University of Sheffield

Experimental Evolution of Vancomycin Resistance in *Clostridioides difficile*: Pathways and Mechanistic Insights

By Jessica Emily Buddle

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Science
School of Biosciences

July 2024

Declaration

I, Jessica Emily Buddle, confirm that this Thesis and the work presented herein is my own work, except where duly acknowledged in the text and/or publication declaration. I am aware of the University's Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means). This work has not been previously been presented for an award at this, or any other, university.

Jessica Buddle, July 2024

Abstract

Clostridioides difficile is the leading cause of hospital-associated diarrhoea, and presents an urgent threat to global health – owing, at least in part, to its ability to resist a wide array of antibiotics. Treatments for *C. difficile* infection are limited, and rely largely upon metronidazole, fidaxomicin, and vancomycin. Vancomycin is the current front-line drug of choice in the UK, and is one of the primary treatments for *C. difficile* worldwide. Perhaps surprisingly for such an adaptable organism, vancomycin resistance in *C. difficile* is not widespread, and knowledge of resistance pathways, beyond regulatory changes to the *van* operon, is lacking. However, isolated reports of resistance in clinic, coupled with anecdotal reports of treatment failure, suggest more effort should be placed on understanding the routes, mechanisms and costs of resistance. To bridge this gap, this work aimed to provide a thorough genetic and molecular characterisation of vancomycin resistance in *C. difficile*. Experimental evolution under increasing vancomycin dose demonstrated high-level vancomycin resistance could evolve rapidly, at the cost of pleiotropic phenotypic changes. The genetic basis of resistance was revealed through genome sequencing of resistant strains, highlighting multiple alternative routes to resistance acquisition, including mutations in *van* genes, *comR*, and the previously uncharacterised *dacS*. Through recapitulation of observed mutations in a clean genetic background, the mechanisms of *vanS* and *dacS* mediated resistance – alterations to the terminal D-Ala in nascent peptidoglycan – were determined, and synergistic interactions between these pathways leading to high-level vancomycin resistance were uncovered. Importantly, only two mutations were required for high-level resistance, a vital consideration for clinical monitoring. Overall, these mutational and mechanistic insights provide a solid foundation to guide future genomic surveillance of vancomycin resistance in *C. difficile*.

Acknowledgements

I am so incredibly grateful for how wonderful my years in Sheffield have been. My PhD certainly wouldn't have been possible (or as enjoyable) without a huge number of people. First and foremost, I'd like to thank my supervisor Dr. Rob Fagan, who has been so fantastic that I'm really struggling to comprehend how I'm ever supposed to work for anyone else again. It's hard to express how grateful I am for all your knowledge, ideas, and enthusiasm – it has been a pleasure spending the last 4 years working with you. Thanks for being so supportive, and so encouraging of a good work-life balance – you taught me how to be my most productive self while also taking all my holidays – I've accomplished so much during my PhD and that's down to you. Most of all, thanks for always telling me I was doing a good job.

I would also like to thank Prof. Mike Brockhurst, for all his evolutionary knowledge, support and encouragement the last 4 years. A special thanks to Dr. Claire Turner, for imparting so much of her RNA/qRT-PCR knowledge to me, and also for dealing with all my crazed qRT-PCR-related requests during the manic final 3 lab months. Thanks to Dr. Roy Chaudhuri, who founded my bioinformatics knowledge and helped me so much along the way. I'm incredibly grateful for this skill and for your support. Thanks to Dr. Rosanna Wright for helping me with genomics and circle plots, and to Anne Williams, for the beautiful vancomycin-BODIPY microscopy images in this thesis.

I'm incredibly proud of the paper, and of the mechanistic insights arising from this PhD, and can honestly say the story wouldn't have been complete without Lucy Thompson. To Lucy, who chose this as her master's project and dedicated so much time and enthusiasm to it – I couldn't have asked for a better person to join the vancomycin journey, and I'm so grateful for you and your contributions to the success of this PhD.

I'd also like to express gratitude to Summit Therapeutics, for the initial support in my PhD, and for the iCASE funding. To Azotic Technologies, for hosting my 3 month placement – I'm grateful for such a fantastic and enjoyable industry experience, and for the wonderful friends I made along the way (especially Gabby Lister, Joe Griffiths and Rodrigo Matus). Thanks to DiMeN, for funding said

placement and my PhD, and to Pam at Don Whitley Scientific, who never refused me conference money despite my incessant requests. Thanks also to Dr. Mal Horsburgh, who put the idea of doing a PhD in my head in the first place.

My experience was made so much brighter by the wonderful friends I have met during my PhD. To Anne, Hannah, Jessie, Josh, Becky, Brooks, Marcel, Joe, Laurence, Anirudh, Finn, and all the other lovely people of F floor, past and present – thanks for the years of support, lab help, lunches, coffee breaks and laughs – you have made my PhD so enjoyable. I'm grateful for all the board games, pub trips, football, baking, walks, zumba, watercolour nights, and everything else which has helped make the last few years so great.

