons and cross-contamination

Following the success of inverse PCR for library v1.0, v1.1 and v1.2 transposon mutants, the problem of multiple amplicons became apparent. Ideally, each mutant must only have one affected gene (i.e., one transposon insertion per genome), which is validated by a single inverse PCR amplicon. With more than one amplicon, and each with unique insertion sites; the possibility of multiple transposon insertions within the same genome was considered. This could have been due to the fact that

library v1.1 and v1.2 did undergo an additional sub-culturing step before induction for plasmid loss. With replicative plasmids, an extra sub-culturing step can lead to more than one active transposon cassettes per bacterial cell; and consequently, more than one transposition events.

However, an alternative explanation for this finding was the possibility of cross-contamination among the spo/ger$^-$ and other transposon mutants during the screening process. It was observed that when some spore assays were repeated, the initially screened spo/ger$^-$ mutants (#36 and #25) sporulated. Moreover, pure-culture streaking on agar plates revealed two types of colony morphologies: translucent and opaque. When each type was purified and re-screened, they proved to be different transposon mutants. Conversely, mutants (#12 and #90) which produced multiple amplicons, but showed single colony morphology, could not be purified, despite sub-culturing a single colony three times. Therefore, the subsequent library stocks were preserved in such a way that the frozen culture stock suspends any additional plasmid replication. Besides, mutant cultures were handled carefully throughout the process to prevent any cross-contamination.

3.3.3.3. Labour-intensive screenings

Screening mutant libraries by spore assay included tasks like patch-plating, broth inoculations and plating of the control and heat-treated cultures; all of which demanded substantial amount of labour. To calculate the size of transposon library for identifying all the desired genes in the genome, Poisson formula and Monte Carlo models are used, which are based on the total number of essential and non-essential genes in the host genome (Baym et al., 2016). Also, in one of the previous reverse genetics studies in *Salmonella typhi*, it has been validated that screening a million transposon mutants ensures high chance of transposon insertion in every possible locus in the host genome (Langridge et al., 2009). Hence, in attempts to obtain a representative collection of novel genes for further characterization, it was decided to

screen at least five thousands of mutants. Screening thousands of individual mutants by spore assay, resulted in heavy and time-consuming workload. Therefore, alternate screening methods were considered.

Interestingly, the phenomenon of sporulation- or germination-deficient mutant colonies turning translucent on old plates, allowed the adaptation of the transparency assay for screening for putative sporulation germination mutants prior to spore assay. With this combination, more than four thousand transposon mutants were easily screened for translucent colonies. This greatly reduced the number of mutants from 4,021 to only 54 to be screened by the spore assay protocol. Despite 29 of these 54 translucent transposon mutants showed sporulation or germination defects, only 17 more unique genes could be identified that had not been identified in the previous screenings. The fact that half of the translucent colonies did not show any sporulation defect, suggests that, not all translucent colonies arise due to mutation in the sporulation genes. Likewise, not all sporulation gene mutations result in translucent colonies and one risks the chance to discover other potentially important sporulation genes if such screening methods are solely used.

One explanation as to why some mutant colonies become translucent, comes from an observation during the transparency assay, in which a transposon mutant colony became translucent within 24 hours of incubation and was unable to grow again upon sub-culturing (data not shown). This might be due to the transposon insertion in a gene essential for survival on nutrient media, which lead to cell death. It can be assumed that translucent phenotype reflected a population of non-viable cells among viable cells within the same colony. Based on an assumption, the link of sporulation defects and observed translucent phenotype can be explained by the fact that since sporulation-deficient mutants cannot use sporulation as a means to survive harsh environment, they have higher frequency of cell death under the nutrient-limiting or unfavourable growth conditions, but then on fresh medium or under growth

favourable conditions, the survived cells are able to recover. This translucent phenotype has also been hinted in a previous study (Edwards and Errington, 1997), in which the sporulation-defective mutants could be distinguished from the opaque colonies of *B. subtilis* wild type strain in a mixed population on agar plates. On the other hand, this phenotype may indicate as to how some essential genes are involved in sporulation and vice versa.

3.3.4. Mutants versus Wild type

Throughout the course of this study, the spore count of *C. sporogenes* wild type strain was observed to be roughly in a range of 2.13E+07 – 2.33E+08 CFU/ml, as based on the obtained data of heat-resistant wild type CFUs. On this basis, transposon mutants that gave lower values were designated as sporulation- or germination-deficient. Since the wild type strain has been observed to give variable spore count values in replicate experiments, it was decided to consider <1E+01 difference as likely reflecting an experimental error. Consequently, several mutants that showed lower sporulation or germination efficiencies that fell in the error range were omitted from further studies.

3.3.4.1. spo/ger$^-$ mutants

Among the mutants mentioned in Table 3.1, #36, T15, #138_v2.0, T52, #24, #90 and T19 gave a spo/ger$^-$ phenotype. Mutant #36 was unable to express a protein from the peptidase-M23 family domain, which is responsible for cleaving amide bonds during the formation of spore cortex in *B. subtilis* (Firczuk and Bochtler, 2007). The affected gene in T15 mutant is *sigE*, which is a sigma factor that activates the transcription of stage II sporulation genes in spore formers (Kirk et al., 2014). One gene under *sigE* control is *yabP*, which was also observed to be interrupted in mutant #138_v2.0. The YabP protein is known for its role in spore coat assembly in *B. subtilis* (van Ooij et al., 2004). Another *C. sporogenes* gene from the *B. subtilis sigE* regulon is the homologue of

serine recombinase, which was inactivated in mutant #24, and has also been studied as an important gene for *skin* excision in *C. difficile* (Serrano et al., 2016, Haraldsen and Sonenshein, 2003). The *sigK* in both *B. subtilis* and *C. difficile* is interrupted by a defective prophage called *skin*, which must be excised during sporulation for the σ^K activity (Stragier et al., 1989, Suzuki et al., 2020). It was surprising to see the disruption of DNA translocase gene *ftsK* caused a massive sporulation defect in T52 mutant. The FtsK protein has been known to function in the chromosomal segregation during the cell division of *E. coli* (Liu et al., 1998); and considering the requirement of this process prior to forespore formation, it can be assumed that any abnormalities in DNA translocation can interfere with the progression of sporulation. Interestingly, two more unconventional genes that function in other cellular processes, were also observed to be affected in the spo/ger$^-$ mutants, namely #90 oxidoreductase and T19 ATPase, respectively. This oxidoreductase gene has been previously identified in the original conditional transposon mutagenesis study for its spo/ger$^-$ phenotype in *C. sporogenes* (Zhang et al., 2015). Whereas for T19 mutant, it was challenging to find relevant literature on the basis of the generic annotation representing an ATP-hydrolysing domain in the gene. Regarding the involvement of ATPase in bacterial sporulation, SpoIIIAA ATPase has been known for the activation of σ^G sigma factor in *B. subtilis* (Doan et al., 2009), as well as SpoIIIE ATPase, which is required for transporting DNA to the forespore of *B. subtilis* (Wu and Errington, 1994); (Burton et al., 2007). However, the closest match to this gene was later determined to be *spoVK* of *B. subtilis*, which is possibly responsible for the heat-resistance properties of spores (Foulger and Errington, 1991).

3.3.4.2. Mutants with reduced sporulation

Furthermore, there were some transposon mutants that were able to sporulate, but at lower levels than *C. sporogenes* wild type. The affected genes in this category of mutants are mainly known for their role in biological processes other than

sporulation. One such example is the cell division *ftsA* gene mutation that showed decreased level of heat-resistant CFUs in mutant T29. Interestingly, it has been previously found that mutations in a *B. subtilis ftsA* homologue tend to disrupt the events proceeding stage II sporulation by impairing the Spo0A-based transcription of important sporulation genes (Kemp et al., 2002). In another case, disruption in the *clpX* gene, which encodes an ATP-dependent protease subunit, caused a severe sporulation defect in T43 and T45 mutants. *B. subtilis* Δ*clpX* mutants have been previously observed to be sporulation-deficient (Gerth et al., 1998), because of the irregularity in the σ^H-dependent transcriptional pathway in the absence of ClpX protein (Liu and Zuber, 2000). Moreover, there were several genes detected in the screenings that had unknown functions, but were observed to play some role in sporulation. One such was the RNA-binding S4 domain containing gene that was interrupted in the #42 mutant, which resulted in severe defects in not only sporulation, but also cell division localization.

3.3.5. Selection for mutant characterization

Although the screenings provided a good assortment of potentially interesting genes for further study, and each with their possibly unique roles in sporulation or germination of *C. sporogenes*, only a few of them could be selected for further characterization. The main selection criteria for gene targets was that they must not have been previously studied in any of the *C. sporogenes* strains. Accordingly, some known sporulation genes that have been characterized in *Bacillus* spp., were selected for modelling a good comparison among the endospore-forming bacteria. In addition, it was interesting to know how the currently annotated "hypothetical protein" genes might play a significant role in sporulation or germination of *C. sporogenes*.

Initially, it was decided to study the genes mentioned in Table 3.2. However, some gene targets were later replaced or omitted out from the list during the course of study.

Table 3.2. List of *C. sporogenes* transposon mutants selected for gene characterization

TRANSPOSON MUTANTS	LOCUS TAG	GENE
#64	CLSPO_RS15205	*yqfD*
~~#42; T35~~*	~~CLSPO_RS07555~~	~~*hypP*~~
T43; T45	CLSPO_RS16680	*clpX*
~~#24~~*	~~CLSPO_RS12440~~	~~*serecomb*~~
T19	CLSPO_RS17345	*spoVK*
T4; T26	CLSPO_RS00950	*spoIID*

*Mutants could not be further characterized; CLSPO_RS07555 (*hypP*) was later replaced by CLSPO_RS07560 (*divIVA*); while CLSPO_RS12440 could not be complemented

CHAPTER FOUR

4. CHARACTERIZATION OF *yqfD*

4. CHARACTERIZATION OF *YQFD*

4.1. INTRODUCTION

Bacterial sporulation is a sophisticated process of cellular development to thrive under harsh environmental conditions. This process is executed by a complex regulation cascade, involving several genes that play their own important roles at various stages of the sporogenesis as well as prior its onset. Endospore is originated from a regulated asymmetric division, which is followed by septum formation and engulfment in the mother cell. To ensure the integrity of the emerging spore for future survival under stress, protective layers of peptidoglycan cortex and spore coat are generated around the spore core before it can be released into the environment (Al-Hinai et al., 2015). Mutations in the spore coat or cortex genes can disrupt the first line of defence of bacterial spores against environmental stress.

Isolated from a batch of sporulation-defective mutants, #64 transposon mutant showed the absence of mature, heat-resistant spores, possibly due to the inactivation of CLSPO_RS15205 gene by a *mariner* transposon (Appendix Table A2). Moreover, the microscopic images of the mutant cells showed the presence of forespore in the sporulating mother cells, but no mature spore was observed (Figure 3.7). Hence, the possibility of the very gene to be involved in the progression of one of the later stages of *C. sporogenes* sporulation was considered, possibly in which the spore outer layers are formed.

According to the NCBI genome databases, CLSPO_RS15205 is a 1179 bp long gene in the *C. sporogenes* NCIMB 10696 strain, with GC content of 27%. Based on earlier studies on *Bacillus* spp., its gene product (WP_032884357.1) has been annotated as "sporulation protein YqfD". The YqfD protein in *C. sporogenes* is made up of 392 amino acids. Browsing the neighbouring gene loci, it can be predicted that

CLSPO_RS15205 is the fourth gene in a five genes operon that is operated by a single promoter region. Another known sporulation gene from the *yqfD* operon is the *yqfC* gene (Fawcett et al., 2000, Eichenberger et al., 2003), which was also identified during the transposon library screenings in this study (Appendix A2).

For the ease of explanation, the term "*yqfD*" will be used throughout this thesis to define the CLSPO_RS15205 gene. Figure 4.1 shows some predictions about the conserved domains in the YqfD protein.



Figure 4.1. **Protein domains of YqfD, as predicted by InterPro**. The initial part of YqfD contains the cytoplasmic domain, which is followed by small transmembrane helical domain. The remaining major region is predicted to be non-cytoplasmic. Graphics were edited for a personalized view.

It has been known that *yqfD* is only found in endospore-forming bacteria (Koo et al., 2017, Galperin et al., 2012, Abecasis et al., 2013). Literature review reveals that *yqfD* has been characterized in *Bacillus* spp., whereas not much data is currently available about its homologue in Clostridia. In one such study, the *yqfD* gene was deleted in an attempt to obtain a biologically contained enzyme-producing strain of *B. pumilus* (Wemhoff and Meinhardt, 2013). *B. pumilus* Δ*yqfD* mutants were then characterized for their biotechnologically-accessible phenotype, which included observance of non-functional, phase grey spores under the phase contrast microscope.

In two collaborative studies on *Bacillus megaterium*, a *yqfD* homologue *spoIV* was characterized for its role in sporulation (Meinhardt et al., 1994, Wittchen et al., 1998). The *spoIV* gene was initially identified through a transposon library screening, and

its transposon mutant exhibited non-functional pre-spores that were able to release from their mother cells, but were not able to germinate. In addition, the *spoIV* mutant spores lacked DPA and were found to be sensitive to physical and chemical stress.

Browsing through the literature reviews on stage IV sporulation genes, it can be assumed that majority of the earlier studies have interchangeably used the terms "*spoIV*" and "*yqfD*" to describe genes involved in the formation of spore cortex (Mizuno et al., 1996). Whereas, "*spoIVA*" has been predominantly used to specify the genes involved in spore coat morphogenesis (Roels et al., 1992).

Interestingly, with a query coverage of 64%, *C. sporogenes* YqfD protein shows 29.72% similarity to its protein homologue in *B. subtilis*. So, they are chances of both similarities and contradictions among the members of two genera regarding the exact role of *yqfD*. For instance, *yqfD* has been repeatedly found to be transcribed in the mother cell as a part of the σ^E regulon in *B. pumilus* (Wemhoff and Meinhardt, 2013), *B. megaterium* (Wittchen et al., 1998) and *B. subtilis* (Roels et al., 1992, Eichenberger et al., 2003, Feucht et al., 2003). However, a recent study on the *C. sporogenes* sigma factors revealed YqfD expression to be only slightly downregulated in the *C. sporogenes* Δ*sigF* and Δ*sigE* mutants (Raquel Rodrigues, unpublished), suggesting that *yqfD* might not entirely be controlled by σ^E. It can also be assumed that *yqfD* might be co-regulated by both σ^F and σ^E in *C. sporogenes*. On the other hand, SpoIIID has been reported to downregulate other σ^E–regulated genes, including *yqfD* (Eichenberger et al., 2004).

Surprisingly, the σ^G–activity was found to be downregulated in one of the studies on the *B. subtilis yqfD* mutant (Meeske et al., 2016). Whereas, in another study, both σ^G and σ^K–activities were reported to be unaffected in the *B. subtilis yqfD* mutant (Feucht et al., 2003).

Regarding the exact function of YqfD, it is not entirely known; however, it has been previously hinted that this protein is involved in the success of stage IV of sporulation. In a study on a *yqfD* homologue in *B. subtilis*, mutants of *spoIVA* gene showed impairment of spore cortex as well as disassembly of the spore coat (Roels et al., 1992), suggesting that the normal sporulation process was blocked at stage IV.

With the lack of current data on the exact function of YqfD, it was decided to study the *yqfD* gene of *C. sporogenes* in detail.

4.1.1. Aim of this chapter

This chapter is aimed at studying the CLSPO_RS15205 *yqfD* gene for its potential role in the sporulation of *C. sporogenes*, by the following steps:

i. Constructing in-frame deleted Δ*yqfD* mutants in the *C. sporogenes* NCIMB 10696 wild type strain and later complementing the mutated gene;

ii. Characterizing the Δ*yqfD* mutants by determining the relationship of *yqfD* gene with sporulation, growth, haemolysis, proteolysis, and cell and colony morphology of *C. sporogenes*, and comparing the mutant phenotype with the wild type phenotype;

iii. Finding relevant literature on the prokaryotic *yqfD* gene and its YqfD protein and comparing the phenotypes of the *C. sporogenes* Δ*yqfD* mutants with the ones generated in *Bacillus* and other clostridial species;

iv. Proposing the potential role of *yqfD* in the biology of *C. sporogenes*, based on the data generated in this study, with reference to earlier studies.

4.2. RESULTS

4.2.1. CRISPR/Cas9-based genome editing in the *C. sporogenes yqfD* gene

The first step in characterization of *yqfD* (CLSPO_RS15205) was to generate in-frame deletion in *C. sporogenes*. Therefore, Δ*yqfD* mutants were made in triplicate from the *C. sporogenes* wild type strain using the pRECas1-yqfD knock-out cassette (Figure 4.10) and verified by PCR and DNA sequencing (data not shown). Subsequently, Δ*yqfD* mutant was complemented using pRECas1-YqfDCO knock-in cassette, which generated YqfDCO strain, as described in Section 2.7. Following the induction of Cas9 activity and relevant recombination events, plasmid loss was confirmed prior to any characterization step.

In order to eradicate potential SNPs problem during cloning, the homology arms-independent tool Target-AID was used and *C. sporogenes yqfD*::TA mutants were generated in triplicates from the *C. sporogenes* wild type strain, using pTargetAID-yqfDsgA knock-out cassette, as described in Section 2.8. Based on translucent colony phenotype, *yqfD*::TA mutants were easily isolated from the wild type population on selective TYG-theophylline plates (Figure 4.2). Plasmid-based complementation was done for each *yqfD*::TA mutant using pMTL87121:*yqfD*, which generated strain *yqfD*::TA:ProyqfD, as described in Section 2.10.1. This complementation-plasmid was stably maintained in *C. sporogenes* by thiamphenicol supplementation.

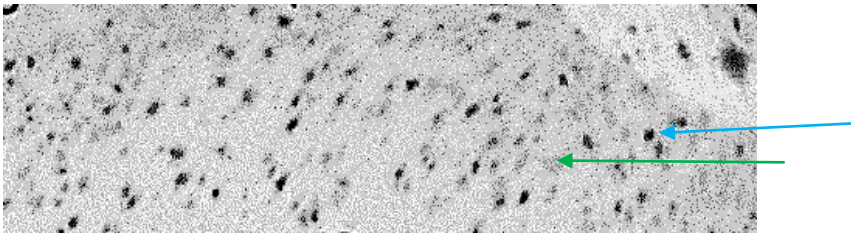


Figure 4.2. **Transconjugant colonies of *C. sporogenes* wild type-pTargetAID-yqfDsgA**. Colonies of *yqfD*::TA mutants appeared faint and translucent (green arrow) among the visible dark colonies of wild type background (blue arrow) on TTC/Theo plate after 48 hour post-conjugation incubation.

4.2.2. Mutations in *yqfD* halts *C. sporogenes* sporulation

For determining the extent of heat-resistant phenotype in *C. sporogenes* strains, relevant cultures were incubated for 120 hours and samples were withdrawn at six time points and heat-treated, as described in Section 2.10.2. Throughout this 5-days sampling, the *spo0A*::CT, Δ*yqfD* and *yqfD*::TA mutants showed a spo/ger$^-$ phenotype, similar to #64, in contrast to the spo/ger+ wild type strain (Figure 4.3).

The YqfDCO strain sporulated to similar levels as the wild type strain, while *yqfD*::TA:ProyqfD strain sporulated slightly less than these two strains. However, the YqfDCO strain not only showed delayed growth, but also delayed sporulation compared to the wild type and *yqfD*::TA:ProyqfD strains. Interestingly, the non-heat treated Δ*yqfD* and *yqfD*::TA mutant cultures showed a higher growth peak than the wild type strain at 24 hours; however, this difference was not apparent by 48 hours. By 120 hours, wild type and YqfDCO strains had higher non-heat treated CFUs than the Δ*yqfD* and *yqfD*::TA mutants.

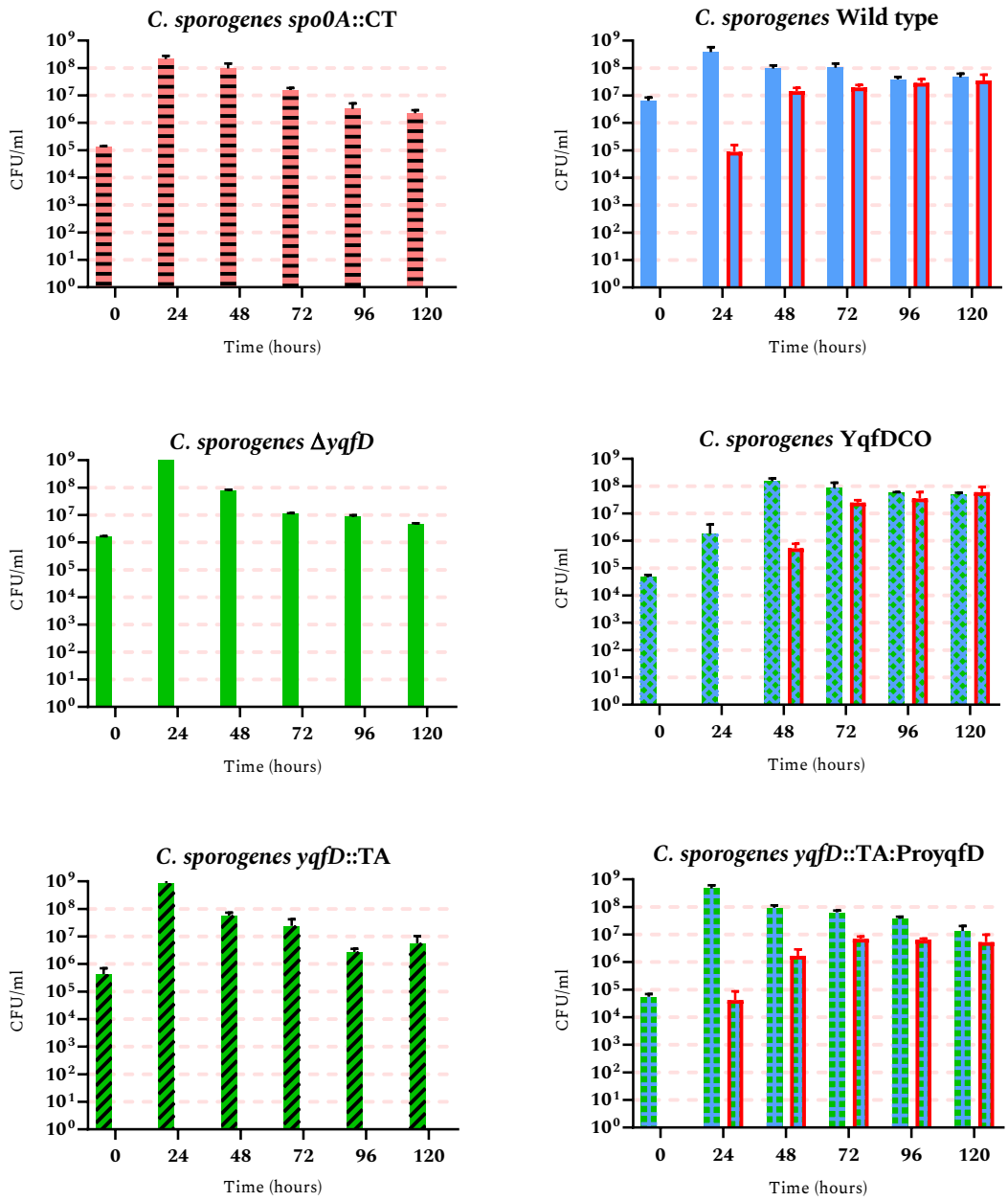


Figure 4.3. **Spore assay of *C. sporogenes yqfD* strains**. Mutants Δ*yqfD* and *yqfD*::TA showed no heat-resistant CFUs, similar to the *spo0A*::CT control strain. Complemented strains YqfDCO and *yqfD*::TA:ProyqfD strains showed comparable heat-resistant CFUs to *C. sporogenes* wild type. This data is based on the 5-day incubation of three biological replicates (except for *spo0A*::CT and YqfDCO) in TYG broth. Error bars show mean with standard deviation. Each strain is represented by a unique colour[1]. Red-bordered bars represent heat-resistant CFUs, whereas coloured bars without red border represent the untreated or the total cell counts. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4×10^9.

[1] Colour key:

Wild type | *spo0A*::CT | Δ*yqfD* | YqfDCO | *yqfD*::TA | *yqfD*::TA:ProyqfD

4.2.3. *yqfD* is not an essential gene in *C. sporogenes*

To determine the influence of *yqfD* mutation on the growth of *C. sporogenes* cells, relevant strains were inoculated in TYG broths and incubated for 24 hours, during which samples were taken at eight time points, as described in Section 2.10.3. In the growth curve of *C. sporogenes* wild type (Figure 4.4), cells were observed to grow exponentially from ~0.01 OD_{600} to a peak cell density of ~3.7 OD_{600} (data not shown) within 6 hours, after which there was stationary phase followed by cell death. All the tested Δ*yqfD* mutants gave a similar growth curve as wild type strain, unlike the slightly defected growth of the *spo0A*::CT mutant. The growth curve of the complemented strain YqfDCO initially appeared to be similar to the wild type. However, when the growth curve was generated using the values in OD_{600} instead, the YqfDCO strain appeared to enter death phase slightly earlier than the wild type strain (data not shown).

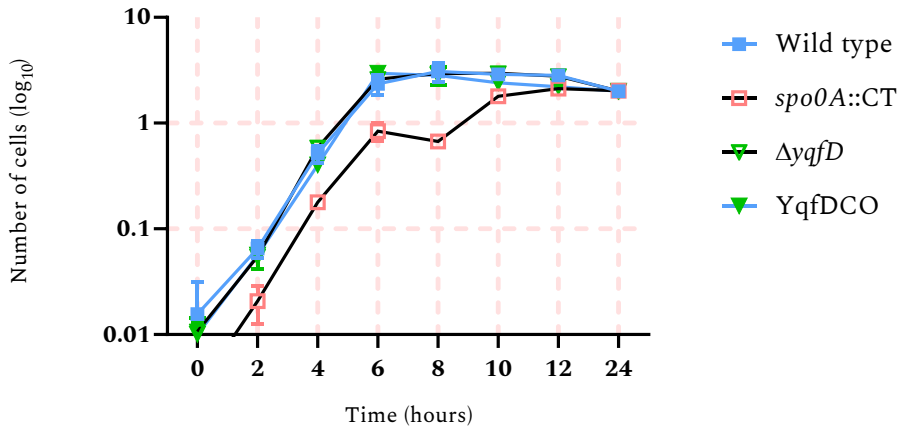


Figure 4.4. **Growth curve of *C. sporogenes yqfD* strains**. Mutant Δ*yqfD* and the complemented YqfDCO strain, showed similar growth patterns to *C. sporogenes* wild type strain, in contrast to the spo/ger$^-$ mutant *spo0A*::CT, which has a slight growth defect. This data is based on the 24 hour incubation of three biological replicates (except for *spo0A*::CT and YqfDCO) in TYG broth. Error bars show mean with standard deviation. Theoretically, 0.01 OD_{600} was the starting inoculum, which was individually calculated for each strain; however, in practice, variable readings were observed at 0 hour.

4.2.4. *C. sporogenes* β-haemolysis does not require *yqfD*

On blood nutrient agar plates, the cultures of *C. sporogenes* wild type strain showed clear zone of β-haemolysis within 48 hours, which was absent on the *spo0A*::CT culture plates (Figure 4.5), identical to a previous study (Maria Zygouropoulou, unpublished). To check if mutation in *yqfD* influenced this haemolytic activity, the cultures of Δ*yqfD* mutants were also spotted on TYG plates, supplemented with horse blood, as described in Section 2.10.4. Wild type level haemolytic activity was observed for Δ*yqfD* mutants. Strangely, no haemolysis was observed for YqfDCO cultures despite gene complementation (Figure 4.5).

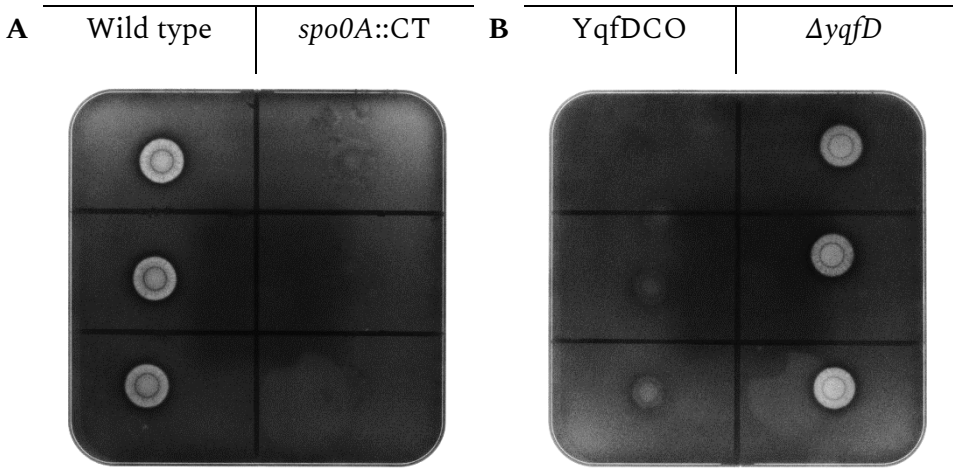


Figure 4.5. **Haemolysis by *C. sporogenes yqfD* strains**. Mutant Δ*yqfD* showed wild type level haemolytic activity, whereas the complemented YqfDCO strain showed very weak β-haemolysis. The spo/ger$^-$ mutant *spo0A*::CT was non-haemolytic and its translucent growth was overshadowed by dark-coloured blood agar. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT and YqfDCO) on TYG agar supplemented with 5% (v/v) horse blood.

4.2.5. Proteolytic activity of *C. sporogenes* is independent of *yqfD*

On milk TYG agar plates, wild type *C. sporogenes* showed proteolytic activity within 24-48 hours and this activity was only slightly reduced by the *spo0A* mutation (Figure 4.6). In order to determine if the mutation in another sporulation gene, like *yqfD* also influences the *C. sporogenes* proteolysis, cultures of Δ*yqfD* mutants were spotted on milk TYG agar plates, as described in Section 2.10.5. Δ*yqfD* mutants showed zones of

proteolysis similar to the *C. sporogenes* wild type strain, while the zones of clearance were slightly smaller in case of the complemented YqfDCO strain (Figure 4.6).

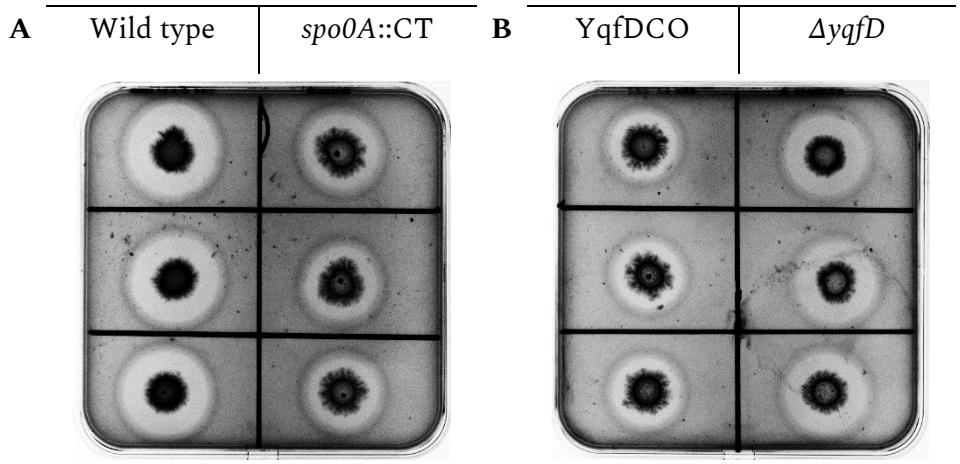


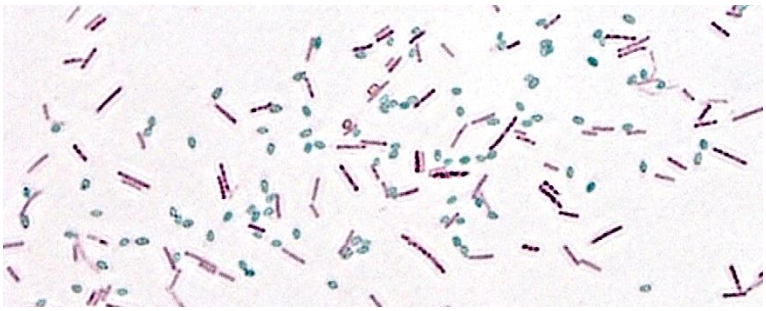
Figure 4.6. **Proteolysis by *C. sporogenes yqfD* strains**. Mutant Δ*yqfD* showed wild type level proteolytic activity, whereas the complemented YqfDCO strain showed comparatively weak proteolysis. The spo/ger$^-$ mutant *spo0A*::CT was also weakly-proteolytic. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT and YqfDCO) on TYG agar supplemented with 4% (v/v) skimmed milk.

4.2.6. Microscopy reveals absence of mature spores of Δ*yqfD* mutants

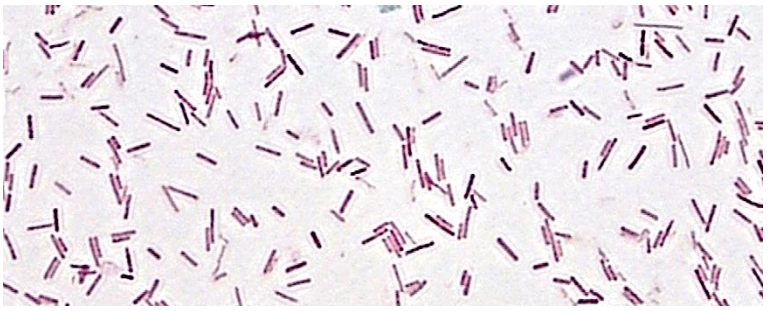
To compare the cell morphologies of *C. sporogenes* strains among the wild type and the genetically-engineered strains, 120 hour old cultures were visualized at high resolutions, as described in Section 2.5.1. *C. sporogenes* wild type strain revealed rod-shaped cells and mature spores, while spores were absent in *spo0A*::CT mutant cultures (Figure 4.7). Although mature spores were also absent in Δ*yqfD* and *yqfD*::TA cultures, forespores that had undergone earlier stages of sporulation were detected inside some of the sporulating mother cells, similar to mutant #64 phenotype (Figure 3.6). In contrast, complemented strains YqfDCO and *yqfD*::TA:ProyqfD cultures showed mature spores.

C. sporogenes

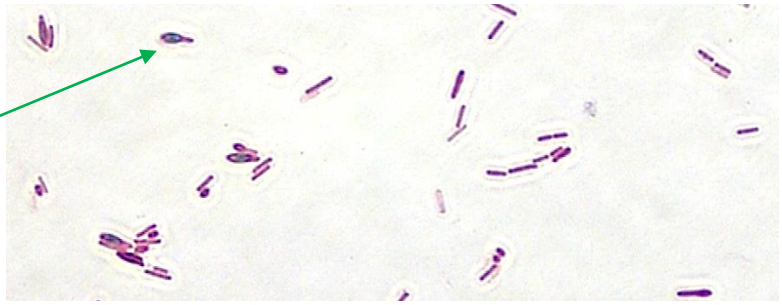
wild type



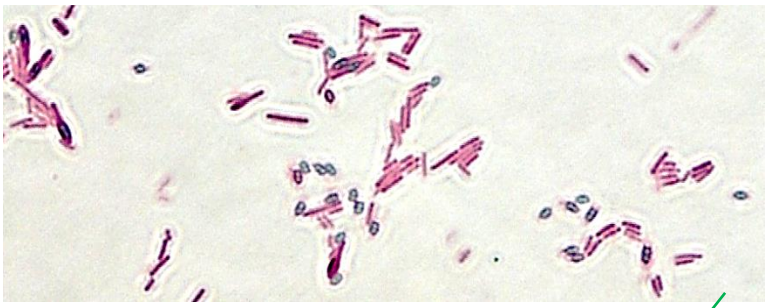
spo0A::CT



Δ*yqfD*



YqfDCO



yqfD::TA



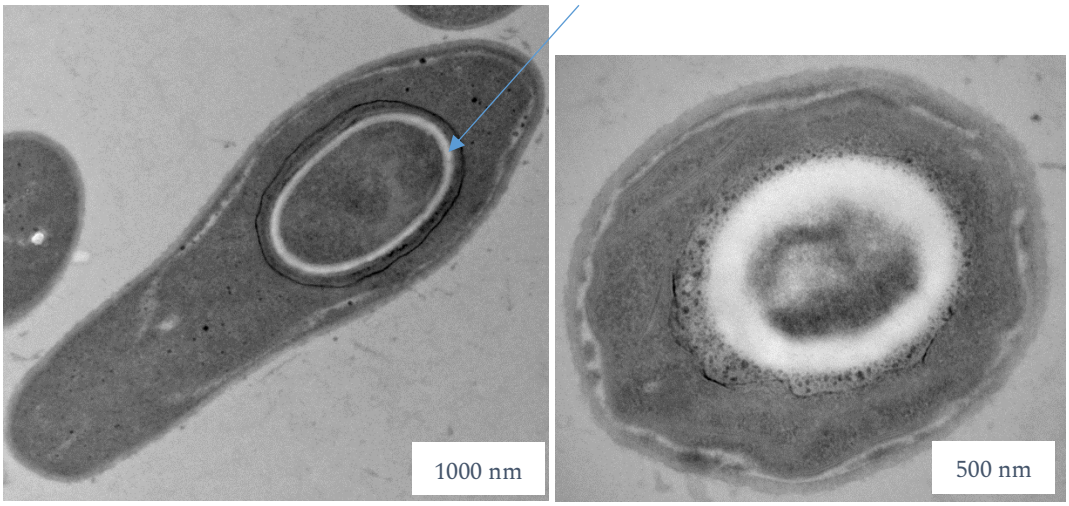
yqfD::TA: ProyqfD



Figure 4.7. **Cell morphology of *C. sporogenes yqfD* strains**. Mutants Δ*yqfD* and *yqfD*::TA showed rods without mature spores, similar to the *spo0A*::CT control strain; however, still displayed forespore in the mother cell, as indicated by an arrow. Complemented strains YqfDCO and *yqfD*::TA:ProyqfD strains showed comparable microscopic phenotype to *C. sporogenes* wild type. Cells were stained pink and spores were stained blue. The images represent 120 hour old TYG broth cultures of *C. sporogenes* strains at 1000 × magnification.

TEM images revealed the detailed view of the *C. sporogenes* wild type and Δ*yqfD* mutant forespore in a sporulating mother cell (Figure 4.8). Since the Δ*yqfD* mutant is unable to produce mature spores, no spores were observed. It can be noted from the micrographs, that the bright and properly defined peptidoglycan cortex of the wild type spore or forespore cannot be clearly observed in the mutant forespore.

C. sporogenes wild type



C. sporogenes Δ*yqfD*

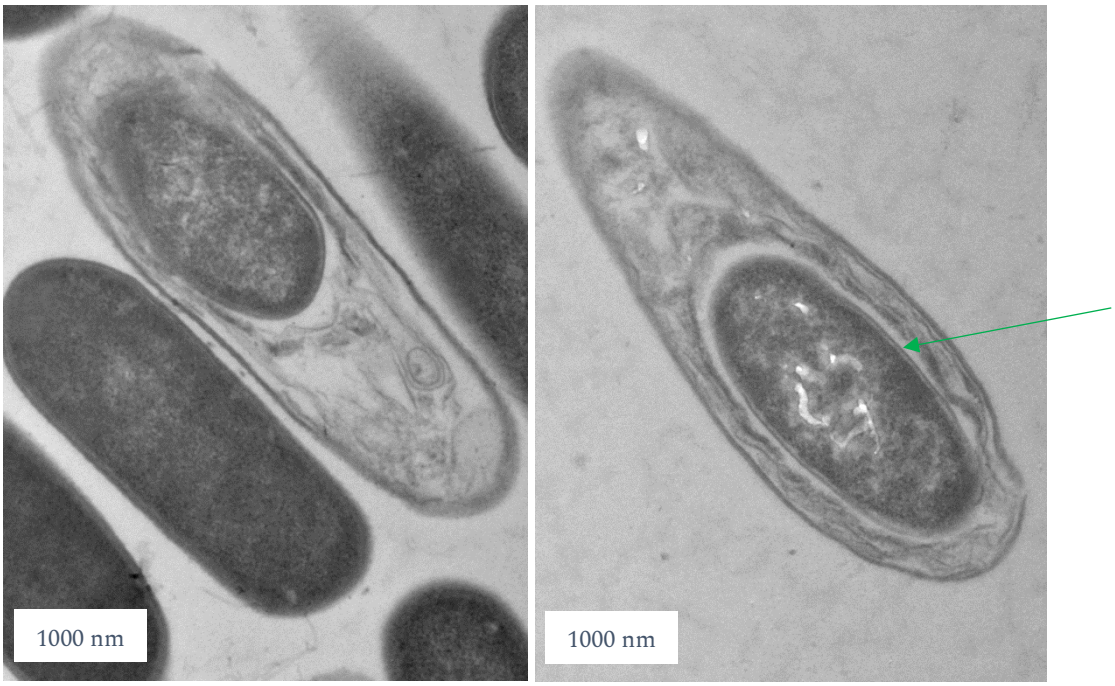


Figure 4.8. **Ultrastructure of *C. sporogenes yqfD* strain**. Forespores of mutant Δ*yqfD* showed the lack of brightly developed spore cortex, which is clearly observed in the wild type strain, as indicated by arrows. The micrographs represent thin section of 72 hour old fixed cultures of *C. sporogenes* strains at high magnifications.

4.2.7. Δ*yqfD* mutant colonies become translucent on old TYG agar plates

C. sporogenes wild type strain produced opaque, cream-coloured colonies on TYG agar plates (Figure 4.9). In contrast, *spo0A*::CT mutant colonies showed an extreme translucent phenotype within 24 hours. To determine if *yqfD* mutation also resulted in translucent colonies, all *C. sporogenes* strains were streaked on TYG agar and compared, as described in Section 2.10.6. It was found that Δ*yqfD* mutant colonies also become translucent; however, it took 48 hours for this mutant to show this phenotype (data not shown). Furthermore, this translucent phenotype was not complemented in the YqfDCO strain and the colonies turned translucent within 27 hours, earlier than the Δ*yqfD* mutant.



Wild type | *spo0A*::CT | Δ*yqfD* | YqfDCO

Figure 4.9. **Colony morphology of *C. sporogenes yqfD* strains**. Mutant colonies of *C. sporogenes* Δ*yqfD* showed translucent phenotype on old agar plates, unlike the opaque growth of wild type, and this unstable growth phenotype was not restored in the complemented YqfDCO strain. However, mutant colonies do appeared smooth, shiny, fimbriated and irregular, just like those of the *C. sporogenes* wild type strain. In contrast, *spo0A*::CT mutant showed highly fimbriated and translucent phenotype within 24 hours. The images were taken after 27 hour incubation of *C. sporogenes* strains on TYG agar.

4.2.8. Illumina sequencing reveals SNPs in the Δ*yqfD* mutant genome

In order to determine the possible presence of spontaneous mutations in addition to that generated by a targeted genome editing, which could have affected sporulation, the Δ*yqfD* mutant was sent for Illumina sequencing, as described in Section 2.3.8. When aligned with the *C. sporogenes* NCIMB 10696 reference strain from the NCBI genome databases, the sequenced data of Δ*yqfD* mutant revealed the potential existence of three more SNPs in addition to the seventeen SNPs that are present in the in-house *C. sporogenes* wild type strain (Appendix Table A3). Among those three

additional SNPs, one was identified in a gene CDS that encodes for the PocR ligand-binding protein; the second SNP was in the gene for LysR transcriptional regulator; and the third one was found in an unknown gene.

4.3. DISCUSSION

In this study, attempts were made to demonstrate the role of *yqfD* in the sporulation process of *C. sporogenes*. As mentioned earlier, the transposon mutant #64 showed the spo/ger$^-$ phenotype because of the interruption of *yqfD* gene with transposon insertion. So, in order to confirm this sporulation phenotype, it was decided to make an in-frame deletion mutant using CRISPR/Cas9-based genome editing tools. Initially, the approach of using "one guide RNA to target all the affected genes in transposon mutants" was adopted. Although this approach could not generate clean mutants, it would still confirm the reproducibility of the sporulation phenotype using a master guide RNA. In this manner, catPsgRNA3 guide RNA was selected for guiding Cas9 nuclease to target the *catP* sequence present in all of the transposon mutants. Mutant #64 was used as a host to delete the *yqfD* gene. As the resulting CRISPR/Cas9 mutant had previously been engineered with ACE (to insert the TcdR CDS) and transposon mutagenesis, it was suggested to check any possibility of spontaneous mutation by sending the final Δ*yqfD*_old mutant's genomic DNA for whole genome sequencing. When the sequenced data was analysed, a SNP appeared at the locus CLSPO_RS01850 (*resE* histidine kinase), which has already been found to be a common target for spontaneous mutations in the in-house engineered *C. sporogenes* strains. Interestingly, this gene has also been identified as an important sporulation gene during a different series of mutant screenings (Rachel White, unpublished). This spontaneous mutation may have occurred due to the repeated sub-culturing during various steps in the three protocols of mutagenesis. Thus, it was decided to make new, cleaner mutants from the wild type strain using the individual gene's target-specific guide RNA.

However, in practice, genome editing in *C. sporogenes* for the *yqfD* gene was not straight forward and many hurdles had to be overcome during mutant generation or mutant gene-complementation.

4.3.1. The problems of cloning knock-out/in cassettes in *E. coli*

The initial *yqfD*-knock-out cassette consisted of catPsgRNA3 and 2 kb homology arms insert (missing the *yqfD* gene). Both inserts were ligated into a previously constructed pMTL83151-CRISPR/Cas9 plasmid, in which the *Cas9* gene was placed under the control of thiolase promoter from *C. acetobutylicum*. However, this plasmid construct (pCRISPR-yqfD) was observed to be toxic in *E. coli* TOP10 competent cells at 30°C and no transformants could be obtained. Replacing the high copy number plasmid replicon ColE1 with the low copy number replicon p15a solved this problem and subsequently, *C. sporogenes* Δ*yqfD*_old mutants were generated in the #64 strain via pCRISPR-yqfD1 plasmid. However, as mentioned above, a SNP was observed in the Δ*yqfD* genome and therefore, attempts were made to re-make the Δ*yqfD*_old mutants in the wild type *C. sporogenes* background.

Although it reduced the potential toxicity of the ligated insert, the use of p15a-based plasmid resulted in problems with inefficient plasmid digestion, i.e., a high number of background (undigested vector backbone) colonies were observed on the *E. coli* transformant plates. Consequently, large numbers of transformant colonies had to be screened by PCR in order to identify the transformants with the desired knock-out cassette. Although no scientific evidence could be found to explain any link between low copy number plasmids and insufficient digestion, multiple colleagues reported experiencing this strange phenomenon. Nevertheless, after several attempts, it was found that digesting and dephosphorylating the plasmids overnight and then repeating both processes the next day, increased the efficiency of plasmid digestion. Thus, plasmids were digested twice for the later cloning attempts.

In the second *yqfD*-knock-out cassette, catPsgRNA3 was replaced by either of the three identified yqfDsgRNAs, and subsequently, homology arms were designed to include a bookmark of catPsgRNA3 sequence, PAM sequence and an extra base. This incorporation of bookmark sequences in between the truncated target gene would provide an effective Cas9 target site for future gene complementation of the engineered mutant. The resultant pCRISPR-yqfD2 cassette was ligated into the pMTL83121-CRISPR/Cas9 (with p15a replicon) plasmid. However, toxicity in previously competent *E. coli* TOP10 cells was re-observed and the transformants could not be obtained. Interestingly, in a further attempt at cloning the *yqfD*-knock-out cassette in *E. coli*, in which a partially digested CRISPR/Cas9 plasmid preparation was accidently used, it was found that using a different guide RNA, i.e., botsgRNA1 rather than any of the three yqfDsgRNAs, allowed the success of obtaining some *E. coli* transformants. Therefore, other protocol customizations were considered to improve efficiency of the mutagenesis protocol.

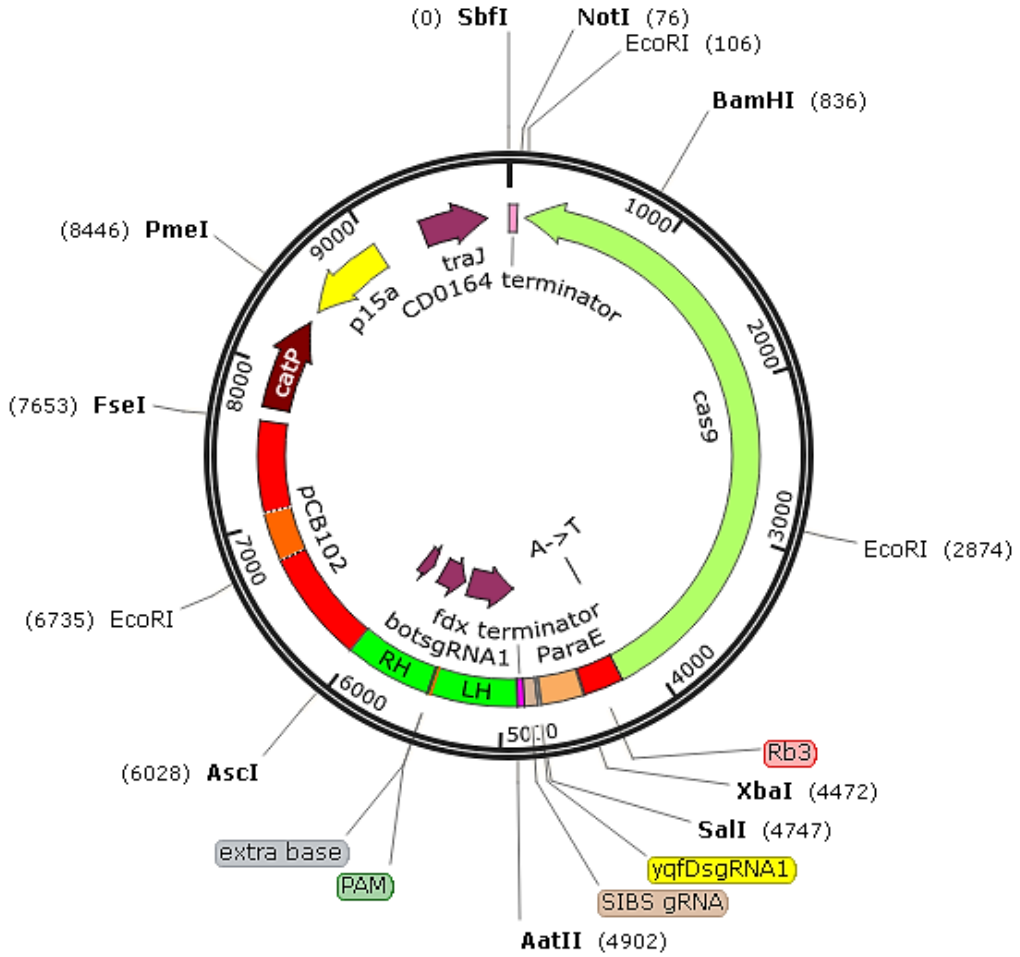


Figure 4.10 Plasmid map of pRECas1 with the *yqfD*-knock-out cassette. Key features: CD0164 transcriptional terminator from *C. difficile*; *cas9* for DSB; Rb3 riboswitch for Cas9 expression; P_{araE} promoter from *C. acetobutylicum* for guide RNA expression; yqfDsgRNA1 and SIBS sgRNA scaffold; fdx transcriptional terminator from *C. pasteurianum*; botsgRNA1-bookmark-based homology arm for homologous recombination; pCB102 Gram positive replicon; *catP* plasmid marker; p15a Gram negative replicon; and *traJ* conjugation gene.

Firstly, a new CRISPR/Cas9 plasmid pRECas1 was adopted (Figure 4.10) that has Cas9 under the tight control of a riboswitch Rb3 (Canadas et al., 2019), which finally resulted in some *E. coli* transformant colonies. However, the sequenced data of those transformants revealed SNPs in the homology arm cassette, which were either deletions, insertions or nucleotide base alterations. The homology arms were therefore shortened from 1 kb per arm to 0.54 kb per arm; the knock-out cassette was also ligated in the reversed orientation compared to the previous version; competent *E. coli* TOP10 cells were replaced by competent *E. coli* 5-alpha F' Iq cells; and the

inoculated plates were incubated at room temperature, following the transformation and recovery of competent cells. Eventually, verified transformants were obtained and *C. sporogenes* Δ*yqfD* mutants were generated from the wild type strain.

4.3.2. Inefficient CRISPR/Cas9 homology-directed repair in *C. sporogenes*

For complementing transposon mutant #64, the first *yqfD*-knock-in cassette was designed to include catPsgRNA3 to guide Cas9 to target the transposon, as well as the 2 kb homology arms cassette (including the *yqfD* CDS); and both of which were ligated into the original pMTL83121-CRISPR/Cas9 plasmid. This plasmid construct was used to transform *E. coli* TOP10 and later *E. coli* CA434 competent cells, which were conjugated with *C. sporogenes* #64 overnight. The two resultant transconjugants were screened by PCR (using yqfdoLH_F and yqfdoRH_R primers), to detect the occurrence of recombination between plasmid and host genome by homology-directed repair or HDR. However, only Δ*yqfD*_old mutant amplicon was observed after gel electrophoresis. The transconjugants were sub-cultured on fresh TYG plates and finally one complemented strain *C. sporogenes* YqfDcomp was obtained. Unfortunately, this strain failed to lose the plasmid, despite repeated sub-culturing in the absence of plasmid antibiotic marker; and therefore, could not be used for characterization.

In order to complement the catPsgRNA3-bookmark-based *C. sporogenes* Δ*yqfD*_catP mutants, the second *yqfD*-knock-in cassette was designed to include catPsgRNA3 to target the previously inserted bookmark, as well as the 1.1 kb watermark-based homology arms (that contained the *yqfD* CDS, but with two altered Isoleucine codons), which were ligated into the new pMTL83121-pRECas1 plasmid. After cloning this plasmid construct in *E. coli* 5-alpha F' Iq and *E. coli* sExpress strains, it was transferred to *C. sporogenes* Δ*yqfD*_catP mutants. However, PCR analysis showed that the mutants had failed to complement after the induction of Cas9 with theophylline. The

transconjugants were then sub-cultured on selective TYG-theophylline plates multiple times, until one verified watermark-based *C. sporogenes* YqfDCO strain was obtained. The YqfDCO strain regained the wild type sporulation phenotype, but showed a translucent phenotype, along with the loss of haemolysis. These findings raised the suspicion of the possibility of SNPs in the YqfDCO genome and re-generation of the strain was considered.

Initially, it was assumed that inefficiency of catPsgRNA3 to guide Cas9 to the target locus in Δ*yqfD*_catP mutant was the reason why the HDR in *C. sporogenes* mutants was not efficient. Therefore, the catPsgRNA3 bookmark was replaced by the botsgRNA1 bookmark and Δ*yqfD* mutants were re-generated. The next *yqfD*-knock-in cassette was designed to include botsgRNA1 and a different 1.1 kb watermark-based homology arm insert (including the *yqfD* CDS, but with an altered STOP codon). Both inserts were ligated into the pRECas1 plasmid and cloned in *E. coli* strains. However, no verified *C. sporogenes* YqfDCOs strain was obtained, despite several attempts. As no other option was available, the previously generated solo YqfDCO strain was therefore selected for further characterization, alongside the botsgRNA1-bookmark-based *C. sporogenes* Δ*yqfD* mutants.

4.3.3. The success of Target-AID base editing

While the problem of SNPs in cloning in *E. coli* was being sorted out, it was decided to consider alternate tools for forward genetic studies that are independent of HDR. Based on CRISPR/Cas9 system, Target-AID is one of the promising genome editing tools primarily used to generate mutants with ease in prokaryotes (Nishida et al., 2016); (Tanaka et al., 2018). Therefore, an in-house Target-AID design tool was used to determine the best guide RNA for inserting a premature STOP codon in the *C. sporogenes yqfD* gene, and yqfDsgA guide RNA was then ligated into pTargetAID plasmid. This pTargetAID-yqfD plasmid construct was easily cloned in *E. coli* strains,

and *C. sporogenes yqfD*::TA mutants were easily generated, following the conjugation with *C. sporogenes* wild type strain, along with the immediate induction of Cas9 activity. Mutants were verified by recognizing the introduction of a STOP codon in the start of the *yqfD* gene sequence that was obtained from Sanger sequencing data of the corresponding PCR amplicons. As expected, sporulation, cell and colony morphology phenotypes of the resultant *yqfD*::TA mutants matched with those for mutants #64 and Δ*yqfD*.

In order to complement this mutation, pMTL87121:*yqfD* was transferred into Δ*yqfD* mutant strains and was constantly maintained in the cells using thiamphenicol as the plasmid selection marker. The presence of *yqfD* CDS, RBS and promoter on the plasmid, potentially compensated the YqfD production, which was absent in *yqfD*::TA mutants and the observed phenotypes of complemented strain *yqfD*::TA:ProyqfD were similar to *C. sporogenes* wild type, while dissimilar to spo/ger$^-$ #64 and Δ*yqfD* mutants. However, the maintenance of plasmid in mutant cells throughout the 5-day spore assay was challenging because of the instability of thiamphenicol during prolonged incubation at 37°C. Consequently, *yqfD*::TA:ProyqfD cultures were re-supplemented with antibiotic after 48 hours. In practice, the plasmid loss was evidently observed for one replicate, which may have affected the observed heat-resistant counts.

4.3.4. The unexpected phenotype of *yqfD* complementation

Apart from the compensation of sporulation-efficiency, the phenotypes of *C. sporogenes* YqfDCO strain appeared to differ from those of the wild type strain. Unexpectedly, YqfDCO strain showed: i) slightly early decline phase, ii) delay in sporulation, iii) translucent colonies, and iv) loss of haemolysis. Since Δ*yqfD*_catP mutants, from which the YqfDCO strain was generated, grew well and showed haemolytic activity, suggests that the presence of other spontaneous mutations in the

YqfDCO genome were being reflected in these phenotypes. While the whole genome sequencing of YqfDCO could not be performed due to time constraints, repeating the haemolysis and proteolysis assays with *yqfD*::TA:ProyqfD strain revealed wild type level lytic properties (data not shown). Since YqfD expression was shown to be possible on a plasmid in the *yqfD*::TA:ProyqfD strain, it could serve as an equivalent of YqfDCO strain. Moreover, this strain underwent fewer sub-culturing steps compared to the YqfDCO strain, which may have lessen the chances of any SNPs being introduced in the genome. Due to time constraints, these strains could not be sent for whole genome sequencing.

4.3.5. SNPs in the Δ*yqfD* mutant

The sequenced data of one Δ*yqfD* mutant revealed the presence of SNPs in three additional loci of *C. sporogenes* mutant genome (Appendix Table A3). One of the affected gene encodes for a hypothetical protein, which lacks any conserved domain and also could not be identified through the NCBI BLAST tool. Another SNP was found in the gene for a PocR ligand-binding protein, which is known for binding to hydrocarbon derivatives and sensing principle substrates for bacterial growth (Anantharaman and Aravind, 2005). The other affected gene was identified to encode for the LysR transcriptional regulator protein, which is commonly known to control various bacterial processes including, virulence, quorum sensing, cell division, motility, nitrogen fixation, toxin production and secretion (Maddocks and Oyston, 2008, Lu et al., 2007). However, none of the affected genes have been previously identified to play a role in sporulation. Surprisingly, these three SNPs were also identified in the sequenced Δ*clpX* (Appendix Table A5) and Δ*spoVK* mutant strains (Appendix Table A6), which might suggest the occurrence of spontaneous mutations during the overnight growth of *C. sporogenes* mutant cultures prior to DNA extraction. It might also reveal different sets of SNPs within the populations of same bacterial frozen stock, which could not be selected for the extraction of genomic DNA.

4.3.6. Characterization of *yqfD* and its role in sporulation

The role of *yqfD* gene in sporulation of *C. sporogenes* was revealed by the inactivation or deletion of the gene, which inhibited sporulation, as represented by the spo/ger$^-$ phenotype on TYG plates (Figure 4.2). The rod-shaped cells of the mutants appeared similar to wild type, under the microscope (Figure 4.7). However, no mature spores were observed. Instead, immature spores or forespores were detected in some mother cells. Expression of YqfD in mutant cells compensated the loss of wild type sporulation/germination phenotypes.

The homologue of *C. sporogenes yqfD* has been characterized in *B. pumilus* in attempts to generate a spo/ger$^-$ strain suitable for industrial production of extracellular enzymes (Wemhoff and Meinhardt, 2013). Considering from an industrial point of view, endospore formation can be seen as bothersome when it comes to the biotechnological workhorse bacteria that are employed for mass production of valuable biochemicals, like enzymes and anti-microbial compounds. Therefore, such genetic studies are required for the development of more efficient and industrially-friendly strains. Nevertheless, both *B. pumilus* Δ*yqfD* and *C. sporogenes* Δ*yqfD* mutants have been observed to produce immature spores, without affecting the growth of their host strains. Interestingly, in a previous transposon mutagenesis study that identified sporulation genes in *B. subtilis* (Le Breton et al., 2006), a homologue equivalent of mutant #64, mutant Spo3 was isolated, which also showed the spo/ger$^-$ phenotype due to the transposon insertion in the *yqfD* gene. A potential *yqfD* gene homologue *spoIV* has also been proposed to be involved in the sporulation cascade of *C. difficile* (Pettit et al., 2014). Additionally, another *spoIV* homologue has been identified as a result of transposon-mediated mutagenesis in *B. megaterium*; however, the mutant phenotype also showed the release of non-functional, immature pre-spores from the mother cells, but those pre-spores were unable to germinate (Meinhardt et al., 1994, Wittchen et al., 1998). Moreover, no DPA production was observed during sporulation and the

pre-spores were observed to be sensitive to heat, lysozyme and chloroform. In this study, no pre-spores were observed to be released from the mother cells, while no work on DPA quantification could be done to compare both phenotypes.

C. sporogenes Δ*yqfD* mutants in this study showed a sporulation frequency of 0%, compared to the wild type strain. Similarly, *yqfD* mutants in *B. subtilis* have been reported to show sporulation efficiencies ranging from 0 to < 0.1 % in various studies (Eichenberger et al., 2003, Feucht et al., 2003, Meeske et al., 2016). The slight difference in various studies might be due to the variability in the detection limits of sporulation assay protocols.

On the other hand, growth of mutants in TYG broth was similar to that of the wild type strain during the first 48 hours incubation, which matched the phenotype of a previously studied *B. subtilis yqfD* mutant (Wittchen et al., 1998). Since Δ*yqfD* mutants are incapable of producing functional spores, unlike the wild type strain, some cell lysis was observed after 48 hours, as revealed by reduced CFUs in Figure 4.3. This suggests that *yqfD* is not directly involved in the vegetative growth of *C. sporogenes* and the observed phenotype was indeed a consequence of a sporulation defect.

However, during the extraction of genomic DNA from all the Δ*yqfD* mutant replicates for whole genome sequencing, the genomic DNA was observed to be extremely sensitive to degradation. After several attempts at Phenol/Chloroform extraction, it was found that only a fresh (4 hours old) *C. sporogenes* Δ*yqfD* culture gave stable genomic DNA preparation. The reason behind this phenomenon is still not clear, but may be linked to an increased susceptibility to cell lysis in this mutant.

C. sporogenes is a strongly proteolytic organism (Allison and Macfarlane, 1990) and has also been previously known to show β-haemolysis on sheep blood agar (Poehlein et al., 2015). Since inactivation of *spo0A*, the master regulator of sporulation, results in loss of haemolytic activity (Figure 4.5), it was interesting to know if deletion of other

sporulation genes affect haemolysis or even proteolysis. But as revealed by the wild type level haemolytic (Figures 4.5) and proteolytic activities (Figure 4.6) of Δ*yqfD* mutants on blood agar and milk agar plates, respectively, it is safe to assume that *yqfD* does not play any role in the ability of *C. sporogenes* to lyse blood cells or ferment amino acids.

4.3.7. Proposed function of *yqfD* in *C. sporogenes*

Further characterization of the *yqfD* gene could not be done using transcriptomics and high specification microscopic tools due to time constraints, but the data generated in this study does support the theory about importance of *yqfD* in the sporulation process of *C. sporogenes*. Also, BLASTP analysis revealed YqfD to be annotated as a sporulation protein in all of the spore-forming bacteria, while the exact function of YqfD has not been clearly identified yet. Since slightly phase-bright, i.e., well-developed forespores were detected in the mother cells of Δ*yqfD* mutants, but no functional spores were observed, suggests that sporulation process was still initiated in the absence of YqfD. Moreover, as understood from the previous literature on sporulation stages in endospore-forming Firmicutes, it can be suggested that sporulation proceeded in Δ*yqfD* mutants, at least until stage III, in which the engulfment of forespore occurs. This observation also matches with the earlier proposed theories in previous studies, which regard to *yqfD* homologues as one of the stage IV sporulation genes.

CHAPTER FIVE

5. CHARACTERIZATION OF *divIVA* and *hypP*

5. CHARACTERIZATION OF *DIVIVA* AND *HYPP*

5.1. INTRODUCTION

Earlier studies have shown that the induction of sporulation requires the starvation stimulus to be detected during DNA replication that is occurring in a nutrient-limiting environment, otherwise the bacterial cell enters another cycle of symmetric division (Dawes et al., 1971); (Veening et al., 2009). Moreover, it has been previously shown that mutants that don't undergo chromosomal replication, are also unable to sporulate (Mandelstam et al., 1971). Hence, the occurrence of asymmetric division is designated as stage I of sporulation. Studies have shown that RacA protein in *B. subtilis* facilitates this asymmetrical division at the cell pole during sporulation, where RacA is sequestered to the poles by the cell division localization protein DivIVA (Ben-Yehuda et al., 2003). However, no RacA homologue has ever been found in clostridial species.

Transposon mutagenesis screening assays with *C. sporogenes* libraries delivered several sporulation-defective mutants, among which transposon mutant #42 revealed a consistently reduced sporulation phenotype on TYG plates (Figure 3.3), and peculiar cell division-defective morphology under the microscope (Figure 3.7), as briefly mentioned in Chapter three of this thesis. Sequence analysis of the transposon insertion site revealed the possible inactivation of the CLSPO_RS07555 gene. On the assumption that the interrupted gene might be involved in the sporulation or germination of *C. sporogenes*, further studies were considered.

5.1.1. The *hypP* gene

According to the NCBI genome databases, CLSPO_RS07555 is a small 771 bp gene in the *C. sporogenes* NCIMB 10696 strain, with 26% GC content. Its gene product (WP_033059171.1) is annotated as "hypothetical protein", which is made up of 256 amino acids. Browsing the neighbouring gene loci, it can be predicted that CLSPO_RS07555 is the fourth gene in a five-gene operon that is operated by a single promoter region. For the ease of explanation, the term "*hypP*" (from hypothetical protein) will be used throughout this thesis to define the CLSPO_RS07555 gene.

The identification of a novel sporulation gene was an interesting finding, but without a functional annotation, it is difficult to find homologues for better understanding the role of an unknown gene; especially if the matched proteins from the BLAST results are also annotated as hypothetical protein. Nevertheless, the protein sequence of *hypP* gene product was analysed with BLASTP and some of the results revealed the similarity of the *C. sporogenes* hypothetical protein to the S4 RNA-binding protein of *Bacillus* spp., based on the presence of the S4 domain in the protein sequence.

RNA-binding proteins are generally multifunctional proteins responsible for binding with 16S rRNA and regulating the translation process. Some RNA-binding domains are known to bind to enzyme domain sites and deliver enzymes to their site of action present on nucleic acids (Aravind and Koonin, 1999). The role of S4 RNA-binding proteins has not been previously described in the context of bacterial sporulation or germination.

5.1.2. The *divIVA* gene

As mentioned in the earlier section, the *hypP* operon has five genes. Directly downstream the *hypP* gene, CLSPO_RS07560 is another small 660 bp gene in the genome of *C. sporogenes* NCIMB 10696 strain, having 26% GC content. Its gene product (WP_003491073.1) is annotated as "cell division protein DivIVA", which is

made up of 219 amino acids. CLSPO_RS07560 is the fifth and the last gene in the *hypP* operon. All the remaining genes of this operon are unknown: their gene products are annotated as hypothetical proteins at the NCBI genome database, and none have been previously documented for characterization.

For the ease of explanation, the term "*divIVA*" will be used throughout this thesis to define the *C. sporogenes* CLSPO_RS07560 gene.

Widely-conserved in Gram positive prokaryotes, *divIVA* gene has a position of being a well-characterized gene in various bacterial species for its versatile role in the cell division localization, virulence as well as sporulation (Kaval and Halbedel, 2012).

Studies have revealed that during vegetative cell division, DivIVA protein gets recruited to the cell division site where it binds to the MinD protein, an ATPase of the cell division inhibitor MinC (de Boer et al., 1991, Jacobs and Shapiro, 1999, Marston and Errington, 1999). In *B. subtilis*, *DivIVA* binds to MinD through a cell division topology factor MinJ (Patrick and Kearns, 2008); however, no such interaction has been observed in *C. beijerinckii* (Valencikova et al., 2018). Anchoring of DivIVA with the MinCD complex allows the localization of the cell division inhibitor to the cell poles; thereby, allowing the formation of the FtsZ-ring only at mid-cell (Marston et al., 1998, Hu and Lutkenhaus, 1999, Thomaides et al., 2001). During the onset of sporulation, DivIVA protein displaces from the cell pole-localized MinCD complex and attracts the chromosome-anchoring protein RacA protein, which allows Z-ring and asymmetric septum formation at the cell pole, along with the efficient trapping of DNA into the forespore of *Bacillus* spp. (Ben-Yehuda et al., 2003). Previous literature also hints the existence of the DivIVA-SpoIIE complex near polar septum for the successful activation of compartment-specific σ^F and eventually the progression of sporulation (Eswaramoorthy et al., 2014).

Since DivIVA regulates the site of cell division, it also maintains the length of rod shaped bacteria. Deletion of the cell division localization protein results in the arise of daughter cells of unequal sizes (Mendelson and Cole, 1972). If the cell division takes place at the cell poles during the vegetative growth of *E. coli* and *B. subtilis*, tiny spherical daughter cells called minicells are formed, most of which lack parental DNA (Adler et al., 1967, Reeve et al., 1973). Although minicells are incapable of cell division, they can still respire (Adler et al., 1967) and act as recipients during bacterial conjugation (Cohen et al., 1968).

The *divIVA* promoter remains active throughout the vegetative growth of *B. subtilis* and DivIVA is constitutively expressed until the late stationary phase (Edwards and Errington, 1997). The *divIVA* gene has been known in the literature as a part of the *B. subtilis spo0A* regulon, in which the protein is actually downregulated after the onset of sporulation (Vishnoi et al., 2013), perhaps to allow Z-ring formation and asymmetric septum at the cell pole. However, the *spo0A* regulon in *C. difficile* has not been found to be involved in the regulation of *divIVA* expression (Pettit et al., 2014). Moreover, a transcriptomics study revealed the positive regulation of DivIVA, MinC, MinD, MinE and FtsZ by the *C. difficile sigH* regulon (Saujet et al., 2013). Additionally, DivIVA expression was also not observed to be controlled by any of the four important sporulation-specific sigma factors σ^F, σ^E, σ^G and σ^K in a recent study on *C. sporogenes* sporulation (Raquel Rodrigues, unpublished).

Apart from its function in the chromosomal segregation of Gram positive bacteria, *divIVA* gene has also been observed to be involved in the virulence of *Listeria monocytogenes*, due to its potential role in biofilm formation and autolysin secretion (Halbedel et al., 2012). The overexpression of *divIVA* has also been observed in the antibiotic-resistant biofilms of *Enterococcus faecalis* (Sharma and Khan, 2018).

Additionally, it is known that DivIVA of *B. subtilis* directly interacts with lipid bilayer of the cytoplasmic membrane and preferably accumulates to the most strongly curved areas of the membrane (Lenarcic et al., 2009).

Nevertheless, the function of *hypP* and *divIVA* has not been previously characterized in *C. sporogenes* or any other clostridial species. Therefore, experiments were designed to study their role in depth.

5.1.3. Aim of this chapter

This chapter is aimed at studying the CLSPO_RS07555 *hypP* and CLSPO_RS07560 *divIVA* genes for their potential role in the sporulation of *C. sporogenes*, by the following steps:

i. Constructing in-frame deleted Δ*hypP* and Δ*divIVA* mutants in the *C. sporogenes* NCIMB 10696 wild type strain and later complementing the mutated genes;

ii. Characterizing the Δ*hypP* and Δ*divIVA* mutants by determining the relationship of *hypP* and *divIVA* genes with sporulation, growth, haemolysis, proteolysis, and cell and colony morphology of *C. sporogenes*, and comparing the mutant phenotype with the wild type phenotype;

iii. Finding relevant literature on the prokaryotic *hypP* and *divIVA* genes and their respective protein products and comparing the phenotypes of the *C. sporogenes* Δ*hypP* and Δ*divIVA* mutants with the ones generated in *Bacillus* and other clostridial species;

iv. Proposing the potential role of *hypP* and *divIVA* in the biology of *C. sporogenes*, based on the data generated in this study, with reference to earlier studies.

5.2. RESULTS

Initially, *hypP* was intended for characterization; however it was speculated that the observed sporulation phenotype of the #42 or *hypP* transposon mutant might be a consequence of polar effect of transposon insertion on neighbouring genes of the same operon. Therefore, the discovery of the *divIVA* gene changed the focus of research from *hypP* to *divIVA*.

5.2.1. Some characterization of *hypP*

As mentioned in Section 3.2.3, the initial characterization of mutant #42 (with the transposon-interrupted *hypP* gene) showed interesting phenotypes such as, reduced sporulation and cell division defects (Figures 5.1 and 5.4), and multiple morphologies of feathery or fimbriated colonies on TYG plates that turned translucent within 48 hours (data not shown). The fact that the gene product of *hypP* is annotated as "hypothetical protein", made it interesting to know how a potentially uncharacterized gene could play roles in various cellular processes.

5.2.1.1. CRISPR/Cas9-based genome editing in the *C. sporogenes hypP* gene

C. sporogenes Δ*hypP* mutants were successfully and readily generated from the *C. sporogenes* wild type strain using pCRISPR-hypP knock-out cassette, as described in Section 2.7. Mutants were the used for further characterization.

Conversely, characterization assays of the two verified Δ*hypP* mutants revealed phenotypes dissimilar to mutant #42, but instead similar to the *C. sporogenes* wild type strain. Moreover, heat-resistant CFUs from the two spore assays gave ~1×10 higher value than the wild type strain (Figure 5.4). Similarly, numerous spores were observed under the phase-contrast microscope, unlike mutant #42 that showed a very small number of spores (Figure 5.1). Furthermore, the colony morphology was the same as wild type (data not shown).

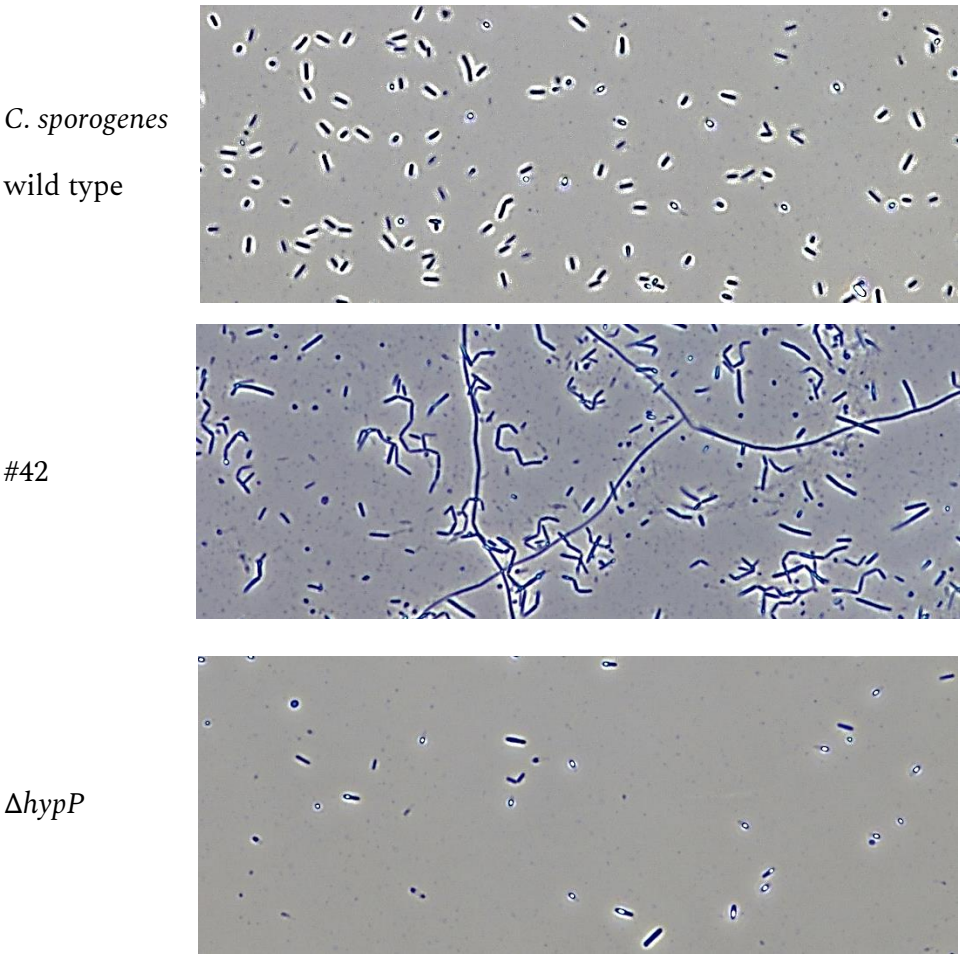


Figure 5.1. **Phase contrast micrographs of *C. sporogenes hypP* strains**. Mutant Δ*hypP* showed the microscopic morphology of *C. sporogenes* wild type strain, unlike its counterpart transposon mutant #42. Numerous spores were observed and the rods were sufficiently branched in Δ*hypP*. The images represent 180 hour old TYG broth cultures of *C. sporogenes* strains at 400 × magnification.

Upon the investigation of the *hypP* locus and its neighbouring genes, it was found that *hypP* is actually the fourth gene in a five genes operon (Figure 5.2). Since the insertion of transposon causes an out-of-frame mutation, polar effects of transposon insertion in mutant #42 were considered. Published literature on the known homologues of genes in the *C. sporogenes hypP* operon was reviewed and the mutant phenotypes of #42 were compared with those in the previous studies. It was then revealed that *Bacillus divIVA* mutants from the previous literature (Perry and Edwards, 2006); (Edwards and Errington, 1997); (Muchova et al., 2002); (Thomaides et al., 2001) matched the phenotype of *C. sporogenes* #42 mutant.

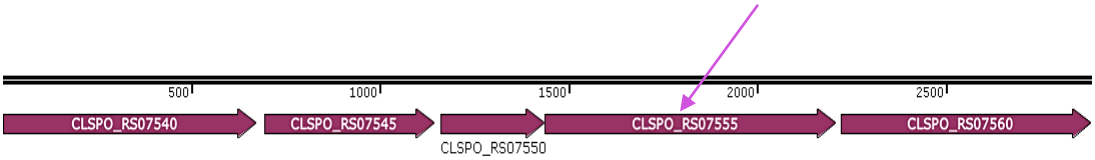


Figure 5.2. **The *divIVA* operon**. There are five genes in the *divIVA* operon, with only one known gene CLSPO_RS07560 *divIVA*. The *hypP* gene at the locus CLSPO_RS07555 is the fourth gene, as indicated by an arrow. Polar effects of *hypP* gene interruption are likely on the downstream gene *divIVA*.

5.2.1.2. Examining polar effects of transposon insertion on *divIVA*

Interestingly, *divIVA* is the fifth gene in the *C. sporogenes hypP* operon, directly downstream of *hypP* gene (Figure 5.2), which suggested the possibility of polar effects of transposon insertion on the expression of DivIVA proteins.

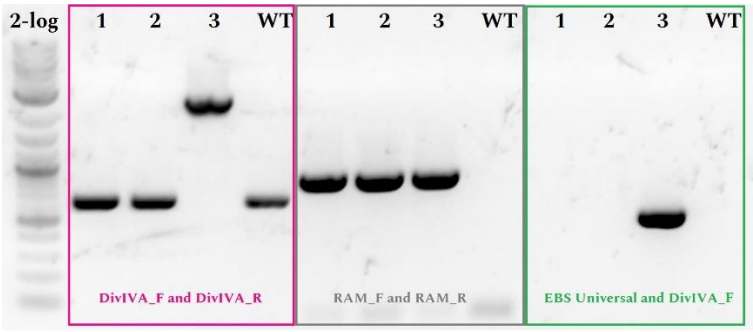


Figure 5.3. **Agarose gel of *divIVA* ClosTron mutant**. Following conjugation and sub-culturing, erythromycin-resistant colonies 1, 2 and 3 were screened by colony PCR to identify the correct *divIVA*::CT mutants in *C. sporogenes*. *C. sporogenes* wild type (WT) strain was used as a control. Three sets of primers (Appendix Table A1) were used: i) DivIVA_F/R flanking primers (pink box), with which, colonies 1 and 2 showed ~0.7 kb *divIVA* amplicon, similar to wild type gene, while colony 3 showed an activated RAM insertion in *divIVA* gene; ii) RAM_F/R primers (grey box), with which, all colonies showed RAM insertion, except wild type strain; and iii) EBS Universal and DivIVA_F primers (green box), with which, only colony 3 was generated an amplicon. Thus, non-specific RAM insertion was observed in colonies 1 and 2, whereas, colony 3 was a true *divIVA*::CT mutant.

Therefore, to investigate this theory, *C. sporogenes divIVA*::CT mutant was generated (Figure 5.3) using the ClosTron-based plasmid pMTL007::*Csp-divIVA*-273s, as described in Section 2.9. The resulting mutant phenotypes were compared with mutants #42 and Δ*hypP*. As suspected, heat-resistant CFUs of the *divIVA*::CT mutant were observed to be similar to that of the transposon mutant #42 (Figure 5.4).

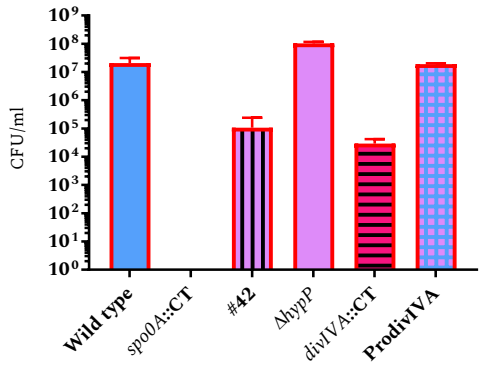
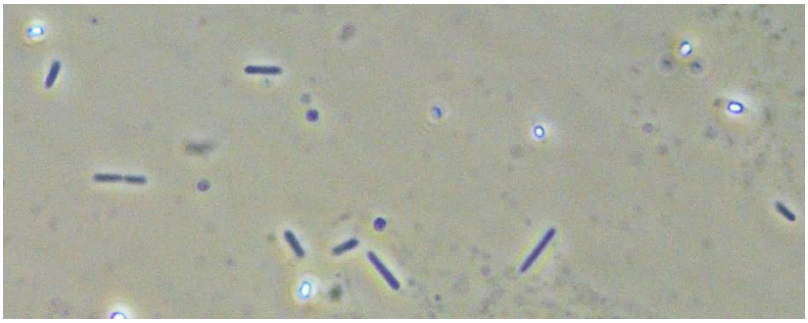


Figure 5.4. **Heat-resistant CFUs of *C. sporogenes hypP* and *divIVA* strains**. In-frame deleted mutant Δ*hypP* showed sporulation-efficient phenotype after heat-shock, unlike its counterpart transposon mutant #42. In contrast, mutant *divIVA*::CT showed the sporulation-deficient phenotype of #42, dissimilar to both Δ*hypP* and wild type strains. Moreover, the complementation of #42 with the *divIVA* CDS restored the sporulation efficiency in the generated ProdivIVA strains. Sporulation-efficient *C. sporogenes* wild type and spo/ger⁻ *spo0A*::CT strains were used as a control. This data is based on 120 hour incubation of two to three biological replicates (except for #42 and *spo0A*::CT) in TYG broth. Error bars show mean with standard deviation. Red-bordered bars represent heat-resistant CFUs. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4×10^9.

Additionally, attempts were made to express DivIVA in mutant #42 using the pMTL82251:*divIVA* plasmid, as described in Section 2.10.1, which complemented the sporulation-defective phenotype of mutant #42 in the generated *C. sporogenes* ProdivIVA strain, as shown in Figures 5.4. Copious amounts of ProdivIVA spores were observed under the microscope, which not only appeared slightly larger than the wild type spores, but also had a shadow of their respective mother cell (Figure 5.5).

C. sporogenes wild type



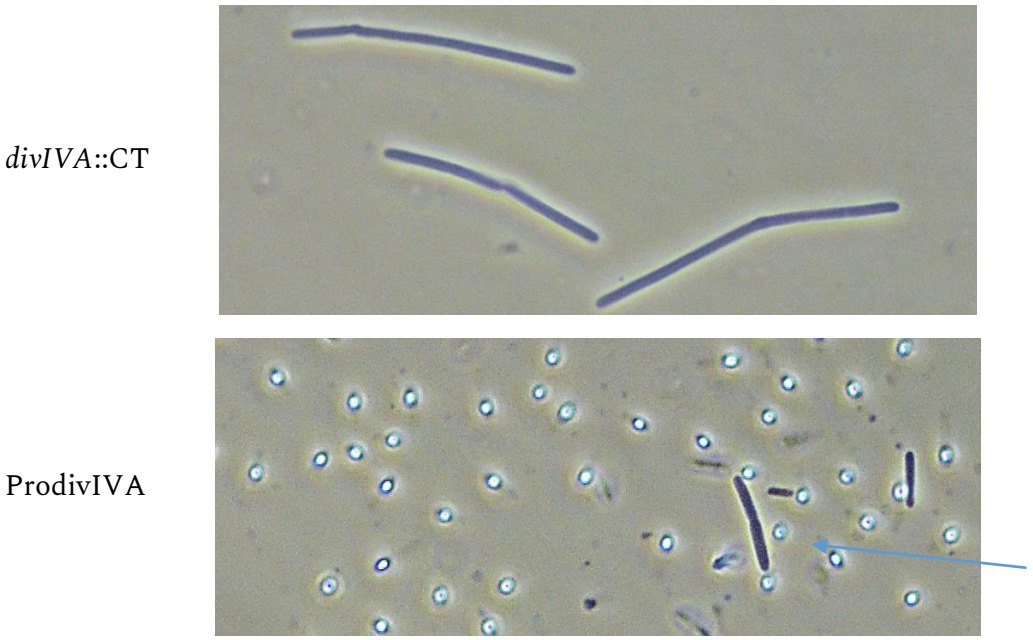


Figure 5.5. **Phase contrast micrograph of *C. sporogenes* ProdivIVA strain**. Complementing #42 *hypP* gene mutant with *divIVA* CDS restored wild type sporulation efficiency in the resulting *C. sporogenes* ProdivIVA strains. However, spores in ProdivIVA strain appeared larger compared to the wild type spores. Moreover, mother cell shadows were strangely observed behind phase-bright spores, as indicated by an arrow. The images represent 120 hour old TYG broth cultures of *C. sporogenes* strains at 1000 × magnification.

This perceived influence of *divIVA* on the sporulation and cell division cascade of *C. sporogenes* made it a promising target gene for further characterization.

5.2.2. CRISPR/Cas9-based genome editing in the *C. sporogenes divIVA* gene

Eventually, *C. sporogenes* Δ*divIVA* mutants were generated in triplicates from the wild type strain using the pRECas1-divIVA knock-out cassette and all three Δ*divIVA* mutants were later complemented using the watermark-based pRECas1-DivCO knock-in cassette to generate complemented *C. sporogenes* DivCO strains, as described in Section 2.7. All strains were used for characterization assays after the confirmation of plasmid loss.

5.2.3. Mutation in *divIVA* alleviates *C. sporogenes* sporulation

C. sporogenes strains were grown for five days in TYG broth and samples were daily heat-treated to compare their heat-resistance profiles, as described in Section 2.10.2. Unlike the spo/ger⁻ mutant *spo0A*::CT, Δ*divIVA* mutants showed heat-resistant CFUs (Figure 5.6). However, the observed titre was ~1×10³ lower than that of the *C. sporogenes* wild type strain. Moreover, Δ*divIVA* sporulation was delayed by 24 hours. This sporulation defect was observed to be repaired in the DivCO strain (Figure 5.6).

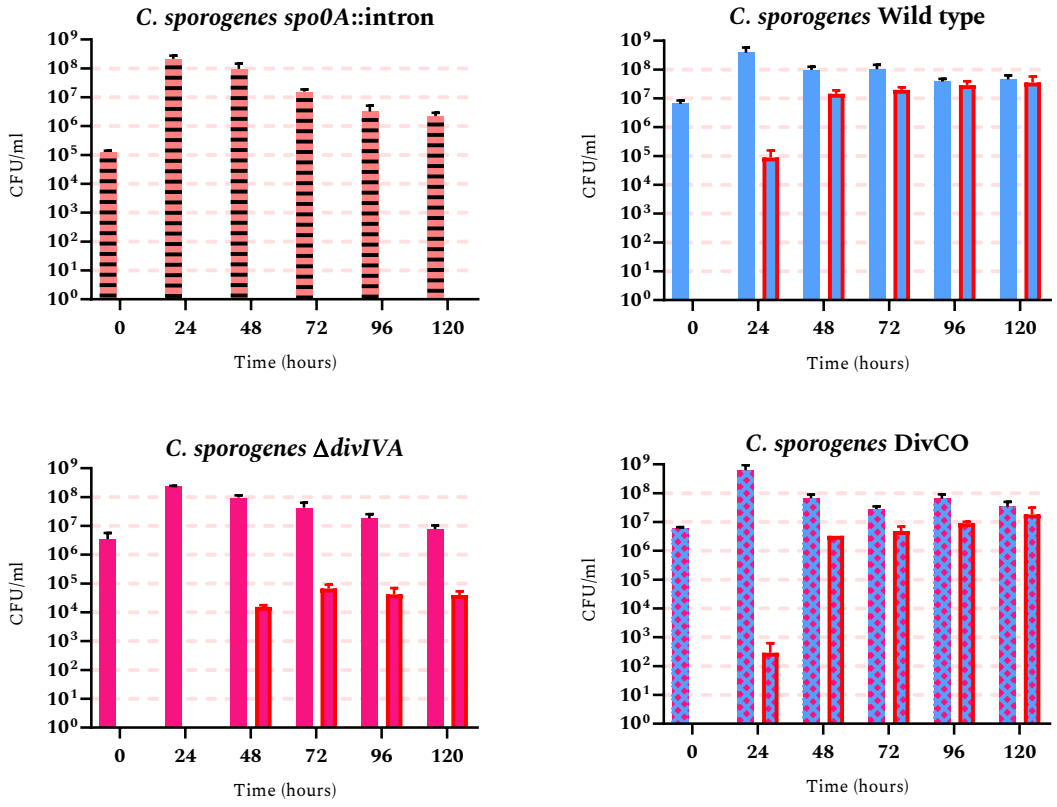


Figure 5.6. **Spore assay of *C. sporogenes divIVA* strains**. Mutant Δ*divIVA* showed ~1×10³ times lower heat-resistant CFUs compared to *C. sporogenes* wild type strain, while the complemented strain DivCO showed similar CFUs as wild type. This data is based on the 5-day incubation of three biological replicates in TYG broth. Error bars show mean with standard deviation. Each strain is represented by a unique colour[2]. Red-bordered bars represent heat-resistant CFUs, whereas coloured bars without red border represent the untreated or the total cell counts. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4 × 10⁹.

[2] Colour key:



5.2.4. *divIVA* is not an essential gene in *C. sporogenes*

The growth curve of *C. sporogenes* strains was determined by growing them in TYG broths and measuring the optical density at various time points over 24 hours, as described in Section 2.10.3. As briefly mentioned in Section 4.2.3, the *C. sporogenes* wild type strain grew to an exponential peak of ~3 OD_{600} (data not shown) before the cell density declined, unlike mutant *spo0A*::CT, which always show a growth defect. Although no growth defects were detected for Δ*divIVA* mutants and complemented DivCO strains (Figure 5.7), it was also observed that these strains took more time to reach the wild type peak cell density, suggesting that the slower growth rate phenotype in Δ*divIVA* mutants was partially complemented in the DivCO strains.

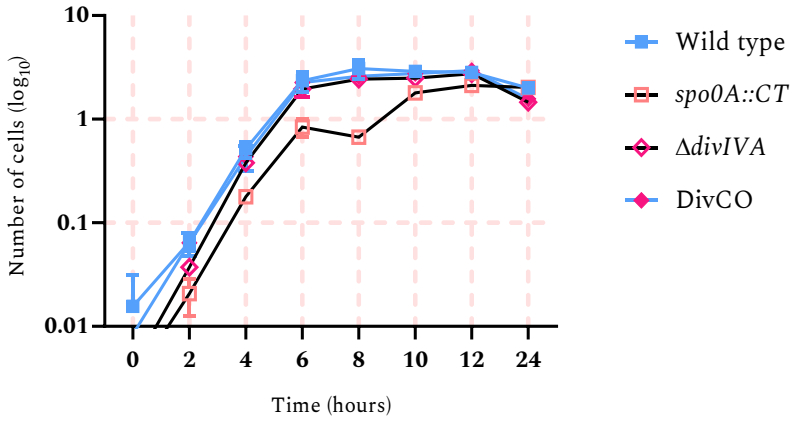


Figure 5.7. **Growth curve of *C. sporogenes divIVA* strains**. Mutant Δ*divIVA* and the complemented DivCO strain, showed similar growth patterns to *C. sporogenes* wild type strain, in contrast to the spo/ger$^-$ mutant *spo0A*::CT, which has a slight growth defect. This data is based on the 24 hour incubation of three biological replicates (except for *spo0A*::CT) in TYG broth. Error bars show mean with standard deviation. Theoretically, 0.01 OD_{600} was the starting inoculum, which was individually calculated for each strain; however, in practice, variable readings were observed at 0 hour.

5.2.5. *C. sporogenes* β-haemolysis does not require *divIVA*

On blood TYG agar plates, cultures of *C. sporogenes* wild type strain showed β-haemolysis, unlike *spo0A*::CT mutant cultures (Figure 5.8). Therefore, the effect of *divIVA* mutation on *C. sporogenes* haemolysis was also investigated on blood agar plates, as described in Section 2.10.4. It was found that mutant Δ*divIVA* retained this wild type haemolytic characteristic (Figure 5.8). Although the complemented strain DivCO was also haemolytic, it appeared to produce smaller zones of β-haemolysis than the highly fimbriated Δ*divIVA* mutant, possibly because of its non-fimbriated colony morphology after the CRISPR-Cas9 gene knock-in, which is further described in Sections 5.2.7 and 5.3.2.

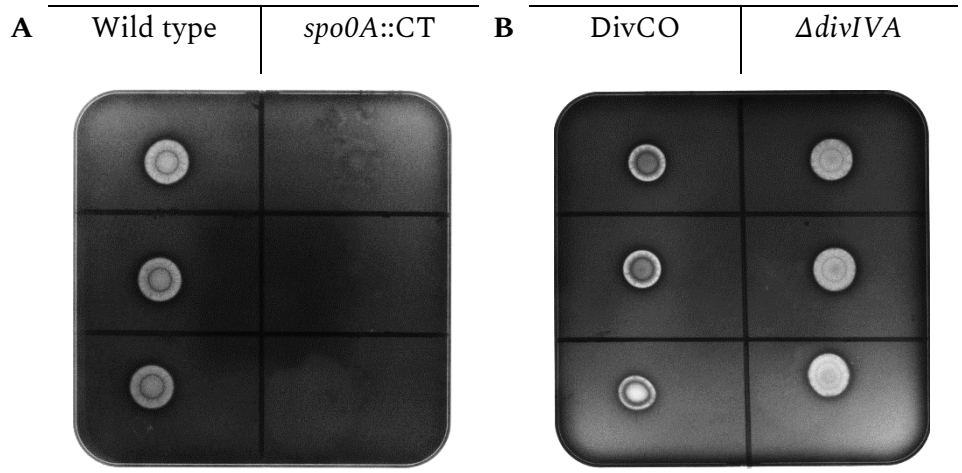


Figure 5.8. **Haemolysis by *C. sporogenes divIVA* strains**. Mutant Δ*divIVA* showed wild type level haemolytic activity, whereas the complemented DivCO strain showed slightly narrower zone of β-haemolysis. The spo/ger$^-$ mutant *spo0A*::CT was non-haemolytic and its translucent growth was overshadowed by dark-coloured blood agar. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT) on TYG agar supplemented with 5% (v/v) horse blood.

5.2.6. Proteolytic activity of *C. sporogenes* is independent of *divIVA*

On milk TYG agar plates, the strongly proteolytic cultures of *C. sporogenes* wild type strain showed zones of clearance within 24-48 hours (Figure 5.9). To determine if this proteolysis can be affected by the deletion of *divIVA* gene, *C. sporogenes* Δ*divIVA* mutants were also tested for proteolytic activity, as described in Section 2.10.5. As expected, no loss of proteolytic activity was observed for both the Δ*divIVA* mutants as well as the complemented DivCO strains (Figure 5.9).

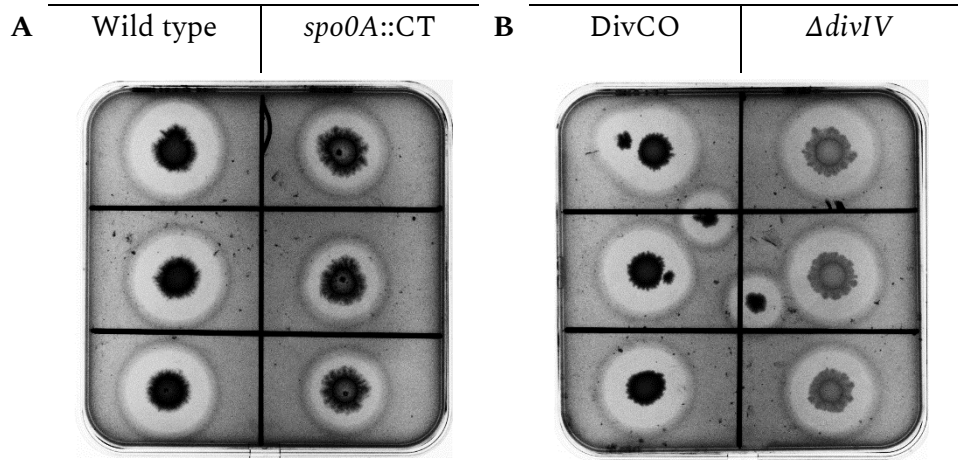


Figure 5.9. **Proteolysis by *C. sporogenes divIVA* strains**. Mutant Δ*divIVA* and the complemented YqfDCO strain showed wild type level proteolytic activity. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT) on TYG agar supplemented with 4% (v/v) skimmed milk.

5.2.7. Microscopy reveals filamentous rods and minicells of Δ*divIVA* mutants

To confirm whether *divIVA* mutation in *C. sporogenes* gave the same phenotype as *hypP* mutation in mutant #42 (Figure 3.6), 120 hour old cultures of *C. sporogenes* strains were visualized under the microscope, as described in Section 2.5.1. Similar to mutant #42, Δ*divIVA* mutants showed cells of numerous sizes: bigger, similar or smaller than the wild type cells (Figure 5.10). These very small mutant cells are known as minicells. Moreover, mutant cells were observed in chains, unlike the wild type and *spo0A*::CT control strains, and very few spores could be detected. On the other hand, the

complemented *C. sporogenes* DivCO strain visually appeared similar to the wild type strain, in terms of cell and spore morphology.

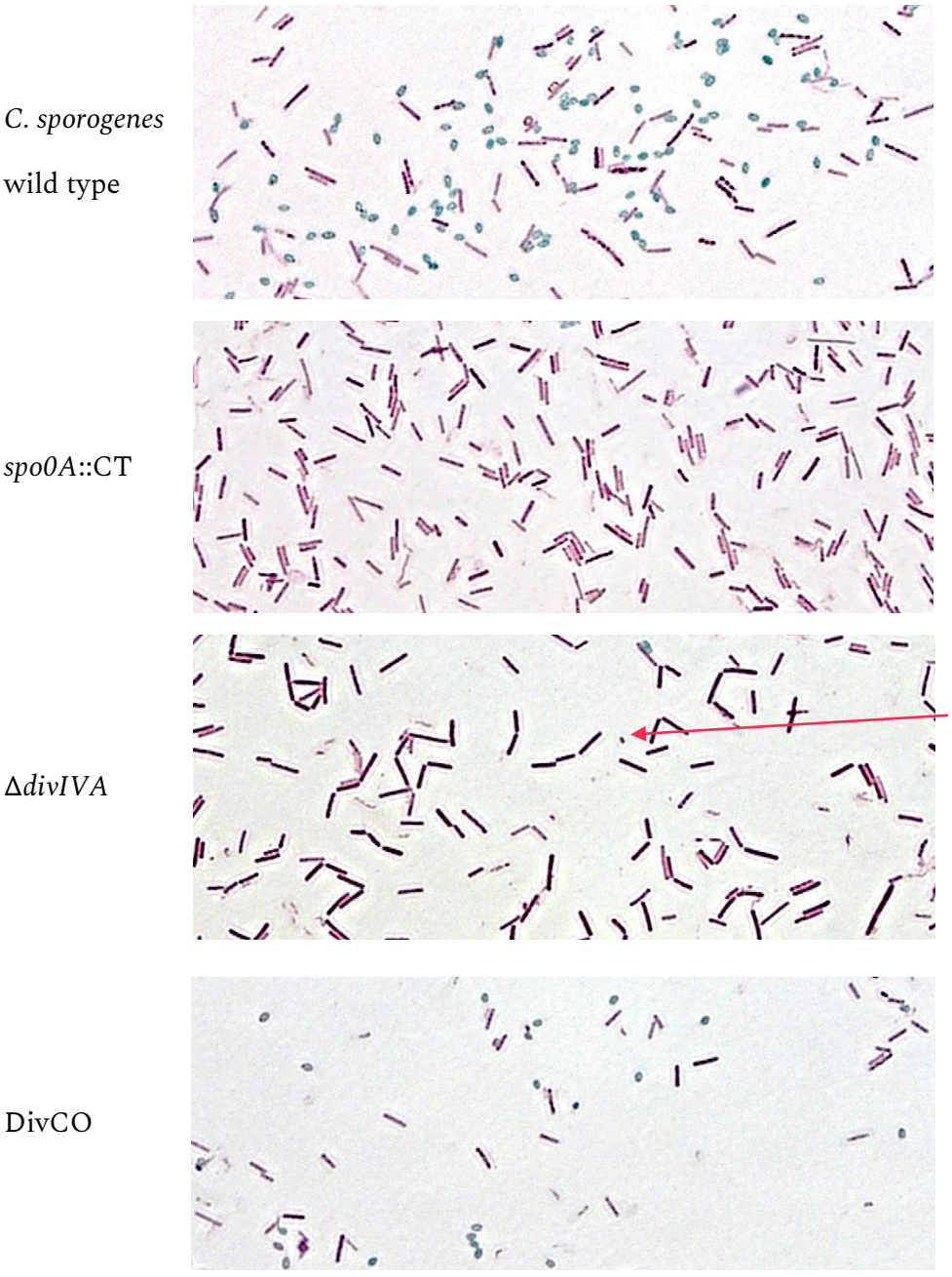
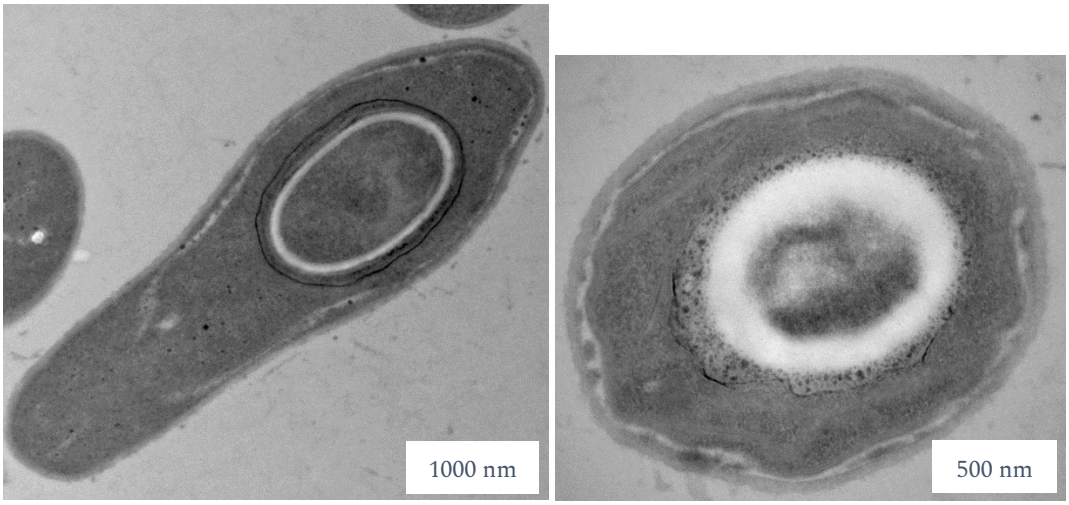


Figure 5.10. **Cell morphology of *C. sporogenes divIVA* strains**. Mutant Δ*divIVA* showed filamentous rods with few mature spores, unlike the *C. sporogenes* wild type strain. Moreover, minicells were observed, as indicated by an arrow. The complemented strains DivCO showed comparable microscopic phenotype to *C. sporogenes* wild type. Cells were stained pink and spores were stained blue. The images represent 120 hour old TYG broth cultures of *C. sporogenes* strains at 1000 × magnification.

Despite attempts at TEM processing and imaging of Δ*divIVA* mutant samples, useful information regarding sporulation could not be retrieved from the TEM micrograph, as shown in Figure 5.11. Some mutant cells were accidently washed away during the fixation step and perhaps the resin embedding technique was not suitable to visualize the Δ*divIVA* mutant cells.

C. sporogenes wild type



C. sporogenes Δ*divIVA*

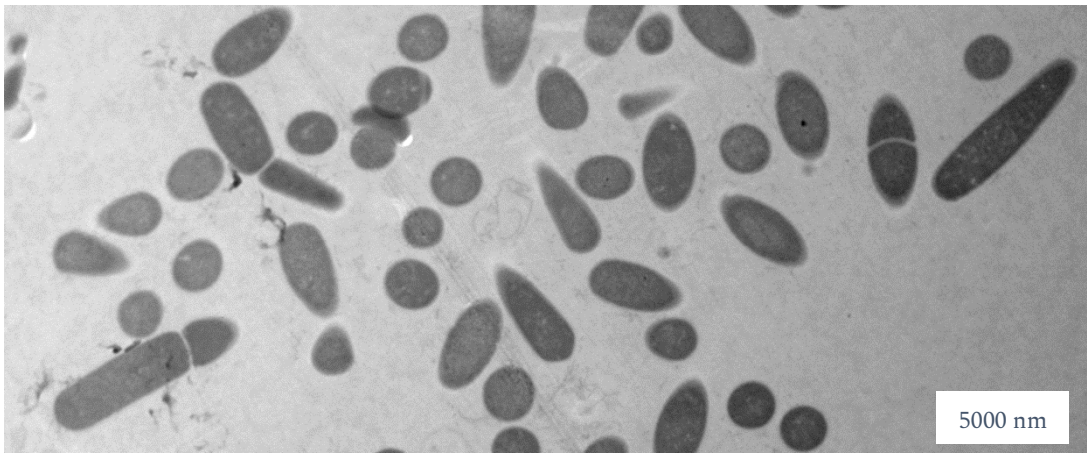


Figure 5.11. **Ultrastructure of *C. sporogenes divIVA* strain**. No spore structure could be visualized for mutant Δ*divIVA*. The micrographs represent thin section of 72 hour old fixed cultures of *C. sporogenes* strains at high magnifications.

5.2.8. Δ*divIVA* mutant colonies have a unique morphology and become translucent on old TYG agar plates

When streaked on TYG agar plates for individual colonies, the Δ*divIVA* mutants not only became translucent within 24 hours, but also make unique feathery colonies of various sizes, similar to mutant #42 (data not shown), unlike the wild type strain (Figure 5.12). Although this phenotype was complemented in the DivCO strain, colonies were observed to be more symmetrical and slightly whiter in colour than those of the cream-coloured wild type strain.



Wild type | *spo0A*::CT | Δ*divIVA* | DivCO

Figure 5.12. **Colony morphology of *C. sporogenes divIVA* strains**. Mutant colonies of *C. sporogenes* Δ*divIVA* showed translucent phenotype on old agar plates, similar to *spo0A*::CT mutant and dissimilar to the opaque growth of wild type. Moreover, the mutant colonies were observed to be more fimbriated than the wild type. Although the opaque phenotype was restored in the complemented strain DivCO, colonies appeared rounded and non-fimbriated, unlike the fimbriated and irregular colonies of *C. sporogenes* wild type strain. The images were taken after 27 hour incubation of *C. sporogenes* strains on TYG agar.

5.2.9. Illumina sequencing reveals SNPs in the Δ*divIVA* mutant genome

To determine the occurrence of spontaneously mutated loci in the genome of *C. sporogenes* Δ*divIVA* mutant, genomic DNA was sent for Illumina sequencing and the sequenced data was then compared with that of the in-house *C. sporogenes* wild type strain, as described in Section 2.3.8. When aligned with the *C. sporogenes* NCIMB 10696 reference strain obtained from the NCBI genome databases, the sequenced data of Δ*divIVA* mutant revealed some more affected genes in addition to the SNP-affected genes in the in-house wild type strain of *C. sporogenes* (Appendix Table A4). One affected gene encodes for the PTS transporter subunit EIIC; another gene expresses a diacyglycerol kinase; and one identified SNP locus was in an unknown gene.

5.3. DISCUSSION

As demonstrated in Chapter Three, mutant #42 had transposon insertion in the *hypP* gene (CLSPO_RS07555), which resulted in a sporulation- and cell division-defective phenotype (Figures 3.2, 3.6, 5.2 and 5.5). Because initially it was assumed that the disruption of gene expression of this "hypothetical" protein could potentially lead to such serious impairments of two of the important biological processes in *C. sporogenes*, attempts were made to characterize the promising *hypP* gene. However, the in-frame gene deletion of *hypP* using CRISPR/Cas9-based genome editing did not show any cell division or sporulation deficiency, unlike the original transposon mutant #42 that has an out-of-frame gene insertion. Henceforth, the study focus had to be switched from *hypP* to its neighbouring gene *divIVA*, for characterizing its role on the sporulation of *C. sporogenes*.

5.3.1. Challenges in processes and problem-solving techniques

There were some problems encountered during the course of this study, which not only changed the understanding regarding research techniques used in general, but it also taught valuable lessons about the biology of the tested bacterial strains.

5.3.1.1. Polar effects of transposon insertion

The first confusion arose when the *hypP* in-frame gene deletion in *C. sporogenes* resulted in a wild type sporulation phenotype (Figure 5.4), dissimilar to transposon mutant #42. Initially, this result was disregarded and the spore assay was repeated; however, the same sporulation-efficiency was observed for these PCR- and sequence-verified *C. sporogenes* Δ*hypP* mutants. This result was surprising, but it was then hypothesized that transposon insertion in the *hypP* gene may have caused a polar effect on the expression of neighbouring genes on the same operon.

As mentioned earlier, *hypP* gene is the fourth gene in a five gene operon, arranged in a forward orientation in the genome of *C. sporogenes* NCIMB 10696. With a 15 bp intergenic space, *hypP* faces *divIVA*, which is the last gene, also in forward orientation in the operon (Figure 5.2). Naturally, *divIVA* was thought to be a potential target of polar effects of transposon insertion. Coincidently, the previously characterized homologues of *C. sporogenes divIVA* mutants were found to have similar sporulation and cell division phenotypes in *Bacillus* spp. (Cha and Stewart, 1997); (Perry and Edwards, 2004), as those for mutant #42 in this study. Although *divIVA* has been proved to be a non-essential gene in some of the previously studied *B. subtilis* strains (Cha and Stewart, 1997), it was also learned that deletion of *divIVA* in other tested *B. subtilis* strains lead to cell death, which makes it almost impossible to have an in-frame *divIVA* gene deletion in the wild type background (Perry and Edwards, 2006). However, Δ*divIVA* mutants can be generated in *Bacillus* strains lacking the expression of MinC or MinD proteins, in which the unrestricted inhibition of cell division by the MinCD complex is prohibited (Perry and Edwards, 2006); (Edwards and Errington, 1997).

Since mutant #42, with a transposon insertion in *hypP* grows well without inactivation of *minC* or *minD* genes, it was decided to first confirm the *divIVA* mutant phenotype in *C. sporogenes*. In order to do a similar type of insertion-based mutagenesis like transposon mutagenesis in mutant #42, the ClosTron gene knock-out system was used to insert the RAM-based intron in the *divIVA* gene of the in-house *C. sporogenes* wild type strain, as described in Section 2.9. The resultant *divIVA*::CT mutant was tested for sporulation frequency, upon which a similar sporulation-defective phenotype was observed as that of mutant #42, as shown in Figure 5.4. Moreover, the expression of DivIVA protein in the complemented strain ProdivIVA restored the wild type sporulation efficiency (Figures 5.5 and 5.6), which was dissimilar to when the "hypothetical" proteins were expressed on plasmid in the ProhypP strain to

complement mutant #42 (data not shown). This proved that the phenotypes of mutant #42 could only be complemented by DivIVA and not the gene product of the interrupted *hypP* itself. On the basis of this data, it was decided to target the *divIVA* gene using the CRISPR/Cas9-based genome editing and characterize it instead of *hypP* for future studies.

5.3.1.2. Non-specific ClosTron RAM insertion

The group II intron sequence can be easily designed to target any locus in the clostridial genome and make use of the RAM insertion to inactivate gene expression (Kuehne and Minton, 2012). For targeting a desired gene, the intron design is based on the nucleotide sequence as determined by a computer-based tool that uses the Perutka algorithm to find the best on- and off-target scores (Perutka et al., 2004). Since the target-specific ClosTron plasmid is commercially designed and stable, it is easier to clone into *E. coli* and *C. sporogenes* to obtain transformants and transconjugants, than with the CRISPR/Cas9-based plasmids. Following successful conjugation, the ClosTron plasmid harbouring *C. sporogenes* transconjugants are allowed to grow on erythromycin-supplemented TYG plates, as also described in Section 2.9. The occurrence of erythromycin-resistant colonies reveals the activation of *ermB* gene marker upon successful insertion into the *C. sporogenes* genome; however, it does not provide data of insertion in the targeted locus or *divIVA* in this case. Therefore, the erythromycin-resistant colonies were verified by PCR using three sets of primers, as described in Section 2.9.3. Despite RAM insertion in all of the 11 tested *C. sporogenes* colonies, only colony #3 was a verified *divIVA*::CT mutant, as revealed by the amplicon obtained using EBS Universal and DivIVA_F forward flanking primers, along with the amplicons generated by DivIVA_F/R flanking primers, as shown in Figure 5.3. The chances of non-specific intron targeting are possible due to the use of mobile genetic elements in *C. sporogenes* having sequence similarities across the genome. Also, because the algorithm is only applied to the gene being targeted and not the

whole genome, other intron insertion sites might be present in other parts of the genome that are not looked by the program.

5.3.1.3. Problems in cloning knock-out/in cassettes in *E. coli*

As mentioned in Section 4.3.1.1, regarding cloning of the CRISPR/Cas9-based homology arms cassettes of *yqfD* in *E. coli* strains, cloning *divIVA* constructs in *E. coli* was problematic as well. Both homology arms with and without the *divIVA* CDS were toxic to *E. coli* TOP10 cells. When pCRISPR-divIVA plasmid that had 2 kb homology arm-based knock-out cassette was used to transform competent *E. coli* TOP10 cells, no transformants could be achieved. The size of homology arms was then shortened to 1.4 kb in pCRISPR-divIVA1 plasmid, which gave transformants in *E. coli*, but at the expense of mutations in the knock-out cassette and vector backbone, as shown in Figure 5.13. Although the pCRISPR-divIVA1 cassette was successfully able to generate *C. sporogenes* Δ*divIVA*_old mutants, the presence of a SNP in the homology arm at the *hypP* CLSPO_RS07555 locus had the potential to affect the data for *divIVA* characterization.

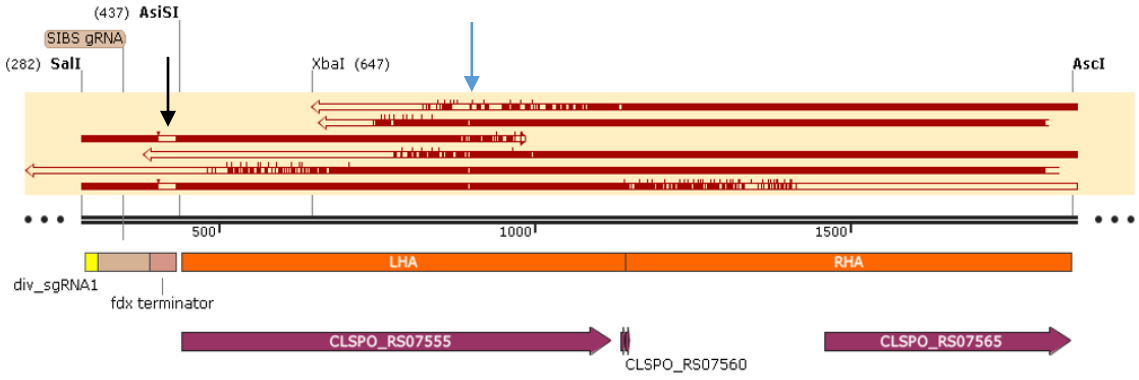


Figure 5.13. **SNPs in the *divIVA*-knock-out cassette**. PCR amplicons of the pCRISPR/Cas9-divIVA1 construct in *E. coli* TOP10 and *E. coli* CA434 transformants were sequenced with sgRNA_fdx_fl_fwd, 83xxxL_R and divRH_R2 primers. Aligning the resultant data with the corresponding sequence of *in silico* plasmid construct in SnapGene revealed the presence of two unwanted mutations. A SNP (G→T) was found in the homology arm, as indicated by a blue arrow; and in the other mutation, fdx terminator got deleted, as indicated by a black arrow.

Additionally, the nucleotide sequence of *divIVA* was not suitable for introducing a premature STOP codon to inactivate the gene using the homology arm-independent Target-AID base editor. Therefore, mutant generation for *divIVA* gene could only be relied on the conventional CRISPR/Cas9 protocol.

Therefore, cloning protocols were customized and after shortening the homology arms to 1.1 kb, ligating the knock-out cassette in the reverse orientation, using competent *E. coli* 5-alpha F' Iq cells and growing transformants at room temperature; SNP-free pRECas1-divIVA was obtained, which was ultimately, used to generate the new Δ*divIVA* mutants in *C. sporogenes*. Similarly, pRECas1-DivCO construct was successfully cloned and then used for complementing Δ*divIVA* mutants; thereby, generating the *C. sporogenes* DivCO strains. The hypPsgRNA1 in the *divIVA*-knock-in cassette was efficient in guiding Cas9 to the chromosomal homology arm sequence and the HDR occurred without any issues, unlike the *yqfD*-knock-in cassette, which has been described in Section 4.3.1.2.

5.3.2. Other unexpected phenotypes

Like other sporulation-deficient mutants, *C. sporogenes* Δ*divIVA* mutants showed the translucent phenotype on old TYG agar plates. Edwards (1997) has also referred to this phenotype in their study on *B. subtilis* Δ*divIVA*, suggesting that this is a common finding for sporulation-deficient colonies. Moreover, the mutant colonies appeared more feathery or fimbriated and irregular than the wild type colonies, possibly due to some defect in chromosomal segregation. Complementing Δ*divIVA* mutants should have produced wild type colonies, which are asymmetric and fimbriated; however, it was observed that colonies of DivCO appeared round, symmetric and non-feathery (Figure 5.13). The reason could have been that, the *hypP* gene in the DivCO strain contains a watermark, in which five codons were altered in the gene sequence. This was done in order to prevent the hypPsgRNA1 from targeting the homology arm

sequence on the plasmid, but to instead allow the guide RNA to target the *hypP* sequence on host's chromosome. Ideally, the truncated *divIVA* gene region would have been the target site to incorporate the gene for complementation. But in the absence of a previously incorporated bookmark in the Δ*divIVA* mutants, no other efficient target site was available except for the neighbouring *hypP* gene. Although all the homology arm cassettes for gene-knock-outs were initially designed to include an efficient guide RNA-targeting site, i.e., a bookmark, but some constructs had to be re-designed without a bookmark in order to avoid any potential toxicity in the *E. coli* donor strains.

Hence, it was decided to further insert a pMTL87121:*hypP* complementation plasmid to allow the expression of *hypP*'s (CLSPO_RS07555) "hypothetical" protein in the previously complemented DivCO strain. However, no change in the symmetric colony phenotype was observed in the DivCO strain containing this plasmid (data not shown), perhaps due to some unexpected SNP occurrence in the genome.

5.3.3. SNPs in the Δ*divIVA* mutant

It was important to know the presence of undesirable mutations in the generated *C. sporogenes* Δ*divIVA* mutant, which was why one of the engineered strains was sent for whole genome sequencing. Three genes were identified, which showed the presence of SNPs additional to those in the in-house wild type strain. One of the affected gene product was annotated as phosphotransferase or PTS transporter subunit EIIC, which is known in the literature to specifically transport glucose molecules into the cell membrane and prevent the efflux of the imported glucose outside the membrane via sugar phosphorylation (McCoy et al., 2015). Another affected gene product was the diacylglycerol kinase, which is an important intermediate in the biosynthesis and breakdown of membrane phospholipids (Amiteye et al., 2003). Interestingly, the role of diacylglycerol kinase in *B. subtilis*

sporulation has been previously characterized, in which the mutants produced unstable spores due to the altered lipid composition of forespore membranes (Amiteye et al., 2003). However, the sporulation phenotype of Δ*divIVA* mutant was fully complemented in the DivCO strains (Figure 5.6), suggesting diacylglycerol kinase might be dispensable for *C. sporogenes* sporulation. Furthermore, one of the unknown gene locus that showed the presence of SNP could not be characterized by BLAST and InterPro tools.

5.3.4. Characterization of *divIVA* and its role in sporulation

Since very few spores were observed for *divIVA*::CT and Δ*divIVA* mutants under the microscope (Figure 5.10), it was apparent that *divIVA* gene is involved in the sporulation of *C. sporogenes*. This role was further confirmed by observing comparatively small number of heat-resistant CFUs on TYG plates after heat-treating cultures of *C. sporogenes divIVA*::CT (Figure 5.4) and Δ*divIVA* mutants (Figure 5.6). However, in the previous literature it has also been shown that whilst the existing DivIVA proteins remain at the cell poles without interrupting the asymmetric forespore septum (Thomaides et al., 2001), further expression of DivIVA in *B. subtilis* is actually repressed by Spo0A after the onset of sporulation (Fujita et al., 2005). While *divIVA* has not been found in the *spo0A* regulon of *C. difficile* (Pettit et al., 2014), it was still assumed that the observed sporulation defect might just be a side effect of an impaired cell division. Because *divIVA* mutant cells undergo abnormal chromosomal segregation, heterogeneously-sized cells are generated with apparently unequal distribution of chromosomal DNA. This could have been a possible explanation of sporulation defect; but, it has also been demonstrated in *B. subtilis* that Δ*minD* mutants have only a slight sporulation defect, despite having a higher frequency of irregular cells than Δ*divIVA* mutants (Thomaides et al., 2001).

Interestingly, it has been demonstrated that the absence of DivIVA in *B. subtilis* results in premature activation of σ^F by an impaired SpoIIE activity (Eswaramoorthy et al., 2014, Thomaides et al., 2001). Although additional data about transcriptomics of *C. sporogenes divIVA* is critical for verifying this interruption of sporulation cascade in the obtained *divIVA* mutants, we can also assume that this mechanism might be similar in other spore-forming Firmicutes, like *C. sporogenes*.

As highlighted throughout this chapter, *divIVA* gene has been well-characterized in several *B. subtilis* strains (Thomaides et al., 2001, Edwards and Errington, 1997, Eswaramoorthy et al., 2014, Perry and Edwards, 2004), and the previous and current data clearly supports the importance of *divIVA* in endospore-formers. BLASTP analysis of *C. sporogenes* DivIVA reveals 38.04% identity with the protein homologue in *B. subtilis*. Combining the BLASTP data along with that in several previous studies (Cha and Stewart, 1997, Muchova et al., 2002, Perry and Edwards, 2006, Edwards and Errington, 1997), the severely-reduced sporulation, large filamentous rods and minicell phenotypes for the previous *B. subtilis* Δ*divIVA* mutants indeed match with those observed for *C. sporogenes* Δ*divIVA* mutants in this study.

C. sporogenes Δ*divIVA* mutants were generated in the wild type background, without inactivating other genes or halting other biological processes, which makes *divIVA* a non-essential gene in *C. sporogenes*. All the mutants grew similar to the wild type strain (Figure 5.7), which indicates that *divIVA* does not particularly play a major role in the exponential growth of *C. sporogenes*. Moreover, the translucent colony phenotype of Δ*divIVA* mutants observed in this study (Figure 5.12), has also been mentioned in the previous studies on *B. subtilis divIVA* mutants (van Baarle et al., 2013, Edwards and Errington, 1997). This phenotype might be related to the growth and survival of the bacterium in the presence of suppressors that arise in old cultures.

In some of the previous studies, the vegetative growth of *B. subtilis* strains with an inactivated *divIVA* gene was not affected (Cha and Stewart, 1997), which is similar to

the data obtained from the growth assay of *C. sporogenes* Δ*divIVA* mutants in this study. However, the deletion of *divIVA* in wild type background was also found to be lethal to *B. subtilis* cells in some other studies, because of the indiscriminative inhibition of cell division by MinCD proteins (Perry and Edwards, 2004, Edwards and Errington, 1997, Thomaides et al., 2001). Perhaps, the presence of the *minE* gene in the *C. sporogenes* genome assures the viability of Δ*divIVA* mutants. MinE is a cell division topological specificity factor in *E. coli* and other Gram negative bacteria, and has an analogous cell division site localization function to DivIVA of *B. subtilis* (de Boer et al., 1989, Edwards and Errington, 1997). Since *minE* is absent in the genome of *B. subtilis* and *divIVA* is absent in *E. coli*, they can be considered as gene homologues (Cha and Stewart, 1997). Surprising, *minE* is found in the genomes of many clostridial species, including *C. sporogenes* (data not shown), despite absent in many *Bacillus* species.

Deletion of *divIVA* has been known to affect the virulence of *L. monocytogenes* by impairing the autolysin secretion of the pathogenic host bacterium (Halbedel et al., 2012), which was why it was interesting to know if *divIVA* also effects haemolysin secretion in (apparently) non-pathogenic strain of *C. sporogenes*. No difference in the haemolytic activity was observed between the wild type and mutant strains, as shown in Figure 5.8. Hence, *divIVA* does not appear to be involved in *C. sporogenes* haemolysis. Moreover, no correlation between *divIVA* and protease activity (Figure 5.9) could be identified in this study.

5.3.5. Proposed function of *divIVA* in *C. sporogenes*

Based on the obtained data in this study, along with the understandings from previous literature, it has been proposed that *divIVA* plays an important role in the sporulation and symmetric cell division segregation of *C. sporogenes*. The *divIVA* gene also shares its promoter with four other uncharacterized genes on the *C. sporogenes* chromosome. It is assumed that in the absence of DivIVA protein during the vegetative growth of *C. sporogenes*, the distribution of the cell division-inhibition complex MinCD is possibly not completely controlled, because of which the aberrant cell division septa occur throughout the dividing cell instead of a single septum at mid-cell. This explains why cells of uneven sizes, such as minicells arise within the same population. However, *divIVA* deletion can be made in the wild type background, which hints the possible existence of a divIVA gene homologue within the *C. sporogenes* genome, which might be *minE*.

CHAPTER SIX

6. CHARACTERIZATION OF *clpX*

6. CHARACTERIZATION OF *CLPX*

6.1. INTRODUCTION

Proteolysis is an important process of cell metabolism, which recycles amino acids and proteins for the generation of energy, especially during cell starvation (Goldberg and St John, 1976, Konovalova et al., 2014). It is brought about by the ATP-dependent two-component ring complex, which consists of a ClpP protease for degrading the peptide chains, and an AAA+ family ATPase subunit to serve as a chaperone for this protein degradation (Singh et al., 2001, Gottesman, 2003, LaBreck et al., 2017). This process is illustrated briefly in Figure 6.1.

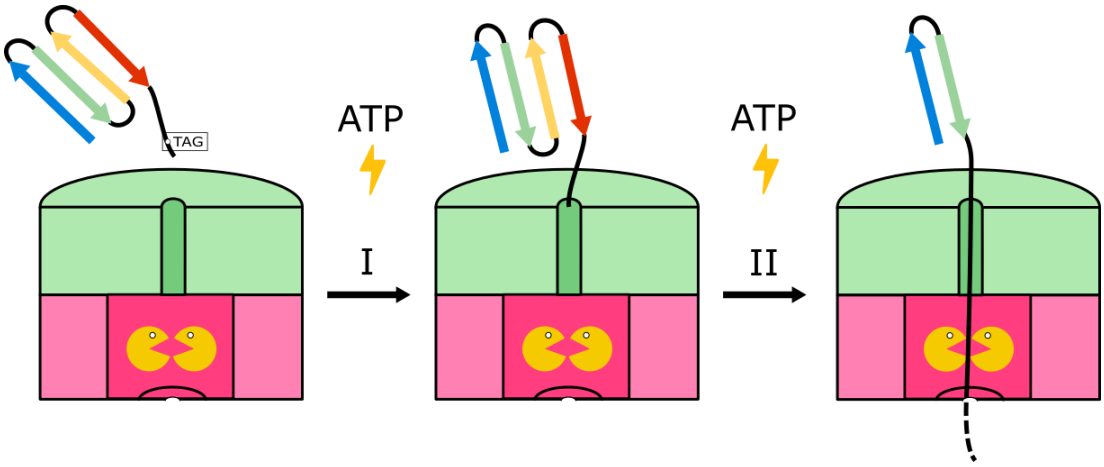


Figure 6.1. **The ClpXP proteasome**. Peptide substrates with relevant tags are recognized by ClpX subunit, which unfolds them and translocate to ClpP protease for degradation. This diagram was inspired by (Sauer et al., 2004, Sivertsson et al., 2019).

One such ATPase subunit is ClpX, which functions in the recognition and unwinding of protein substrates for proteolysis (Baker and Sauer, 2012, Hanson and Whiteheart, 2005, Sauer et al., 2004, Sauer and Baker, 2011). The transposon mutant library screenings in Chapter three of this study also identified the *C. sporogenes* ClpX-encoding gene CLSPO_Rs16880, which showed a possible involvement in sporulation or germination, as shown in Figure 3.3 and its corresponding Table 3.1.

According to the NCBI genome databases, CLSPO_RS16880 is a 1290 bp long gene in the genome of *C. sporogenes* NCIMB 10696 strain, with 31% GC content. Its gene product (WP_003490561.1) is annotated as "ATP-dependent Clp protease ATP-binding subunit ClpX", which is made up of a stretch of 429 amino acids. Browsing the neighbouring gene loci, it can be predicted that CLSPO_RS16880 is the second gene in a two-gene operon along with the gene for ClpP protease. ClpP has also been previously known to be involved in the sporulation of *B. subtilis* (Gerth et al., 1998, Msadek et al., 1998, Nanamiya et al., 2000).

For the ease of explanation, the term "*clpX*" will be used throughout this thesis to define the CLSPO_RS16880 gene.

Possible protein domains in *clpX* are illustrated in Figure 6.2.

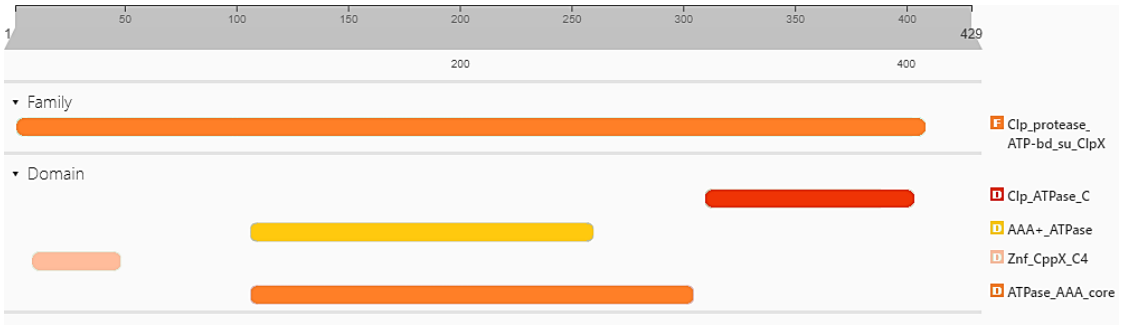


Figure 6.2. **Protein domains of the ClpX predicted by InterPro**. ClpX protein consists of Zinc or metal-binding motif, along with the ATP-binding domain. Graphics were edited for a personalized view.

ClpX recognizes its peptide substrates by the presence of specific peptide sequences called degradation tags or degrons, after which it binds with the substrates and unfolds the tertiary structure and translocate the unfolded protein substrates to the ClpP protease for degradation (Baker and Sauer, 2012). ClpXP-binding, protein unfolding and translocation processes are all powered by ATP binding and ATP hydrolysis (Joshi et al., 2004, Sauer and Baker, 2011). One of the mechanisms through which bacterial cells maintain and regulate protein levels is by ClpXP degradation

(Baker and Sauer, 2012). For example, the DNA repair protein RecN, which also contains a ClpX-recognisable degron, becomes accumulated in the cell during SOS transcriptional response following a DNA damage (Odsbu and Skarstad, 2014). Once the SOS event ceases and RecN is no more required by the cell, ClpXP degrades this stress response protein (Neher et al., 2006).

ClpX is widely known as a component of the ATP-dependent proteasome in all the eubacteria, as well as the mitochondria and chloroplasts of eukaryotes (Frees et al., 2005, Olivares et al., 2016). Moreover, mitochondrial ClpX is better conserved than ClpP, suggesting a ClpP-independent role of ClpX in cell biology (Kardon et al., 2015). In addition to being a chaperone to Clp protease for protein degradation, several roles of ClpX and ClpXP complex have been reported in the previous literature. The importance of ClpX also varies species to species (Frees et al., 2007, Jensen et al., 2019). Deletion of *clpX* has previously induced heat-sensitivity in *B. subtilis* (Gerth et al., 1998), whereas improved heat-resistance in *S. aureus* (Frees et al., 2003), but not the cold-tolerance (Jensen et al., 2019). Additionally, *clpX* alone is an essential gene in *Streptococcus pneumonia* (Robertson et al., 2003) and *Caulobacter crescentus* (Jenal and Fuchs, 1998). In another interesting study, ClpX was discovered to be an inhibitor of the FtsZ assembly in *B. subtilis*, without the requirement for ClpP protease or ATP hydrolysis (Weart et al., 2005, Haeusser et al., 2009). Since the bacterial cytoplasm maintains a good reservoir of cell division protein FtsZ for Z-ring formation, the vital role of ClpX is required to maintain the balance between the assembled and unassembled cell division proteins (Weart and Levin, 2003, Weart et al., 2005). The ClpXP complex is also involved in the downregulation of flagellin protein synthesis in *Salmonella enterica* serovar Typhimurium, since the deletion of *clpX* in the bacterium causes overexpression of *fliC* gene, and therefore, a hyperflagellate or highly motile phenotype is observed in the mutant (Tomoyasu et al., 2002).

Interestingly, *clpX* has also been previously documented to be involved in the virulence of *S. aureus*. Deletion of the gene, not only reduced the visibility of abscess in a mouse skin model, but also reduced concentrations of *S. aureus* α-haemolysin or *hla*, several extracellular proteases and cell-surface associated Protein A (Frees et al., 2003, Frees et al., 2005). Similarly, *clpX* has also been found to be involved in the haemolysis and proteolysis of *Bacillus anthracis*, as well as its antibiotic resistance property (McGillivray et al., 2009). The *clpX* gene has also been proven to be involved in the virulence of *Xanthomonas campestris* pv. *campestris*, since the *clpX* mutant exhibited reduced cell adherence to polystyrene plates, borosilicate glass and cabbage leaf; reduced enzymatic activity; reduced black rot disease lesions on cabbage foliage; and reduced tolerance to physical and chemical stress stimuli (Lo et al., 2020). The same study detected the altered expression of 211 proteins, which included the downregulation of several proteins involved in bacterial adherence, enzyme production and virulence in *clpX* mutants.

Literature review reveals how ClpX is linked with the bacterial sporulation through various cascades. It has been observed that during sporulation, ClpXP seems to preferentially locate in the mother cell region to commence proteolytic activity (Kain et al., 2008). *clpX* has been identified as a heat-inducible gene in *B. subtilis*, which is not only over-expressed under stress conditions, but the lack of *clpX* expression also results in sporulation defects. Moreover, ClpX and ClpP are also involved in the regulation of other survival genes during the stationary phase, which may suggest their possible role in the starvation-induced process of sporulation (Gerth et al., 1998).

One of the cascades through which sporulation in *B. subtilis* is initiated, is through the activation of sigma factor σ^H, which is responsible for the activation of several other sporulation genes (Carter and Moran, 1986). σ^H concentration and regulatory activity depends on the pH of the environment: low pH has negative effects, whereas high pH increases the σ^H activity and hence the sporulation frequency (Cosby and

Zuber, 1997). In previous studies on *B. subtilis*, *clpX* has been observed to stimulate the σ^H–dependent transcriptional expression, under nutrient-deprived conditions, by driving the accumulation of phosphorylated Spo0A, which in turn displaces the house-keeping sigma factor σ^A from the core RNA polymerase to allow σ^H RNA polymerase activity (Liu et al., 1999, Liu and Zuber, 2000). The same study observed the overexpression of *rpsD* stress response gene in the *clpX* mutant during stationary phase, which suggests the possibility of a negative role of ClpX on the σ^A RNA-based transcriptional activity.

In another study, it has been observed that ClpXP complex in *B. subtilis* also serves as the AAA+ protease that utilizes CmpA as an adaptor for mediating the degradation of SpoIVA proteins in those sporulating mother cells, which carry a misassembled spore envelop in their forespore (Tan et al., 2015). Since SpoIVA is an important structural protein that polymerizes to form the spore envelop (Price and Losick, 1999), the defected forespores are therefore regulated and eliminated by this quality control pathway to ensure the integrity of the spore.

The *clpX* gene expression has not been found to be regulated by any of the sporulation-specific sigma factorc, according to a recent study on sporulation of *C. sporogenes* (Raquel Rodrigues, unpublished). Previous studies reveal that *clpP* and other stress response *clp* genes might be members of the *sigA* or *sigB* regulon in *B. subtilis* (Gerth et al., 1998, Kruger et al., 1996, Gerth et al., 1996). However, the transcription of *clpX* is not majorly dependent on the CtsR repressor protein, unlike other *clp* genes (Kruger and Hecker, 1998, Derre et al., 1999, Nair et al., 2000, Frees et al., 2007).

6.1.1. Aim of this chapter

This chapter is aimed at studying the CLSPO_RS16680 *clpX* gene for its potential role in the sporulation of *C. sporogenes*, by the following steps:

i. Constructing in-frame deleted Δ*clpX* mutants in the *C. sporogenes* NCIMB 10696 wild type strain and later complementing the mutated gene;

ii. Characterizing the Δ*clpX* mutants by determining the relationship of *clpX* gene with sporulation, growth, haemolysis, proteolysis, and cell and colony morphology of *C. sporogenes*, and comparing the mutant phenotype with the wild type phenotype;

iii. Finding relevant literature on the prokaryotic *clpX* gene and its ClpX protein and comparing the phenotypes of the *C. sporogenes* Δ*clpX* mutants with the ones generated in *Bacillus* and other clostridial species;

iv. Proposing the potential role of *clpX* in the biology of *C. sporogenes*, based on the data generated in this study, with reference to earlier studies.

6.2. RESULTS

6.2.1. CRISPR/Cas9-based genome editing in the *clpX* gene

To confirm that the in-frame deletion of the *clpX* gene (CLSPO_RS16680) causes a reduced sporulation phenotype, as seen with the *clpX* gene-inactivation in transposon mutants T43 and T45, attempts were made to generate three *C. sporogenes* Δ*clpX* mutants (with botsgRNA1 bookmark) from the *C. sporogenes* wild type strain using the pRECas1-clpX knock-out cassette, as described in Section 2.7. Δ*clpX* mutants were later, complemented with the pRECas1-ClpXcomp knock-in cassette. The resulting *C. sporogenes* ClpXcomp strains and their mutant variants were characterized after the confirmation of plasmid loss.

To reduce the chances of possible SNPs during cloning of the knock-out cassette in *E. coli*, Target-AID base editing was also considered. Therefore, several isogenic *C. sporogenes clpX*::TA mutants were generated from the wild type strain, using the

pTargetAID-clpXsgC construct, as described in Section 2.8. In this way a premature STOP codon was successfully introduced in the start of the *clpX* gene. Mutants were initially identified by their translucent phenotype on selective TYG/Theo agar plates, as shown in Figure 6.3, and were later sequence-verified by DNA sequencing.

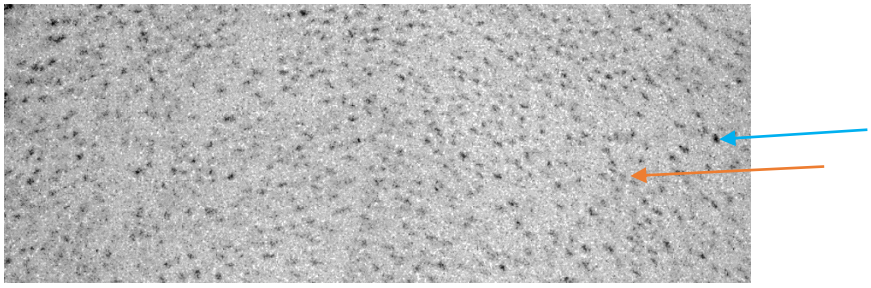


Figure 6.3. **Transconjugant colonies of *C. sporogenes* wild type-pTargetAID-clpXsgC**. Colonies of *clpX*::TA mutants appeared faint and translucent (orange arrow) among the visible dark colonies of wild type background (blue arrow) on TTC/Theo plate after 48 hour post-conjugation incubation.

6.2.2. Mutation in *clpX* alleviates *C. sporogenes* sporulation

To investigate the sporulation profile of Δ*clpX* mutants, relevant *C. sporogenes* strains were grown in TYG broth for 5 days and samples were withdrawn daily and heat-treated to determine heat-resistant CFUs, as described in Section 2.10.2. Similar to mutants T43 and T45 (Figure 3.2), *C. sporogenes* Δ*clpX* mutants were observed to have severe sporulation or germination defect, with more than 1×10^3 reduction in heat-resistant CFUs compared to the wild type (Figure 6.4). However, sporulation was not completely abolished, unlike that observed for the spo/ger$^-$ *spo0A*::CT mutant. This sporulation defect of Δ*clpX* mutants was successfully repaired in ClpXcomp strains.

Unexpectedly, the verified *clpX*::TA mutants did not show sporulation-defective phenotypes of mutants T43, T45 (Figure 3.3) and Δ*clpX* (Figure 6.4). Despite the presence of a STOP codon in the start of the *clpX* gene, which would have halted ClpX expression and reduced the sporulation efficiency to an expected level; wild type level sporulation frequency was still observed in the Target-AID-generated *clpX* mutants.

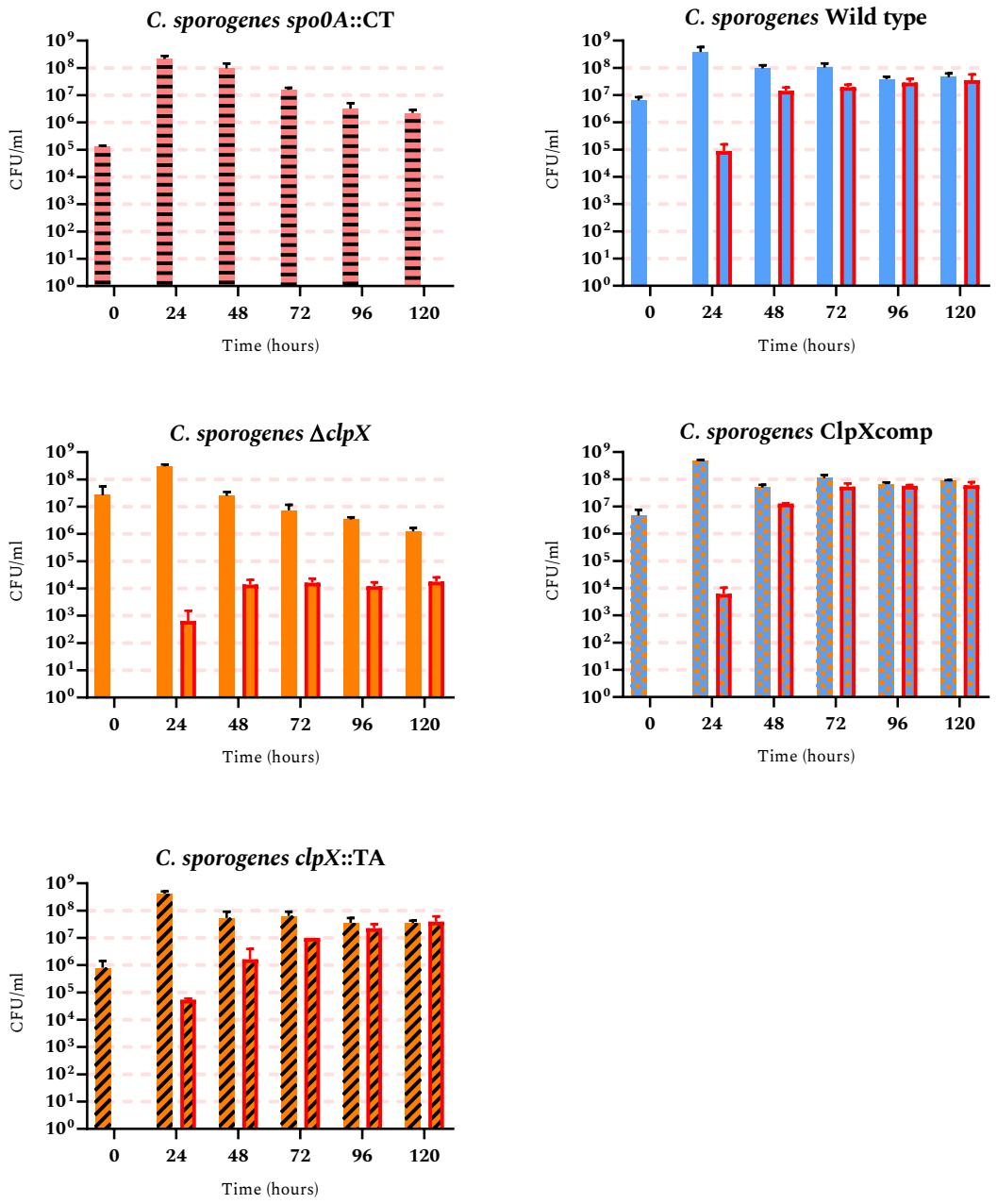


Figure 6.4. **Spore assay of *C. sporogenes clpX* strains**. Mutant Δ*clpX* showed ~1×10^3 times lower heat-resistant CFUs compared to *C. sporogenes* wild type strain, while mutant *clpX*::TA and the complemented strain ClpXcomp showed similar heat-resistant CFUs as wild type. This data is based on the 5-day incubation of three biological replicates (except for *spo0A*::CT) in TYG broth. Error bars show mean with standard deviation. Each strain is represented by a unique colour[3]. Red-bordered bars represent heat-resistant CFUs, whereas coloured bars without red border represent the untreated or the total cell counts. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4 × 10^9.

[3] Colour key:



6.2.3. *clpX* is not an essential gene in *C. sporogenes*

In order to determine the effect of *clpX* deletion on the growth of *C. sporogenes*, relevant strains were grown in TYG broth and samples were withdrawn to measure the optical density at various time points, as described in Section 2.10.3. It was observed that unlike the slightly growth-defective *spo0A*::CT mutant, the growth curve for *C. sporogenes* Δ*clpX* mutants was observed similar to that of the wild type strain, as shown in Figure 6.5. Likewise, the growth of complemented ClpXcomp strains also appeared similar to the wild type strain.

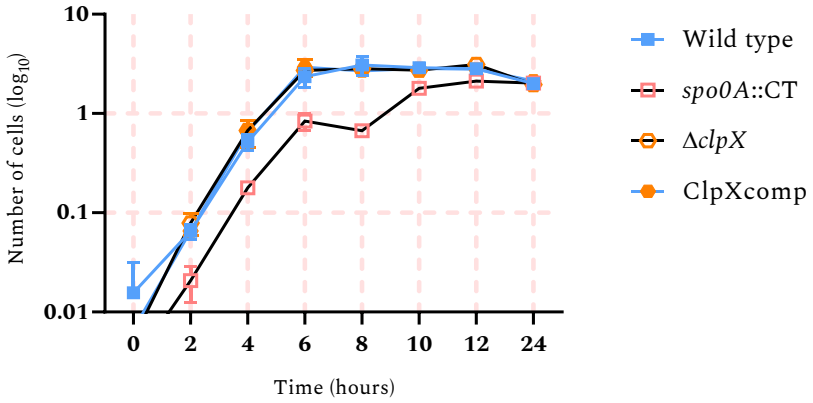


Figure 6.5. **Growth curve of *C. sporogenes clpX* strains**. Mutant Δ*clpX* and the complemented ClpXcomp strain, showed similar growth patterns to *C. sporogenes* wild type strain, in contrast to the spo/ger$^-$ mutant *spo0A*::CT, which has a slight growth defect. This data is based on the 24 hour incubation of three biological replicates (except for *spo0A*::CT) in TYG broth. Error bars show mean with standard deviation. Theoretically, 0.01 OD_{600} was the starting inoculum, which was individually calculated for each strain; however, in practice, variable readings were observed at 0 hour.

6.2.4. *C. sporogenes* β-haemolysis does not require *clpX*

Since *C. sporogenes* wild type is able to lyse blood cells, while the spo/ger⁻ mutant *spo0A*::CT lose this property (Figure 6.6), it was interesting to know if mutation in the *C. sporogenes clpX* gene can also affect haemolysis. It was observed that when cultures of *C. sporogenes* Δ*clpX* mutants were spotted on blood-supplemented TYG agar, as described in Section 2.10.4, zones of β-haemolysis were formed, as shown in Figure 6.6. Similar haemolytic activity was demonstrated by the complemented ClpXcomp strains.

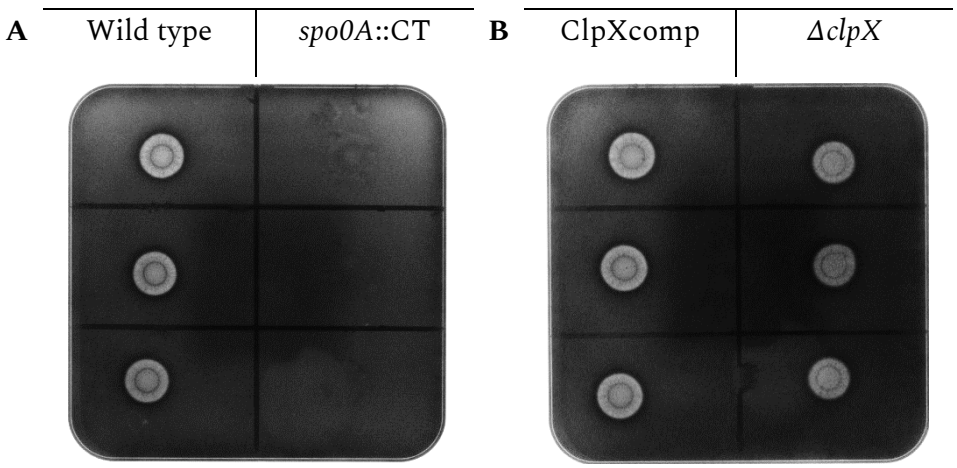


Figure 6.6. **Haemolysis by *C. sporogenes clpX* strains**. Mutant Δ*clpX* and the complemented DivCO strain showed wild type level haemolytic activity, in terms of β-haemolysis zones. The spo/ger⁻ mutant *spo0A*::CT was non-haemolytic and its translucent growth was overshadowed by dark-coloured blood agar. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT) on TYG agar supplemented with 5% (v/v) horse blood.

6.2.5. Proteolytic activity of *C. sporogenes* is independent of *clpX*

On milk-supplemented TYG agar plates, the proteolytic *C. sporogenes* wild type strain showed zones of clearance within 24-48 hours of anoxic incubation (Figure 6.7). It was interesting to know if the lack of *ClpX* protease in Δ*clpX* mutants could affect the strongly proteolytic nature of *C. sporogenes*. Therefore, Δ*clpX* mutants and their complemented ClpXcomp strains were grown on milk TYG agar plates, as described in Section 2.10.5. However, the *C. sporogenes* proteolysis remained unaltered regardless of *clpX* expression in the tested strains (Figure 6.7).

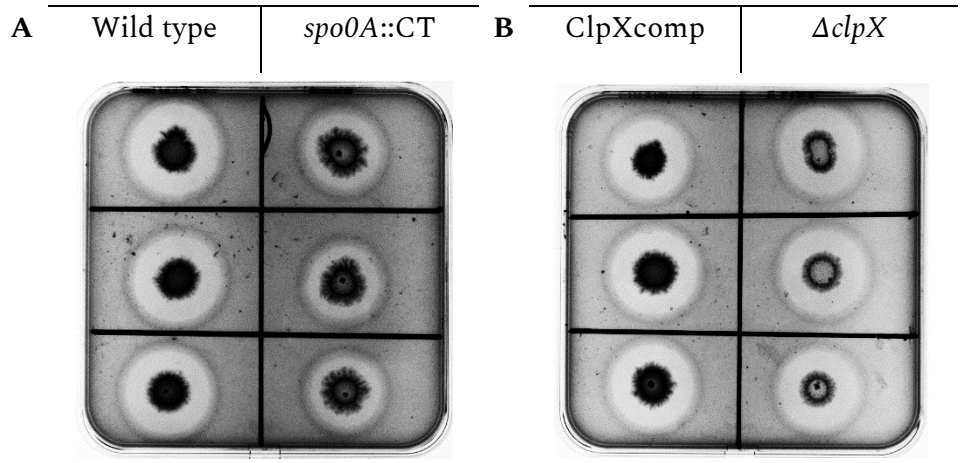


Figure 6.7. **Proteolysis by *C. sporogenes clpX* strains**. Mutant Δ*clpX* and the complemented strain ClpXcomp showed wild type level proteolytic activity. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT) on TYG agar supplemented with 4% (v/v) skimmed milk.

6.2.6. Microscopy reveals filamentous or swollen cells of Δ*clpX* mutants

To determine if *clpX* mutation affected cell morphology, 120 hour old *C. sporogenes* TYG broth cultures were examined under the microscope, as described in Section 2.5.1. Compared to the cells of *C. sporogenes* wild type and *spo0A*::CT strains, Δ*clpX* mutant cells appeared quite different, with various cell morphologies visible at high magnification, as shown in Figure 6.8. Moreover, spores could not be detected, despite searching many fields under the microscope. In contrast, *C. sporogenes* ClpXcomp cultures showed the restoration of both cell morphology and sporulation efficiency. Surprisingly, the complemented mutants repeatedly showed more mature spores under the microscope, than the in-house wild type strain.

Although functional spores were observed for *clpX*::TA mutants, the vegetative cells appeared bigger than those of the wild type and *spo0A*::CT strains. Moreover, spores of the later strain were variable in size and shape, and some were stained pink instead of blue.

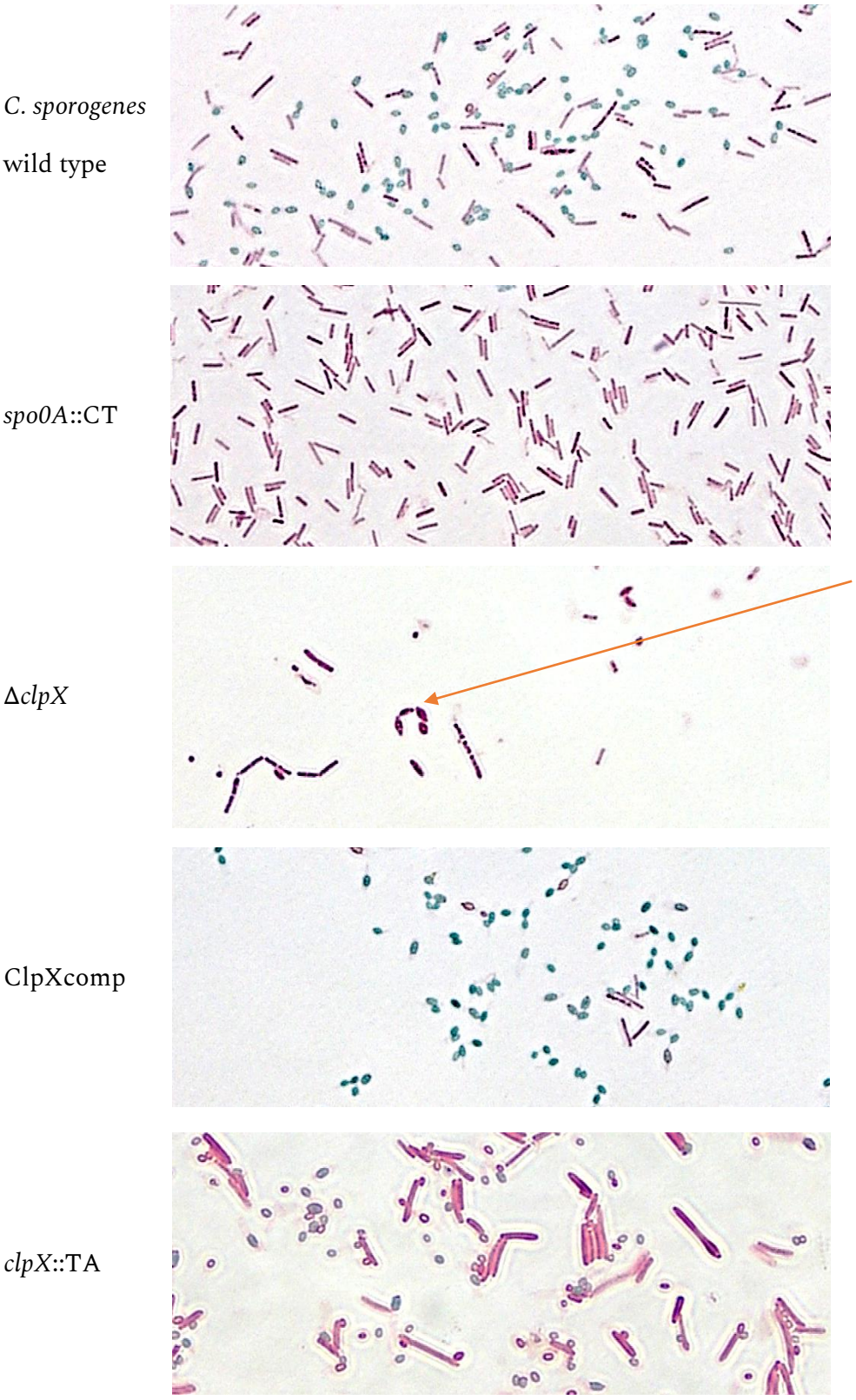
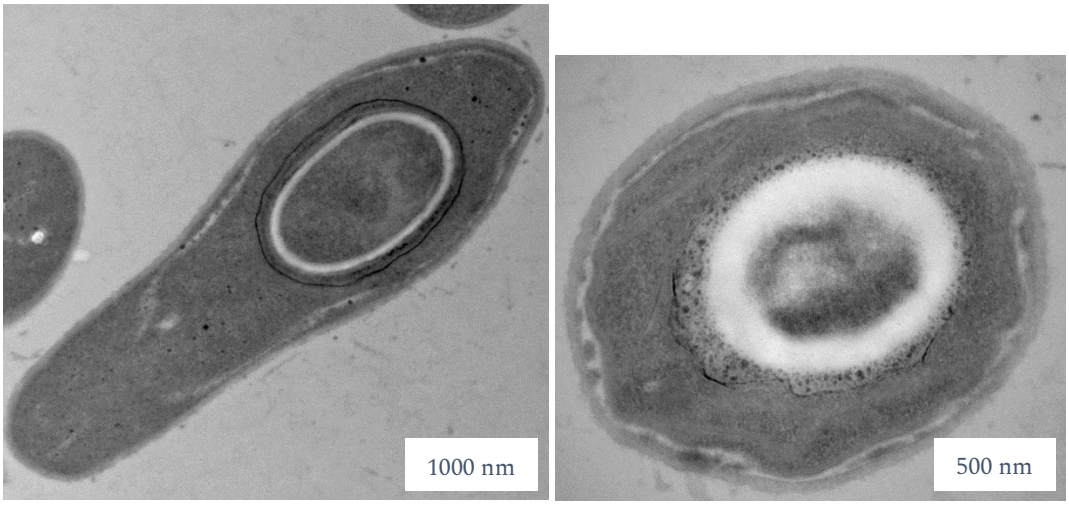


Figure 6.8. **Cell morphology of *C. sporogenes clpX* strains**. Mutant Δ*clpX* showed filamentous rods and cells of various shapes and sizes with few mature spores, unlike the *C. sporogenes* wild type strain. The complemented strain ClpXcomp showed restoration of wild type microscopic phenotype. In case of mutant *clpX*::TA, several spores of various sizes were observed, and some of them were poorly stained. Moreover, the *clpX*::TA rods appeared larger than the wild type.

Cells were stained pink and spores were stained blue. The images represent 120 hour old TYG broth cultures of *C. sporogenes* strains at 1000 × magnification.

TEM images of the *C. sporogenes* Δ*clpX* mutant showed appearance of normal forespore in the sporulating mother cell, as shown in Figure 6.9. Moreover, the potentially swollen cell phenotype was also perhaps observed.

C. sporogenes wild type



C. sporogenes Δ*clpX*

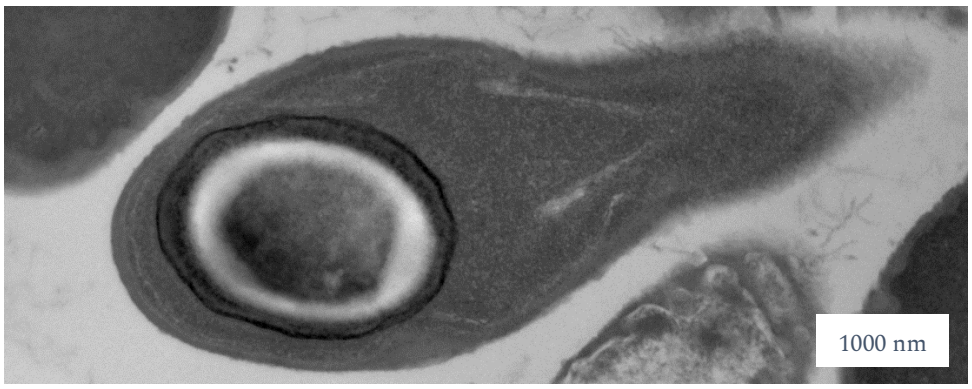


Figure 6.9. **Ultrastructure of *C. sporogenes clpX* strain**. Forespore of mutant Δ*clpX* showed no visible sporulation defect. The micrographs represent thin section of 72 hour old fixed cultures of *C. sporogenes* strains at high magnifications.

6.2.7. Δ*clpX* mutant colonies become translucent on old TYG agar plates

Since the colonies of transposon-based *clpX* mutants T43 and T45 become translucent on TYG plates (data not shown), it was decided to confirm this phenotype by streaking *C. sporogenes* Δ*clpX* mutants on TYG agar plates, as described in Section 2.10.6. As expected, Δ*clpX* mutants showed translucent phenotype on old agar plates, as shown in Figure 6.10, similar to their corresponding transposon mutants. Moreover, the Δ*clpX* mutant colonies started to become translucent within 24 hours, similar to the spo/ger$^-$ mutant *spo0A*::CT. ClpXcomp colonies were observed to have the same opaque colonies as the wild type strain, proving the complementation of the mutant phenotype.



Wild type | *spo0A*::CT | Δ*clpX* | ClpXcomp

Figure 6.10. **Colony morphology of *C. sporogenes clpX* strains**. Mutant colonies of *C. sporogenes* Δ*clpX* showed translucent phenotype within 24 hours, like the *spo0A*::CT mutant. The opaque growth phenotype was restored in the complemented ClpXcomp strain and the colonies appeared smooth, shiny, fimbriated and irregular, similar to the colonies of the *C. sporogenes* wild type strain. The images were taken after 27 hour incubation of *C. sporogenes* strains on TYG agar.

6.2.8. Illumina sequencing reveals SNPs in the Δ*clpX* mutant genome

In order to exclude the chances of misinterpretation of *clpX* gene's function in sporulation, it was necessary to know if there were any undesirable SNPs in other important gene loci of *C. sporogenes* Δ*clpX* mutant. Therefore, the mutant was sent for whole genome sequencing and the acquired sequenced data was analysed, as described in Section 2.3.8. When aligned with the *C. sporogenes* NCIMB 10696 reference strain from the NCBI genome databases, the sequenced data of Δ*clpX* mutant revealed the potential existence of additional SNPs along with the ones present in the in-house *C. sporogenes* wild type strain (Appendix Table A5). Three of

the additional SNPs were same as the ones found in Δ*yqfD* mutant, as mentioned in Section 4.2.8. These SNPs potentially affected the PocR ligand-binding protein encoding gene; the LysR regulator gene; and an unknown gene. Furthermore, locus for L-Aspargine-binding tRNA was also identified to contain three SNPs.

6.3. DISCUSSION

In this study, attempts were made to characterize the *clpX* gene for its role in the sporulation cascade of *C. sporogenes*. As already mentioned in Chapter Three, two of the sporulation-deficient mutants T43 and T45 had a transposon insertion in *clpX* gene, which reduced the wild type sporulation efficiency by >1000 fold (Figure 3.2). Therefore, three isogenic in-frame deletion Δ*clpX* mutants were generated by CRISPR/Cas9-based tools, to follow the basic characterization assay protocols.

6.3.1. Challenges in cloning and problem-solving techniques

As stated in the previous chapters, cloning CRISPR/Cas9-based knock-in cassettes in *E. coli* strains had its own share of problems, most of which were eventually eliminated by trial and error.

Although the use of low copy plasmid replicon p15a and short homology arms (0.54 kb/arm) enabled cloning of the *clpX*-knock-out cassette in the competent *E. coli* TOP10 and CA434 cells without any SNP problem, the *clpX* CDS-containing knock-in cassettes could not be successfully cloned in *E. coli* TOP10 cells without involving cell toxicity or undesirable mutations in the construct. Eventually, the use of 5-alpha F'Iq Competent *E. coli* cells for cloning, and incubation of transformants at room temperature solved the SNPs problem.

In addition, the SOEing PCR reaction to combine the watermark-based left and right homology arms for *clpX*-knock-in cassette did not generate a correct sized amplicon, despite several attempts at re-designing the primers as well as changing the primer

annealing temperatures. It was one of the two PCR reactions among all the many SOEing PCRs done in this study that consistently did not work and no cause could be identified. In the end, wild type homology arms, without an altered-codon watermark, were amplified by a simple PCR reaction, using the flanking primers clpxLHaat_F and clpxRH_R. This *clpX*-knock-in construct was then used to complement *C. sporogenes* Δ*clpX* mutants. Thus, the absence of a watermark in the resultant *C. sporogenes* ClpXcomp strains means that the possibility of the complemented strains to be contaminated with wild type or any other *C. sporogenes* cannot be completely ruled out.

6.3.2. The unexpected mutant phenotype from the Target-AID system

The size of *clpX* gene in the *C. sporogenes* genome is 1.29 kb with 31% GC content and 429 amino acids. The presence of a good proportion of purine and pyrimidine bases in the start of the gene made it an ideal sequence for the Target-AID base editing. The high efficacy of clpXsgC guide RNA paired with Cas9-AID complex successfully changed the 24th codon (Arg-CGA) to a STOP codon (TGA). As a result, several *clpX*::TA mutants were generated in *C. sporogenes* strain, as shown in Figure 6.3. However, surprisingly, none of the three tested isogenic strains of *clpX*::TA showed a sporulation defect (Figure 6.4), unlike T43 (Figure 3.2), T45 and Δ*clpX* mutants (Figure 6.4). Authenticity of the obtained results was also verified by sequencing the PCR amplicons of the heat-resistant colonies from spore assay plates to avoid any false interpretation of results.

The transposon insertion in sporulation-deficient mutants T43 and T45 is at 994^995 (anti-sense strand) and 812^813 bp (sense strand), respectively, as opposed to the SNP at the 70th base (sense strand) in the *clpX* gene of *clpX*::TA mutants. An explanation for the lack of effects seen on the sporulation profile of *clpX*::TA mutants could be the possible presence of another functional ORF in the middle of the *clpX* gene sequence.

From Figure 6.1, it can be seen that the main functional ATPase domain for ClpX activity starts from the 107th amino acid (~319th base), which made sense; but, the results obtained using the NCBI ORFfinder tool (Wheeler et al., 2003) did not show relevant evidence of another ORF (data not shown). The possibility of activation of a different gene, as a consequence of *clpX* inactivation, is also unlikely, since the Δ*clpX* mutants consistently show sporulation deficiency.

Regardless of the sporulation-efficient phenotype of the *clpX*::TA mutants, the translucent colony phenotype (Figure 6.3) and elongated vegetative cells, unlike the wild type strain were indeed observed (Figure 6.8). This indicates that *clpX* was somehow partially inactivated and that the protein domain that plays a role in the sporulation cascade of *C. sporogenes* was still functional in *clpX*::TA mutants.

6.3.3. Other unexpected mutant phenotypes

When 120 hour old broth cultures of *C. sporogenes* Δ*clpX* mutants were observed at high magnification, the nutrient-starved mutant cells appeared considerably different from the wild type cells. Cells of various shapes were observed, which included many elongated cells, ovoid cells and some filamentous rods (Figure 6.8). While elongated and filamentous phenotypes have been recorded in the previous studies of *clpX* mutants in *Bacillus* (Gerth et al., 1998), the observation of these peculiar swollen cells in isogenic *C. sporogenes* Δ*clpX* mutants has not previously been reported. The mechanisms involved in generation of this phenotype are presently unclear, but presumably are due to effects on mechanisms involved in cell division and branching of cells.

6.3.4. SNPs in the Δ*clpX* mutant

As mentioned in Section 4.3.5, three SNPs that were identified in the PocR gene, LysR regulator gene and an unknown gene in the genomes of Δ*yqfD* and Δ*spoVK* mutants, were also identified in the sequenced Δ*clpX* mutant. This phenomenon of common

SNPs might suggest the occurrence of spontaneous mutations during the overnight growth of *C. sporogenes* mutant cultures prior to DNA extraction. It might also reveal different sets of SNPs within the populations of same bacterial frozen stock, which could not be selected for genomic DNA extraction. Additionally, three more SNPs were identified in a $tRNA^{Asn}$, which can possibly be a false positive result due to low quality of submitted samples or any DNA damage during the transit.

6.3.5. Characterization of *clpX* and its role in sporulation

The role of *clpX* in the sporulation cascade of *C. sporogenes* was suggested after the observation of few mature spores by microscopy (Figure 6.8), as well as from the low number of heat-resistant *clpX* mutants CFUs on TYG plates compared to the wild type strain (Figure 6.4). As briefly described in Section 6.1, ClpX and ClpP proteins have been studied for their role in the regulation of stress tolerance in *B. subtilis* for the survival of bacterial species under unfavourable growth conditions (Gerth et al., 1998). Since sporulation is a mode of bacterial survival upon starvation stimuli, it is possible that the stress proteins, which are critical for survival under starvation, are inactive or repressed due to the absence of ClpX expression in *C. sporogenes* Δ*clpX* mutants; and consequently, sporulation defects are observed under such unfavourable conditions. In addition, the genomic DNA of Δ*clpX* mutants was observed to be more prone to degradation as compared to that of the sporulation-efficient strains (data not shown). This may have been due to the early decline of mutant cell population in broth cultures within 7 hours of inoculation. Therefore, 4 hour old fresh cultures were used to minimise DNA degradation prior its extraction from the mutated cells.

Inactivation of *clpX* in *C. sporogenes* generated metabolically active, viable cells. Although the log phase of *C. sporogenes* in the absence of functional ClpX proteins was observed to be similar to the wild type strain during the first 24 hours (Figure 6.5), the viable cell count started to gradually decline afterwards, and CFUs were visibly

lower than sporulation-efficient *C. sporogenes* strains after 120 hour incubation (Figure 6.4). This data indicates that *clpX* is neither an essential gene, nor plays a critical role in the exponential growth of *C. sporogenes*.

Moreover, the transition of opaque colonies of Δ*clpX* mutants into translucent colonies occurred earlier than the other sporulation-deficient mutants, which was a similar phenotype to that of the spo/ger$^-$ mutant *spo0A*::CT (Figure 6.8). This similarity might indicate additional roles of *clpX* in other important biological cascades of *C. sporogenes* as those recognized for *spo0A* (Fujita et al., 2005). Perhaps, the best reason to explain this translucent phenotype is the lack of tolerance of mutants to the suppressors that emerge in old cultures (Gerth et al., 1998).

Although no data about the haemolytic domain of ClpX protein has been documented for prokaryotes, it has still been known that mutations in the *clpX* gene of *S. aureus* affects virulence by impairing haemolysis and proteolysis gene cascades of the host bacterium (Frees et al., 2003). The haemolytic attributes of the pathogenic endospore-forming *B. anthracis* have also been shown to get nullified by the deletion of *clpX* in a previously tested strain (McGillivray et al., 2009). Therefore, similar impairment in the haemolysis was expected in *C. sporogenes* Δ*clpX* mutants, and also for which, relevant assay was conducted on blood-supplemented TYG agar plates. However, the strong β-haemolysis of *C. sporogenes* was observed to be unaffected even in the absence of a functional copy of *clpX* gene in mutant strains (Figure 6.6).

While the link between *clpX* and haemolysis is unknown; interestingly, it has been known that mitochondrial ClpX is involved in the efficient biosynthesis of haeme in *Saccharomyces cerevisiae*, via ClpP-independent pathway (Kardon et al., 2015). The same study also analysed a previously published data and observed the upregulation of *clpX* mRNA during the synthesis of red blood cells or erythropoiesis in human.

Similar was the case with the *C. sporogenes* proteolysis on milk-supplemented agar plates in this study (Figure 6.7). This data indicates that neither the haemolysis nor the proteolysis of *C. sporogenes* are dependent on the *clpX* gene. The strongly proteolytic *C. sporogenes* strains possess several other protease genes, which was why it was unsurprising to find no effects of *clpX* deletion on *C. sporogenes* proteolysis (Allison and Macfarlane, 1990).

From a query cover of 94% of protein sequence, the ClpX protein of *C. sporogenes* has 70.10% similar identity to the protein homologue in *B. subtilis*, as specified by the NCBI BLASTP tool. The sporulation-defective phenotype of *C. sporogenes* Δ*clpX* mutants in this study appeared to be similar to those observed in *B. subtilis* Δ*clpX* mutants (Gerth et al., 1998, Liu et al., 1999). Although Gerth (1998) observed no defects in the exponential growth of *B. subtilis* Δ*clpX* mutants, they noted colony lysis on plates following prolonged-incubation. Similar growth defects were also observed in a different study on *B. subtilis clpX* null mutants (Pan et al., 2001). The growth curve data and translucent colony phenotype generated for *C. sporogenes clpX* mutants in this study are in agreement with the equivalent *B. subtilis* mutants in the previous literature.

Intriguingly, a recent study identified the dependence of ClpX activity on the ClpXP proteasome, along with its independent role as a chaperone protein in *S. aureus* (Jensen et al., 2019). ClpX regulates the *sigB* transcriptional activity, as well as the regulation of several other genes involved in amino acid metabolism and transport or various other biological processes, via ClpXP-dependent and -independent cascades. While ClpX influences virulence of *S. aureus* solely in the form of the ClpXP complex, the ClpP-independent transcriptional role of ClpX has also been identified. A small number of genes involved in peptidoglycan biosynthesis, type VII secretion system or T7SS, and three genes of unknown function have also been identified in the *S. aureus clpX* regulon (Jensen et al., 2019).

6.3.6. Proposed function of *clpX* in *C. sporogenes*

Relevant characterization studies in the previous literature have enabled a better understanding of common functions of *clpX* in bacteria. In the light of the obtained data in this study, it can be assumed that *clpX* plays a role in the sporulation of *C. sporogenes* via a ClpXP-dependent or –independent pathway. Transcriptomics and SNAP-tag data of *C. sporogenes clpX* mutants would potentially reveal the actual mode of action of ClpX in sporulation of *C. sporogenes*. In addition, characterization of *clpX* mutants that are only deficient in ClpP-binding can further provide valuable data about the independent role of ClpX in various biological processes of clostridial species.

CHAPTER SEVEN

7. CHARACTERIZATION OF *spoVK*

7. CHARACTERIZATION OF *SPOVK*

7.1. INTRODUCTION

ATPases or adenosine 5'-triphosphatases are enzymes that catalyse the hydrolysis of ATP to ADP in order to free a phosphate ion. This process releases energy to drive DNA replication, DNA repair, gene expression and other protein machineries of the cell (Geider and Hoffmann-Berling, 1981, Kornberg and Baker, 1992, Koonin, 1993a). ATPases also play a large role in assisting with the protein assembly, proteolysis and transport in bacteria (Gottesman, 1996, Wawrzynow et al., 1996, Suzuki et al., 1997, Neuwald et al., 1999).

Among the newly isolated mutants from the *C. sporogenes* transposon library screenings, mutant T19 showed no heat-resistant colonies on TYG agar, as briefly mentioned in Figure 3.3 and Table 3.1. Transposon insertion site was later identified in CLSPO_RS17345, an ATPase encoding gene.

According to the NCBI genome databases, CLSPO_RS17345 is a 1092 bp long gene in the *C. sporogenes* NCIMB 10696 strain, with 27% GC content. Its gene product (WP_033061400.1) is annotated as "ATPase", which is made up of 363 amino acids, and possesses an alternate start codon TTG instead of ATG. Browsing the neighbouring gene loci, it can be predicted that CLSPO_RS17345 is a monocistronic gene.

Common annotations from the BLASTP results of ATPase protein sequence against several *Bacillus* and clostridial species included: AAA family ATPase, ATPase, ATP- or GTP-binding protein; mother cell sporulation ATPase, SpoVJ, SpoVK; DNA helicase, RuvB, RecD, DnaC; LonB, ClpA-like; and hypothetical protein. All the suggested annotations obtained from the protein BLAST results of the

CLSPO_RS17345 gene product correspond to the available literature with respect to the conserved domains and motifs of the ATP- or GTP-binding AAA family NTPases.

AAA stands for "ATPases associated with a variety of cellular activities" (Kunau et al., 1993). AAA family NTPase chaperons regulate and assemble protein complexes (Neuwald et al., 1999). Generally, the P-loop NTP-binding site of NTPases consists of conserved α and β subunits, which are responsible for ATP-binding and catalysis reactions (Walker et al., 1982). Moreover, the AAA family NTPases have other motif signatures and clusters, in addition to these Walker A and B motifs. Such NTPases constitute major groups namely, ATPases, GTPases, helicases, etc. (Koonin, 1993a, Koonin, 1993b, Gorbalenya and Koonin, 1989). Some regulatory NTPases, like RuvB helicase, Lon and Clp proteases have a human replication factor C or RFC box motif (Cullmann et al., 1995), and are known to be involved in DNA replication, recombination, restriction and transcription (Gottesman, 1996, Wawrzynow et al., 1996, Suzuki et al., 1997, Guenther et al., 1997, Neuwald et al., 1999). Generally, AAA NTPase families are well-conserved in all bacteria; however, some families, like the NtrC family that includes SpoVJ or SpoVK protein, are only found in some bacteria (Neuwald et al., 1999, Galperin et al., 2012).

Some examples can be found in the literature regarding the role of ATPases in the sporulation cascade of some of the known endospore-forming bacteria. For instance, during chromosomal segregation of *B. subtilis*, MinD ATPase is involved in the localization and stimulation of the cell division inhibitory activity of MinC protein (de Boer et al., 1991). The ATPase domain in SpoIIIAA protein has been found to be responsible for σ^G activation and the maintenance of forespore integrity in *B. subtilis* (Doan et al., 2009). Similarly, ATPase domain of *B. subtilis* SpoIIIE is required for transporting the remaining 70% of chromosomal DNA from the mother cell to the emerging forespore (Wu and Errington, 1994, Burton et al., 2007). Additionally, ClpX is a regulatory ATPase, which not only functions as a chaperone for the ClpP protease

to aid its proteolytic activity, but also has a role in the initiation of *B. subtilis* sporulation (Liu et al., 1999).

Since the T19 mutant phenotype was closest to that of *spoVK* or *spoVJ* mutants in *B. subtilis*, it was decided to consider *C. sporogenes* CLSPO_RS17345 gene to be a homologue of *spoVK* gene of *B. subtilis*. Therefore, the term "*spoVK*" will be used throughout this thesis to define *C. sporogenes* CLSPO_RS17345 gene.

The *spoVK* gene is present in all endospore-forming *Bacillus* and clostridial spp. (Galperin et al., 2012). Interestingly, it is absent in non-sporulating bacteria (Koo et al., 2017). Moreover, *spoVK* has also been known to be absent in the endospore-forming *Candidatus Arthromitus* (Kuwahara et al., 2011), which is a filamentous member of Firmicutes, and identified to contain a minimalistic set of sporulation gene set in its small-sized genome (Galperin et al., 2012).

Despite the insufficiency of the currently available data on *spoVK* ATPase, a recently obtained RNA sequencing data revealed the downregulation of the *spoVK* gene expression in *C. sporogenes* Δ*sigE*, Δ*sigG* and Δ*sigK* mutants (Raquel Rodrigues, unpublished), possibly hinting about the existence of *spoVK* in the *sigE*, *sigG* and *sigK* regulons. Some part of this data corresponds with the previous transcriptomic studies, which found *spoVK* to be a part of the *sigE* regulon in *B. subtilis* (Feucht et al., 2003, Eichenberger et al., 2003), as well as in *C. difficile* (Saujet et al., 2013). Moreover, *spoVK* transcription is compartment-specific in *B. subtilis* and dependent on the SpoIIID protein (Eichenberger et al., 2003). However, *spoVK* transcription has also been found to occur in a SpoIIID-independent manner (Eichenberger et al., 2004). Two promoters of *spoVK* gene have been identified in *B. subtilis*, in which the first promoter P1 is controlled by σ^E and the second promoter P2 is controlled by σ^K (Foulger and Errington, 1991). The existence of dual promoters might indicate the consistent requirement of the protein throughout a biological process, despite the change in the presently active transcriptional factor (Foulger and Errington, 1991). In

general, the expression of some gene products gets halted after a certain stage or process, because the accumulation of that protein might otherwise become inhibitory at a different stage (Foulger and Errington, 1991), and the production of an unrequired protein is also a waste of valuable resources. Nevertheless, the transcriptomics data of *C. sporogenes* mutants in the previous study potentially validates the existence of *spoVK* in both *sigE* and *sigK* regulons.

In order to discover similar and more accurate gene homologues of *spoVK* in other endospore-formers, experiments were designed to characterize gene mutants for better understanding the role of *spoVK* gene in *C. sporogenes*.

7.1.1. Aim of this chapter

This chapter is aimed at studying the CLSPO_RS17345 *spoVK* gene for its potential role in the sporulation of *C. sporogenes*, by the following steps:

i. Constructing in-frame deleted Δ*spoVK* mutants in the *C. sporogenes* NCIMB 10696 wild type strain and later complementing the mutated gene;

ii. Characterizing the Δ*spoVK* mutants by determining the relationship of *spoVK* gene with sporulation, growth, haemolysis, proteolysis, and cell and colony morphology of *C. sporogenes*, and comparing the mutant phenotype with the wild type phenotype;

iii. Finding relevant literature on the prokaryotic *spoVK* gene and its SpoVK protein and comparing the phenotypes of the *C. sporogenes* Δ*spoVK* mutants with the ones generated in *Bacillus* and other clostridial species;

iv. Proposing the potential role of *spoVK* in the biology of *C. sporogenes*, based on the data generated in this study, with reference to earlier studies.

7.2. RESULTS

7.2.1. CRISPR/Cas9-based genome editing in the *spoVK* gene

To confirm the phenotype of mutant T19 and characterize the *spoVK* gene, Δ*spoVK* mutants were generated in triplicates from the *C. sporogenes* wild type strain, using the botsgRNA1 bookmark-based pRECas1-spoVK knock-out cassette, as described in Section 2.7. Δ*spoVK* mutants were then complemented using the watermark-based pRECas1-SpoVKCO knock-in cassette. The resulting complementing *C. sporogenes* SpoVKCOs strains and their respective Δ*spoVK* mutants were used for characterization after confirmation of plasmid loss.

The use of Target-AID base editor for mutant generation was also considered in order to avoid the potential risk of SNP occurrence in the knock-out cassette during cloning in *E. coli*. However, no guide RNA sequence could be determined for targeted mutagenesis, due to the lack of suitable pyrimidine-based codons in the start of the *spoVK* gene.

7.2.2. Mutation in *spoVK* affects *C. sporogenes* sporulation

The sporulation profile of *C. sporogenes* Δ*spoVK* mutants was studied by incubating the strains in TYG broth for 5 days, during which samples were withdrawn daily and checked for heat-resistance, as described in Section 2.10.2. Similar to mutants *spo0A*::CT and T19, Δ*spoVK* gave no heat-resistant CFU (Figure 7.1). Conversely, complemented strain SpoVKCOs sporulated ~1×10 times lower than the *C. sporogenes* wild type strain. Moreover, SpoVKCOs sporulation was 24 hours delayed.

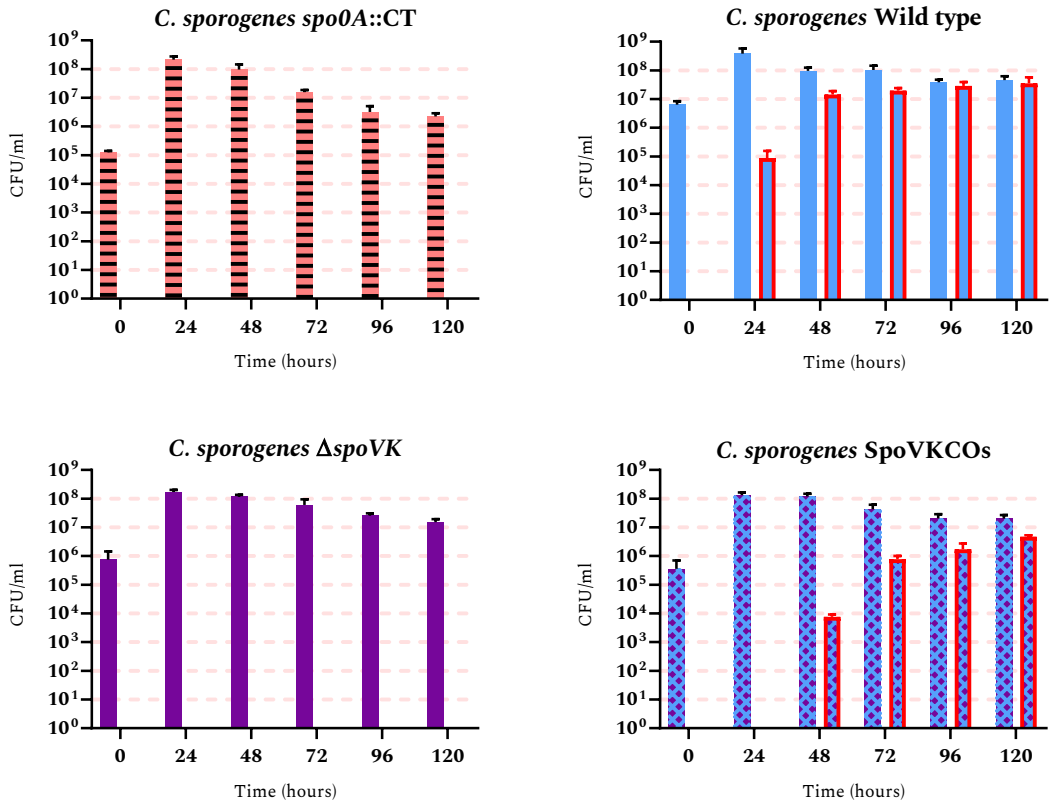


Figure 7.1. **Spore assay of *C. sporogenes spoIID* strains**. Mutant Δ*spoVK* showed no heat-resistant CFUs, similar to the *spo0A*::CT control strain. Whereas, the complemented strain SpoVKCOs showed comparable heat-resistant CFUs to *C. sporogenes* wild type. This data is based on the 5-day incubation of three biological replicates (except for *spo0A*::CT and SpoVKCOs) in TYG broth. Error bars show mean with standard deviation. Each strain is represented by a unique colour[4]. Red-bordered bars represent heat-resistant CFUs, whereas coloured bars without red border represent the untreated or the total cell counts. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4×10^9.

[4] Colour key:



7.2.3. *spoVK* is not an essential gene in *C. sporogenes*

To determine the effect of mutation in *spoVK* gene on the growth of *C. sporogenes*, cultures were grown in fresh TYG broths and optical density was measured regularly at several time points, as described in Section 2.10.3. It was observed that both Δ*spoVK* and SpoVKCOs strains entered the exponential growth around the same time as the *C. sporogenes* wild type strain; however, they both entered the cell death phase earlier without going through a distinctive stationary phase (Figure 7.2).

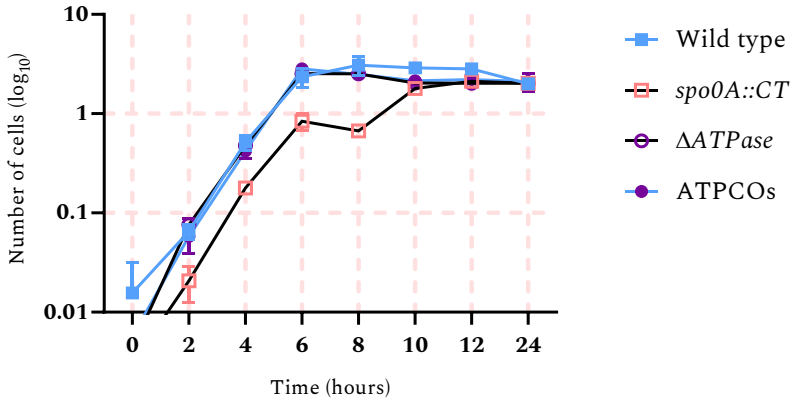


Figure 7.2. **Growth curve of *C. sporogenes spoVK* strains**. Mutant Δ*spoVK* and the complemented strain SpoVKCOs, showed similar exponential growth to *C. sporogenes* wild type strain, in contrast to the spo/ger⁻ mutant *spo0A*::CT, which has a slight growth defect. However, both *spoVK* strains started to decline earlier than the wild type. This data is based on the 24 hour incubation of three biological replicates (except for *spo0A*::CT and SpoVKCOs) in TYG broth. Error bars show mean with standard deviation. Theoretically, 0.01 OD_{600} was the starting inoculum, which was individually calculated for each strain; however, in practice, variable readings were observed at 0 hour.

7.2.4. *C. sporogenes* β-haemolysis might require *spoVK*

It was interesting to know if the lack of *spoVK* gene expression in *C. sporogenes* strains also causes any defects in the β-haemolysis on blood agar plates, as seen for the *spo0A*::CT mutant. Except for one replicate, the Δ*spoVK* mutants showed weaker haemolytic activity than the wild type strain when spotted on TYG-blood agar plates, as mentioned in Section 2.10.4. Unexpectedly, this activity was even weaker for the cultures of SpoVKCOs strains (Figure 7.3), possibly due to additional SNPs.

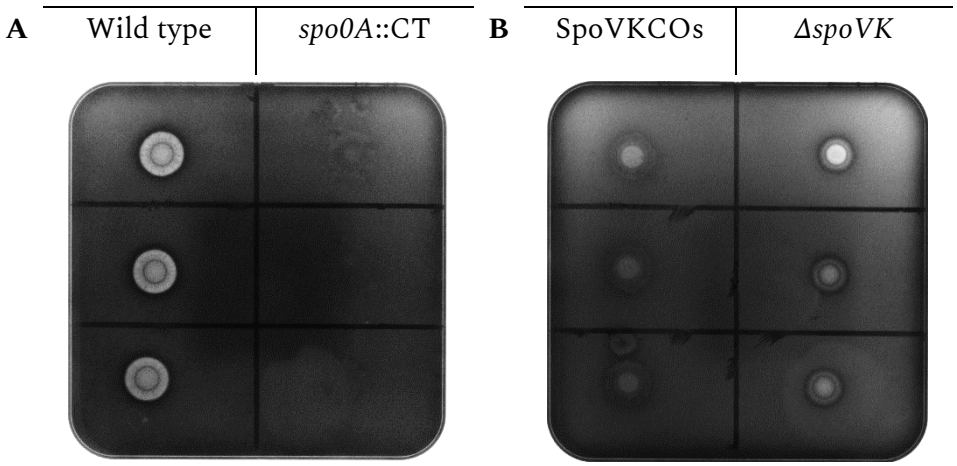


Figure 7.3. **Haemolysis by *C. sporogenes spoVK* strains**. Mutant Δ*spoVK* showed weaker haemolytic activity than the *C. sporogenes* wild type strain. In addition, the complemented strain SpoVKCOs showed the weakest β-haemolysis. The spo/ger⁻ mutant *spo0A*::CT was non-haemolytic and its translucent growth was overshadowed by dark-coloured blood agar. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT and SpoVKCOs) on TYG agar supplemented with 5% (v/v) horse blood.

7.2.5. Proteolytic activity of *C. sporogenes* might require *spoVK*

On milk TYG agar plates, wild type *C. sporogenes* strain showed zones of clearance within 24-48 hours (Figure 7.4). To determine if the proteolytic activity of *C. sporogenes* was affected in Δ*spoVK* mutants, all the relevant strains were grown on milk-based TYG agar plates, as described in Section 2.10.5. Intriguingly, Δ*spoVK* mutant replicates as well as SpoVKCOs strains showed slightly smaller zones of clearance than the *C. sporogenes* wild type strain, as shown in Figure 7.4, but similar to the spo/ger⁻ mutant *spo0A*::CT.

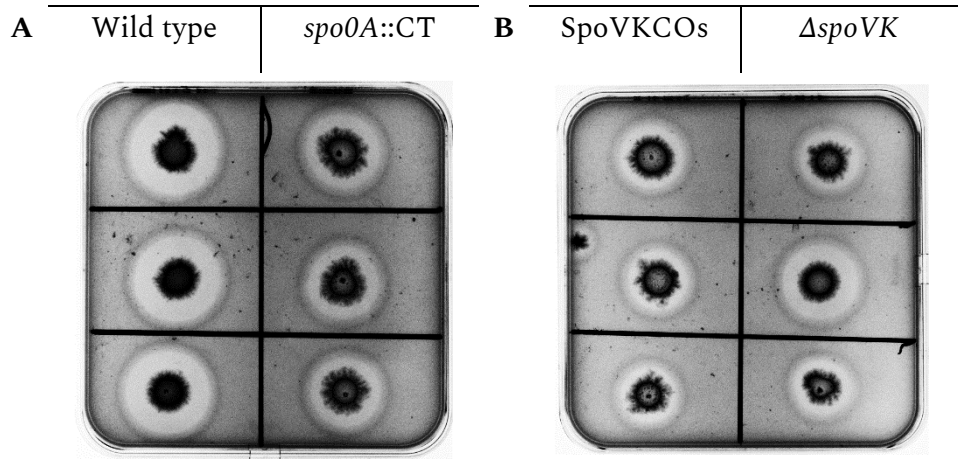


Figure 7.4. **Proteolysis by *C. sporogenes spoVK* strains**. Mutant Δ*spoVK* and the complemented strain SpoVKCOs showed comparatively weak proteolysis, in contrast to *C. sporogenes* wild type, but similar to the spo/ger$^-$ mutant *spo0A*::CT. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT) on TYG agar supplemented with 4% (v/v) skimmed milk.

7.2.6. Microscopy reveals numerous enlarged spores of Δ*spoVK* mutants

In order to compare the effects of presence and absence of *spoVK* gene on the morphology of *C. sporogenes* cells, 120 hour old TYG broth cultures were investigated under the microscope, as described in Section 2.5.1. Unexpectedly, all three Δ*spoVK* mutants showed a myriad of large-sized spores, with unique morphology, as shown in Figure 7.5, despite the lack of any heat-resistant CFU on TYG agar. Interestingly, SpoVKCOs cultures had the same spore morphology as the Δ*spoVK* mutants, but were somehow able to resist heat-shock.

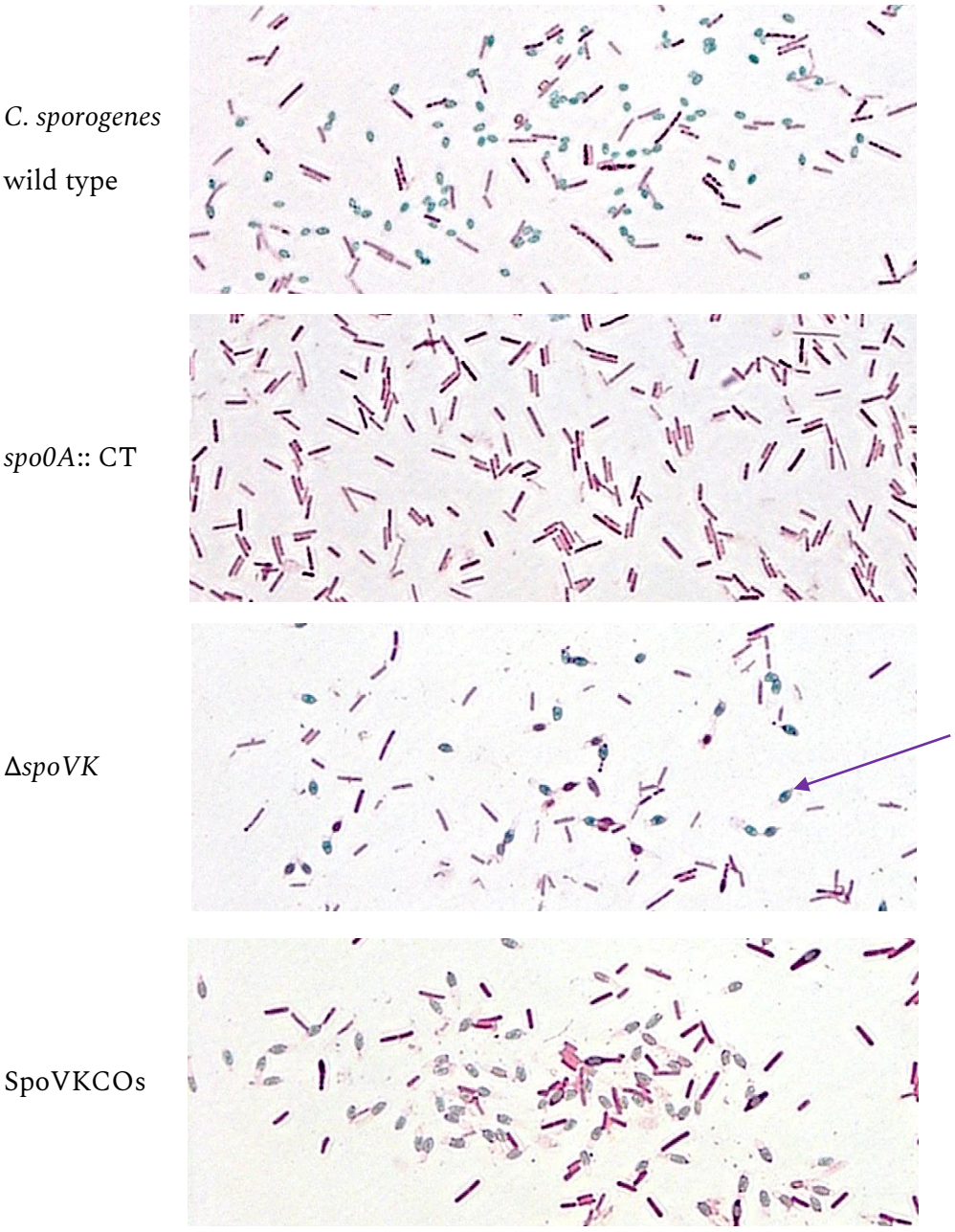
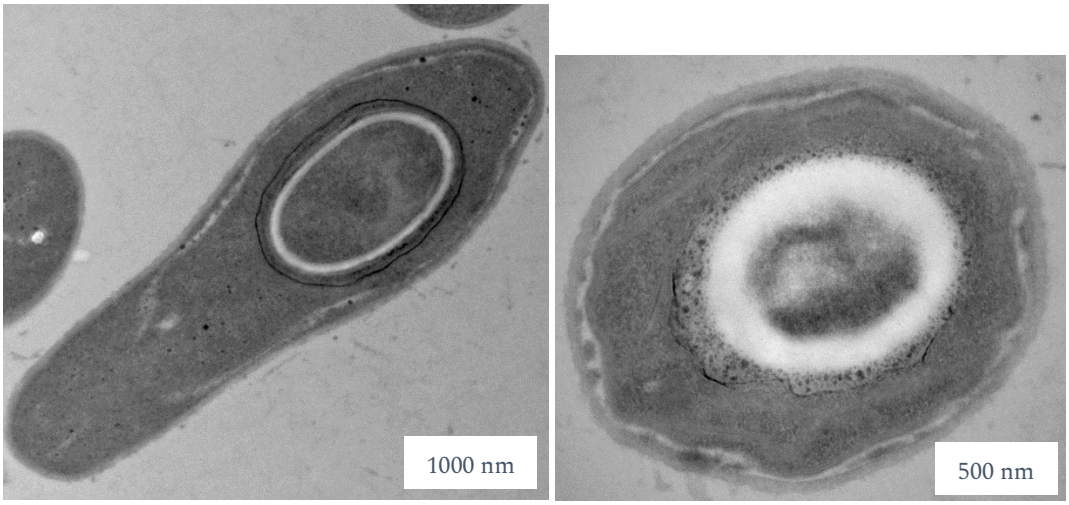


Figure 7.5. **Cell morphology of *C. sporogenes spoVK* strains**. Mutant Δ*spoVK* showed rods with numerous spores, similar to the *C. sporogenes* wild type strain. However, the mutant spores not only appeared larger than wild type spores, but gave a characteristic "pupil in a human eye" phenotype, as indicated by an arrow. Complemented strain SpoVKCOs also showed numerous large-sized spores. Cells were stained pink and spores were stained blue. The images represent 120 hour old TYG broth cultures of *C. sporogenes* strains at 1000 × magnification.

TEM images of *C. sporogenes* Δ*spoVK* mutant showed the same oval-shaped cell and spore phenotype, as observed with the light microscope (Figure 7.6). The mutant forespore and spore appeared to have a misassembled layer of coat proteins compared to the wild type forespore and spore, as shown in Figure 7.6.

C. sporogenes wild type



C. sporogenes Δ*spoVK*

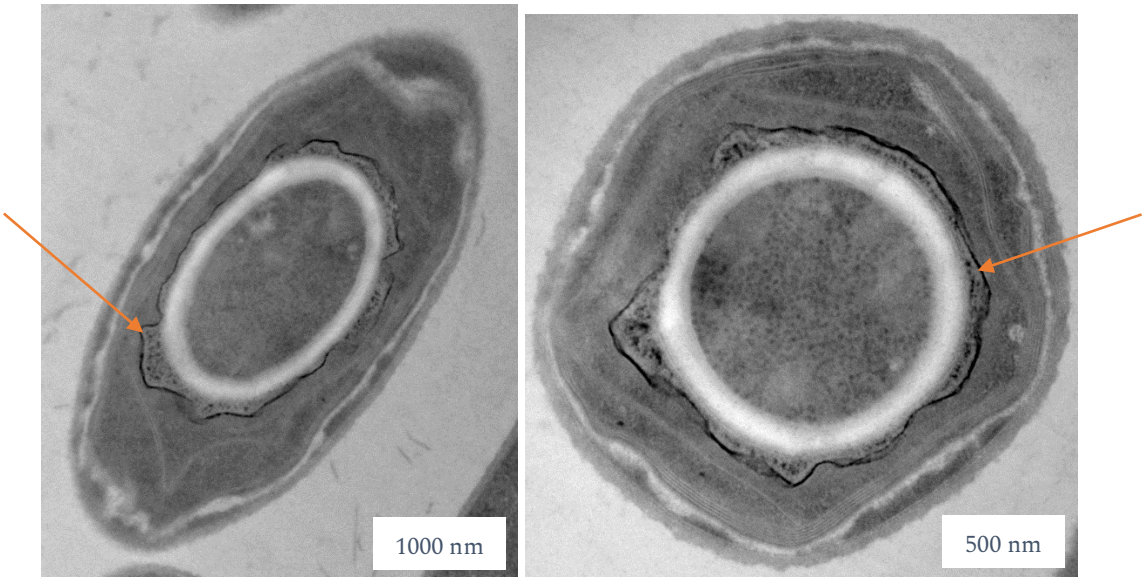


Figure 7.6. **Ultrastructure of *C. sporogenes spoVK* strain**. The spore and forespore of mutant Δ*spoVK* showed misassembled spore coat proteins, as indicated by arrows. The micrographs represent thin section of 72 hour old fixed cultures of *C. sporogenes* strains at high magnifications.

7.2.7. Δ*spoVK* mutant colonies become translucent on old TYG agar plates

Similar to mutant T19 (data not shown), *C. sporogenes* Δ*spoVK* mutants become translucent within 72 hours on old agar plates. However, this translucent phenotype was not observed to be fully complemented in the SpoVKCOs strain, as shown in Figure 7.7.



Wild type | *spo0A*::CT | Δ*spoVK* | SpoVKCOs

Figure 7.7. **Colony morphology of *C. sporogenes spoVK* strains**. Mutant colonies of *C. sporogenes* Δ*spoVK* showed translucent phenotype on old agar plates, similar to the *spo0A*::CT mutant. The wild type colony morphology could not restored in the complemented SpoVKCOs strain. However, colonies of both strains appeared smooth, fimbriated and irregular, just like those of the *C. sporogenes* wild type strain. The images were taken after 27 hour incubation of *C. sporogenes* strains on TYG agar.

7.2.8. Illumina sequencing reveals SNPs in the Δ*spoVK* mutant genome

To detect the presence of SNPs in the genome of *C. sporogenes* Δ*spoVK* mutant, genomic DNA from the mutant strain was sent for whole genome sequencing, as described in Section 2.3.8, and the sequenced data was compared with that of the in-house *C. sporogenes* wild type strain. When aligned with the *C. sporogenes* NCIMB 10696 reference strain from the NCBI genome databases, the sequenced data of Δ*spoVK* mutant revealed the potential existence of SNPs in $tRNA^{Asn}$, PocR ligand-binding protein-encoding gene, the LysR regulator gene, and two unknown genes, in addition to the seventeen SNPs that are present in the in-house *C. sporogenes* wild type strain (Appendix Table A6).

7.3. DISCUSSION

In this study, attempts were made to characterize another newly identified gene, which resides on the CLSPO_RS17345 locus in the *C. sporogenes* chromosome and encodes for an ATPase. Since there are at least 22 unknown *C. sporogenes* genes whose products are annotated in the NCBI genome databases as "AAA family ATPase" or simply "ATPase", it was decided to designate the gene target of this study with a new name, in the first step of its characterization. The sporulation phenotype of the *C. sporogenes* mutants, generated in this study, matched with the phenotype of *B. subtilis spoVK* mutants generated in the earlier studies. Therefore, the gene nomenclature was decided accordingly and the targeted ATPase gene was given the annotation "*spoVK*". This *spoVK* gene was initially discovered during the screenings for sporulation-defective phenotypes in *C. sporogenes* transposon libraries, when a spo/ger$^-$ mutant T19 showed transposon insertion in the CLSPO_RS17345 locus (Figure 3.2). To prevent the chances of polar effects in mutant T19 that could lead to false interpretation of this finding, it was decided to make "clean" in-frame deletion mutants. Hence, three isogenic strains of *C. sporogenes* Δ*spoVK* mutants were generated using CRISPR/Cas9-based genome editing, which were characterized without the presence of gene CDS, as well as after knocking the *spoVK* CDS back into the *C. sporogenes* genome.

7.3.1. Challenges in cloning and problem-solving techniques

Similar to the problems faced with the other characterized gene targets mentioned in the preceding chapters of this thesis, cloning the *spoVK*-knock-in cassette for CRISPR/Cas9-based genome editing was not the easiest tasks to accomplish. While the first attempt at using shorter homology arms and low copy number plasmid enabled the cloning of pRECas1-Erm-spoVK construct in the competent *E. coli* TOP10 cells for *C. sporogenes* Δ*spoVK* mutant generation, the same protocol was not

efficient for constructing a SNP-free *spoVK*-knock-in cassette. Only one *E. coli* TOP10 transformant could be obtained that harboured the pRECas1-Erm-SpoVKcomp construct, but with SNPs. Although this plasmid construct was able to generate the *C. sporogenes* SpoVKcomp strains through efficient botsgRNA1-bookmark-based targeting and homology-directed repair or HDR (Figure 7.8.A), the SpoVKcomp strains did not show wild type level of sporulation efficiency, as shown in Figure 7.8.B. This may have been due to the presence of SNPs in the SpoVKcomp construct, as identified by DNA sequencing.

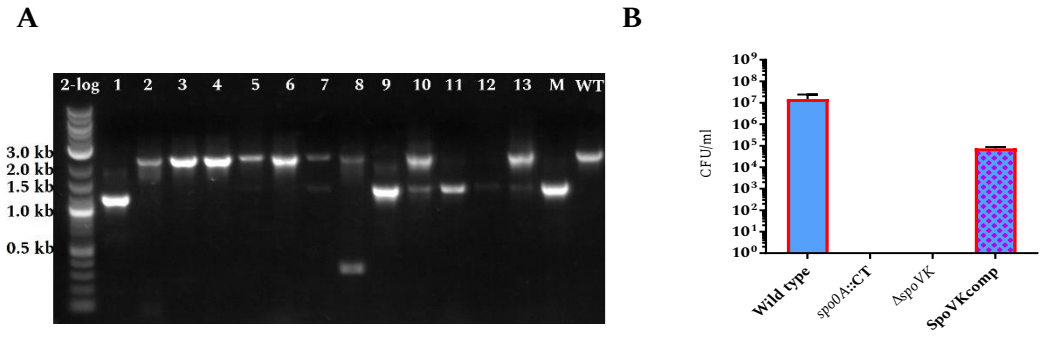


Figure 7.8. **Complementation of *spoVK* gene and restoration of sporulation in *C. sporogenes* mutants**. A) Agarose gel of PCR amplicons of the theophylline-induced *C. sporogenes* Δ*spoVK*-pRECas1-Erm-SpoVKcomp transconjugants (lanes 1-13) as compared to the amplicons of Δ*spoVK* (lane M: 1.4 kb) and wild type (lane WT: 2.4 kb) strains, and the 2-log DNA ladder. Colonies analyzed in lanes 2-7, 10 and 13 showed successful partial/complete HDR-based gene complementation; whereas, colonies in lanes 1, 9, 11 and 12 showed no HDR. B) Spore assay of *C. sporogenes spoVK* strains revealed partial complementation of sporulation defect of mutant Δ*spoVK* in SpoVKcomp strains. Sporulation-efficient *C. sporogenes* wild type and spo/ger$^-$ mutant *spo0A*::CT were used as control strains. This data is based on 120 hour incubation of two to three biological replicates (except for *spo0A*::CT) in TYG broth. Error bars show mean with standard deviation. Red-bordered bars represent heat-resistant CFUs. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4×10^9.

Therefore, cloning conditions were improved to obtain *E. coli* transformants lacking any undesirable mutation in the plasmid construct. 5-alpha F'Iq Competent *E. coli* cells and room temperature incubation were integrated into the protocol. The resulting *C. sporogenes* SpoVKCOs strains showed higher heat-resistant CFUs (Figure 7.1) than the previously-generated SpoVKcomp strains (Figure 7.8.B). However,

SpoVKCOs heat-resistant CFUs were still ~1×10 lower than the wild type CFUs, and sporulation in the engineered strains was also delayed for 24 hours. Moreover, the slight growth defect of Δ*spoVK* mutants was not complemented in SpoVKCOs strains, as portrayed in Figure 7.2. Because the HDR for the watermark-based *spoVK*-knock-in cassette pRECas1-SpoVKCOs was inefficient with the botsgRNA1 bookmark-based CRISPR/Cas9 system, it took several sub-culturing events on theophylline-supplemented TYG plates in order to induce Cas9 activity. This procedure might have led to unwanted SNPs in the genome of SpoVKCOs strains. To further verify the complementation of SpoVK protein function, pMTL87121:*spoVK* plasmid with the CLSPO_RS17345 CDS + RBS+ promoter cassette was used to allow expression of SpoVK in *C. sporogenes* Δ*spoVK* mutants. The sporulation profile of the resultant *C. sporogenes* ProSpoVK strains fully matched with those of the SpoVKCOs strains (data not shown), which suggests the possibility of additional SNPs in Δ*spoVK* mutants rather than problems with the complemented SpoVKCOs strains.

7.3.2. SNPs in the Δ*spoVK* mutant

As mentioned in Section 4.3.5, SNPs that were identified in the PocR- and LysR regulator-encoding genes of Δ*yqfD*, along with the $tRNA^{Asn}$ loci of Δ*clpX* mutant, were also identified in the sequenced Δ*spoVK* mutant strain. This phenomenon of similarly occurring SNPs might suggest the occurrence of spontaneous mutations during the overnight growth of *C. sporogenes* mutant cultures prior to DNA extraction. It might also reveal different sets of SNPs within the populations of same bacterial frozen stock, which could not be selected for extraction of genomic DNA. Nevertheless, none of the identified genes have been known for their role in the sporulation process.

7.3.3. Characterization of *spoVK* and its role in sporulation

The absence of heat-resistant CFUs in Δ*spoVK* mutants (Figure 7.1) suggested a defect in either sporulation or germination. Interestingly, numerous spores were observed for these mutant cultures under the microscope (Figure 7.5), supporting a germination defect. However, the morphology of the Δ*spoVK* mutant spores appeared slightly different than that of the wild type spores. When examined more closely, the Δ*spoVK* mutant spores appeared to have an oval shape with sharp edges and a dot in the centre, which almost resembled the pupil in a human eye. In contrast, mature spores of *C. sporogenes* wild type are generally more rounded and appear hollow in the centre. This suggests the possibility of disruption in either the later stages of sporulation or some defects in the germination. Therefore, attempts were made to germinate the Δ*spoVK* mutant spores; however, no change was observed in response to L-alanine (data not shown). This might have reflected some defects in the germinant receptor for L-alanine binding; however this theory was disregarded due to visual defects in the outer layers of the mutant spores, as portrayed by the TEM micrographs (Figure 7.6).

The spore assay data (Figure 7.1) showed no growth defects, unlike the CFUs of *spo0A*::CT and other sporulation-defective mutants, as mentioned in the former chapters. Since cells are able to sporulate, cell lysis is avoided in old cultures, similar to the wild type strain. Moreover, these higher cell counts might also suggest that the untreated spores were able to germinate as well, but not after being heat-shocked. On the other hand, the OD_{600} values during the exponential growth of the Δ*spoVK* mutants started to decline some hours before the wild type strain (Figure 7.2), and this phenotype could not be complemented in SpoVKCOs by incorporating the *spoVK* gene back into the genome. This suggests the presence of other spontaneous mutations in the genome of Δ*spoVK* mutants, which were also confirmed by the whole genome sequencing results that are listed in Appendix Table A6. Nevertheless, on the

basis of this data, it can be assumed that *spoVK* is neither an essential gene nor involved is involved in the exponential growth of *C. sporogenes*.

With the limited data on the exact function of the protein product of the *spoVK* gene, it was challenging to find protein homologues in *Bacillus* spp. So, a known sporulation ATPase, SpoVJ or SpoVK protein was analysed to compare the data generated in this study with that available in the literature. In previous studies, *spoVJ* or *spoVK* mutants in *B. subtilis* showed phase-dark spores under the microscope, and were sensitive to heat-treatment without any defects in the vegetative growth, similar to *C. sporogenes* Δ*spoVK* mutants in this study (Hill, 1983, Fan et al., 1992). Moreover, in a different study on *B. subtilis spoVJ* it was found that deletion of *spoVJ* promoter also interfered with the heat-resistance of the mutants (Foulger and Errington, 1991). It was inferred that SpoVJ is important in the heat-resistance of *B. subtilis* spores and its activity is ATP-dependent. Moreover, no DPA was detected in the forespores, despite the production of DPA in the sporulating mother cell, which hints a potential role of SpoVJ or SpoVK protein in the DPA transport (Foulger and Errington, 1991).

The weak ability of Δ*spoVK* mutant cultures as well as of SpoVKCOs strains to show β-haemolytic activity in the haemolysis assay were initially considered to be not representing true phenotypes of Δ*spoVK* mutants. This assay, along with proteolysis was repeated with the T19 mutant instead, to see if the disruption of *spoVK* gene with transposon insertion can affect haemolytic and proteolytic efficiencies of mutant *C. sporogenes*. Inactivation of *spoVK* gene in mutant T19 also resulted in the reduction of both haemolytic and proteolytic activities, shown in Figure 7.9. This finding indicates that the CLSPO_RS17345 *spoVK* may have some role in the *C. sporogenes* haemolysis and proteolysis. ATPases are generally known to chaperone ATP-dependent proteases in protein degradation (Neuwald et al., 1999). The observance of smaller zones of clearance on milk TYG agar, compared to the *C. sporogenes* wild type strain appears to be an expected finding. Previous studies on the deletion of gene for

ATPase subunit ClpX have revealed the negative influence on both haemolysis and proteolysis of the host bacteria (Frees et al., 2005, McGillivray et al., 2009), which might have explained a little why both of these processes were observed to be affected in this study. However, *clpX* deletion did not affect *C. sporogenes* haemolysis (Figure 6.6) or proteolysis (Figure 6.7) in this study, which indicates an unknown mechanism.

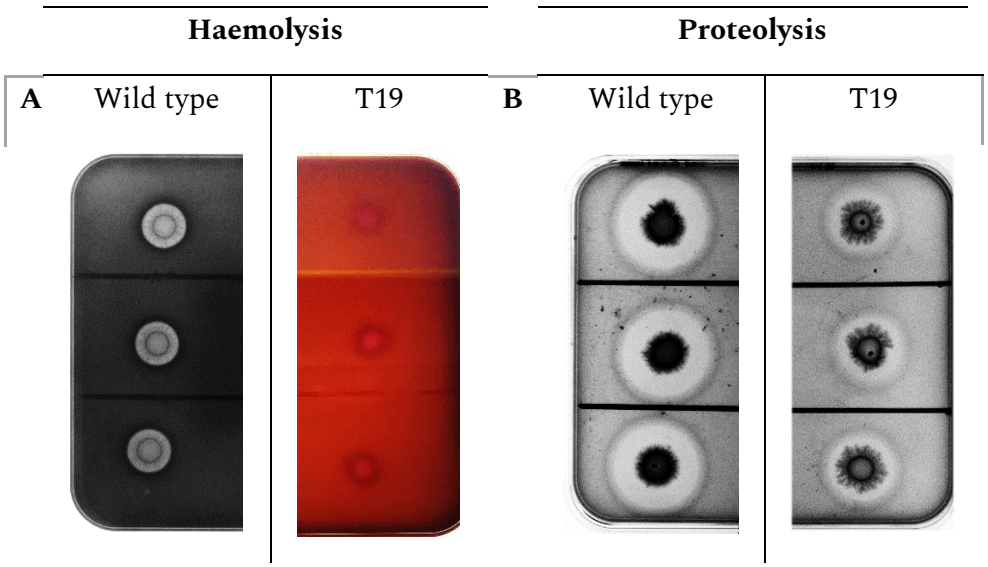


Figure 7.9. **Haemolysis and proteolysis by *C. sporogenes* T19 strain**. Transposon mutant T19 showed level weaker haemolytic (A) and proteolytic (B) activities, compared to the *C. sporogenes* wild type strain. This data is based on the 48 hour incubation of three biological replicates (except for T19) on TYG agar supplemented with 5% (v/v) horse blood (A) or 4% (v/v) skimmed milk (B), respectively.

7.3.4. Proposed function of *spoVK* in *C. sporogenes*

While ATPases have been previously known to function in the sporulation process of endospore-formers (Section 7.1), the micrographs of *C. sporogenes* Δ*spoVK* mutant spore cortex and coat also revealed some impairment, which indicates the possible role of the *spoVK* gene during the sporulation stages succeeding stage IV. This might be the reason why spores were sensitive to high temperature in spore assays, despite their apparent mature forms under the microscope.

The DPA assay on mutants will further reveal the function of SpoVK in the heat-resistance of *C. sporogenes* spores.

CHAPTER EIGHT

8. CHARACTERIZATION OF *spoIID*

8. CHARACTERIZATION OF *SPOIID*

8.1. INTRODUCTION

As mentioned in the earlier chapters, the complexity of the sporulation process in the endospore-forming bacteria can be divided in the form of progressive stages, in which the forespore originates within the mother cell and ultimately matures and releases into the environment. Following polar cell division, the next important phase in sporogenesis is the formation of an asymmetric septum that divides the forespore from the mother cell. However, this forespore septum soon gets engulfed as whole by the sporulating mother cell (Al-Hinai et al., 2015).

Governed by the mother cell-specific sigma factor σ^E during early sporulation (Pogliano et al., 1999), the process of forespore engulfment in *B. subtilis* involves the synergy of several important proteins, including SpoIIIAH, SpoIIQ, SpoIIM, SpoIIP and SpoIID, which bring about peptidoglycan hydrolysis of the forespore septum (Lopez-Diaz et al., 1986, Gutierrez et al., 2010), as illustrated in Figure 8.1. Understood from the sporulation studies in *B. subtilis*, the cell membrane protein SpoIIIAH and the forespore membrane protein SpoIIQ interacts to form the basis of the peptidoglycan-hydrolysing complex (Rodrigues et al., 2013). SpoIIM is another membrane protein, which recruits a cell wall-degrading enzyme SpoIIP to the forespore septum, which cleaves the peptide chains and later another lytic enzyme SpoIID is recruited for degrading the glycan strands (Chastanet and Losick, 2007, Gutierrez et al., 2010). SpoIID displaces SpoIIP to the new uncleaved peptide chains and moves forward along the forespore septum. SpoIID releases itself into the membrane pool and then localizes SpoIIP to further uncleaved peptide chains. This cycle of protein-binding, peptidoglycan degradation and movement of SpoIIP-SpoIID

complex across the forespore septum is repeated until the forespore engulfment process is completed in the sporulating mother cell (Gutierrez et al., 2010).

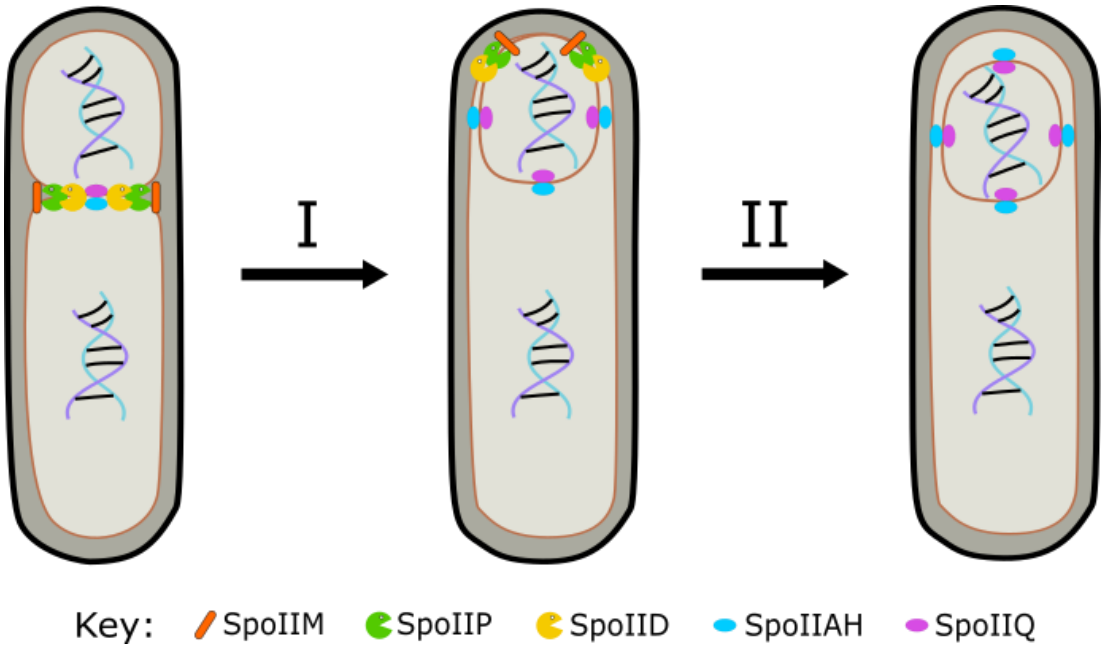


Figure 8.1. **Engulfment of forespore by peptidoglycan-degrading complex**. SpoIIM (orange rod) recruits SpoIIP (green pac-man), which further recruits SpoIID (yellow pac-man) at the centre of the forespore septum and execute peptidoglycan hydrolysis. Whereas, SpoIIQ (magenta oval) and SpoIIIAH (cyan oval) co-localizes at the forespore-mother cell junction to anchor SpoIIP and SpoIID and later nurture the spore. This diagram was inspired by (Gutierrez et al., 2010, Dembek et al., 2018).

Interestingly, some studies have suggested the existence of bipolar septa following asymmetric cell division, from which one of the septa gets dissolved by the SpoIIMPD complex, and the other septum becomes the forespore (Pogliano et al., 1999, Eichenberger et al., 2001). However, it has also been documented that the presence of mother cell transcription pathways in *B. subtilis* prevents cytokinesis on the distal pole of the forespore (Piggot and Coote, 1976, Sonenshein et al., 2002). Perhaps, the SpoIIMPD complex is one of the mother cell-controlled machinery that allows only one polar septum formation.

As briefly mentioned earlier, SpoIID is a membrane-anchored protein involved in the forespore engulfment process in endospore-forming bacteria (Piggot and Losick,

2001, Abanes-De Mello et al., 2002, Galperin et al., 2012, Gutierrez et al., 2010, Kelly and Salgado, 2019). It has been known that *spoIID* is only found in endospore-forming bacteria (Koo et al., 2017, Galperin et al., 2012, Abecasis et al., 2013). SpoIID is a lytic transglycolsylase or LT (Nocadello et al., 2016). LTs are enzymes, which act on bacterial cell wall and cleave the glycosidic bonds between the alternating chains of *N*-acetylglucosamine and *N*-acetylmuramic acid within the cell wall peptidoglycan (Scheurwater et al., 2008). The SpoIID protein has been demonstrated *in vitro* to cleave glycan strands, after the peptides have been pre-cleaved from the strands by the amidase and endopeptidase activities of SpoIIP protein (Chastanet and Losick, 2007, Morlot et al., 2010, Nocadello et al., 2016). SpoIID, in turn, also stimulates the lytic activity of SpoIIP (Gutierrez et al., 2010). Potential protein domains in the *C. sporogenes* SpoIID protein are illustrated in Figure 8.2.

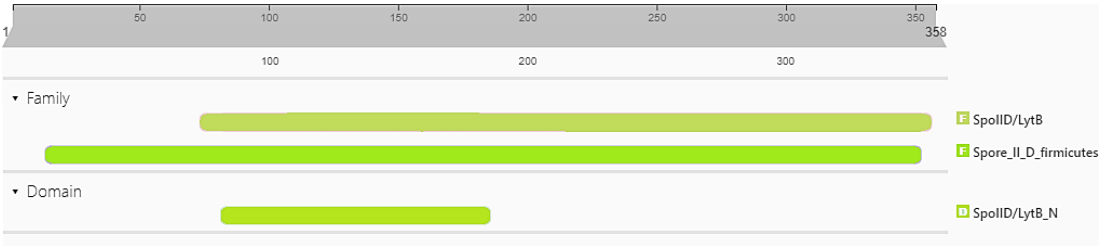


Figure 8.2. Protein domains of the SpoIID predicted by the InterPro. The SpoIID protein contains the LytB peptidoglycan-hydrolysing domain. Graphics edited for a personalized view.

Screening for sporulation-defective mutants in *C. sporogenes* transposon libraries resulted in the discovery of two asporulous mutants that had unique transposon insertions at the CLSPO_RS00950 locus (Appendix Table A2), which is known to be involved in this forespore engulfment process.

According to the NCBI genome databases, CLSPO_RS00950 is a small 1077 bp long gene in the genome of *C. sporogenes* NCIMB 10696 strain, with 27% GC content. Its gene product (WP_033058091.1) is annotated as “stage II sporulation protein D”, which is made up of a stretch of 358 amino acids. Browsing the neighbouring gene

loci, it can be predicted that CLSPO_RS00950 is potentially a monocistronic gene in *C. sporogenes*, similar to its homologue in *B. subtilis* (Lopez-Diaz et al., 1986). For the ease of explanation, the term "*spoIID*" will be used throughout this thesis to define the CLSPO_RS00950 gene.

Studies have revealed the positive regulation of *spoIID* transcription by σ^E in *B. subtilis* (Eichenberger et al., 2004, Meeske et al., 2016) and *C. difficile* (Saujet et al., 2013, Ribis et al., 2018), and this transcription has been shown to strictly occur in the sporulating mother cell, immediately after the formation of polar septum (Driks and Losick, 1991). Similar to these transcriptomic studies, *C. sporogenes spoIID* has also been observed to be in the *sigE* regulon, as identified in a recent study on sporulation-specific sigma factors of *C. sporogenes* sporulation (Raquel Rodrigues, unpublished). Additionally, the transcription of *B. subtilis spoIID* has been reported to get switched off by the SpoIIID protein, which is another protein in the *sigE* regulon (Eichenberger et al., 2004). This repression has been found to be achieved by the direct binding of SpoIIID to the *spoIID* promoter region, due to which binding of the σ^E-RNA polymerase to the *spoIID* promoter becomes prohibited.

Another interesting study revealed the absence of SpoIID transcript in some of the sporulation mutants in *B. subtilis*, which were engineered to contain mutations in the *spo0B*, *sigH*, *spoIIA*, *spoIIE*, *spoIIG* or *spoIIIA* gene (Rong et al., 1986, Clarke et al., 1986). Whereas, the SpoIID mRNA was still observed in *spo0J* and *spoIIB* mutants, along with the wild type *B. subtilis* strain. This is because *spoIIA* and *spoIIG* operons have been previously documented to participate in the transcription of many other sporulation genes in *B. subtilis* (Stragier, 1986, Lopez-Diaz et al., 1986, Stragier et al., 1984).

Therefore, it was interesting to characterize the autolysin mutant in *C. sporogenes* for the better understanding of the sporulation process.

8.1.1. Aim of this chapter

This chapter is aimed at studying the CLSPO_RS00950 *spoIID* gene for its potential role in the sporulation of *C. sporogenes*, by the following steps:

i. Constructing in-frame deleted Δ*spoIID* mutants in the *C. sporogenes* NCIMB 10696 wild type strain and later complementing the mutated gene;

ii. Characterizing the Δ*spoIID* mutants by determining the relationship of *spoIID* gene with sporulation, growth, haemolysis, proteolysis, and cell and colony morphology of *C. sporogenes*, and comparing the mutant phenotype with the wild type phenotype;

iii. Finding relevant literature on the prokaryotic *spoIID* gene and its SpoIID protein and comparing the phenotypes of the *C. sporogenes* Δ*spoIID* mutants with the ones generated in *Bacillus* and other clostridial species;

iv. Proposing the potential role of *spoIID* in the biology of *C. sporogenes*, based on the data generated in this study, with reference to earlier studies.

8.2. RESULTS

8.2.1. CRISPR/Cas9-based genome editing in *spoIID*

To confirm that the spo/ger$^-$ phenotype is due to the disruption of *spoIID* gene CLSPO_RS00950 in mutants T4 and T26, in-frame deletion Δ*spoIID* mutants were generated in triplicates from the *C. sporogenes* wild type strain, using the botsgRNA1 bookmark-based pRECas1-spoIID knock-out cassette, as described in Section 2.7. Δ*spoIID* mutants were then complemented using the watermark-based pRECas1-SpoIIDCOs knock-in cassette; and so, *C. sporogenes* SpoIIDCOs complemented strains were generated. All strains were characterized following the confirmation of plasmid loss.

Meanwhile, to combat with the cloning problems in *E. coli*, Target-AID base editor was also used, which generated three *C. sporogenes spoIID*::TA mutants, using the pTargetAID-spoIIDsgA knock-out cassette, as described in Section 2.8. However, a different pyrimidine base was changed instead of the one originally targeted despite three individual attempts. Although, no STOP codon could be pre-maturely introduced into the start of *spoIID*, translucent transconjugants colonies were observed on the TTC/Theo plates (Figure 8.3). Nevertheless, it was still interesting to know if a SNP in *spoIID* was capable of disrupting the sporulation of *C. sporogenes*.

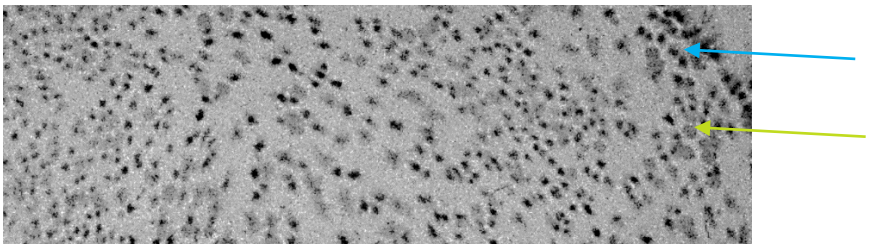


Figure 8.3. **Transconjugant colonies of *C. sporogenes* wild type-pTargetAID-spoIIDsgA**. Colonies of *spoIID*::TA mutants appeared faint and translucent (lime arrow) among the visible dark colonies of wild type background (blue arrow) on TTC/Theo plate after 48 hour post-conjugation incubation.

8.2.2. Mutation in *spoIID* halts *C. sporogenes* sporulation

Hence, in order to determine the effect of mutation in the *spoIID* gene on the survival of heat-treated *C. sporogenes* mutant strains, a 5-day spore assay was done, as mentioned in Section 2.10.2. As expected, *C. sporogenes* Δ*spoIID* mutants showed the spo/ger$^-$ phenotype, similar to the mutants T4, T26 and *spo0A*::CT; whereas the complemented *C. sporogenes* SpoIIDCOs strains, and unexpectedly, *spoIID*::TA mutants sporulated similar to the wild type strain (Figure 8.4).

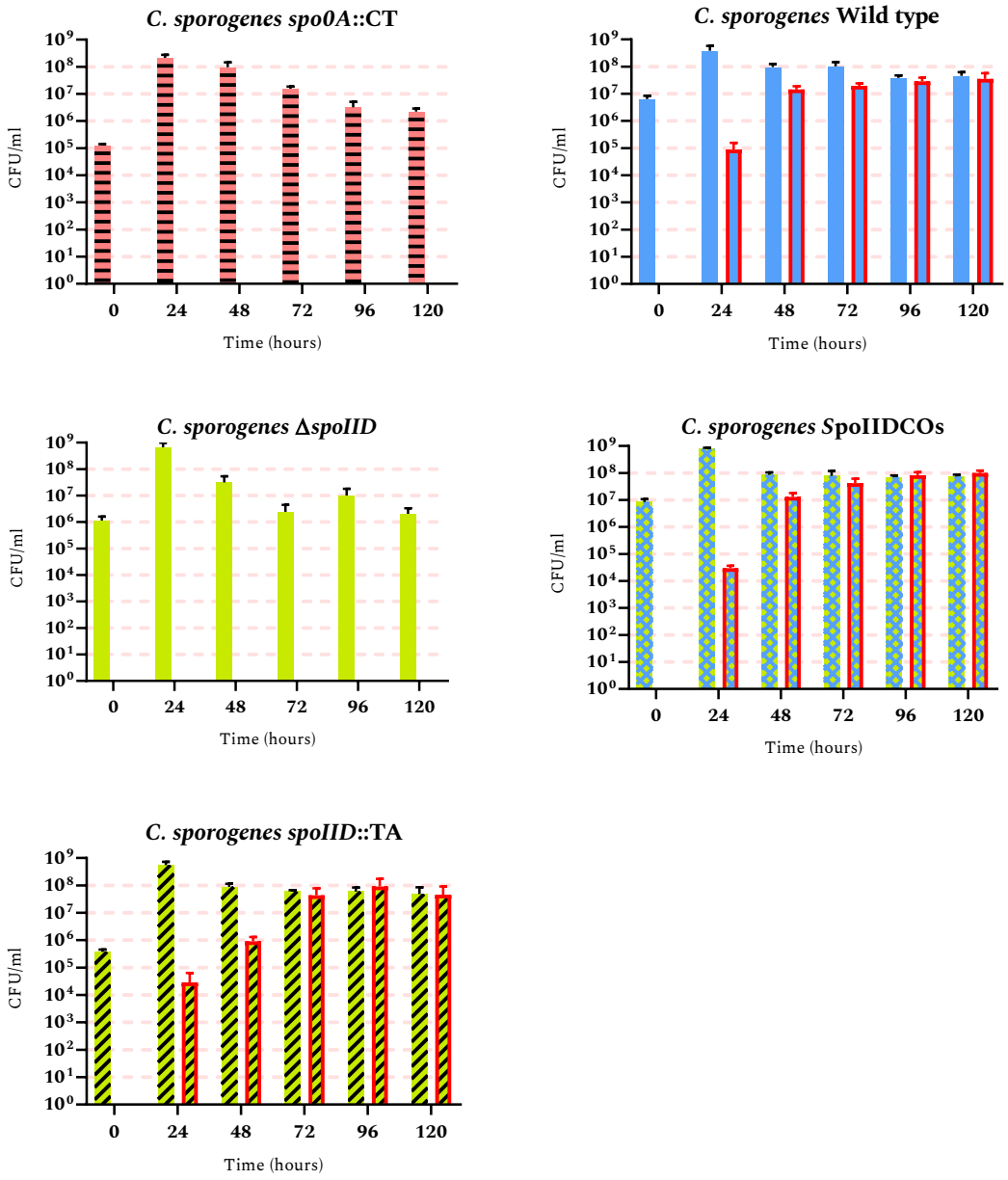


Figure 8.4. **Spore assay of *C. sporogenes spoIID* strains**. Mutant Δ*spoIID* showed no heat-resistant CFUs, similar to the *spo0A*::CT control strain. Whereas, mutant *spoIID*::TA and the complemented strain SpoIIDCOs showed comparable heat-resistant CFUs to *C. sporogenes* wild type. This data is based on the 5-day incubation of three biological replicates (except for *spo0A*::CT) in TYG broth. Error bars show mean with standard deviation. Each strain is represented by a unique colour[5]. Red-bordered bars represent heat-resistant CFUs, whereas coloured bars without red border represent the untreated or the total cell counts. Assay detection limit in CFU/ml was: lower limit = 0 or 66; upper limit = 4×10^9.

[5] Colour key:



8.2.3. *spoIID* is not an essential gene in *C. sporogenes*

To determine differences in the growth pattern among the SpoIID-producing and non-producing strains, *C. sporogenes* mutants and their relevant complemented strains were grown in TYG broth, as described in Section 2.10.2. Unlike the growth-impaired *spo0A*::CT mutant, the growth curves of *C. sporogenes* Δ*spoIID* mutants and their respective complemented SpoIIDCOs strains were observed to be comparable with that of the *C. sporogenes* wild type strain, as shown in Figure 8.5. Moreover, SpoIIDCOs strains showed the peak cell density of ~4.3 OD_{600} (data not shown), which was slightly higher than the highest (~3.7 OD_{600}) observed for the wild type strain.

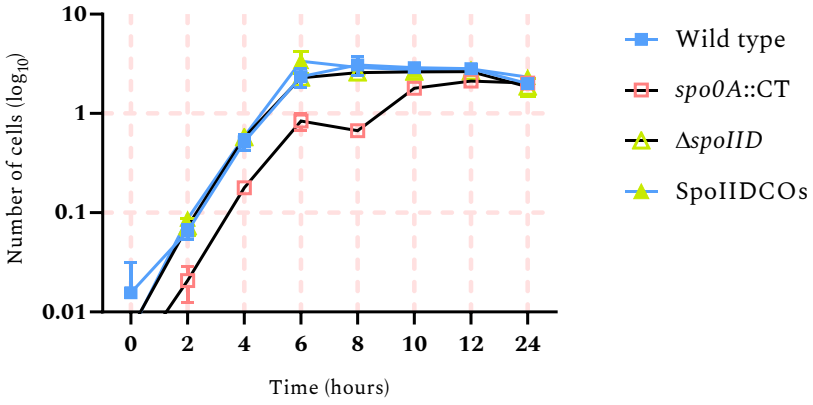


Figure 8.5. **Growth curve of *C. sporogenes spoIID* strains**. Mutant Δ*spoIID* and the complemented SpoIIDCOs strain, showed similar growth patterns to *C. sporogenes* wild type strain, in contrast to the spo/ger$^-$ mutant *spo0A*::CT, which has a slight growth defect. This data is based on the 24 hour incubation of three biological replicates (except for *spo0A*::CT) in TYG broth. Error bars show mean with standard deviation. Theoretically, 0.01 OD_{600} was the starting inoculum, which was individually calculated for each strain; however, in practice, variable readings were observed at 0 hour.

8.2.4. *C. sporogenes* β-haemolysis does not require *spoIID*

Since haemolytic activity of *C. sporogenes* wild type is lost due to mutation in the *spo0A* gene, as shown in Figure 8.6, it was interesting to know if another important sporulation gene is also capable of controlling *C. sporogenes* haemolysis. Therefore, relevant *C. sporogenes* cultures were spotted on TYG blood agar and examined for zones of clearance after incubation, as described in Section 2.10.4. After 48 hours of anaerobic incubation, all the *C. sporogenes* Δ*spoIID* and SpoIIDCOs strains showed similar β-haemolysis to the wild type strain (Figure 8.6).

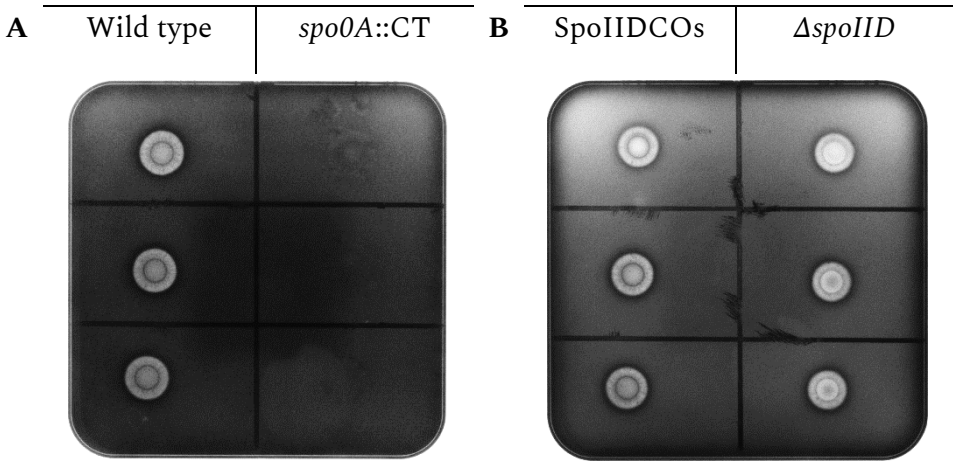


Figure 8.6. **Haemolysis by *C. sporogenes spoIID* strains**. Mutant Δ*spoIID* and the complemented strain SpoIIDCOs showed wild type level haemolytic activity, in terms of β-haemolysis zones. The spo/ger⁻ mutant *spo0A*::CT was non-haemolytic and its translucent growth was overshadowed by dark-coloured blood agar. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT) on TYG agar supplemented with 5% (v/v) horse blood.

8.2.5. Proteolytic activity of *C. sporogenes* is independent of *spoIID*

On milk TYG agar plates, wild type *C. sporogenes* strain showed large zones of clearance within 24-48 hours, whereas the cultures of *spo0A*::CT mutant gave slightly smaller zone of clearance than the wild type strain, shown in Figure 8.7. To study the effect of *spoIID* gene deletion on the proteolysis of *C. sporogenes*, Δ*spoIID* mutants along with their complemented counterparts were grown on milk TYG agar plates, as described in Section 2.10.5. Wild type level proteolytic activity was observed for all the tested *C. sporogenes spoIID* strains, with an exception to one of the complemented SpoIIDCOs replicate, as also shown in Figure 8.7.

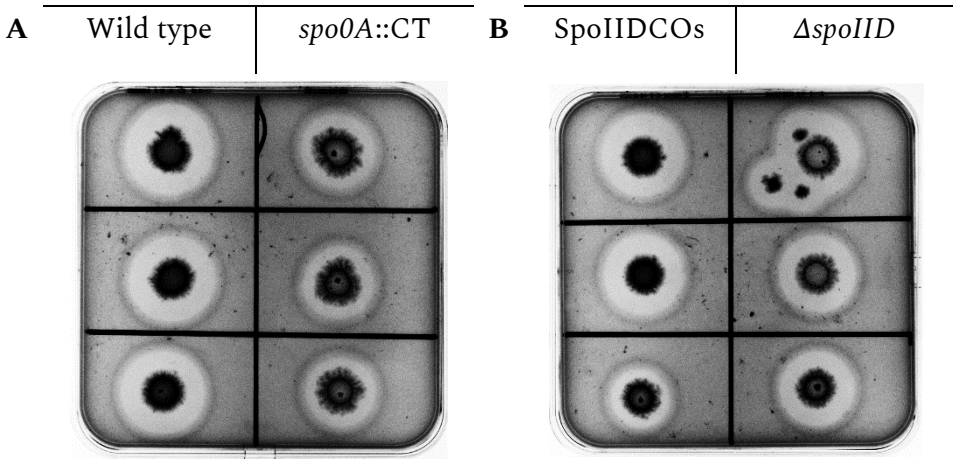


Figure 8.7. **Proteolysis by *C. sporogenes spoIID* strains**. Mutant Δ*spoIID* and the complemented strain SpoIIDCOs showed wild type level proteolytic activity. This data is based on the 48 hour incubation of three biological replicates (except for *spo0A*::CT) on TYG agar supplemented with 4% (v/v) skimmed milk.

8.2.6. Microscopy reveals Δ*spoIID* mutants lacking forespore and mature spores

To further confirm the spo/ger$^-$ phenotype of *C. sporogenes* Δ*spoIID* mutants, 120 hour old relevant *C. sporogenes* broth cultures were spore-stained and observed under the microscope, as described in Section 2.5.1. Similar to the spo/ger$^-$ mutant *spo0A*::CT, spores were not visible for Δ*spoIID* mutant cultures under the microscope as shown in Figure 8.8, supporting the lack of heat-resistant CFU counts observed on agar

plates (Figure 8.4). This defect was completely restored in the complemented SpoIIDCOs strain.

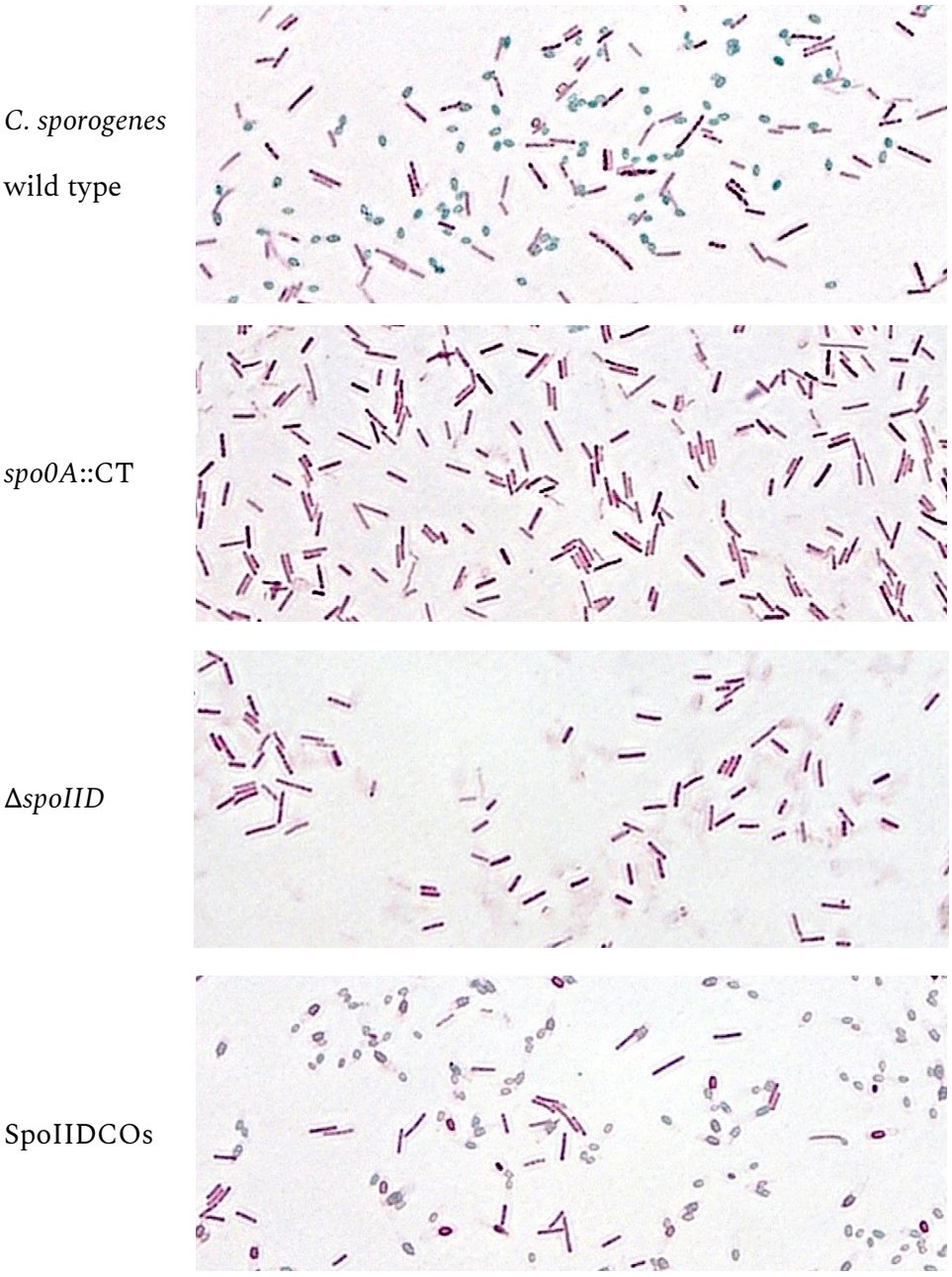
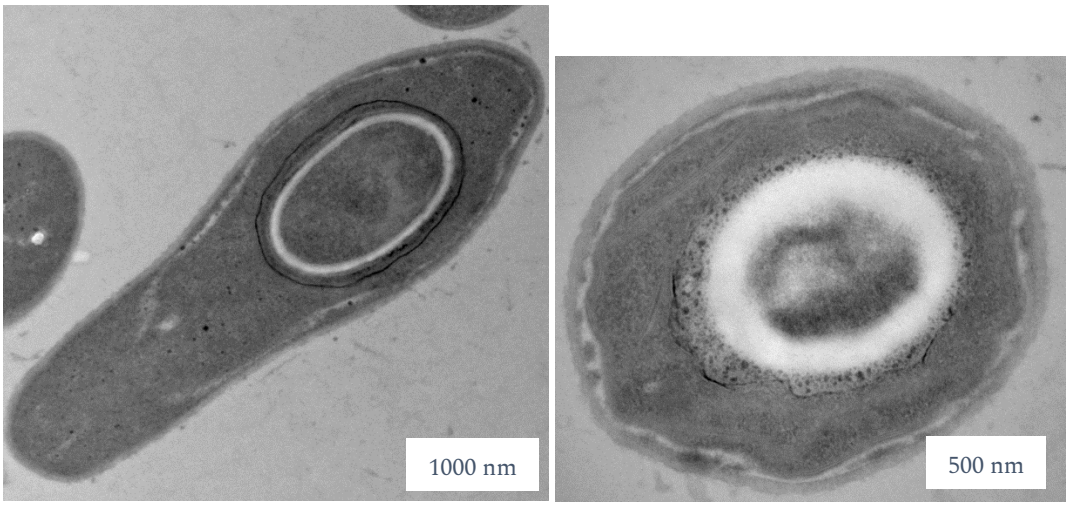


Figure 8.8. **Cell morphology of *C. sporogenes spoIID* strains**. Mutant Δ*spoIID* showed rods without mature spores, similar to the *spo0A*::CT control strain. The complemented strain SpoIIDCOs showed restoration of the microscopic phenotype of *C. sporogenes* wild type strain. Cells were stained pink and spores were stained blue. The images represent 120 hour old TYG broth cultures of *C. sporogenes* strains at 1000 × magnification.

Despite attempts at TEM processing and imaging of *C. sporogenes* Δ*spoIID* mutant, informative TEM micrographs could not be obtained due to the low quality of the processed mutant culture, as seen in Figure 8.9. Some mutant cells were accidently washed away during the fixation step and perhaps the resin embedding technique was not suitable to visualize Δ*spoIID* mutant cells.

***C. sporogenes* wild type**



***C. sporogenes* ΔspoIID**

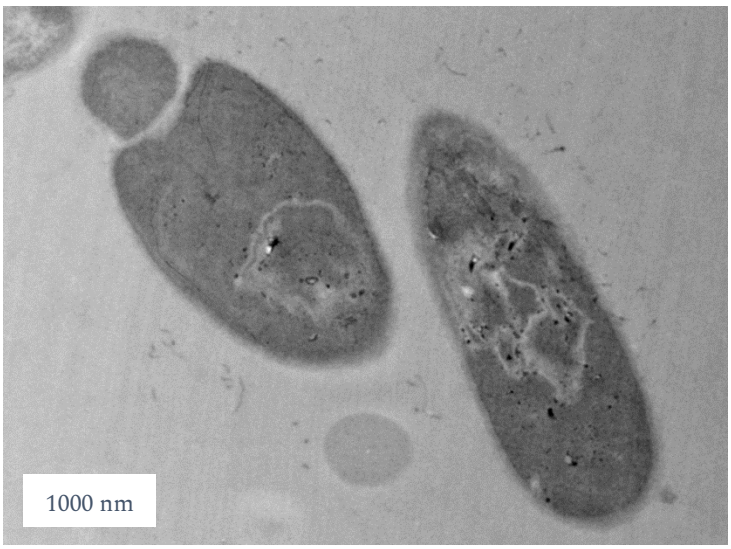


Figure 8.9. **Ultrastructure of *C. sporogenes spoIID* strain**. No detailed cell structure could be visualized for mutant Δ*spoIID*. The micrographs represent thin section of 72 hour old fixed cultures of *C. sporogenes* strains at high magnifications.

8.2.7. Δ*spoIID* mutant colonies become translucent on old TYG agar plates

Two of the initially isolated mutants, T4 and T26, which had transposon insertion in the *spoIID* gene, gave translucent colonies on TYG agar (data not shown). And as expected, colonies of in-frame deletion Δ*spoIID* mutants become translucent within 27-72 hours after incubation on TYG agar plates (data not clearly shown). Conversely, the complemented SpoIIDCOs strains were found to produce opaque colonies, similar to those of the *C. sporogenes* wild type strain (Figure 8.10).



Wild type | *spo0A*::CT | Δ*spoIID* | SpoIIDCOs

Figure 8.10. **Colony morphology of *C. sporogenes spoIID* strains**. Mutant colonies of *C. sporogenes* Δ*spoIID* showed translucent phenotype on 2-3 days old agar plates, unlike the sustained opaque growth of wild type. The wild type colony morphology was restored in the complemented SpoIIDCOs strain and colonies appeared smooth, shiny, fimbriated and irregular, just like those of the *C. sporogenes* wild type strain. In contrast, *spo0A*::CT mutant showed highly fimbriated and translucent phenotype within 24 hours. The images were taken after 27 hour incubation of *C. sporogenes* strains on TYG agar.

8.2.8. Illumina sequencing reveals SNPs in the Δ*spoIID* mutant genome

It was important to confirm that the observed spo/ger$^-$ phenotype of the *C. sporogenes* Δ*spoIID* mutants was manifested due to the defect in SpoIID expression and not because of undesirable mutations in other loci in the engineered *C. sporogenes* genome. So, one of the Δ*spoIID* mutant was sent for whole genome sequencing and analysed for the presence of additional SNPs, as described in in Section 2.3.8. When aligned with the *C. sporogenes* NCIMB 10696 reference strain from the NCBI genome databases, the sequenced data of Δ*spoIID* mutant revealed the potential existence of SNPs in PTS transporter gene, diacylglycerol gene and some unknown gene loci,

along with the seventeen SNPs that are present in the in-house *C. sporogenes* wild type strain (Appendix Table A7).

8.3. DISCUSSION

In this study, attempts were made to characterize the *spoIID* gene of *C. sporogenes* NCIMB 10696 strain. This gene was found to be inactivated in mutants T4 and T26, which were isolated from a transposon mutant library through the transparency assay. Although mutant T4 has transposon insertion in the middle of *spoIID* gene and T26 near the 3' end of the gene, both showed spo/ger$^-$ phenotype. To confirm that this sporulation-deficient phenotype was due to the absence of *spoIID* or its protein product, in-frame deletion Δ*spoIID* mutants were made using CRISPR/Cas9-based genome editing tools. These mutants were later complemented to restore the wild type sporulation properties, while still carrying a watermark on the *spoIID* gene's STOP codon to distinguish the resultant complemented SpoIIDCOs strains from the *C. sporogenes* wild type strain.

Cloning events for *spoIID*-knock-out and -knock-in cassettes were fairly trouble-free as improved tools were incorporated and protocols were already optimised to limit any gene toxicity in the transformed *E. coli* strains harbouring the engineered plasmid constructs. Also, the bookmark-based gene complementation of Δ*spoIID* mutants was accomplished due to the combined action of botsgRNA1 and efficient homology-directed repair in *C. sporogenes*.

8.3.1. Target-AID base editor and the off-targeting

Despite the success of generating SNP-free plasmid constructs for *spoIID* gene deletion and gene complementation, attempts were still made to mutate *spoIID* of *C. sporogenes* using the latest Target-AID tool. It was thought that mutants with a point mutation might further show some other interesting phenotypes that might not

be apparent in the absence of the gene CDS. The in-house guide RNA design tool suggested only one spacer sequence that could possibly guide the Cas9-AID complex to the targeted *spoIID* gene, and edit a C base in the sense strand, such that the Gln-147 becomes a STOP codon. In practice, however, the Ser-146 was targeted instead, resulting in a change to Phe instead of a STOP codon. This off-target effect is not an unusual finding, but exemplifies one of the limitations of Target-AID base editing, in which the presence of several cytosine bases near the target cytosine base caused off-targeted genome editing (Eid et al., 2018). It has already been demonstrated in yeast that the -18 position upstream of PAM is more likely to be targeted by AID for base-editing than the -16 position (Nishida et al., 2016), and coincidently, the exact phenomenon was also observed in this study. This explains why mutation kept occurring at the same position, despite several attempts at mutant generation.

Based on the InterPro tool's prediction on SpoIID protein domains (Figure 8.2), this altered Ser-146 codon lies within the SpoIID/LytB domain, which covers amino acids 81st-185th. Also the transconjugant colonies with *spoIID*::TA mutation showed translucent phenotype (Figure 8.3). So, a sporulation-defective phenotype was still expected. However, the numbers of heat-resistant CFUs of *spoIID*::TA mutant cultures were observed to be similar to the wild type strain, in contrast to the spo/ger$^-$ mutants T4, T26 and Δ*spoIID*. This indicates that Ser-146 is not involved in the catalytic activity of SpoIID in *C. sporogenes*, unlike the Cys-146 of *C. difficile*, which completely inactivated the enzyme after being altered to Ala-146 in a previous study (Nocadello et al., 2016). Cys-146 of *C. difficile* SpoIID protein is an important amino acid of the zinc-binding site, which is crucial for the catalytic activity in *C. difficile*, yet absent in the *B. subtilis* SpoIID protein. It is assumed that prolonged incubation of the mutated *C. sporogenes* pTargetAID-spoIIDsgA transconjugants on plates might have affected the vegetative cell structure and contributed to a change in the autolysin activity.

8.3.2. SNPs in the Δ*spoIID* mutant

Analysing the Illumina sequencing data revealed the existence of additional unintentional mutations in the Δ*spoIID* mutant, some of which were also shared by the sequenced Δ*divIVA* mutant, such as in the PTS transporter EIIC gene and diacylglycerol kinase gene, as briefly mentioned in Section 5.3.3. Similar to the DivCO complemented strain, the SpoIIDCOs strain was fully complemented to match the wild type phenotype, despite the potential role of diacylglycerol kinase in the cortex biosynthesis and assembly in *B. subtilis* (Amiteye et al., 2003).

8.3.3. Characterization of *spoIID* and its role in sporulation

The role of *spoIID* gene in the sporulation of *C. sporogenes* is very obvious from the observations, in which not only heat-resistant CFUs were non-existent on TYG plates (Figure 8.4), but spores were also absent from the broth cultures when examined under the microscope (Figure 8.8). The gene product of *spoIID* has been reported to play an essential role in sporulation by hydrolysis of the peptidoglycan cell wall during the engulfment process of the forespore in the sporulating mother cell of *B. subtilis* and *C. difficile* (Driks and Losick, 1991, Lopez-Diaz et al., 1986, Abanes-De Mello et al., 2002, Chastanet and Losick, 2007, Morlot et al., 2010, Gutierrez et al., 2010, Nocadello et al., 2016, Kelly and Salgado, 2019).

BLASTP results show the presence of conserved domains in SpoIID protein among the endospore-forming bacteria. For a query cover of 95%, *C. sporogenes* SpoIID is 38.26% similar to LytB and 37.22% to the SpoIID protein homologues in *B. subtilis*. Similar to the *C. sporogenes* genome, *spoIID* in *B. subtilis* is also is monocistronic and the only gene its operon (Lopez-Diaz et al., 1986). As expected, the spo/ger$^-$ phenotype of the generated *C. sporogenes* Δ*spoIID* mutants in this study, matches with the phenotype demonstrated for *B. subtilis spoIID* mutants (Lopez-Diaz et al., 1986, Rong et al., 1986, Driks and Losick, 1991). In the absence of LytM, another protein

homologue of *C. sporogenes* SpoIID, the growth and haemolytic profiles of *S. aureus* were not altered (O'Halloran et al., 2015). This finding matches with the data obtained for Δ*spoIID* mutants in this study. In addition, the translucent colony phenotype of the *C. sporogenes spoIID* mutants, observed in the present study, has also been reported for the spo/ger^{-} *spoIID* mutants in *B. subtilis* (Atkinson and Fisher, 1991).

As mentioned earlier, SpoIID expression in *C. sporogenes* is under the control of σ^{E} regulon (Raquel Rodrigues, unpublished), similar to what is documented for *C. difficile* (Ribis et al., 2018). Regarding the spo/ger^{-} phenotype of *C. sporogenes* Δ*spoIID* mutants, different observations have been documented for the equivalent mutants in *C. difficile*. In one study, the Δ*spoIID* mutants still produced some heat-resistant spores (Ribis et al., 2018). Whereas, in a different study, *C. difficile* Δ*spoIID* mutants not only showed similar growth phenotype, but also the absence of spores (Dembek et al., 2018). This contradiction between the two studies has been assumed to be due to the differences in the spore assay protocols, which used different nutrient media (Ribis et al., 2018).

In *C. sporogenes*, *spoIID* is a non-essential gene dispensable for vegetative cell growth, since null mutation of *spoIID* did not affect the exponential growth of the host organism (Figure 8.5). The reduced CFUs of untreated Δ*spoIID* mutant cultures, which were observed after 48 hours of incubation (Figure 8.4) is a common consequence of the severe sporulation-defect, based on the data observed in this study.

Although the role of the SpoIID/LytB domain in the haemolysis of bacteria has not been previously documented, it was still decided to check if Δ*spoIID* mutants show a different phenotype on TYG blood agar. As expected, no impairment of *C. sporogenes* haemolytic activity was observed on plates (Figure 8.6). For a peptidoglycan-degrading gene, *spoIID* was initially assumed to be somewhat related to the proteolytic activity. However, literature shows the SpoIID activity to commence only after the peptide chains have been degraded by SpoIIP. Additionally, the presence of

other potentially important protease genes in *C. sporogenes* might have rendered the effect of *spoIID* deletion on proteolysis to be unaffected (Figure 8.7).

8.3.4. Proposed function of *spoIID* in *C. sporogenes*

In the light of the currently available information, it is safe to say that *spoIID* is a sporulation-specific gene in *C. sporogenes*. Although the obtained TEM images (data not shown) could not provide much evidence regarding which stage of *C. sporogenes* sporulation was blocked in Δ*spoIID* mutants, it can still be assumed that *spoIID* is possibly involved in the earlier stages of sporulation, on the basis of complete absence of phase-bright forespores in the mutant cultures.

The transcriptional profiling of *C. sporogenes* Δ*spoIIID*, in addition to Δ*spoIID* mutants will provide more insights about the *spoIID* repression cascade and the possibility of finding additional genes that are influenced by the SpoIID protein.

CHAPTER NINE

9. GENERAL DISCUSSION AND CONCLUSION

9. GENERAL DISCUSSION AND CONCLUSION

In this study, several potential sporulation or germination genes were identified, which can be considered novel to the currently available literature on sporulation of *C. sporogenes*. Stage IV sporulation gene *yqfD*, cell division gene *divIVA*, protease chaperone gene *clpX*, AAA+ family gene *spoVK* and stage II sporulation gene *spoIID* were briefly characterized by genome editing in *C. sporogenes*.

Genetic studies are not particularly challenging to do in *C. sporogenes*, due to both the availability of innovative gene manipulation tools, as well as the presence of minimum restriction modification systems in this clostridial host species, which allows the ease of gene insertions and deletions. The accessibility of using conjugation and electroporation protocols also help in the success of genetic tools like ClosTron, Allelic Exchange and Allele-Coupled Exchange (Minton et al., 2016), as well as the CRISPR/Cas9-based genome editing (Canadas et al., 2019, Ingle et al., 2019).

However, the use of such genetic tools come with some limitations as well. In the beginning, the conditional transposon mutagenesis tool was integrated in this study for generating libraries of random transposon mutants. The transposon cassette in the initial transposon plasmid pMTL-YZ14 was unstable in the *E. coli* CA434 donor strain, for which a negative selection marker had to be integrated in the new transposon plasmid pMTL-CW20, as described in Section 3.3.1. Secondly, screening large number of mutants by spore assay required intensive labour. Generally, robot workstations are available for screening transposon mutant libraries (Kristich et al., 2008); however, the current unavailability of such workstations for growing anaerobic bacteria, demands the requirement of manual labour. Eventually, screening large mutant libraries on the basis of translucent colonies made the protocol more user-friendly, but on the cost of some biased results, since not all translucent colonies represent sporulation defects, as mentioned briefly in Section 3.3.3.3.

Genetic studies such as, TraDIS (Dembek et al., 2015), are based on sequencing the whole genome under the desired growth conditions, which efficiently generates data about all the essential genes, using less manual labour. However, the development of a workable TraDIS protocol for the host organism requires ample amount of research and literature reviews. In the absence of a workable TraDIS protocol for *C. sporogenes*, conditional transposon mutagenesis gave a good representation of potential sporulation genes that had not been previously characterized in *C. sporogenes*. TraDIS protocol has been developed and exemplified in a recent study on *C. sporogenes* (Raquel Rodrigues, unpublished), which not only matches with the results obtained in this study, but also shows the entire list of genes potentially involved in the sporulation or germination of *C. sporogenes*.

The polar effects of transposon insertion on neighbouring gene loci can also not be overlooked, as described in the case with *hypP* and *divIVA* genes in chapter five of this thesis. Similarly, a previous gene target CLSPO_RS18590 (*yabP*) was initially considered due to the observance of asporulous phenotype in the relevant transposon mutant #138_v2. (Appendix Table A2). However, the homologue of this gene and equivalent transposon mutant has already been studied in a previous study on sporulation genes of *C. difficile*, which characterized this asporogenous or spo/ger$^-$ phenotype to actually be a consequence of transposon insertion that affected the neighbouring genes on the same operon (Dembek et al., 2015). The identification of transposon insertion site in *C. sporogenes* #138_v2 mutant revealed its presence in the promoter region upstream the *yabP* gene, instead of insertion within the gene itself. Due to the possibility of one promoter controlling several genes in the same operon, it was later decided to find a different gene target for a more accurate sporulation mutant characterization.

It is important to mention the existence of misinterpreted data in high-throughput sequencing studies, like TraDIS. False results due to polar effects of transposon insertion is a common problem and has also been exemplified in the Chapter four of this thesis. It has been suggested to use flanking transcriptional terminators with the transposon cassette, as a possible solution to overcome this problem; however, relevant research is still in progress (Craig Woods, unpublished).

Another weakness of sequencing a mutant library by TraDIS is that one million differently-affected mutants are pooled in one culture vial, which puts the burden of competition among the mutants during the tested conditions. Because of the "survival of fittest" phenomenon, some slow-growers or growth-defective mutants are unable to keep up with the growth-efficient strains and get eliminated from the library, but would otherwise, be capable of growth and survival if grown in pure cultures. This can lead to false labelling of genes as "essential" for the host strain. For instance, the recent TraDIS data on *C. sporogenes* identified the ferric uptake regulator *fur* gene as an essential gene. However, in a side project of this study, *fur* mutants were easily generated in *C. sporogenes* using ClosTron system (data not shown). Additionally, ClosTron-generated *fur* mutants have also been characterized in an earlier study in *C. difficile* (Magdalena Katarzyna Fit, unpublished).

Additionally, the development of CRISPR/Cas9-based gene editing tools was in progress throughout the course of this study, which resulted in using different techniques as well as CRISPR/Cas9 plasmids for different gene targets. Following the development of the RiboCas tool (Canadas et al., 2019), all the genomic modifications were re-generated to ensure homogeneity among the gene targets for a more accurate characterization of the modified strains. This also contributed to the lack of available time for a much detailed mutant characterization.

As mentioned in the earlier chapters, dependence of the current CRISPR/Cas9-based tools on homology-directed recombination or HDR resulted in encountering cloning problems due to gene toxicity in *E. coli* strains. This gene toxicity has also been briefly mentioned in other recent studies on Clostridia (Ribis et al., 2018, Dembek et al., 2018). For mutant generation, Target-AID is independent of the use of HDR; however, it has its own limitations due to the short mutagenesis frame for the activity of the cytidine deaminase enzyme or AID. Additionally, the lack of gene complementation protocols is another major limitation in the AT-rich clostridial strains. Nevertheless, majority of the precious research time was spent in testing out various approaches to prevent the possibility of any undesirable mutations in the *E. coli* cloning-based plasmid constructs, as mentioned in the previous chapters of this study.

Regardless of the precautionary measurements, the occurrence of additional SNPs in the *C. sporogenes* mutant genome could not be avoided. Based on the results of whole genome sequencing, the identification of such SNPs apparently indicated the presence of additional spontaneous mutations in the current in-house wild type frozen stock, along with the mutations occurring during the sub-culturing steps in the mutagenesis protocols. However, the presence of some matching SNPs among the *C. sporogenes* mutants might also indicate their occurrence during the overnight incubation of broth cultures prior to the DNA purification steps for whole genome sequencing.

Because of the persistence of undesirable SNPs occurring in plasmid constructs, there was an accumulation of several knock-out cassettes, which couldn't have a purpose. It was then considered to make use of such point mutations in genes for the identification of important protein domains in the function of the gene product of such genes. However, the locations of SNPs were analysed to be too discrete to get an accurate representation of the mutant phenotype that has been altered only due to a single corresponding SNP. Nevertheless, generating *C. sporogenes* mutants by

inducing targeted point mutations can provide valuable data on various functional domains in a protein.

Another interesting assay would be characterising different types of mutant colonies within the population of mutant cells that have the same targeted mutation, but appear to have various phenotypes. For example, the pure cultures of each replicate of mutant Δ*divIVA*, gave different types of colonies on TYG plates, as well as different sizes of cells under the microscope, as described in Chapter five of this thesis. Literature explains that this difference among the same mutant population is a reflection of cell division Z-ring formation at different sites in the mutant cell, in contrast to the normal Z-ring formation that occurs only at the mid-cell in a growing culture of metabolically-active cells. Therefore, it would be intriguing to isolate the morphologically different colonies, such as those of the minicells and characterize them individually.

Beside the six mentioned gene targets for characterization, CLSPO_RS12440 was another gene, identified from the transposon library screenings and was initially meant for in-depth characterization in this study. It has been known that a homologue of CLSPO_RS12440 is a part of the skin element in *C. difficile* and has also been demonstrated to be essential for the sporulation cascade of *C. difficile* (Serrano et al., 2016). The earlier study provided a good reason to target the CLSPO_RS12440 gene in *C. sporogenes*. The knock-out and knock-in cassettes of this "serine recombinase" gene were initially very toxic in *E. coli*. Despite overcoming the cloning problems later on, a complemented strain of CLSPO_RS12440 null mutants could not be obtained, due to the complete absence of an efficient CRISPR/Cas9-based HDR in the mutant strains. After several attempts in vain, the characterization of this gene was excluded from this study.

Some additional mutant characterization protocols were initially planned to conduct during the course of this study. For instance, the transcriptomics of the generated mutants could provide valuable data regarding their role in the regulation of other important genes, including the ones for *C. sporogenes* growth and sporulation. For the *hypP* gene characterization, SNAP-tag seemed an interesting tool, which could have identified the localization of the hypothetical protein and perhaps helped in better understanding the mechanism by which sporulation was being affected in the relevant mutant. Since the obtained phenotype was a result of transposon insertional polar effects, it was no longer relevant to this study on sporulation. However, the SNAP-tag data for *clpX* gene product could have revealed more helpful data in terms of its function during the sporulation of *C. sporogenes*. Moreover, DPA quantitation assay could have explained the *spoVK* mutant phenotype in terms of the heat-sensitivity of mutant spores.

Nevertheless, higher characterization studies could not be carried out in the given time frame. But, on the brighter side, it also provides several new project proposals for future research.

REFERENCES

ABANES-DE MELLO, A., SUN, Y. L., AUNG, S. & POGLIANO, K. 2002. A cytoskeleton-like role for the bacterial cell wall during engulfment of the *Bacillus subtilis* forespore. *Genes Dev,* 16, 3253-64.

ABECASIS, A. B., SERRANO, M., ALVES, R., QUINTAIS, L., PEREIRA-LEAL, J. B. & HENRIQUES, A. O. 2013. A genomic signature and the identification of new sporulation genes. *J Bacteriol,* 195, 2101-15.

ABUSNINA, W., SHEHATA, M., KAREM, E., KOC, Z. & KHALIL, E. 2019. *Clostridium sporogenes* bacteremia in an immunocompetent patient. *IDCases,* 15, e00481.

ADLER, H. I., FISHER, W. D., COHEN, A. & HARDIGREE, A. A. 1967. Miniature *Escherichia coli* cells deficient in DNA. *Proc Natl Acad Sci U S A,* 57, 321-6.

AGREN, J. A. & CLARK, A. G. 2018. Selfish genetic elements. *PLoS Genet,* 14, e1007700.

AL-HINAI, M. A., JONES, S. W. & PAPOUTSAKIS, E. T. 2015. The *Clostridium* sporulation programs: diversity and preservation of endospore differentiation. *Microbiol Mol Biol Rev,* 79, 19-37.

ALDAPE, M. J., BRYANT, A. E. & STEVENS, D. L. 2006. *Clostridium sordellii* infection: Epidemiology, clinical findings, and current perspectives on diagnosis and treatment. *Clin Infect Dis,* 43, 1436-46.

ALLISON, C. & MACFARLANE, G. T. 1990. Regulation of protease production in *Clostridium sporogenes*. *Appl Environ Microbiol,* 56, 3485-90.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215, 403-10.

AMITEYE, S., KOBAYASHI, K., IMAMURA, D., HOSOYA, S., OGASAWARA, N. & SATO, T. 2003. *Bacillus subtilis* diacylglycerol kinase (DgkA) enhances efficient sporulation. *J Bacteriol,* 185, 5306-9.

ANANTHARAMAN, V. & ARAVIND, L. 2005. MEDS and PocR are novel domains with a predicted role in sensing simple hydrocarbon derivatives in prokaryotic signal transduction systems. *Bioinformatics,* 21, 2805-11.

ANDREESEN, J. R., BAHL, H. & GOTTSCHALK, G. 1989. Introduction to the physiology and biochemistry of the genus *Clostridium*. *Clostridia.* Springer.

ANNAN, F. J., AL-SINAWI, B., HUMPHREYS, C. M., NORMAN, R., WINZER, K., KOPKE, M., SIMPSON, S. D., MINTON, N. P. & HENSTRA, A. M. 2019. Engineering of vitamin prototrophy in *Clostridium ljungdahlii* and *Clostridium autoethanogenum*. *Appl Microbiol Biotechnol,* 103, 4633-4648.

ARAVIND, L. & KOONIN, E. V. 1999. Novel predicted RNA-binding domains associated with the translation machinery. *J Mol Evol,* 48, 291-302.

ARNON, S. S., SCHECHTER, R., INGLESBY, T. V., HENDERSON, D. A., BARTLETT, J. G., ASCHER, M. S., EITZEN, E., FINE, A. D., HAUER, J. & LAYTON, M. 2001. Botulinum toxin as a biological weapon: medical and public health management. *Jama,* 285, 1059-1070.

ATARASHI, K., TANOUE, T., SHIMA, T., IMAOKA, A., KUWAHARA, T., MOMOSE, Y., CHENG, G., YAMASAKI, S., SAITO, T., OHBA, Y., TANIGUCHI, T., TAKEDA, K., HORI, S., IVANOV, II, UMESAKI, Y., ITOH, K. & HONDA, K. 2011. Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science,* 331, 337-41.

ATKINSON, M. R. & FISHER, S. H. 1991. Identification of genes and gene products whose expression is activated during nitrogen-limited growth in *Bacillus subtilis*. *J Bacteriol,* 173**,** 23-7.

ATMADJAJA, A. N., HOLBY, V., HARDING, A. J., KRABBEN, P., SMITH, H. K. & JENKINSON, E. R. 2019. CRISPR-Cas, a highly effective tool for genome editing in *Clostridium saccharoperbutylacetonicum* N1-4(HMT). *FEMS Microbiol Lett,* 366.

ATRIH, A., ZOLLNER, P., ALLMAIER, G. & FOSTER, S. J. 1996. Structural analysis of *Bacillus subtilis* 168 endospore peptidoglycan and its role during differentiation. *J Bacteriol,* 178**,** 6173-83.

AWAN, K. H. 2017. The therapeutic usage of botulinum toxin (Botox) in non-cosmetic head and neck conditions - An evidence based review. *Saudi Pharm J,* 25**,** 18-24.

BAKER, T. A. & SAUER, R. T. 2012. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim Biophys Acta,* 1823**,** 15-28.

BARKER, H. A. 1981. Amino acid degradation by anaerobic bacteria. *Annu Rev Biochem,* 50**,** 23-40.

BARONZIO, G., FREITAS, I., BARONZIO, A., BARONZIO, M., CRESPI, E. & NETTI, P. A. 2009. Barriers to drug delivery in cancer: Clinical implications. *In:* BARONZIO, G., FIORENTINI, G. & COGLE, C. R. (eds.) *Cancer Microenvironment and Therapeutic Implications: Tumor Pathophysiology Mechanisms and Therapeutic Strategies.* Dordrecht: Springer Netherlands.

BAYM, M., SHAKET, L., ANZAI, I. A., ADESINA, O. & BARSTOW, B. 2016. Rapid construction of a whole-genome transposon insertion collection for *Shewanella oneidensis* by Knockout Sudoku. *Nat Commun,* 7**,** 13270.

BELLIDO, C., LOUREIRO PINTO, M., COCA, M., GONZALEZ-BENITO, G. & GARCIA-CUBERO, M. T. 2014. Acetone-butanol-ethanol (ABE) production by *Clostridium beijerinckii* from wheat straw hydrolysates: efficient use of penta and hexa carbohydrates. *Bioresour Technol,* 167**,** 198-205.

BEN-YEHUDA, S., RUDNER, D. Z. & LOSICK, R. 2003. RacA, a bacterial protein that anchors chromosomes to the cell poles. *Science,* 299**,** 532-6.

BERGSTROM, S. L. & FOUTCH, G. L. 1983. *Effect of cocultures on the production of butanol by Clostridium sp. [C. butyricum, C. pasteurianum, C. butylicum, and C. acetobutylicum]*, ; Jet Propulsion Lab., Pasadena, CA.

BIEMONT, C. 2010. A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics,* 186**,** 1085-93.

BINDER, W. J., BRIN, M. F., BLITZER, A. & POGODA, J. M. 2002. Botulinum toxin type A (BOTOX) for treatment of migraine. *Dis Mon,* 48**,** 323-35.

BORRIELLO, S. P. 1998. Pathogenesis of *Clostridium difficile* infection. *J Antimicrob Chemother,* 41 Suppl C**,** 13-9.

BRADBURY, M., GREENFIELD, P., MIDGLEY, D., LI, D., TRAN-DINH, N., VRIESEKOOP, F. & BROWN, J. L. 2012. Draft genome sequence of *Clostridium sporogenes* PA 3679, the common nontoxigenic surrogate for proteolytic *Clostridium botulinum*. *J Bacteriol,* 194**,** 1631-2.

BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R. L., TORRE, L. A. & JEMAL, A. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin,* 68**,** 394-424.

BREZGIN, S., KOSTYUSHEVA, A., KOSTYUSHEV, D. & CHULANOV, V. 2019. Dead Cas systems: Types, principles, and applications. *Int J Mol Sci,* 20.

BRIN, M. F., TAN, K. & GALLAGHER, C. 2009. Differences in botulinum toxins: time to end the confusion. *J Cosmet Laser Ther,* 11, 180-1.

BROWN, J. L., TRAN-DINH, N. & CHAPMAN, B. 2012. *Clostridium sporogenes* PA 3679 and its uses in the derivation of thermal processing schedules for low-acid shelf-stable foods and as a research model for proteolytic *Clostridium botulinum*. *J Food Prot,* 75, 779-92.

BROWN, J. M. & WILSON, W. R. 2004. Exploiting tumour hypoxia in cancer treatment. *Nat Rev Cancer,* 4, 437-47.

BRUGGEMANN, H., BAUMER, S., FRICKE, W. F., WIEZER, A., LIESEGANG, H., DECKER, I., HERZBERG, C., MARTINEZ-ARIAS, R., MERKL, R., HENNE, A. & GOTTSCHALK, G. 2003. The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. *Proc Natl Acad Sci U S A,* 100, 1316-21.

BRUNT, J., CROSS, K. L. & PECK, M. W. 2015. Apertures in the *Clostridium sporogenes* spore coat and exosporium align to facilitate emergence of the vegetative cell. *Food Microbiol,* 51, 45-50.

BURKE, K. E. & LAMONT, J. T. 2014. *Clostridium difficile* infection: a worldwide disease. *Gut Liver,* 8, 1-6.

BURNS, D. A., HEAP, J. T. & MINTON, N. P. 2010. *Clostridium difficile* spore germination: an update. *Res Microbiol,* 161, 730-4.

BURTON, B. M., MARQUIS, K. A., SULLIVAN, N. L., RAPOPORT, T. A. & RUDNER, D. Z. 2007. The ATPase SpoIIIE transports DNA across fused septal membranes during sporulation in *Bacillus subtilis*. *Cell,* 131, 1301-12.

BUTLER, R. R., 3RD, SCHILL, K. M., WANG, Y. & POMBERT, J. F. 2017. Genetic characterization of the exceptionally high heat resistance of the non-toxic surrogate *Clostridium sporogenes* PA 3679. *Front Microbiol,* 8.

CANADAS, I. C., GROOTHUIS, D., ZYGOUROPOULOU, M., RODRIGUES, R. & MINTON, N. P. 2019. RiboCas: A universal CRISPR-based editing tool for *Clostridium*. *ACS Synth Biol,* 8, 1379-1390.

CANO, R. J. & BORUCKI, M. K. 1995. Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber. *Science,* 268, 1060-4.

CAO, J., WU, L., ZHANG, S. M., LU, M., CHEUNG, W. K., CAI, W., GALE, M., XU, Q. & YAN, Q. 2016. An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting. *Nucleic Acids Res,* 44, e149.

CARTER, H. L., 3RD & MORAN, C. P., JR. 1986. New RNA polymerase sigma factor under spo0 control in *Bacillus subtilis*. *Proc Natl Acad Sci U S A,* 83, 9438-42.

CARTMAN, S. T. 2011. Time to consider *Clostridium* probiotics? *Future Microbiol,* 6, 969-71.

CARTMAN, S. T. & MINTON, N. P. 2010. A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. *Appl Environ Microbiol,* 76, 1103-9.

CHA, J. H. & STEWART, G. C. 1997. The divIVA minicell locus of *Bacillus subtilis*. *J Bacteriol,* 179, 1671-83.

CHARNEY, J., FISHER, W. P. & HEGARTY, C. P. 1951. Managanese as an essential element for sporulation in the genus *Bacillus*. *J Bacteriol,* 62, 145-8.

CHASTANET, A. & LOSICK, R. 2007. Engulfment during sporulation in *Bacillus subtilis* is governed by a multi-protein complex containing tandemly acting autolysins. *Mol Microbiol,* 64, 139-52.

CHAUDHURI, R. R., ALLEN, A. G., OWEN, P. J., SHALOM, G., STONE, K., HARRISON, M., BURGIS, T. A., LOCKYER, M., GARCIA-LARA, J., FOSTER, S. J., PLEASANCE, S. J., PETERS, S. E., MASKELL, D. J. & CHARLES, I. G. 2009. Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics,* 10, 291.

CLARKE, S., LOPEZ-DIAZ, I. & MANDELSTAM, J. 1986. Use of lacZ gene fusions to determine the dependence pattern of the sporulation gene spoIID in spo mutants of *Bacillus subtilis*. *J Gen Microbiol,* 132, 2987-94.

COHEN, A., FISHER, W. D., CURTISS, R., 3RD & ADLER, H. I. 1968. The properties of DNA transferred to minicells during conjugation. *Cold Spring Harb Symp Quant Biol,* 33, 635-41.

COHN, F. 1896. *Beiträge zur Biologie der Pflanzen*, Duncker & Humblot.

COOTE, J. G. 1972. Sporulation in *Bacillus subtilis*. Characterization of oligosporogenous mutants and comparison of their phenotypes with those of asporogenous mutants. *J Gen Microbiol,* 71, 1-15.

COSBY, W. M. & ZUBER, P. 1997. Regulation of *Bacillus subtilis* sigmaH (spo0H) and AbrB in response to changes in external pH. *J Bacteriol,* 179, 6778-87.

CUI, L. & BIKARD, D. 2016. Consequences of Cas9 cleavage in the chromosome of Escherichia coli. *Nucleic Acids Res,* 44, 4243-51.

CULLMANN, G., FIEN, K., KOBAYASHI, R. & STILLMAN, B. 1995. Characterization of the five replication factor C genes of *Saccharomyces cerevisiae*. *Mol Cell Biol,* 15, 4661-71.

DANG, L. H., BETTEGOWDA, C., HUSO, D. L., KINZLER, K. W. & VOGELSTEIN, B. 2001. Combination bacteriolytic therapy for the treatment of experimental tumors. *Proc Natl Acad Sci U S A,* 98, 15155-60.

DANIEL, R. A. & ERRINGTON, J. 1993. Cloning, DNA sequence, functional analysis and transcriptional regulation of the genes encoding dipicolinic acid synthetase required for sporulation in *Bacillus subtilis*. *J Mol Biol,* 232, 468-83.

DAWES, I. W., KAY, D. & MANDELSTAM, J. 1971. Determining effect of growth medium on the shape and position of daughter chromosomes and on sporulation in *Bacillus subtilis*. *Nature,* 230, 567-9.

DE BOER, P. A., CROSSLEY, R. E., HAND, A. R. & ROTHFIELD, L. I. 1991. The MinD protein is a membrane ATPase required for the correct placement of the *Escherichia coli* division site. *EMBO J,* 10, 4371-80.

DE BOER, P. A., CROSSLEY, R. E. & ROTHFIELD, L. I. 1989. A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in *Escherichia coli*. *Cell,* 56, 641-9.

DEAKIN, L. J., CLARE, S., FAGAN, R. P., DAWSON, L. F., PICKARD, D. J., WEST, M. R., WREN, B. W., FAIRWEATHER, N. F., DOUGAN, G. & LAWLEY, T. D. 2012. The *Clostridium difficile* spo0A gene is a persistence and transmission factor. *Infect Immun,* 80, 2704-11.

DEMBEK, M., BARQUIST, L., BOINETT, C. J., CAIN, A. K., MAYHO, M., LAWLEY, T. D., FAIRWEATHER, N. F. & FAGAN, R. P. 2015. High-throughput analysis of gene essentiality and sporulation in *Clostridium difficile*. *MBio,* 6, e02383.

DEMBEK, M., KELLY, A., BARWINSKA-SENDRA, A., TARRANT, E., STANLEY, W. A., VOLLMER, D., BIBOY, J., GRAY, J., VOLLMER, W. & SALGADO, P. S. 2018. Peptidoglycan degradation machinery in *Clostridium difficile* forespore engulfment. *Mol Microbiol,* 110, 390-410.

DERRE, I., RAPOPORT, G. & MSADEK, T. 1999. CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in gram-positive bacteria. *Mol Microbiol,* 31, 117-31.

DING, Y., LI, H., CHEN, L. L. & XIE, K. 2016. Recent advances in genome editing using CRISPR/Cas9. *Front Plant Sci,* 7, 703.

DOAN, T., MORLOT, C., MEISNER, J., SERRANO, M., HENRIQUES, A. O., MORAN, C. P., JR. & RUDNER, D. Z. 2009. Novel secretion apparatus maintains spore integrity and developmental gene expression in *Bacillus subtilis*. *PLoS Genet,* 5, e1000566.

DORNBUSCH, K., NORD, C. E. & DAHLBACK, A. 1975. Antibiotic susceptibility of *Clostridium* species isolated from human infections. *Scand J Infect Dis,* 7, 127-34.

DRIKS, A. 1999. *Bacillus subtilis* spore coat. *Microbiol Mol Biol Rev,* 63, 1-20.

DRIKS, A. & LOSICK, R. 1991. Compartmentalized expression of a gene under the control of sporulation transcription factor sigma E in *Bacillus subtilis*. *Proc Natl Acad Sci U S A,* 88, 9934-8.

DUNCAN, C. L. & STRONG, D. H. 1968. Improved medium for sporulation of *Clostridium perfringens*. *Appl Microbiol,* 16, 82-9.

DÜRRE, P. 2005. *Handbook on Clostridia*, CRC press.

EDWARDS, D. H. & ERRINGTON, J. 1997. The *Bacillus subtilis* DivIVA protein targets to the division septum and controls the site specificity of cell division. *Mol Microbiol,* 24, 905-15.

EICHENBERGER, P., FAWCETT, P. & LOSICK, R. 2001. A three-protein inhibitor of polar septation during sporulation in *Bacillus subtilis*. *Mol Microbiol,* 42, 1147-62.

EICHENBERGER, P., FUJITA, M., JENSEN, S. T., CONLON, E. M., RUDNER, D. Z., WANG, S. T., FERGUSON, C., HAGA, K., SATO, T., LIU, J. S. & LOSICK, R. 2004. The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol,* 2, e328.

EICHENBERGER, P., JENSEN, S. T., CONLON, E. M., VAN OOIJ, C., SILVAGGI, J., GONZALEZ-PASTOR, J. E., FUJITA, M., BEN-YEHUDA, S., STRAGIER, P., LIU, J. S. & LOSICK, R. 2003. The sigmaE regulon and the identification of additional sporulation genes in *Bacillus subtilis*. *J Mol Biol,* 327, 945-72.

EID, A., ALSHAREEF, S. & MAHFOUZ, M. M. 2018. CRISPR base editors: genome editing without double-stranded breaks. *Biochem J,* 475, 1955-1964.

EL JAAM, O., FLISS, I. & AIDER, M. 2017. Effect of electro-activated aqueous solutions, nisin and moderate heat treatment on the inactivation of *Clostridium sporogenes* PA 3679 spores in green beans puree and whole green beans. *Anaerobe,* 47, 173-182.

ELLNER, P. D. 1956. A medium promoting rapid quantitative sporulation in *Clostridium perfringens*. *J Bacteriol,* 71, 495-6.

ESWARAMOORTHY, P., WINTER, P. W., WAWRZUSIN, P., YORK, A. G., SHROFF, H. & RAMAMURTHI, K. S. 2014. Asymmetric division and differential gene expression during a bacterial developmental program requires DivIVA. *PLoS Genet,* 10, e1004526.

FAN, N., CUTTING, S. & LOSICK, R. 1992. Characterization of the *Bacillus subtilis* sporulation gene spoVK. *J Bacteriol,* 174, 1053-4.

FAWCETT, P., EICHENBERGER, P., LOSICK, R. & YOUNGMAN, P. 2000. The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A,* 97, 8063-8.

FEUCHT, A., EVANS, L. & ERRINGTON, J. 2003. Identification of sporulation genes by genome-wide analysis of the sigmaE regulon of *Bacillus subtilis*. *Microbiology,* 149, 3023-3034.

FIMLAID, K. A. & SHEN, A. 2015. Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Curr Opin Microbiol,* 24, 88-95.

FINEGOLD, S. M., SONG, Y. & LIU, C. 2002. Taxonomy--General comments and update on taxonomy of Clostridia and Anaerobic cocci. *Anaerobe,* 8, 283-5.

FINNE, G. & MATCHES, J. R. 1974. Low-temperature-growing clostridia from marine sediments. *Can J Microbiol,* 20, 1639-45.

FIRCZUK, M. & BOCHTLER, M. 2007. Folds and activities of peptidoglycan amidases. *FEMS Microbiol Rev,* 31, 676-91.

FLYTHE, M. D. & RUSSELL, J. B. 2004. The effect of pH and a bacteriocin (bovicin HC5) on *Clostridium sporogenes* MD1, a bacterium that has the ability to degrade amino acids in ensiled plant materials. *FEMS Microbiol Ecol,* 47, 215-22.

FORTI, F., MAURI, V., DEHO, G. & GHISOTTI, D. 2011. Isolation of conditional expression mutants in *Mycobacterium tuberculosis* by transposon mutagenesis. *Tuberculosis (Edinb),* 91, 569-78.

FOULGER, D. & ERRINGTON, J. 1991. Sequential activation of dual promoters by different sigma factors maintains spoVJ expression during successive developmental stages of *Bacillus subtilis*. *Mol Microbiol,* 5, 1363-73.

FRANK, D. N., ST AMAND, A. L., FELDMAN, R. A., BOEDEKER, E. C., HARPAZ, N. & PACE, N. R. 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A,* 104, 13780-5.

FRANKS, A. H., HARMSEN, H. J., RAANGS, G. C., JANSEN, G. J., SCHUT, F. & WELLING, G. W. 1998. Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. *Appl Environ Microbiol,* 64, 3336-45.

FREES, D., QAZI, S. N., HILL, P. J. & INGMER, H. 2003. Alternative roles of ClpX and ClpP in Staphylococcus aureus stress tolerance and virulence. *Mol Microbiol,* 48, 1565-78.

FREES, D., SAVIJOKI, K., VARMANEN, P. & INGMER, H. 2007. Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Mol Microbiol,* 63, 1285-95.

FREES, D., SORENSEN, K. & INGMER, H. 2005. Global virulence regulation in *Staphylococcus aureus*: pinpointing the roles of ClpP and ClpX in the sar/agr regulatory network. *Infect Immun,* 73, 8100-8.

FUJITA, M., GONZALEZ-PASTOR, J. E. & LOSICK, R. 2005. High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis*. *J Bacteriol,* 187, 1357-68.

GALPERIN, M. Y., MEKHEDOV, S. L., PUIGBO, P., SMIRNOV, S., WOLF, Y. I. & RIGDEN, D. J. 2012. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ Microbiol,* 14, 2870-90.

GEIDER, K. & HOFFMANN-BERLING, H. 1981. Proteins controlling the helical structure of DNA. *Annu Rev Biochem,* 50, 233-60.

GERHARDT, P. 1989. Spore thermoresistance mechanisms. *Regulation. of Procaryotic Development*.

GERTH, U., KRUGER, E., DERRE, I., MSADEK, T. & HECKER, M. 1998. Stress induction of the *Bacillus subtilis* clpP gene encoding a homologue of the proteolytic component of the Clp protease and the involvement of ClpP and ClpX in stress tolerance. *Mol Microbiol,* 28, 787-802.

GERTH, U., WIPAT, A., HARWOOD, C. R., CARTER, N., EMMERSON, P. T. & HECKER, M. 1996. Sequence and transcriptional analysis of clpX, a class-III heat-shock gene of *Bacillus subtilis*. *Gene,* 181, 77-83.

GOLDBERG, A. L. & ST JOHN, A. C. 1976. Intracellular protein degradation in mammalian and bacterial cells: Part 2. *Annu Rev Biochem,* 45, 747-803.

GONZALEZ, D. J., LEE, S. W., HENSLER, M. E., MARKLEY, A. L., DAHESH, S., MITCHELL, D. A., BANDEIRA, N., NIZET, V., DIXON, J. E. & DORRESTEIN, P. C. 2010. Clostridiolysin S, a post-translationally modified biotoxin from *Clostridium botulinum*. *J Biol Chem,* 285, 28220-8.

GOORHUIS, A., BAKKER, D., CORVER, J., DEBAST, S. B., HARMANUS, C., NOTERMANS, D. W., BERGWERFF, A. A., DEKKER, F. W. & KUIJPER, E. J. 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction ribotype 078. *Clin Infect Dis,* 47, 1162-70.

GORBALENYA, A. E. & KOONIN, E. V. 1989. Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res,* 17, 8413-40.

GOTTESMAN, S. 1996. Proteases and their targets in *Escherichia coli*. *Annu Rev Genet,* 30, 465-506.

GOTTESMAN, S. 2003. Proteolysis in bacterial regulatory circuits. *Annu Rev Cell Dev Biol,* 19, 565-87.

GOTTUMUKKALA, L. D., PARAMESWARAN, B., VALAPPIL, S. K., MATHIYAZHAKAN, K., PANDEY, A. & SUKUMARAN, R. K. 2013. Biobutanol production from rice straw by a non acetone producing *Clostridium sporogenes* BE01. *Bioresour Technol,* 145, 182-7.

GUENTHER, B., ONRUST, R., SALI, A., O'DONNELL, M. & KURIYAN, J. 1997. Crystal structure of the delta' subunit of the clamp-loader complex of *Escherichia coli* DNA polymerase III. *Cell,* 91, 335-45.

GUTIERREZ, J., SMITH, R. & POGLIANO, K. 2010. SpoIID-mediated peptidoglycan degradation is required throughout engulfment during *Bacillus subtilis* sporulation. *J Bacteriol,* 192, 3174-86.

HAEUSSER, D. P., LEE, A. H., WEART, R. B. & LEVIN, P. A. 2009. ClpX inhibits FtsZ assembly in a manner that does not require its ATP hydrolysis-dependent chaperone activity. *J Bacteriol,* 191, 1986-91.

HALBEDEL, S., HAHN, B., DANIEL, R. A. & FLIEGER, A. 2012. DivIVA affects secretion of virulence-related autolysins in *Listeria monocytogenes*. *Mol Microbiol,* 83, 821-39.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: The next generation. *Cell,* 144, 646-74.

HANSON, P. I. & WHITEHEART, S. W. 2005. AAA+ proteins: Have engine, will work. *Nat Rev Mol Cell Biol,* 6, 519-29.

HAQUE, A., SUGIMOTO, N., HORIGUCHI, Y., OKABE, T., MIYATA, T., IWANAGA, S. & MATSUDA, M. 1992. Production, purification, and characterization of botulinolysin, a thiol-activated hemolysin of *Clostridium botulinum*. *Infect Immun,* 60, 71-8.

HARA-KUDO, Y., OGURA, A., NOGUCHI, Y. & KUMAGAI, S. 1997. Characteristics of toxicity and haemorrhagic toxin produced by *Clostridium sporogenes* in various animals and cultured cells. *J Med Microbiol,* 46, 270-5.

HARA-KUDO, Y., YAMAKAWA, Y. & KUMAGAI, S. 1996. Purification and some properties of *Clostridium sporogenes* hemorrhagic toxin. *Biochem Biophys Res Commun,* 227, 413-8.

HARALDSEN, J. D. & SONENSHEIN, A. L. 2003. Efficient sporulation in *Clostridium difficile* requires disruption of the sigmaK gene. *Mol Microbiol,* 48, 811-21.

HATHEWAY, C. L. 1990. Toxigenic Clostridia. *Clin Microbiol Rev,* 3, 66-98.

HEAP, J. T., KUEHNE, S. A., EHSAAN, M., CARTMAN, S. T., COOKSLEY, C. M., SCOTT, J. C. & MINTON, N. P. 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J Microbiol Methods,* 80, 49-55.

HEAP, J. T., PENNINGTON, O. J., CARTMAN, S. T., CARTER, G. P. & MINTON, N. P. 2007. The ClosTron: A universal gene knock-out system for the genus *Clostridium*. *J Microbiol Methods,* 70, 452-64.

HEAP, J. T., PENNINGTON, O. J., CARTMAN, S. T. & MINTON, N. P. 2009. A modular system for *Clostridium* shuttle plasmids. *J Microbiol Methods,* 78, 79-85.

HEAP, J. T., THEYS, J., EHSAAN, M., KUBIAK, A. M., DUBOIS, L., PAESMANS, K., MELLAERT, L. V., KNOX, R., KUEHNE, S. A., LAMBIN, P. & MINTON, N. P. 2014. Spores of *Clostridium* engineered for clinical efficacy and safety cause regression and cure of tumors in vivo. *Oncotarget,* 5, 1761-9.

HELLER, H. H. 1922. Certain genera of the Clostridiaceae: Studies in pathogenic anaerobes. V. *J Bacteriol,* 7, 1-38.

HENDERSON, D. A. 1999. The looming threat of bioterrorism. *Science,* 283, 1279-82.

HENRIQUES, A. O. & MORAN, C. P., JR. 2007. Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol,* 61, 555-88.

HILL, K. K., SMITH, T. J., HELMA, C. H., TICKNOR, L. O., FOLEY, B. T., SVENSSON, R. T., BROWN, J. L., JOHNSON, E. A., SMITH, L. A., OKINAKA, R. T., JACKSON, P. J. & MARKS, J. D. 2007. Genetic diversity among Botulinum Neurotoxin-producing clostridial strains. *J Bacteriol,* 189, 818-32.

HILL, S. H. 1983. SpoVH and spoVJ--new sporulation loci in *Bacillus subtilis* 168. *J Gen Microbiol,* 129, 293-302.

HODGKISS, W., ORDAL, Z. J. & CANN, D. C. 1967. The morphology and ultrastructure of the spore and exosporium of some *Clostridium* species. *J Gen Microbiol,* 47, 213-25.

HOFFMAN, R. M., TAN, Y., LI, S., HAN, Q., YAGI, S., TAKAKURA, T., TAKIMOTO, A., INAGAKI, K. & KUDOU, D. 2019. Development of recombinant methioninase for cancer treatment. *Methods Mol Biol,* 1866, 107-131.

HOLD, G. L., PRYDE, S. E., RUSSELL, V. J., FURRIE, E. & FLINT, H. J. 2002. Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. *FEMS Microbiol Ecol,* 39**,** 33-9.

HSU, P. D., LANDER, E. S. & ZHANG, F. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell,* 157**,** 1262-1278.

HU, Z. & LUTKENHAUS, J. 1999. Topological regulation of cell division in *Escherichia coli* involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE. *Mol Microbiol,* 34**,** 82-90.

HUANG, H., CHAI, C., LI, N., ROWE, P., MINTON, N. P., YANG, S., JIANG, W. & GU, Y. 2016. CRISPR/Cas9-based efficient genome editing in *Clostridium ljungdahlii*, an autotrophic gas-fermenting bacterium. *ACS Synth Biol,* 5**,** 1355-1361.

HUDSON, K. D., CORFE, B. M., KEMP, E. H., FEAVERS, I. M., COOTE, P. J. & MOIR, A. 2001. Localization of GerAA and GerAC germination proteins in the *Bacillus subtilis* spore. *J Bacteriol,* 183**,** 4317-22.

HUGHES, J. M., BLUMENTHAL, J. R., MERSON, M. H., LOMBARD, G. L., DOWELL, V. R., JR. & GANGAROSA, E. J. 1981. Clinical features of types A and B food-borne botulism. *Ann Intern Med,* 95**,** 442-5.

HUSSAIN, H. A., ROBERTS, A. P. & MULLANY, P. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol,* 54**,** 137-41.

INGLE, P., GROOTHUIS, D., ROWE, P., HUANG, H., COCKAYNE, A., KUEHNE, S. A., JIANG, W., GU, Y., HUMPHREYS, C. M. & MINTON, N. P. 2019. Generation of a fully erythromycin-sensitive strain of *Clostridioides difficile* using a novel CRISPR-Cas9 genome editing system. *Sci Rep,* 9**,** 8123.

INGRAM, M. & ROBINSON, R. A Discussion of the Literature on Botulism in relation to Acid Foods. Proceedings of the Society for Applied Bacteriology, 1951. 73-84.

ISA, K., OKA, K., BEAUCHAMP, N., SATO, M., WADA, K., OHTANI, K., NAKANISHI, S., MCCARTNEY, E., TANAKA, M., SHIMIZU, T., KAMIYA, S., KRUGER, C. & TAKAHASHI, M. 2016. Safety assessment of the *Clostridium butyricum* MIYAIRI 588(R) probiotic strain including evaluation of antimicrobial sensitivity and presence of *Clostridium* toxin genes in vitro and teratogenicity in vivo. *Hum Exp Toxicol,* 35**,** 818-32.

IVICS, Z., HACKETT, P. B., PLASTERK, R. H. & IZSVAK, Z. 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell,* 91**,** 501-10.

IVICS, Z. & IZSVAK, Z. 2010. The expanding universe of transposon technologies for gene and cell engineering. *Mob DNA,* 1**,** 25.

JACOBS, C. & SHAPIRO, L. 1999. Bacterial cell division: a moveable feast. *Proc Natl Acad Sci U S A,* 96**,** 5891-3.

JACOBS, M. A., ALWOOD, A., THAIPISUTTIKUL, I., SPENCER, D., HAUGEN, E., ERNST, S., WILL, O., KAUL, R., RAYMOND, C., LEVY, R., CHUN-RONG, L., GUENTHNER, D., BOVEE, D., OLSON, M. V. & MANOIL, C. 2003. Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A,* 100**,** 14339-44.

JACOBSON, J. W., MEDHORA, M. M. & HARTL, D. L. 1986. Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proc Natl Acad Sci U S A,* 83, 8684-8.

JANGANAN, T. K., MULLIN, N., TZOKOV, S. B., STRINGER, S., FAGAN, R. P., HOBBS, J. K., MOIR, A. & BULLOUGH, P. A. 2016. Characterization of the spore surface and exosporium proteins of *Clostridium sporogenes*; implications for *Clostridium botulinum* group I strains. *Food Microbiol,* 59, 205-12.

JENAL, U. & FUCHS, T. 1998. An essential protease involved in bacterial cell-cycle control. *EMBO J,* 17, 5658-69.

JENSEN, C., FOSBERG, M. J., THALSO-MADSEN, I., BAEK, K. T. & FREES, D. 2019. *Staphylococcus aureus* ClpX localizes at the division septum and impacts transcription of genes involved in cell division, T7-secretion, and SaPI5-excision. *Sci Rep,* 9, 16456.

JOSHI, S. A., HERSCH, G. L., BAKER, T. A. & SAUER, R. T. 2004. Communication between ClpX and ClpP during substrate processing and degradation. *Nat Struct Mol Biol,* 11, 404-11.

KAIN, J., HE, G. G. & LOSICK, R. 2008. Polar localization and compartmentalization of ClpP proteases during growth and sporulation in *Bacillus subtilis*. *J Bacteriol,* 190, 6749-57.

KARBERG, M., GUO, H., ZHONG, J., COON, R., PERUTKA, J. & LAMBOWITZ, A. M. 2001. Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat Biotechnol,* 19, 1162-7.

KARDON, J. R., YIEN, Y. Y., HUSTON, N. C., BRANCO, D. S., HILDICK-SMITH, G. J., RHEE, K. Y., PAW, B. H. & BAKER, T. A. 2015. Mitochondrial ClpX activates a key enzyme for heme biosynthesis and erythropoiesis. *Cell,* 161, 858-67.

KATOR, H. & RHODES, M. 2003. Detection, enumeration and identification of environmental microorganisms of public health significance. *The handbook of water and wastewater microbiology*, 113-126.

KAUSHAL, M., AHLAWAT, S., MAKUT, B. B., GOSWAMI, G. & DAS, D. 2019. Dual substrate fermentation strategy utilizing rice straw hydrolysate and crude glycerol for liquid biofuel production by *Clostridium sporogenes* NCIM 2918. *Biomass and Bioenergy,* 127, 105257.

KAUSHAL, M., AHLAWAT, S., MUKHERJEE, M., MUTHURAJ, M., GOSWAMI, G. & DAS, D. 2017. Substrate dependent modulation of butanol to ethanol ratio in non-acetone forming *Clostridium sporogenes* NCIM 2918. *Bioresour Technol,* 225, 349-358.

KAVAL, K. G. & HALBEDEL, S. 2012. Architecturally the same, but playing a different game: the diverse species-specific roles of DivIVA proteins. *Virulence,* 3, 406-7.

KELLY, A. & SALGADO, P. S. 2019. The engulfasome in *Clostridium difficile*: Variations on protein machineries. *Anaerobe,* 60, 102091.

KEMP, J. T., DRIKS, A. & LOSICK, R. 2002. FtsA mutants of *Bacillus subtilis* impaired in sporulation. *J Bacteriol,* 184, 3856-63.

KENG, V. W., VILLANUEVA, A., CHIANG, D. Y., DUPUY, A. J., RYAN, B. J., MATISE, I., SILVERSTEIN, K. A., SARVER, A., STARR, T. K., AKAGI, K., TESSAROLLO, L., COLLIER, L. S., POWERS, S., LOWE, S. W., JENKINS, N. A., COPELAND, N. G., LLOVET, J. M. & LARGAESPADA, D. A. 2009. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nat Biotechnol,* 27, 264-74.

KENNEDY, M. J., READER, S. L. & SWIERCZYNSKI, L. M. 1994. Preservation records of micro-organisms: evidence of the tenacity of life. *Microbiology (Reading),* 140 (Pt 10), 2513-29.

KIHM, D. J., HUTTON, M. T., HANLIN, J. H. & JOHNSON, E. A. 1990. Influence of transition metals added during sporulation on heat resistance of *Clostridium botulinum* 113B spores. *Appl Environ Microbiol,* 56, 681-5.

KIM, C. H., CHENEY, R. & WOODBURN, M. 1967. Sporulation of *Clostridium perfringens* in a modified medium and selected foods. *Appl Microbiol,* 15, 871-6.

KIRK, D. G., DAHLSTEN, E., ZHANG, Z., KORKEALA, H. & LINDSTROM, M. 2012. Involvement of *Clostridium botulinum* ATCC 3502 sigma factor K in early-stage sporulation. *Appl Environ Microbiol,* 78, 4590-6.

KIRK, D. G., ZHANG, Z., KORKEALA, H. & LINDSTROM, M. 2014. Alternative sigma factors SigF, SigE, and SigG are essential for sporulation in *Clostridium botulinum* ATCC 3502. *Appl Environ Microbiol,* 80, 5141-50.

KLOBUTCHER, L. A., RAGKOUSI, K. & SETLOW, P. 2006. The *Bacillus subtilis* spore coat provides "eat resistance" during phagocytic predation by the protozoan *Tetrahymena thermophila*. *Proc Natl Acad Sci U S A,* 103, 165-70.

KOCH, A. 1888. *Über morphologie und entwickelungsgeschichte einiger endosporer bacterienformen*.

KOMOR, A. C., BADRAN, A. H. & LIU, D. R. 2018. Editing the genome without double-stranded DNA breaks. *ACS Chem Biol,* 13, 383-388.

KONOVALOVA, A., SOGAARD-ANDERSEN, L. & KROOS, L. 2014. Regulated proteolysis in bacterial development. *FEMS Microbiol Rev,* 38, 493-522.

KOO, B. M., KRITIKOS, G., FARELLI, J. D., TODOR, H., TONG, K., KIMSEY, H., WAPINSKI, I., GALARDINI, M., CABAL, A., PETERS, J. M., HACHMANN, A. B., RUDNER, D. Z., ALLEN, K. N., TYPAS, A. & GROSS, C. A. 2017. Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell Syst,* 4, 291-305 e7.

KOONIN, E. V. 1993a. A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res,* 21, 2541-7.

KOONIN, E. V. 1993b. A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J Mol Biol,* 229, 1165-74.

KOPREK, T., MCELROY, D., LOUWERSE, J., WILLIAMS-CARRIER, R. & LEMAUX, P. G. 1999. Negative selection systems for transgenic barley (*Hordeum vulgare* L.): comparison of bacterial codA- and cytochrome P450 gene-mediated selection. *Plant J,* 19, 719-26.

KORNBERG, A. & BAKER, T. A. 1992. DNA replication 2nd edition. Freeman, WD and Co.(Ed.), University Science Books, California.

KREIS, W. & HESSION, C. 1973. Biological effects of enzymatic deprivation of L-methionine in cell culture and an experimental tumor. *Cancer Res,* 33, 1866-9.

KRISTICH, C. J., NGUYEN, V. T., LE, T., BARNES, A. M., GRINDLE, S. & DUNNY, G. M. 2008. Development and use of an efficient system for random mariner transposon mutagenesis to identify novel genetic determinants of biofilm formation in the core Enterococcus faecalis genome. *Appl Environ Microbiol,* 74, 3377-86.

KRUGER, E. & HECKER, M. 1998. The first gene of the *Bacillus subtilis* clpC operon, ctsR, encodes a negative regulator of its own operon and other class III heat shock genes. *J Bacteriol,* 180, 6681-8.

KRUGER, E., MSADEK, T. & HECKER, M. 1996. Alternate promoters direct stress-induced transcription of the *Bacillus subtilis* clpC operon. *Mol Microbiol,* 20, 713-23.

KUBIAK, A. M. & MINTON, N. P. 2015. The potential of clostridial spores as therapeutic delivery vehicles in tumour therapy. *Res Microbiol,* 166, 244-54.

KUBIAK, A. M., POEHLEIN, A., BUDD, P., KUEHNE, S. A., WINZER, K., THEYS, J., LAMBIN, P., DANIEL, R. & MINTON, N. P. 2015. Complete genome sequence of the nonpathogenic soil-dwelling bacterium *Clostridium sporogenes* strain NCIMB 10696. *Genome announcements,* 3, e00942-15.

KUEHNE, S. A. & MINTON, N. P. 2012. ClosTron-mediated engineering of *Clostridium*. *Bioengineered,* 3, 247-54.

KUNAU, W. H., BEYER, A., FRANKEN, T., GOTTE, K., MARZIOCH, M., SAIDOWSKY, J., SKALETZ-ROROWSKI, A. & WIEBEL, F. F. 1993. Two complementary approaches to study peroxisome biogenesis in *Saccharomyces cerevisiae*: forward and reversed genetics. *Biochimie,* 75, 209-24.

KUWAHARA, T., OGURA, Y., OSHIMA, K., KUROKAWA, K., OOKA, T., HIRAKAWA, H., ITOH, T., NAKAYAMA-IMAOHJI, H., ICHIMURA, M., ITOH, K., ISHIFUNE, C., MAEKAWA, Y., YASUTOMO, K., HATTORI, M. & HAYASHI, T. 2011. The lifestyle of the segmented filamentous bacterium: a non-culturable gut-associated immunostimulating microbe inferred by whole-genome sequencing. *DNA Res,* 18, 291-303.

LAABERKI, M. H. & DWORKIN, J. 2008. Role of spore coat proteins in the resistance of *Bacillus subtilis* spores to *Caenorhabditis elegans* predation. *J Bacteriol,* 190, 6197-203.

LABRECK, C. J., MAY, S., VIOLA, M. G., CONTI, J. & CAMBERG, J. L. 2017. The protein chaperone ClpX targets native and non-native aggregated substrates for remodeling, disassembly, and degradation with ClpP. *Front Mol Biosci,* 4, 26.

LAMBIN, P., THEYS, J., LANDUYT, W., RIJKEN, P., VAN DER KOGEL, A., VAN DER SCHUEREN, E., HODGKISS, R., FOWLER, J., NUYTS, S., DE BRUIJN, E., VAN MELLAERT, L. & ANNE, J. 1998. Colonisation of *Clostridium* in the body is restricted to hypoxic and necrotic areas of tumours. *Anaerobe,* 4, 183-8.

LANGRIDGE, G. C., PHAN, M. D., TURNER, D. J., PERKINS, T. T., PARTS, L., HAASE, J., CHARLES, I., MASKELL, D. J., PETERS, S. E., DOUGAN, G., WAIN, J., PARKHILL, J. & TURNER, A. K. 2009. Simultaneous assay of every *Salmonella typhi* gene using one million transposon mutants. *Genome Res,* 19, 2308-16.

LAWSON, P. A., CITRON, D. M., TYRRELL, K. L. & FINEGOLD, S. M. 2016. Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prevot 1938. *Anaerobe,* 40, 95-9.

LE BRETON, Y., MOHAPATRA, N. P. & HALDENWANG, W. G. 2006. In vivo random mutagenesis of *Bacillus subtilis* by use of TnYLB-1, a mariner-based transposon. *Appl Environ Microbiol,* 72, 327-33.

LEE, J. 2015. *CRISPR-Cas9: Targeted Genome Editing* [Online]. Available: https://knowingneurons.com/2015/03/18/crispr-cas9-targeted-genome-editing/ [Accessed March 18 2015].

LEMMON, M. J., VAN ZIJL, P., FOX, M. E., MAUCHLINE, M. L., GIACCIA, A. J., MINTON, N. P. & BROWN, J. M. 1997. Anaerobic bacteria as a gene delivery system that is controlled by the tumor microenvironment. *Gene Ther,* 4, 791-6.

LENARCIC, R., HALBEDEL, S., VISSER, L., SHAW, M., WU, L. J., ERRINGTON, J., MARENDUZZO, D. & HAMOEN, L. W. 2009. Localisation of DivIVA by targeting to negatively curved membranes. *EMBO J,* 28, 2272-82.

LI, Y., BUTZIN, X. Y., DAVIS, A., SETLOW, B., KORZA, G., USTOK, F. I., CHRISTIE, G., SETLOW, P. & HAO, B. 2013. Activity and regulation of various forms of CwlJ, SleB, and YpeB proteins in degrading cortex peptidoglycan of spores of *Bacillus* species in vitro and during spore germination. *J Bacteriol,* 195, 2530-40.

LINDSTROM, M. & KORKEALA, H. 2006. Laboratory diagnostics of botulism. *Clin Microbiol Rev,* 19, 298-314.

LIU, G., DRAPER, G. C. & DONACHIE, W. D. 1998. FtsK is a bifunctional protein involved in cell division and chromosome localization in *Escherichia coli*. *Mol Microbiol,* 29, 893-903.

LIU, H., BOUILLAUT, L., SONENSHEIN, A. L. & MELVILLE, S. B. 2013. Use of a mariner-based transposon mutagenesis system to isolate *Clostridium perfringens* mutants deficient in gliding motility. *J Bacteriol,* 195, 629-36.

LIU, J., COSBY, W. M. & ZUBER, P. 1999. Role of lon and ClpX in the post-translational regulation of a sigma subunit of RNA polymerase required for cellular differentiation in *Bacillus subtilis*. *Mol Microbiol,* 33, 415-28.

LIU, J. & ZUBER, P. 2000. The ClpX protein of *Bacillus subtilis* indirectly influences RNA polymerase holoenzyme composition and directly stimulates sigma-dependent transcription. *Mol Microbiol,* 37, 885-97.

LIU, S. C., MINTON, N. P., GIACCIA, A. J. & BROWN, J. M. 2002. Anticancer efficacy of systemically delivered anaerobic bacteria as gene therapy vectors targeting tumor hypoxia/necrosis. *Gene Ther,* 9, 291-6.

LO, H. H., LIAO, C. T., LI, C. E., CHIANG, Y. C. & HSIAO, Y. M. 2020. The clpX gene plays an important role in bacterial attachment, stress tolerance, and virulence in *Xanthomonas campestris* pv. *campestris*. *Arch Microbiol,* 202, 597-607.

LOPETUSO, L. R., SCALDAFERRI, F., PETITO, V. & GASBARRINI, A. 2013. Commensal Clostridia: Leading players in the maintenance of gut homeostasis. *Gut Pathog,* 5, 23.

LOPEZ-DIAZ, I., CLARKE, S. & MANDELSTAM, J. 1986. spoIID operon of *Bacillus subtilis:* cloning and sequence. *J Gen Microbiol,* 132, 341-54.

LU, Z., TAKEUCHI, M. & SATO, T. 2007. The LysR-type transcriptional regulator YofA controls cell division through the regulation of expression of ftsW in *Bacillus subtilis*. *J Bacteriol,* 189, 5642-51.

LUECHAPATTANAPORN, K., WANG, Y., WANG, J., AL-HOLY, M., KANG, D. H., TANG, J. & HALLBERG, L. M. 2004. Microbial safety in radio-frequency processing of packaged foods. *Journal of Food Science,* 69, 201-206.

LYND, L. R., VAN ZYL, W. H., MCBRIDE, J. E. & LASER, M. 2005. Consolidated bioprocessing of cellulosic biomass: an update. *Curr Opin Biotechnol,* 16, 577-83.

MADDOCKS, S. E. & OYSTON, P. C. F. 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology (Reading),* 154, 3609-3623.

MAGLENNON, G. A., COOK, B. S., DEENEY, A. S., BOSSE, J. T., PETERS, S. E., LANGFORD, P. R., MASKELL, D. J., TUCKER, A. W., WREN, B. W., RYCROFT, A. N. & CONSORTIUM, B. R. T. 2013. Transposon mutagenesis in *Mycoplasma hyopneumoniae* using a novel mariner-based system for generating random mutations. *Vet Res,* 44, 124.

MAH, J. H., KANG, D. H. & TANG, J. 2008. Effects of minerals on sporulation and heat resistance of *Clostridium sporogenes*. *Int J Food Microbiol,* 128, 385-9.

MAH, J. H., KANG, D. H. & TANG, J. 2009. Comparison of viability and heat resistance of *Clostridium sporogenes* stored at different temperatures. *J Food Sci,* 74, M23-7.

MANDELSTAM, J., STERLINI, J. M. & KAY, D. 1971. Sporulation in *Bacillus subtilis*. Effect of medium on the form of chromosome replication and on initiation to sporulation in Bacillus subtilis. *Biochem J,* 125, 635-41.

MARSTON, A. L. & ERRINGTON, J. 1999. Selection of the midcell division site in *Bacillus subtilis* through MinD-dependent polar localization and activation of MinC. *Mol Microbiol,* 33, 84-96.

MARSTON, A. L., THOMAIDES, H. B., EDWARDS, D. H., SHARPE, M. E. & ERRINGTON, J. 1998. Polar localization of the MinD protein of *Bacillus subtilis* and its role in selection of the mid-cell division site. *Genes Dev,* 12, 3419-30.

MCCOY, J. G., LEVIN, E. J. & ZHOU, M. 2015. Structural insight into the PTS sugar transporter EIIC. *Biochim Biophys Acta,* 1850, 577-85.

MCDONEL, J. L. 1980. *Clostridium perfringens* toxins (type A, B, C, D, E). *Pharmacol Ther,* 10, 617-55.

MCGILLIVRAY, S. M., EBRAHIMI, C. M., FISHER, N., SABET, M., ZHANG, D. X., CHEN, Y., HASTE, N. M., AROIAN, R. V., GALLO, R. L., GUINEY, D. G., FRIEDLANDER, A. M., KOEHLER, T. M. & NIZET, V. 2009. ClpX contributes to innate defense peptide resistance and virulence phenotypes of *Bacillus anthracis*. *J Innate Immun,* 1, 494-506.

MEESKE, A. J., RODRIGUES, C. D., BRADY, J., LIM, H. C., BERNHARDT, T. G. & RUDNER, D. Z. 2016. High-throughput genetic screens identify a large and diverse collection of new sporulation genes in *Bacillus subtilis*. *PLoS Biol,* 14, e1002341.

MEI, J. M., NOURBAKHSH, F., FORD, C. W. & HOLDEN, D. W. 1997. Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Mol Microbiol,* 26, 399-407.

MEINHARDT, F., BUSSKAMP, M. & WITTCHEN, K. D. 1994. Cloning and sequencing of the leu C and npr M genes and a putative spo IV gene from *Bacillus megaterium* DSM319. *Appl Microbiol Biotechnol,* 41, 344-51.

MENDELSON, N. H. & COLE, R. M. 1972. Genetic regulation of cell division initiation in *Bacillus subtilis*. *J Bacteriol,* 112, 994-1003.

MICHEL, F. & FERAT, J. L. 1995. Structure and activities of group II introns. *Annu Rev Biochem,* 64, 435-61.

MILLS, D. A., MCKAY, L. L. & DUNNY, G. M. 1996. Splicing of a group II intron involved in the conjugative transfer of pRS01 in lactococci. *J Bacteriol,* 178, 3531-8.

MINTON, N. P. 2003. Clostridia in cancer therapy. *Nat Rev Microbiol,* 1, 237-42.

MINTON, N. P., BROWN, J. M., LAMBIN, P. & ANNÉ, J. 2001. Clostridia in cancer therapy. *Clostridia: biotechnology and medical applications.*

MINTON, N. P. & CLARKE, D. J. 2013. *Clostridia*, Springer US.

MINTON, N. P., EHSAAN, M., HUMPHREYS, C. M., LITTLE, G. T., BAKER, J., HENSTRA, A. M., LIEW, F., KELLY, M. L., SHENG, L., SCHWARZ, K. & ZHANG, Y. 2016. A roadmap for gene system development in *Clostridium*. *Anaerobe*.

MITCHELL, A. L., ATTWOOD, T. K., BABBITT, P. C., BLUM, M., BORK, P., BRIDGE, A., BROWN, S. D., CHANG, H. Y., EL-GEBALI, S., FRASER, M. I., GOUGH, J., HAFT, D. R., HUANG, H., LETUNIC, I., LOPEZ, R., LUCIANI, A., MADEIRA, F., MARCHLER-BAUER, A., MI, H., NATALE, D. A., NECCI, M., NUKA, G., ORENGO, C., PANDURANGAN, A. P., PAYSAN-LAFOSSE, T., PESSEAT, S., POTTER, S. C., QURESHI, M. A., RAWLINGS, N. D., REDASCHI, N., RICHARDSON, L. J., RIVOIRE, C., SALAZAR, G. A., SANGRADOR-VEGAS, A., SIGRIST, C. J. A., SILLITOE, I., SUTTON, G. G., THANKI, N., THOMAS, P. D., TOSATTO, S. C. E., YONG, S. Y. & FINN, R. D. 2019. InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res,* 47**,** D351-D360.

MIZUNO, M., MASUDA, S., TAKEMARU, K., HOSONO, S., SATO, T., TAKEUCHI, M. & KOBAYASHI, Y. 1996. Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the *Bacillus subtilis* genome containing the skin element and many sporulation genes. *Microbiology,* 142 (Pt 11)**,** 3103-11.

MOIR, A. & SMITH, D. A. 1990. The genetics of bacterial spore germination. *Annu Rev Microbiol,* 44**,** 531-53.

MOLONGOSKI, J. J. & KLUG, M. J. 1976. Characterization of anaerobic heterotrophic bacteria isolated from freshwater lake sediments. *Appl Environ Microbiol,* 31**,** 83-90.

MORLOT, C., UEHARA, T., MARQUIS, K. A., BERNHARDT, T. G. & RUDNER, D. Z. 2010. A highly coordinated cell wall degradation machine governs spore morphogenesis in *Bacillus subtilis*. *Genes Dev,* 24**,** 411-22.

MÖSE, J. R. & MÖSE, G. 1964. Oncolysis by Clostridia. I. Activity of *Clostridium butyricum* (M-55) and other nonpathogenic Clostridia against the Ehrlich Carcinoma. *Cancer Res,* 24**,** 212-6.

MSADEK, T., DARTOIS, V., KUNST, F., HERBAUD, M. L., DENIZOT, F. & RAPOPORT, G. 1998. ClpP of *Bacillus subtilis* is required for competence development, motility, degradative enzyme synthesis, growth at high temperature and sporulation. *Mol Microbiol,* 27**,** 899-914.

MUCHOVA, K., KUTEJOVA, E., SCOTT, D. J., BRANNIGAN, J. A., LEWIS, R. J., WILKINSON, A. J. & BARAK, I. 2002. Oligomerization of the *Bacillus subtilis* division protein DivIVA. *Microbiology,* 148**,** 807-13.

MUNOZ-LOPEZ, M. & GARCIA-PEREZ, J. L. 2010. DNA transposons: nature and applications in genomics. *Curr Genomics,* 11**,** 115-28.

NABEL, C. S., JIA, H., YE, Y., SHEN, L., GOLDSCHMIDT, H. L., STIVERS, J. T., ZHANG, Y. & KOHLI, R. M. 2012. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol,* 8**,** 751-8.

NAGARAJU, S., DAVIES, N. K., WALKER, D. J., KOPKE, M. & SIMPSON, S. D. 2016. Genome editing of *Clostridium autoethanogenum* using CRISPR/Cas9. *Biotechnol Biofuels,* 9**,** 219.

NAIR, S., DERRE, I., MSADEK, T., GAILLOT, O. & BERCHE, P. 2000. CtsR controls class III heat shock gene expression in the human pathogen *Listeria monocytogenes*. *Mol Microbiol,* 35**,** 800-11.

NAKAMURA, S., OKADO, I., NAKASHIO, S. & NISHIDA, S. 1977. *Clostridium sporogenes* isolates and their relationship to *Clostridium botulinum* based on deoxyribonucleic acid reassociation. *J Gen Microbiol,* 100, 395-401.

NAKAMURA, S., TANABE, N., YAMAKAWA, K. & NISHIDA, S. 1983. Cytotoxin production by *Clostridium sordellii* strains. *Microbiol Immunol,* 27, 495-502.

NAKAYAMA, S., KIYOSHI, K., KADOKURA, T. & NAKAZATO, A. 2011. Butanol production from crystalline cellulose by cocultured *Clostridium thermocellum* and *Clostridium saccharoperbutylacetonicum* N1-4. *Appl Environ Microbiol,* 77, 6470-5.

NANAMIYA, H., TAKAHASHI, K., FUJITA, M. & KAWAMURA, F. 2000. Deficiency of the initiation events of sporulation in *Bacillus subtilis* clpP mutant can be suppressed by a lack of the Spo0E protein phosphatase. *Biochem Biophys Res Commun,* 279, 229-33.

NEHER, S. B., VILLEN, J., OAKES, E. C., BAKALARSKI, C. E., SAUER, R. T., GYGI, S. P. & BAKER, T. A. 2006. Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon. *Mol Cell,* 22, 193-204.

NEUWALD, A. F., ARAVIND, L., SPOUGE, J. L. & KOONIN, E. V. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res,* 9, 27-43.

NG, Y. K., EHSAAN, M., PHILIP, S., COLLERY, M. M., JANOIR, C., COLLIGNON, A., CARTMAN, S. T. & MINTON, N. P. 2013. Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using pyrE alleles. *PLoS One,* 8, e56051.

NISHIDA, K., ARAZOE, T., YACHIE, N., BANNO, S., KAKIMOTO, M., TABATA, M., MOCHIZUKI, M., MIYABE, A., ARAKI, M., HARA, K. Y., SHIMATANI, Z. & KONDO, A. 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science,* 353.

NOCADELLO, S., MINASOV, G., SHUVALOVA, L. S., DUBROVSKA, I., SABINI, E. & ANDERSON, W. F. 2016. Crystal structures of the SpoIID lytic transglycosylases essential for bacterial sporulation. *J Biol Chem,* 291, 14915-26.

NOLAND, M. E., LALONDE, D. H., YEE, G. J. & ROHRICH, R. J. 2016. Current uses of botulinum neurotoxins in plastic surgery. *Plast Reconstr Surg,* 138, 519e-30e.

O'HALLORAN, D. P., WYNNE, K. & GEOGHEGAN, J. A. 2015. Protein A is released into the Staphylococcus aureus culture supernatant with an unprocessed sorting signal. *Infect Immun,* 83, 1598-609.

ODSBU, I. & SKARSTAD, K. 2014. DNA compaction in the early part of the SOS response is dependent on RecN and RecA. *Microbiology (Reading),* 160, 872-882.

OLIVARES, A. O., BAKER, T. A. & SAUER, R. T. 2016. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat Rev Microbiol,* 14, 33-44.

PAIDHUNGAT, M., RAGKOUSI, K. & SETLOW, P. 2001. Genetic requirements for induction of germination of spores of *Bacillus subtilis* by Ca(2+)-dipicolinate. *J Bacteriol,* 183, 4886-93.

PAIDHUNGAT, M., SETLOW, B., DRIKS, A. & SETLOW, P. 2000. Characterization of spores of *Bacillus subtilis* which lack dipicolinic acid. *J Bacteriol,* 182, 5505-12.

PAIDHUNGAT, M. & SETLOW, P. 2000. Role of ger proteins in nutrient and nonnutrient triggering of spore germination in *Bacillus subtilis*. *J Bacteriol,* 182, 2513-9.

PAIDHUNGAT, M. & SETLOW, P. 2002. Spore germination and outgrowth. *Bacillus subtilis and its relatives: from genes to cells. American Society for Microbiology, Washington, DC*, 537-548.

PAN, Q., GARSIN, D. A. & LOSICK, R. 2001. Self-reinforcing activation of a cell-specific transcription factor by proteolysis of an anti-sigma factor in *Bacillus subtilis*. *Mol Cell,* 8, 873-83.

PAPOUTSAKIS, E. T. 2008. Engineering solventogenic Clostridia. *Curr Opin Biotechnol,* 19, 420-9.

PATRICK, J. E. & KEARNS, D. B. 2008. MinJ (YvjD) is a topological determinant of cell division in *Bacillus subtilis*. *Mol Microbiol,* 70, 1166-79.

PEREIRA, F. C., SAUJET, L., TOME, A. R., SERRANO, M., MONOT, M., COUTURE-TOSI, E., MARTIN-VERSTRAETE, I., DUPUY, B. & HENRIQUES, A. O. 2013. The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. *PLoS Genet,* 9, e1003782.

PERRY, S. E. & EDWARDS, D. H. 2004. Identification of a polar targeting determinant for *Bacillus subtilis* DivIVA. *Mol Microbiol,* 54, 1237-49.

PERRY, S. E. & EDWARDS, D. H. 2006. The *Bacillus subtilis* DivIVA protein has a sporulation-specific proximity to Spo0J. *J Bacteriol,* 188, 6039-43.

PERUTKA, J., WANG, W., GOERLITZ, D. & LAMBOWITZ, A. M. 2004. Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *J Mol Biol,* 336, 421-39.

PETTIT, L. J., BROWNE, H. P., YU, L., SMITS, W. K., FAGAN, R. P., BARQUIST, L., MARTIN, M. J., GOULDING, D., DUNCAN, S. H., FLINT, H. J., DOUGAN, G., CHOUDHARY, J. S. & LAWLEY, T. D. 2014. Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism. *BMC Genomics,* 15, 160.

PICARDEAU, M. 2010. Transposition of fly mariner elements into bacteria as a genetic tool for mutagenesis. *Genetica,* 138, 551-8.

PIGGOT, P. J. & COOTE, J. G. 1976. Genetic aspects of bacterial endospore formation. *Bacteriol Rev,* 40, 908-62.

PIGGOT, P. J. & LOSICK, R. 2001. Sporulation genes and intercompartmental regulation. *Bacillus subtilis and its Closest Relatives: from Genes to Cells*, 483-517.

PLASTERK, R. H., IZSVAK, Z. & IVICS, Z. 1999. Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet,* 15, 326-32.

PLOMP, M., MCCAFFERY, J. M., CHEONG, I., HUANG, X., BETTEGOWDA, C., KINZLER, K. W., ZHOU, S., VOGELSTEIN, B. & MALKIN, A. J. 2007. Spore coat architecture of *Clostridium novyi* NT spores. *J Bacteriol,* 189, 6457-68.

POEHLEIN, A., RIEGEL, K., KONIG, S. M., LEIMBACH, A., DANIEL, R. & DURRE, P. 2015. Genome sequence of *Clostridium sporogenes* DSM 795(T), an amino acid-degrading, nontoxic surrogate of neurotoxin-producing *Clostridium botulinum*. *Stand Genomic Sci,* 10, 40.

POGLIANO, J., OSBORNE, N., SHARP, M. D., ABANES-DE MELLO, A., PEREZ, A., SUN, Y. L. & POGLIANO, K. 1999. A vital stain for studying membrane dynamics in bacteria: a novel mechanism controlling septation during *Bacillus subtilis* sporulation. *Mol Microbiol,* 31, 1149-59.

POMPEO, F., FOULQUIER, E. & GALINIER, A. 2016. Impact of serine/threonine protein kinases on the regulation of sporulation in *Bacillus subtilis*. *Front Microbiol,* 7, 568.

PRICE, K. D. & LOSICK, R. 1999. A four-dimensional view of assembly of a morphogenetic protein during sporulation in *Bacillus subtilis*. *J Bacteriol,* 181, 781-90.

PRICE, M. N., WETMORE, K. M., WATERS, R. J., CALLAGHAN, M., RAY, J., LIU, H., KUEHL, J. V., MELNYK, R. A., LAMSON, J. S., SUH, Y., CARLSON, H. K., ESQUIVEL, Z., SADEESHKUMAR, H., CHAKRABORTY, R., ZANE, G. M., RUBIN, B. E., WALL, J. D., VISEL, A., BRISTOW, J., BLOW, M. J., ARKIN, A. P. & DEUTSCHBAUER, A. M. 2018. Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature,* 557, 503-509.

PRINCEWILL, T. 1978. Differences in colony morphology and carbohydrate fermentation of *Clostridium sporogenes*. *Microbiology,* 108, 315-319.

QURESHI, N. & BLASCHEK, H. P. 2000. Butanol production using *Clostridium beijerinckii* BA101 hyper-butanol producing mutant strain and recovery by pervaporation. *Appl Biochem Biotechnol,* 84-86, 225-35.

RAJAGOPALAN, G., HE, J. & YANG, K.-L. 2013. A highly efficient NADH-dependent butanol dehydrogenase from high-butanol-producing *Clostridium* sp. BOH3. *BioEnergy Research,* 6, 240-251.

RAN, F. A., HSU, P. D., WRIGHT, J., AGARWALA, V., SCOTT, D. A. & ZHANG, F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc,* 8, 2281-2308.

REEVE, J. N., MENDELSON, N. H., COYNE, S. I., HALLOCK, L. L. & COLE, R. M. 1973. Minicells of *Bacillus subtilis*. *J Bacteriol,* 114, 860-73.

REZNIKOFF, W. S. & WINTERBERG, K. M. 2008. Transposon-based strategies for the identification of essential bacterial genes. *Methods Mol Biol,* 416, 13-26.

RIBIS, J. W., FIMLAID, K. A. & SHEN, A. 2018. Differential requirements for conserved peptidoglycan remodeling enzymes during *Clostridioides difficile* spore formation. *Mol Microbiol,* 110, 370-389.

RICH, T., ALLEN, R. L. & WYLLIE, A. H. 2000. Defying death after DNA damage. *Nature,* 407, 777-83.

ROBERTS, S. A., SHORE, K. P., PAVIOUR, S. D., HOLLAND, D. & MORRIS, A. J. 2006. Antimicrobial susceptibility of anaerobic bacteria in New Zealand: 1999-2003. *J Antimicrob Chemother,* 57, 992-8.

ROBERTSON, G. T., NG, W. L., GILMOUR, R. & WINKLER, M. E. 2003. Essentiality of clpX, but not clpP, clpL, clpC, or clpE, in *Streptococcus pneumoniae* R6. *J Bacteriol,* 185, 2961-6.

RODRIGUES, C. D., MARQUIS, K. A., MEISNER, J. & RUDNER, D. Z. 2013. Peptidoglycan hydrolysis is required for assembly and activity of the transenvelope secretion complex during sporulation in *Bacillus subtilis*. *Mol Microbiol,* 89, 1039-52.

ROELS, S., DRIKS, A. & LOSICK, R. 1992. Characterization of spoIVA, a sporulation gene involved in coat morphogenesis in *Bacillus subtilis*. *J Bacteriol,* 174, 575-85.

RONG, S., ROSENKRANTZ, M. S. & SONENSHEIN, A. L. 1986. Transcriptional control of the *Bacillus subtilis* spoIID gene. *J Bacteriol,* 165, 771-9.

ROOD, J. I. & COLE, S. T. 1991. Molecular genetics and pathogenesis of *Clostridium perfringens*. *Microbiol Rev,* 55, 621-48.

RUBIN, E. J., AKERLEY, B. J., NOVIK, V. N., LAMPE, D. J., HUSSON, R. N. & MEKALANOS, J. J. 1999. In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proc Natl Acad Sci U S A,* 96, 1645-50.

RUPNIK, M., WILCOX, M. H. & GERDING, D. N. 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol,* 7, 526-36.

RYTER, A. 1965. Morphologic study of the sporulation of *Bacillus subtilis*. *Ann Inst Pasteur (Paris),* 108, 40-60.

SAMUL, D., WORSZTYNOWICZ, P., LEJA, K. & GRAJEK, W. 2013. Beneficial and harmful roles of bacteria from the *Clostridium* genus. *Acta Biochim Pol,* 60, 515-21.

SANDER, J. D. & JOUNG, J. K. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol,* 32, 347-55.

SARVARI, K. P. & SCHOBLOCHER, D. 2020. The antibiotic susceptibility pattern of gas gangrene-forming *Clostridium* spp. clinical isolates from South-Eastern Hungary. *Infect Dis (Lond),* 52, 196-201.

SASSETTI, C. M., BOYD, D. H. & RUBIN, E. J. 2001. Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A,* 98, 12712-7.

SATLEWAL, A., AGRAWAL, R., BHAGIA, S., DAS, P. & RAGAUSKAS, A. J. 2018. Rice straw as a feedstock for biofuels: availability, recalcitrance, and chemical properties. *Biofuels, Bioproducts and Biorefining,* 12, 83-107.

SAUER, R. T. & BAKER, T. A. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem,* 80, 587-612.

SAUER, R. T., BOLON, D. N., BURTON, B. M., BURTON, R. E., FLYNN, J. M., GRANT, R. A., HERSCH, G. L., JOSHI, S. A., KENNISTON, J. A., LEVCHENKO, I., NEHER, S. B., OAKES, E. S., SIDDIQUI, S. M., WAH, D. A. & BAKER, T. A. 2004. Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell,* 119, 9-18.

SAUJET, L., PEREIRA, F. C., SERRANO, M., SOUTOURINA, O., MONOT, M., SHELYAKIN, P. V., GELFAND, M. S., DUPUY, B., HENRIQUES, A. O. & MARTIN-VERSTRAETE, I. 2013. Genome-wide analysis of cell type-specific gene transcription during spore formation in *Clostridium difficile*. *PLoS Genet,* 9, e1003756.

SCHANTZ, E. J. & JOHNSON, E. A. 1992. Properties and use of botulinum toxin and other microbial neurotoxins in medicine. *Microbiol Rev,* 56, 80-99.

SCHEURWATER, E., REID, C. W. & CLARKE, A. J. 2008. Lytic transglycosylases: Bacterial space-making autolysins. *Int J Biochem Cell Biol,* 40, 586-91.

SCHILL, K. M., WANG, Y., BUTLER, R. R., 3RD, POMBERT, J. F., REDDY, N. R., SKINNER, G. E. & LARKIN, J. W. 2016. Genetic diversity of *Clostridium sporogenes* PA 3679 isolates obtained from different sources as resolved by Pulsed-Field Gel Electrophoresis and High-Throughput Sequencing. *Appl Environ Microbiol,* 82, 384-93.

SCHWARTZMAN, J. D., RELLER, L. B. & WANG, W. L. 1977. Susceptibility of *Clostridium perfringens* isolated from human infections to twenty antibiotics. *Antimicrob Agents Chemother,* 11, 695-7.

SEBAIHIA, M., PECK, M. W., MINTON, N. P., THOMSON, N. R., HOLDEN, M. T., MITCHELL, W. J., CARTER, A. T., BENTLEY, S. D., MASON, D. R., CROSSMAN, L., PAUL, C. J., IVENS, A., WELLS-BENNIK, M. H., DAVIS, I. J., CERDENO-TARRAGA, A. M., CHURCHER, C., QUAIL, M. A., CHILLINGWORTH, T., FELTWELL, T., FRASER, A., GOODHEAD, I., HANCE, Z., JAGELS, K., LARKE, N.,

MADDISON, M., MOULE, S., MUNGALL, K., NORBERTCZAK, H., RABBINOWITSCH, E., SANDERS, M., SIMMONDS, M., WHITE, B., WHITHEAD, S. & PARKHILL, J. 2007. Genome sequence of a proteolytic (Group I) *Clostridium botulinum* strain Hall A and comparative analysis of the clostridial genomes. *Genome Res,* 17, 1082-92.

SERRANO, M., KINT, N., PEREIRA, F. C., SAUJET, L., BOUDRY, P., DUPUY, B., HENRIQUES, A. O. & MARTIN-VERSTRAETE, I. 2016. A recombination directionality factor controls the cell type-specific activation of sigmaK and the fidelity of spore development in *Clostridium difficile*. *PLoS Genet,* 12, e1006312.

SETLOW, P. 1994. Mechanisms which contribute to the long-term survival of spores of *Bacillus* species. *Soc Appl Bacteriol Symp Ser,* 23, 49S-60S.

SETLOW, P. 2003. Spore germination. *Curr Opin Microbiol,* 6, 550-6.

SETLOW, P. 2014. Spore resistance properties. *Microbiol Spectr,* 2.

SHAH, I. M., LAABERKI, M. H., POPHAM, D. L. & DWORKIN, J. 2008. A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell,* 135, 486-96.

SHAPIRO, R. L., HATHEWAY, C. & SWERDLOW, D. L. 1998. Botulism in the United States: A clinical and epidemiologic review. *Ann Intern Med,* 129, 221-8.

SHARMA, D. & KHAN, A. U. 2018. Role of cell division protein divIVA in *Enterococcus faecalis* pathogenesis, biofilm and drug resistance: A future perspective by in silico approaches. *Microb Pathog,* 125, 361-365.

SILHAVY, T. J., KAHNE, D. & WALKER, S. 2010. The bacterial cell envelope. *Cold Spring Harb Perspect Biol,* 2, a000414.

SINAVE, C., LE TEMPLIER, G., BLOUIN, D., LEVEILLE, F. & DELAND, E. 2002. Toxic shock syndrome due to *Clostridium sordellii*: a dramatic postpartum and postabortion disease. *Clin Infect Dis,* 35, 1441-3.

SINGH, S. K., ROZYCKI, J., ORTEGA, J., ISHIKAWA, T., LO, J., STEVEN, A. C. & MAURIZI, M. R. 2001. Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis. *J Biol Chem,* 276, 29420-9.

SINNELA, M. T., PARK, Y. K., LEE, J. H., JEONG, K. C., KIM, Y. W., HWANG, H. J. & MAH, J. H. 2019. Effects of Calcium and Manganese on sporulation of *Bacillus* species involved in food poisoning and spoilage. *Foods,* 8.

SIVERTSSON, E. M., JACKSON, S. E. & ITZHAKI, L. S. 2019. The AAA+ protease ClpXP can easily degrade a 31 and a 52-knotted protein. *Sci Rep,* 9, 2421.

SMALL, R. 2014. Botulinum toxin injection for facial wrinkles. *Am Fam Physician,* 90, 168-75.

SMITH, L. D. 1975. Common mesophilic anaerobes, including *Clostridium botulinum* and *Clostridium tetani*, in 21 soil specimens. *Appl Microbiol,* 29, 590-4.

SMITH, L. D. S. & HOBBS III, G. 1974. Genus III. *Clostridium* Prazmowski 1880, 23. *Bergey's Manual of Determinative Bacteriology, Eighth Edition, The Williams and Wilkins Co., Baltimore*, 551-572.

SMITH, T., WILLIAMSON, C. H. D., HILL, K., SAHL, J. & KEIM, P. 2018. Botulinum neurotoxin-producing bacteria. Isn't it time that we called a species a species? *mBio,* 9.

SONENSHEIN, A. L., LOSICK, R. & HOCH, J. A. 2002. *Bacillus subtilis and its closest relatives*, American Society of Microbiology.

SOREK, R., LAWRENCE, C. M. & WIEDENHEFT, B. 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem,* 82**,** 237-66.

SREEKUMAR, S., BAER, Z. C., PAZHAMALAI, A., GUNBAS, G., GRIPPO, A., BLANCH, H. W., CLARK, D. S. & TOSTE, F. D. 2015. Production of an acetone-butanol-ethanol mixture from *Clostridium acetobutylicum* and its conversion to high-value biofuels. *Nat Protoc,* 10**,** 528-37.

STAEDTKE, V., ROBERTS, N. J., BAI, R. Y. & ZHOU, S. 2016. *Clostridium novyi*-NT in cancer therapy. *Genes Dis,* 3**,** 144-152.

STEINER, E., DAGO, A. E., YOUNG, D. I., HEAP, J. T., MINTON, N. P., HOCH, J. A. & YOUNG, M. 2011. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Mol Microbiol,* 80**,** 641-54.

STEPHENSON, K. & LEWIS, R. J. 2005. Molecular insights into the initiation of sporulation in Gram-positive bacteria: new technologies for an old phenomenon. *FEMS Microbiol Rev,* 29**,** 281-301.

STEVENS, D. L., MAIER, K. A., LAINE, B. M. & MITTEN, J. E. 1987. Comparison of clindamycin, rifampin, tetracycline, metronidazole, and penicillin for efficacy in prevention of experimental gas gangrene due to *Clostridium perfringens*. *J Infect Dis,* 155**,** 220-8.

STEWART, G. C. 2015. The exosporium layer of bacterial spores: A connection to the environment and the infected host. *Microbiol Mol Biol Rev,* 79**,** 437-57.

STRAGIER, P. 1986. Comment on 'Duplicated sporulation genes in bacteria' by J. Errington, P. Fort and J. Mandelstam. *FEBS Lett,* 195**,** 9-11.

STRAGIER, P., BOUVIER, J., BONAMY, C. & SZULMAJSTER, J. 1984. A developmental gene product of *Bacillus subtilis* homologous to the sigma factor of *Escherichia coli*. *Nature,* 312**,** 376-378.

STRAGIER, P., KUNKEL, B., KROOS, L. & LOSICK, R. 1989. Chromosomal rearrangement generating a composite gene for a developmental transcription factor. *Science,* 243**,** 507-512.

STREET, W. 2019. Cancer Facts & Figures 2019. *American Cancer Society: Atlanta, GA, USA*.

SUGIYAMA, H. 1951. Studies on factors affecting the heat resistance of spores of *Clostridium botulinum*. *J Bacteriol,* 62**,** 81-96.

SUZUKI, C. K., REP, M., VAN DIJL, J. M., SUDA, K., GRIVELL, L. A. & SCHATZ, G. 1997. ATP-dependent proteases that also chaperone protein biogenesis. *Trends Biochem Sci,* 22**,** 118-23.

SUZUKI, S., YOSHIKAWA, M., IMAMURA, D., ABE, K., EICHENBERGER, P. & SATO, T. 2020. Compatibility of site-specific recombination units between mobile genetic elements. *Iscience,* 23**,** 100805.

TAKUMI, K., KINOUCHI, T. & KAWATA, T. 1979. Isolation and partial characterization of exosporium from spores of a highly sporogenic mutant of *Clostridium botulinum* type A. *Microbiol Immunol,* 23**,** 443-54.

TAN, I. S., WEISS, C. A., POPHAM, D. L. & RAMAMURTHI, K. S. 2015. A quality-control mechanism removes unfit cells from a population of sporulating bacteria. *Dev Cell,* 34**,** 682-93.

TANAKA, S., YOSHIOKA, S., NISHIDA, K., HOSOKAWA, H., KAKIZUKA, A. & MAEGAWA, S. 2018. In vivo targeted single-nucleotide editing in zebrafish. *Sci Rep,* 8, 11423.

TAYLOR, G. R. 1974. Space Microbiology. *Annual Review of Microbiology,* 28, 121-137.

TAYLOR, R. H., DUNN, M. L., OGDEN, L. V., JEFFERIES, L. K., EGGETT, D. L. & STEELE, F. M. 2013. Conditions associated with *Clostridium sporogenes* growth as a surrogate for *Clostridium botulinum* in nonthermally processed canned butter. *J Dairy Sci,* 96, 2754-64.

THOMAIDES, H. B., FREEMAN, M., EL KAROUI, M. & ERRINGTON, J. 2001. Division site selection protein DivIVA of *Bacillus subtilis* has a second distinct function in chromosome segregation during sporulation. *Genes Dev,* 15, 1662-73.

TOCHEVA, E. I., MATSON, E. G., MORRIS, D. M., MOUSSAVI, F., LEADBETTER, J. R. & JENSEN, G. J. 2011. Peptidoglycan remodeling and conversion of an inner membrane into an outer membrane during sporulation. *Cell,* 146, 799-812.

TOMOYASU, T., OHKISHI, T., UKYO, Y., TOKUMITSU, A., TAKAYA, A., SUZUKI, M., SEKIYA, K., MATSUI, H., KUTSUKAKE, K. & YAMAMOTO, T. 2002. The ClpXP ATP-dependent protease regulates flagellum synthesis in *Salmonella enterica* serovar Typhimurium. *J Bacteriol,* 184, 645-53.

TOWNSEND, C., ESTY, J. & BASELT, F. 1938. Heat - resistance studies on spores of putrefactive anaerobes in relation to determination of safe processes for canned foods. *Journal of Food Science,* 3, 323-346.

TRUBITSYNA, M., MICHLEWSKI, G., FINNEGAN, D. J., ELFICK, A., ROSSER, S. J., RICHARDSON, J. M. & FRENCH, C. E. 2017. Use of mariner transposases for one-step delivery and integration of DNA in prokaryotes and eukaryotes by transfection. *Nucleic Acids Res,* 45, e89.

VALENCIKOVA, R., KRASCSENITSOVA, E., LABAJOVA, N., MAKROCZYOVA, J. & BARAK, I. 2018. Clostridial DivIVA and MinD interact in the absence of MinJ. *Anaerobe,* 50, 22-31.

VAN BAARLE, S., CELIK, I. N., KAVAL, K. G., BRAMKAMP, M., HAMOEN, L. W. & HALBEDEL, S. 2013. Protein-protein interaction domains of *Bacillus subtilis* DivIVA. *J Bacteriol,* 195, 1012-21.

VAN MELLAERT, L., BARBE, S. & ANNE, J. 2006. *Clostridium* spores as anti-tumour agents. *Trends Microbiol,* 14, 190-6.

VAN OOIJ, C., EICHENBERGER, P. & LOSICK, R. 2004. Dynamic patterns of subcellular protein localization during spore coat morphogenesis in *Bacillus subtilis*. *J Bacteriol,* 186, 4441-8.

VAUPEL, P. & HARRISON, L. 2004. Tumor hypoxia: causative factors, compensatory mechanisms, and cellular response. *The oncologist,* 9, 4-9.

VEENING, J. W., MURRAY, H. & ERRINGTON, J. 2009. A mechanism for cell cycle regulation of sporulation initiation in *Bacillus subtilis*. *Genes Dev,* 23, 1959-70.

VERHEYDEN, J. & BLITZER, A. 2002. Other noncosmetic uses of BOTOX. *Dis Mon,* 48, 357-66.

VERONESI, R. & FOCACCIA, R. 1981. Tetanus. Important New Concepts. *Annals of Internal Medicine,* 95, 790-791.

VISHNOI, M., NARULA, J., DEVI, S. N., DAO, H. A., IGOSHIN, O. A. & FUJITA, M. 2013. Triggering sporulation in *Bacillus subtilis* with artificial two-component systems reveals the importance of proper Spo0A activation dynamics. *Mol Microbiol,* 90, 181-94.

VOLLMER, W. 2011. Bacterial outer membrane evolution via sporulation? *Nat Chem Biol,* 8, 14-8.

VREELAND, R. H., ROSENZWEIG, W. D. & POWERS, D. W. 2000. Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. *Nature,* 407, 897-900.

WALKER, J. E., SARASTE, M., RUNSWICK, M. J. & GAY, N. J. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J,* 1, 945-51.

WANG, Y., LIU, X., REN, C., ZHONG, G. Y., YANG, L., LI, S. & LIANG, Z. 2016. Identification of genomic sites for CRISPR/Cas9-based genome editing in the Vitis vinifera genome. *BMC Plant Biol,* 16, 96.

WANG, Y., ZHANG, Z. T., SEO, S. O., CHOI, K., LU, T., JIN, Y. S. & BLASCHEK, H. P. 2015. Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system. *J Biotechnol,* 200, 1-5.

WARTH, A. D. 1978. Molecular structure of the bacterial spore. *Adv Microb Physiol,* 17, 1-45.

WAWRZYNOW, A., BANECKI, B. & ZYLICZ, M. 1996. The Clp ATPases define a novel class of molecular chaperones. *Mol Microbiol,* 21, 895-9.

WEART, R. B. & LEVIN, P. A. 2003. Growth rate-dependent regulation of medial FtsZ ring formation. *J Bacteriol,* 185, 2826-34.

WEART, R. B., NAKANO, S., LANE, B. E., ZUBER, P. & LEVIN, P. A. 2005. The ClpX chaperone modulates assembly of the tubulin-like protein FtsZ. *Mol Microbiol,* 57, 238-49.

WEIZMANN, C. & ROSENFELD, B. 1937. The activation of the butanol-acetone fermentation of carbohydrates by *Clostridium acetobutylicum* (Weizmann). *Biochem J,* 31, 619-39.

WELLS, C. L. & WILKINS, T. D. 1996. Clostridia: Sporeforming anaerobic bacilli. *In:* BARON, S. (ed.) *Medical Microbiology.* 4th ed. Galveston (TX).

WEMHOFF, S. & MEINHARDT, F. 2013. Generation of biologically contained, readily transformable, and genetically manageable mutants of the biotechnologically important *Bacillus pumilus*. *Appl Microbiol Biotechnol,* 97, 7805-19.

WHEELER, D. L., CHURCH, D. M., FEDERHEN, S., LASH, A. E., MADDEN, T. L., PONTIUS, J. U., SCHULER, G. D., SCHRIML, L. M., SEQUEIRA, E., TATUSOVA, T. A. & WAGNER, L. 2003. Database resources of the National Center for Biotechnology. *Nucleic Acids Res,* 31, 28-33.

WILLIAMS, D. R., YOUNG, D. I. & YOUNG, M. 1990. Conjugative plasmid transfer from *Escherichia coli* to *Clostridium acetobutylicum*. *J Gen Microbiol,* 136, 819-26.

WILLIAMS, ELSIE M., LITTLE, RORY F., MOWDAY, ALEXANDRA M., RICH, MICHELLE H., CHAN-HYAMS, JASMINE V. E., COPP, JANINE N., SMAILL, JEFF B., PATTERSON, ADAM V. & ACKERLEY, DAVID F. 2015. Nitroreductase gene-directed enzyme prodrug therapy: insights and advances toward clinical utility. *Biochemical Journal,* 471, 131-153.

WILSON, K. H. & BLITCHINGTON, R. B. 1996. Human colonic biota studied by ribosomal DNA sequence analysis. *Appl Environ Microbiol,* 62, 2273-8.

WITTCHEN, K. D., STREY, J. & BULTMANN, A. 1998. Molecular characterization of the operon comprising the spoIV gene of *Bacillus megaterium* DSM319 and generation of a deletion mutant. *J Gen Appl Microbiol,* 44, 317-326.

WOODS, C., HUMPHREYS, C. M., RODRIGUES, R. M., INGLE, P., ROWE, P., HENSTRA, A. M., KOPKE, M., SIMPSON, S. D., WINZER, K. & MINTON, N. P. 2019. A novel conjugal donor strain for improved DNA transfer into *Clostridium* spp. *Anaerobe,* 59, 184-191.

WU, L. J. & ERRINGTON, J. 1994. *Bacillus subtilis* SpoIIIE protein required for DNA segregation during asymmetric cell division. *Science,* 264, 572-5.

XU, T., LI, Y., SHI, Z., HEMME, C. L., LI, Y., ZHU, Y., VAN NOSTRAND, J. D., HE, Z. & ZHOU, J. 2015. Efficient genome editing in *Clostridium cellulolyticum* via CRISPR-Cas9 nickase. *Appl Environ Microbiol,* 81, 4423-31.

YANG, W. W., CROW-WILLARD, E. N. & PONCE, A. 2009. Production and characterization of pure *Clostridium* spore suspensions. *J Appl Microbiol,* 106, 27-33.

YUTIN, N. & GALPERIN, M. Y. 2013. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol,* 15, 2631-41.

ZHANG, J., ZONG, W., HONG, W., ZHANG, Z. T. & WANG, Y. 2018. Exploiting endogenous CRISPR-Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production. *Metab Eng,* 47, 49-59.

ZHANG, L., ZHANG, L., ZHAN, X., ZENG, X., ZHOU, L., CAO, G., CHEN, A. & YANG, C. 2016. Effects of dietary supplementation of probiotic, *Clostridium butyricum*, on growth performance, immune response, intestinal barrier function, and digestive enzyme activity in broiler chickens challenged with *Escherichia coli* K88. *J Anim Sci Biotechnol,* 7, 3.

ZHANG, Y., GROSSE-HONEBRINK, A. & MINTON, N. P. 2015. A universal mariner transposon system for forward genetic studies in the genus *Clostridium*. *PLoS One,* 10, e0122411.

ZHENG, L., ABHYANKAR, W., OUWERLING, N., DEKKER, H. L., VAN VEEN, H., VAN DER WEL, N. N., ROSEBOOM, W., DE KONING, L. J., BRUL, S. & DE KOSTER, C. G. 2016. *Bacillus subtilis* spore inner membrane proteome. *J Proteome Res,* 15, 585-94.

ZHENG, L. B., DONOVAN, W. P., FITZ-JAMES, P. C. & LOSICK, R. 1988. Gene encoding a morphogenic protein required in the assembly of the outer coat of the *Bacillus subtilis* endospore. *Genes Dev,* 2, 1047-54.

ZU, Y., TONG, X., WANG, Z., LIU, D., PAN, R., LI, Z., HU, Y., LUO, Z., HUANG, P., WU, Q., ZHU, Z., ZHANG, B. & LIN, S. 2013. TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods,* 10, 329-31.

APPENDIX

Table A1. List of oligonucleotide sequences used in this study

PRIMERS	SEQUENCES (5' to 3')	SOURCE
TRANSPOSON MUTAGENESIS		
catP-INV-F1	TAAATCATTTTTAGCAGATTATGAAAGTGATACGCAACGGTATGG	
catP-INV-R1	TATTGTATAGCTTGGTATCATCTCATCATATATCCCCAATTCACC	
catP-INV-R2	TATTTGTGTGATATCCACTTTAACGGTCATGCTGTAGGTACAAGG	
CLOSTRON MUTAGENESIS		
DivIVA_F	ATGAAGATAACATCAATGGATATTACGAA	This study
DivIVA_R	TTAATCCTCTTTTATATAAAAACTCTTTATTTCATTATTATTTTTG	
RAM_F	ACGCGTTATATTGATAAAAATAATAATAGTGGG	
RAM_R	ACGCGTGCGACTCATAGAATTATTTCCTCCCG	
EBS Universal	CGAAATTAGAAACTTGCGTTCAGTAAAC	
CRISPR/Cas9 MUTAGENESIS		
sgRNA-AatII_R	GGCCCCGGGACGTCATAAAAATAAGAAGCCTGCAAATGCAGGCTTCTTATTTTTATAAAAAA AGCACCGACTCGGTGCCACTTTTTCAAGTTG	
sgRNA-PD_Rv	CGCGCGCGGCGATCGCATAAAAATAAGAAGCCTGCAAATGCAGGCTTCTTATTTTTATAAAA AAAGCACCGACTCGGTGCCACTTTTTCAAGTTG	
sgRNA_fdx_fl_fwd	ATATATTCTAGATTTATATTTAGTCCCTTGCC	
83xxxL_R	ACGGCTTGATGTGTTGGTAG	
pCB102-R1	CTGTTATGCCTTTTGACTATC	

sgRNA1_bont/A_F	TTTTCGTCGACGTTAACAGAGATTTACACAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	
catPsgRNA3_F	TTTTCGTCGACTGATGAACTTGAATTGCCAAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	This study
SIBS_F	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG	
yqfD		
yqfDsgA1_F	TTTTCGTCGACTTTTTAGCTACTAAATCACAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	This study
yqfdLHaat_Frv	GTAGTAGACGTCCCTGTGCATTATCCTTTTGAGAATC	
yqfdLHb_Rrv	TTTGATTCCTCTGTGTAAATCTCTGTTAACGACAGCTAAAATAAAGGTGGTATTATAAT	
yqfdRHb_Frv	GCTGTCGTTAACAGAGATTTACACAGAGGAATCAAATTTCATTTGTTTCAGACCC	
yqfdLH_Rrv	TTTGATATGATGAACTTGAATTGCCAAAGGGACAGCTAAAATAAAGGTGGTATTATAAT	
yqfdRH_Frv	GCTGTCCCTTTGGCAATTCAAGTTCATCATATCAAATTTCATTTGTTTCAGACCC	
yqfdRH_Rrv	GTAGTAGGCGCGCCGAAAAAGCAGTTAATGAAACTGGTG	
yqfdoLH_F	GGTATAGATGCTAAAGGAATTAAAAAAGGTG	
yqfdoRH_R	CACTAAGAAGATTTAAGAAGAAATGTGCA	
yqfdiLH_F	CTAATACTCTCTTTAGTTTATTGTCCTTT	
yqfdiRH_R	GATCTGGATTGATATCTATTTATGG	
yqfdLHc_R3	GCTGTAAAATTAGAATGCTAATAATTGCGGAAGAAGACATAGCTG	
yqfdRHc_F3	CAGCTATGTCTTCTTCCGCAATTATTAGCATTCTAATTTTACAGC	
yqfDsgA_F	TTTTCGTCGACAGACAACAATTAAAACAATAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	
yqfD_iR	GATGACATAGCTAGGAGAACAGG	

hyp-p		
hyppsgRNA1_F	TTTTCGTCGACTTTATAGGGGAAACAGCTAAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	This study
hyppLH_F	ATTATTGCGATCGCAAGGGTCTTATGACCATAATACCT	
hyppLH_R	AATAAATCCTTTGGCAATTCAAGTTCATCAGTCCATTTAGAACATTCCTAAAATAA	
hyppRH_F	ATGGACTGATGAACTTGAATTGCCAAAGGATTTATTTGATTTGAGGTGTG	
hyppRH_R	ATTATTGGCGCGCCTTCTTTCATAACTTCACCTACTTTCAT	
hyppoLH_F	GGAAATATGCTAGATGCTATAGAAAAATGTG	
hyppoRH_R	GTACACCCTATGCTTAAGTTCTAAACTTTC	
hyppiLH_F	GATAGATTTTTCACCTATATTTGCTTT	
hyppiRH_R	CCAAAAACTCATCTACTTCATCAC	
divIVA		
divsgRNA1_F	TTTTCGTCGACCAATTCAAAAAATCCATGAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	This study
divLH_Frv	GTAGTAGGCGCGCCGGATTTAATGTTTATGATACTAAAGATTTTCCAAT	
divLH_R	AAACTCTTCTGTTAATCCTCCTTCATAGAACACACCTCAAATCAA	
divRH_F	TTGAGGTGTGTTCTATGAAGGAGGATTAACAGAAGAGTTTTTGCA	
divRH_R2	GTAGTAGGCGCGCCTGTAGTTGGTATATTTATAAATCGGATACCTCT	
divRHaat_Rrv	GTAGTAGACGTCATTTACTGTATTAATATTTTTGTTTTCTTCTCCTG	
divoLH_F	AATTTGGAGGCAAATATGGTATCTAAC	
divoRH_R	GCCTAAAGCTACTGCTTCTCC	
diviLH_F	GAGCAGATTCTATAGTGGCATCT	
diviRH_F	CCGATGATTGACCGCTATAATAAT	
divLHc_R3	TTTAATTCTTCCTTTGGCAGTCTCTCCAATAAAAGAATCCATTTTAAATTTACC	
divRHc_F3	GGTAAATTTAAAATGGATTCTTTTATTGGAGAGACTGCCAAAGGAAGAATTAAA	

clpX		
clpXsgRNA1_F	TTTTCGTCGACGCAGCTGTTCCACCACAGGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	This study
clpxLHaat_F	GTAGTAGACGTCAGGTTGTTGAAGGACAGAATAATAATG	
clpxLHb_R	TGCTGTTCCTCTGTGTAAATCTCTGTTAACAGACATTCCTACACCTCACTAAAG	
clpxRHb_F	ATGTCTGTTAACAGAGATTTACACAGAGGAACAGCATAAAAATAACACCAAGATTT	
clpxRH_R	GTAGTAGGCGCGCCCACTGTAGTAGCATCTATCTCCATAA	
clpxoLH_F	CGGCAAATGCTGAAATAGAAATAC	
clpxoRH_R	GTATCCTTGGAGAATATTGACCATTTT	
clpxiLH_F	GTGCAGAAGAAGCTAAAGAATATGG	
clpxiRH_R	GTCATATCCAAAACCCCTTTAATACTT	
clpXsgC_F	TTTTCGTCGACAGACGACTGATAGCAGGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	
clpX_iR	GGTCCTAAAAGAAGAATATTACTCTTTTGAAG	
spoVK		
ATPsgRNA1_F	TTTTCGTCGACGAGAGAGTAGAAACCTTCAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	This study
AtpLH_F	GTAGTAGACGTCATTACGGTCCTGGGAGATG	
AtpLH_R1	TATGGATCCTCTGTGTAAATCTCTGTTAACGTGTTTTGATGGGTTTATCAAAAAAATC	
AtpRH_F1	AAACACGTTAACAGAGATTTACACAGAGGATCCATAATTGAAGATTTTTTAAGTTTTAAAACAAATTAAT	
AtpRH_R	GTAGTAGGCGCGCCCAACTACACTAGAACAGGGGAAAG	
AtpoLH_F	TGAGTTATGAGCTAATAATGGGGG	
AtpoRH_R	TCAACAAAAGAGATTATGGAGTATGCT	
AtpiLH_F	AGTTACATCCTAAATATAATACGATTATAACCATAG	
AtpiRH_R	GCAGTTCTTTATTTTTGTGGAATTTTCC	

AtpLHs_R1	GCTATAATTAAACTAATTTGTTTTAAAACTTAAAAAATCTTCAATTATGGATTCTC	
AtpRHs_F1	GAGAATCCATAATTGAAGATTTTTTAAGTTTTAAAACAAATTAGTTTAATTATAGC	
spoIID		
spoIIDsgRNA1_F	TTTTCGTCGACAAAAAGTGTAAAAAGTCCAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	This study
spoIIdLH_F	GTAGTAGACGTCGAAAATAGGTTTATGCATGTATCTGAAATG	
spoIIdLH_R	ATTTTTTCCTCTGTGTAAATCTCTGTTAACTCTCATGTTTTCCCTCTTTTCATAATAAAG	
spoIIdRH_F	ATGAGAGTTAACAGAGATTTACACAGAGGAAAAAATTAGTGTTAATAAATTTATGGATTGGC	
spoIIdRH_R	GTAGTAGGCGCGCCACTACTTTATTTTCAATAGGTTCTTTCTTTAC	
spoIIdoLH_F	TTTCCAACAGATATGCAAGCAC	
spoIIdoRH_R	CTTCTTTTTTCACTTCAAGGGC	
spoIIdiLH_F	AGTTTTAGCTTATATTTTAGCATAATAAACCTC	
spoIIdiRH_R	CCACAAAGATAATAAAATACACTTCGC	
spoIIdLHs_R	CCATAAATTTATTAACATTAATTTTTATATTTAAGAATTTCTATATTTACTCC	
spoIIdRHs_F	GGAGTAAATATAGAAATTCTTAAATATAAAAATTAATGTTAATAAATTTATGG	
spoIIDsgA_F	TTTTCGTCGACGTCTCAAGTTTTTATGCCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	
spoIID_iF	GCCACAGGTTAGAGTTTATTTTGT	
COMPLEMENTING GENE ON PLASMID		
proYqfD_F	GTAGTAGTCGACAATACCTTCCTCAGCAGTTGA	This study
yqfD_R	GTAGTAGGATCCTTAGCTGTCTTCTTCTTTATCCCTAAT	
divhyppPro_F	GGTGGTGTCGACGCTTAGAAATGATGGTTATCTTCAAATATTAATGC	
hyppPro_R	GTCCATTTAGAACATTCCTAAGTCATGCCCCCTTACAG	
divPro_R	TATCTTCATAGAACACACCTAGTCATGCCCCCTTACAG	
hyppcomp_F	GGCTGTAAGGGGGCATGACTTAGGAATGTTCTAAATGGACAAAGAAAAATTT	
hyppcomp_R	GGTGGTGGATCCCACACCTCAAATCAAATAAATTTTTTAAATAATATTT	

divcomp_F	GGCTGTAAGGGGGCATGACTAGGTGTGTTCTATGAAGATAACATCAATG	
divcomp_R	GGTGGTGGATCCCTCTTCTGTTAATCCTCTTTTATATAAAAACTCT	

* Colour key: red = restriction sites; blue = 20 bp guide RNA; green = PAM; black = genome sequence and overhangs

Table A2. List of sporulation-/germination-deficient *C. sporogenes* transposon mutants identified in this study

TRANSPOSON MUTANTS		LOCUS TAG	GENE PRODUCT	GENE SIZE	TRANPOSON INSERTION SITE	STRAND AND ORIENTATION*	PHENOTYPES
No.	Name						
1.	#36	CLSPO_RS00710	Peptidase M23	1563 bp	715^716 bp	sense →	Translucent; spo/ger$^-$
2.	#12**	CLSPO_RS00745	Spore germination protein: GerAB/ ArcD/ProY family transporter	1104 bp	339^340 bp	sense →	Reduced sporulation
		CLSPO_RS17470	DUF4097-family protein	921 bp	571^572 bp	anti-sense ←	-
3.	#25	CLSPO_RS09490	Stage III sporulation protein AA	933 bp	36^37 bp	anti-sense ←	Translucent; spo/ger$^-$
4.	#90**	CLSPO_RS11675	Histidine kinase	273 bp	160^161 bp	anti-sense	-
		CLSPO_RS11725	Oxidoreductase	984 bp	954^955 bp	sense →	spo/ger$^-$
5.	#2	CLSPO_RS18530	Stage II sporulation E protein	2388 bp	1206^1207 bp	sense ←	Translucent; spo/ger$^-$
6.	#64	CLSPO_RS15205	Sporulation protein YqfD	1179 bp	485^486 bp	sense ←	Translucent; spo/ger$^-$

7.	#138_v2.	Intergenic (promoter) region CLSPO_RS18585^ CLSPO_RS18590	(potential) Sporulation protein YabP	118 bp	48^49 bp upstream CLSPO_RS18585	-	Translucent; spo/ger$^-$
8.	#1	CLSPO_RS17010	Triple tyrosine motif-containing protein	2004 bp	1501^1502	sense ←	Translucent; reduced sporulation
9.	#42	CLSPO_RS07555	Hypothetical protein	771 bp	419^420	sense →	Translucent; spo/ger$^-$
10.	#24	CLSPO_RS12440	Serine recombinase	1845 bp	131^132	anti-sense →	
11.	#33	CLSPO_RS15795	Thioether cross-link-forming SCIFF peptide maturase	1368 bp	328^329	anti-sense ←	
12.	#138_v5.	CLSPO_RS13160	sigma-E processing peptidase SpoIIGA	798 bp	151^152	sense ←	Translucent; spo/ger$^-$
13.	T1	CLSPO_RS15210	Sporulation protein YqfC	297 bp	186^187 bp	anti-sense ←	Translucent; spo/ger$^-$
14.	T2	CLSPO_RS09470	Stage III sporulation protein AE	1164 bp	245^246 bp	anti-sense ←	Translucent; spo/ger$^-$
15.	T3	CLSPO_RS00590	Methyltransferase: MgtC/SapB family protein		-		Translucent; spo/ger$^-$
16.	T4	CLSPO_RS00950	Stage II sporulation D protein	1077 bp	420^421 bp	anti-sense →	Translucent; spo/ger$^-$
17.	T5	CLSPO_RS09350	Stage II sporulation M protein	654 bp	562^563	anti-sense ←	Translucent; spo/ger$^-$
18.	T6	CLSPO_RS09470	Stage III sporulation protein AE	1164 bp	177^178 bp	anti-sense ←	Translucent; spo/ger$^-$

19.	T10	CLSPO_RS18530	Stage II sporulation E protein	2388 bp	371^372 bp	anti-sense ←	Translucent; spo/ger$^-$
20.	T15	CLSPO_RS13155	RNA polymerase sporulation sigma factor SigE	708 bp	523^524 bp	sense ←	Translucent; spo/ger$^-$
21.	T19	CLSPO_RS17345	ATPase	1092 bp	430^431 bp	sense →	Translucent; spo/ger$^-$
22.	T22	CLSPO_RS17010	Triple tyrosine motif-containing protein	2004 bp	195^196 bp	sense ←	Translucent; reduced sporulation
23.	T26	CLSPO_RS00950	Stage II sporulation D protein	1077 bp	1019^1020 bp	anti-sense →	Translucent; spo/ger$^-$
24.	T28	Intergenic (promoter) region CLSPO_RS13070^ CLSPO_RS13080	(potential) Stage IV sporulation protein A	726 bp	33^34 bp upstream CLSPO_RS13070	-	Translucent; spo/ger$^-$
25.	T29	CLSPO_RS13170	Cell division protein FtsA	1254 bp	490^491 bp	anti-sense ←	Translucent; reduced sporulation
26.	T30	CLSPO_RS09490	Stage III sporulation protein AA	933 bp	367^368 bp	sense ←	Translucent; spo/ger$^-$
27.	T31	CLSPO_RS09490	Stage III sporulation protein AA	933 bp	367^368 bp	sense ←	Translucent; spo/ger$^-$
28.	T35	CLSPO_RS07555	Hypothetical protein	771 bp	61^62 bp	sense →	Translucent; reduced sporulation
29.	T36	CLSPO_RS09350	Stage II sporulation M protein	654 bp	244^245 bp	anti-sense ←	Translucent; spo/ger$^-$
30.	T37	CLSPO_RS08825	Hypothetical protein	975 bp	427^428 bp	sense →	Translucent; reduced sporulation

31.	T38	CLSPO_RS09485	Stage III sporulation protein AB	519 bp	115^116 bp	sense ←	Translucent; spo/ger$^-$
32.	T40	CLSPO_RS18530	Stage II sporulation E protein	2388 bp	290^291 bp	sense ←	Translucent; spo/ger$^-$
33.	T41	CLSPO_RS12595	RIP zinc metalloprotease RseP	1011 bp	739^740 bp	anti-sense ←	Translucent; spo/ger$^-$
34.	T42	CLSPO_RS18530	Stage II sporulation E protein	2388 bp	2093^2094 bp	anti-sense ←	Translucent; spo/ger$^-$
35.	T43	CLSPO_RS16680	ATP-dependent Clp protease ATP-binding subunit ClpX	1290 bp	994^995 bp	anti-sense ←	Translucent; reduced sporulation
36.	T45	CLSPO_RS16680	ATP-dependent Clp protease ATP-binding subunit ClpX	1290 bp	812^813 bp	sense ←	Translucent; reduced sporulation
37.	T46	CLSPO_RS10840	Hypothetical protein	1647 bp	491^492 bp	sense ←	Translucent; spo/ger$^-$
38.	T48	CLSPO_RS00930	ATP synthase subunit beta: AtpD	1389 bp	144^145 bp	anti-sense →	Translucent; spo/ger$^-$
39.	T51	CLSPO_RS03315	MFS transporter	1434 bp	234^235 bp	anti-sense →	Translucent; reduced sporulation
40.	T52	CLSPO_RS12505	DNA translocase FtsK	2277 bp	1676^1677 bp	anti-sense ←	Translucent; spo/ger$^-$
41.	T53	CLSPO_RS17010	Triple tyrosine motif-containing protein	2004 bp	1438^1439 bp	anti-sense ←	Translucent; reduced sporulation

* sense and anti-sense represents the coding and non-coding strands of DNA, respectively; whereas, → and ← represents the forward or reverse orientation of the affected gene in the *C. sporogenes* genome, respectively

** mutants with two transposon insertion sites

Table A3. List of SNPs present in the *C. sporogenes* Δ*yqfD* mutant genome when aligned to the reference NCIMB 10696 strain

Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Average quality	Overlapping annotations	Amino acid change
215214	SNV	G	T	60	60	100	34.62	CLSPO_RS01055	WP_033058072.1:p.Asp25Tyr
1069736	SNV	C	A	42	42	100	36.07	CLSPO_RS04955	WP_003492454.1:p.Glu255*
1093130	MNV	TA	AT	58	61	95.08	36.83		
1289844	Insertion	-	A	62	63	98.41	35.81		
1457212	Insertion	-	T	49	51	96.08	36.41		
1769513	SNV	C	G	41	41	100	35.44	CLSPO_RS08160	WP_003491229.1:p.Val217Leu
1986188	Insertion	-	T	37	41	90.24	36.7	CLSPO_RS08990	WP_046340391.1:p.Asn659fs
2245185	SNV	G	C	53	53	100	35.98	CLSPO_RS10155	WP_046340395.1:p.Arg86Ser
2823829	Deletion	T	-	38	38	100	31.03		
2882363	SNV	C	T	46	46	100	34.52		
2928918	Insertion	-	A	61	63	96.83	35.66	CLSPO_RS13380	CLSPO_RS13380:p.Lys319fs
2933602	Insertion	-	AAA	35	41	85.37	35.19	CLSPO_RS13385	WP_046340411.1: p.Lys233_Phe234insLys
3508435	SNV	A	G	56	58	96.55	33.16	CLSPO_RS16045	
3509139	SNV	A	G	63	66	95.45	35.65	CLSPO_RS16045	WP_046340417.1:p.Tyr364His
3539562	SNV	T	C	52	64	81.25	24.23	CLSPO_RS16145	WP_052728859.1:p.Ser526Pro
4002339	Insertion	-	A	56	59	94.92	32.07		
4002343	Deletion	T	-	55	58	94.83	31.98		
4002348	SNV	G	A	63	63	100	33.49		
4002355	SNV	T	C	63	63	100	33.67		
4107100	Deletion	T	-	101	102	99.02	34.02		

* Colour key: red = new SNPs; blue = new shared SNPs; black = in-house wild type SNPs

Table A4. List of SNPs present in the *C. sporogenes* Δ*divIVA* mutant genome when aligned to the reference NCIMB 10696 strain

Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Average quality	Overlapping annotations	Amino acid change
1093130	MNV	TA	AT	40	41	97.56	36.15		
1289844	Insertion	-	A	62	63	98.41	35.69		
1457212	Insertion	-	T	48	49	97.96	36.48		
1473817	Insertion	-	AG	54	58	93.1	35.09	CLSPO_RS06805	WP_033059120.1:p.Gly363fs
1643035	SNV	A	C	40	40	100	36.35	CLSPO_RS07560	WP_003491073.1:p.Tyr215Ser
1643039	SNV	A	G	44	44	100	36.82	CLSPO_RS07560	WP_003491073.1:p.Ile216Met
1643042	SNV	A	G	44	45	97.78	36.61	CLSPO_RS07560	
1879584	Insertion	-	T	46	47	97.87	34.65		
1986188	Insertion	-	T	45	46	97.83	35.78	CLSPO_RS08990	WP_046340391.1:p.Asn659fs
2245185	SNV	G	C	47	48	97.92	35.32	CLSPO_RS10155	WP_046340395.1:p.Arg86Ser
2823829	Deletion	T	-	51	51	100	34.29		
2882363	SNV	C	T	57	57	100	34.95		
2928918	Insertion	-	A	52	55	94.55	36.92	CLSPO_RS13380	CLSPO_RS13380:p.Lys319fs
2933602	Insertion	-	AAA	51	58	87.93	36.44	CLSPO_RS13385	WP_046340411.1: p.Lys233_Phe234insLys
3338141	SNV	G	T	79	80	98.75	33.85	CLSPO_RS15190	WP_003496719.1:p.Ala140Asp
3427743	SNV	A	T	75	75	100	36.16	CLSPO_RS15635	WP_033061039.1:p.Asp25Glu
3508435	SNV	A	G	52	53	98.11	34.1	CLSPO_RS16045	
3509139	SNV	A	G	69	72	95.83	32.99	CLSPO_RS16045	WP_046340417.1:p.Tyr364His
3539562	SNV	T	C	69	79	87.34	27.59	CLSPO_RS16145	WP_052728859.1:p.Ser526Pro
4002339	Insertion	-	A	59	60	98.33	32		
4002343	Deletion	T	-	53	56	94.64	31.04		

4002348	SNV	G	A	64	65	98.46	33.03		
4002355	SNV	T	C	66	66	100	31.73		
4107100	Deletion	T	-	107	107	100	33.64		

* Colour key: red = new SNPs; blue = new shared SNPs; black = in-house wild type SNPs

Table A5. List of SNPs present in the *C. sporogenes* Δ*clpX* mutant genome when aligned to the reference NCIMB 10696 strain

Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Average quality	Overlapping annotations	Amino acid change
215214	SNV	G	T	75	75	100	35.36	CLSPO_RS01055	WP_033058072.1:p.Asp25Tyr
1069736	SNV	C	A	62	63	98.41	36.29	CLSPO_RS04955	WP_003492454.1:p.Glu255*
1093130	MNV	TA	AT	58	60	96.67	36.58		
1289844	Insertion	-	A	61	62	98.39	36.75		
1457212	Insertion	-	T	46	47	97.87	34.13		
1769513	SNV	C	G	45	46	97.83	35.82	CLSPO_RS08160	WP_003491229.1:p.Val217Leu
1986188	Insertion	-	T	39	39	100	36.41	CLSPO_RS08990	WP_046340391.1:p.Asn659fs
2245185	SNV	G	C	50	51	98.04	36.74	CLSPO_RS10155	WP_046340395.1:p.Arg86Ser
2823829	Deletion	T	-	43	44	97.73	32.81		
2882363	SNV	C	T	48	48	100	35.08		
2928918	Insertion	-	A	61	64	95.31	35.31	CLSPO_RS13380	CLSPO_RS13380:p.Lys319fs
2933602	Insertion	-	AAA	41	45	91.11	37.2	CLSPO_RS13385	WP_046340411.1: p.Lys233_Phe234insLys
3508435	SNV	A	G	59	64	92.19	33.83	CLSPO_RS16045	
3509139	SNV	A	G	60	60	100	33.05	CLSPO_RS16045	WP_046340417.1:p.Tyr364His
3539562	SNV	T	C	56	58	96.55	27.12	CLSPO_RS16145	WP_052728859.1:p.Ser526Pro

3997146	SNV	A	G	90	90	100	35.12	CLSPO_RS18415	
3997195	SNV	G	A	50	51	98.04	37.24	CLSPO_RS18415	
3997204	SNV	T	C	41	43	95.35	36.22	CLSPO_RS18415	
3997232	SNV	T	A	27	27	100	35.26		
3997234	SNV	T	A	27	27	100	36.26		
4002339	Insertion	-	A	67	68	98.53	34.46		
4002343	Deletion	T	-	68	69	98.55	32.29		
4002348	SNV	G	A	71	72	98.61	35.96		
4002355	SNV	T	C	70	71	98.59	33.57		
4107100	Deletion	T	-	106	106	100	34.37		

* Colour key: red = new SNPs; blue = new shared SNPs; black = in-house wild type SNPs

Table A6. List of SNPs present in the *C. sporogenes* Δ*spoVK* mutant genome when aligned to the reference NCIMB 10696 strain

Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Average quality	Overlapping annotations	Amino acid change
215214	SNV	G	T	101	101	100	33.43	CLSPO_RS01055	WP_033058072.1:p.Asp25Tyr
388122	SNV	A	T	87	88	98.86	36.46	CLSPO_RS01835	WP_033057873.1:p.Lys144*
1069736	SNV	C	A	61	61	100	36.69	CLSPO_RS04955	WP_003492454.1:p.Glu255*
1093130	MNV	TA	AT	49	50	98	35.77		
1289844	Insertion	-	A	76	77	98.7	36.54		
1457212	Insertion	-	T	54	56	96.43	34.72		
1769513	SNV	C	G	44	45	97.78	35.66	CLSPO_RS08160	WP_003491229.1:p.Val217Leu
1986188	Insertion	-	T	28	30	93.33	36.07	CLSPO_RS08990	WP_046340391.1:p.Asn659fs
2245185	SNV	G	C	45	45	100	34.16	CLSPO_RS10155	WP_046340395.1:p.Arg86Ser

2823829	Deletion	T	-	55	56	98.21	31.56		
2882363	SNV	C	T	64	64	100	36.44		
2928918	Insertion	-	A	68	71	95.77	35.4	CLSPO_RS13380	CLSPO_RS13380:p.Lys319fs
2933602	Insertion	-	AAA	48	61	78.69	36.92	CLSPO_RS13385	WP_046340411.1: p.Lys233_Phe234insLys
3508435	SNV	A	G	78	79	98.73	32.78	CLSPO_RS16045	
3509139	SNV	A	G	79	80	98.75	34.58	CLSPO_RS16045	WP_046340417.1:p.Tyr364His
3539562	SNV	T	C	60	67	89.55	27.45	CLSPO_RS16145	WP_052728859.1:p.Ser526Pro
3784857	SNV	A	G	51	51	100	34.51	CLSPO_RS17345	WP_033061400.1:p.Glu351Gly
3785048	SNV	A	G	52	82	63.41	35.96		
3997146	SNV	A	G	98	99	98.99	36.34	CLSPO_RS18415	
3997195	SNV	G	A	73	73	100	36.36	CLSPO_RS18415	
3997204	SNV	T	C	68	68	100	34.68	CLSPO_RS18415	
3997232	SNV	T	A	34	34	100	34.56		
3997234	SNV	T	A	33	34	97.06	36.48		
4002339	Insertion	-	A	70	70	100	32.76		
4002343	Deletion	T	-	67	68	98.53	30.81		
4002348	SNV	G	A	70	71	98.59	32.14		
4002355	SNV	T	C	71	72	98.61	32.99		
4107100	Deletion	T	-	134	134	100	33.66		

* Colour key: red = new SNPs; blue = new shared SNPs; black = in-house wild type SNPs

Table A7. List of SNPs present in the *C. sporogenes* Δ*spoIID* mutant genome when aligned to the reference NCIMB 10696 strain

Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Average quality	Overlapping annotations	Amino acid change
997844	SNV	A	G	78	78	100	34.18	CLSPO_RS04565	
1093130	MNV	TA	AT	42	43	97.67	36.19		
1228820	SNV	G	T	60	60	100	32.5	CLSPO_RS05670	
1289844	Insertion	-	A	68	69	98.55	37.13		
1457212	Insertion	-	T	46	47	97.87	34.5		
1473817	Insertion	-	AG	39	44	88.64	35.05	CLSPO_RS06805	WP_033059120.1:p.Gly363fs
1986188	Insertion	-	T	37	41	90.24	36.27	CLSPO_RS08990	WP_046340391.1:p.Asn659fs
2245185	SNV	G	C	52	52	100	32.44	CLSPO_RS10155	WP_046340395.1:p.Arg86Ser
2823829	Deletion	T	-	52	52	100	29.9		
2882363	SNV	C	T	51	52	98.08	33.67		
2928918	Insertion	-	A	61	61	100	34.39	CLSPO_RS13380	CLSPO_RS13380:p.Lys319fs
2933602	Insertion	-	AA	7	63	11.11	36.74	CLSPO_RS13385	WP_046340411.1:p.Phe234fs
2933602	Insertion	-	AAA	47	64	73.44	36.16	CLSPO_RS13385	WP_046340411.1: p.Lys233_Phe234insLys
3338141	SNV	G	T	70	71	98.59	34.44	CLSPO_RS15190	WP_003496719.1:p.Ala140Asp
3508435	SNV	A	G	61	61	100	32.36	CLSPO_RS16045	
3509139	SNV	A	G	81	84	96.43	34.02	CLSPO_RS16045	WP_046340417.1:p.Tyr364His
3539562	SNV	T	C	66	74	89.19	30.08	CLSPO_RS16145	WP_052728859.1:p.Ser526Pro
4002339	Insertion	-	A	79	80	98.75	33.37		
4002343	Deletion	T	-	85	86	98.84	30.25		

4002348	SNV	G	A	89	89	100	33.72		
4002355	SNV	T	C	89	89	100	32.12		
4107100	Deletion	T	-	107	107	100	31.95		

* Colour key: red = new SNPs; blue = new shared SNPs; black = in-house wild type SNPs

The End

"WHEN YOU REACH THE END OF WHAT YOU SHOULD KNOW, YOU WILL BE AT THE BEGINNING OF WHAT YOU SHOULD SENSE."

Kahlil Gibrán,
Sand and Foam

SYEDA RUBAB ZEHRA RIZVI ♥