I'd like to acknowledge all my friends outside of Sheffield, but especially Jorjan Regan, Matthew Ackers, and Jake Hood for their unwavering support during my PhD. I'm truly grateful to have such amazing friends, thank you for listening to me talk about vancomycin more times than you probably wanted to hear about it. An extra special thank you to my family, for their support and love throughout my whole life, and for encouraging me to always follow my heart and try my best. I am truly grateful for everything you have done and continue to do for me, I don't know where I'd be without you all.

Finally, I would like to thank Richard, for your endless support and patience through my PhD. Thanks for never once being cross when I was late out of the lab every time you came to see me, and for sharing my joy at the wins, and cheering me up when things weren't going well. Thanks for all the happiness and laughter you've brought to me, I will always be grateful.

Publications

Buddle, J.E., Fagan, R.P., 2023. Pathogenicity and virulence of *Clostridioides difficile*. Virulence 14, 2150452. https://doi.org/10.1080/21505594.2022.2150452

Buddle, J.E., Wright, R.C.T., Turner, C.E., Chaudhuri, R.R., Brockhurst, M.A., Fagan, R.P., 2023. Multiple evolutionary pathways lead to vancomycin resistance in *Clostridioides difficile*. https://doi.org/10.1101/2023.09.15.557922

Galley, N.F., Greetham, D., Alamán-Zárate, M.G., Williamson, M.P., Evans, C.A., Spittal, W.D., Buddle, J.E., Freeman, J., Davis, G.L., Dickman, M.J., Wilcox, M.H., Lovering, A.L., Fagan, R.P., Mesnage, S., 2024. *Clostridioides difficile* canonical L,D-transpeptidases catalyze a novel type of peptidoglycan cross-links and are not required for beta-lactam resistance. Journal of Biological Chemistry 300. https://doi.org/10.1016/j.jbc.2023.105529

Publications Declaration

Chapters 1, 3, 4, and 5 represent expanded versions of Buddle and Fagan, 2023 (chapter 1) and Buddle et al., 2023 (chapters 3 to 5).

The work in Buddle et al., 2023 was performed in collaboration with Lucy M Thompson, Anne S Williams, Rosanna C T Wright, William M Durham, Claire E Turner, Roy R Chaudhuri, and Michael A Brockhurst. The majority of the experimental work (with the exception of works relating to Bc1Δ*dacJ* and R20291Δ*PaLoc*Δ*dacRS*, and vancomycin-BODIPY imaging and analysis, both acknowledged in-text) was performed by Jessica E Buddle. This includes the majority of *C. difficile* strain construction, microscopy, and qRT-PCR; and all of the evolution, phenotypic assessment and bioinformatic analyses.

Further, all figures presented in this thesis were designed and produced by Jessica E Buddle and edited by Robert P Fagan (with the exception of AlphaFold models, which were produced by Robert P Fagan).

For the most part, the text herein is original to this thesis. Any text that is taken from the published works listed above was written by Jessica E Buddle in collaboration with Robert P Fagan (and in the case of Buddle et al., 2023, edited by Michael A Brockhurst).

Abbreviations

A_{260}	Absorbance at 260 nm
A2pm	Meso Diaminopimelic Acid
ABC transporter	ATP Binding Cassette Transporter
ABCD model	Activity, Binding, Cutting, Delivery Model
AMP	Adenosine Monophosphate
ANCOVA	Analysis Of Covariance
ANOVA	Analysis Of Variance
aTc	Anhydrotetracycline
ATCC	American Type Culture Collection
ATP	Adenosine Triphosphate
AUC	Area Under the Curve
Bc	Barcode
BHI	Brain Heart Infusion
BODIPY	4,4-difluoro-4-bora-3a,4a-diaza-s-indacene
bp	Base Pair
BWA	Burrows-Wheeler Aligner
CDC	Centers for Disease Control and Prevention
CDMM	*C. difficile* Minimal Medium
cDNA	Complementary DNA
CDT	*C. difficile* Binary Toxin
CFU	Colony Forming Unit
CLSI	Clinical and Laboratory Standards Institute
CROPs	Combined Repetitive Oligopeptides
Cyclic di-AMP	Cyclic di-adenosine Monophosphate
Del	Deletion
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
DRBD	Delivery and Receptor Binding Domain
DTT	Dithiothreitol

Dup	Duplication
EB	Elution Buffer
EDTA	Ethylenediaminetetraacetic acid
EUCAST	European Committee on Antimicrobial Susceptibility Testing
FMT	Faecal Microbiota Transplantation
gDNA	Genomic DNA
GFP	Green Fluorescent Protein
GI	Gastrointestinal
GlcNAc	N-acetylglucosamine
GTD	Glucosyltransferase Domain
GTP	Guanosine Triphosphate
IGV	Integrative Genomics Viewer
IL	Interleukin
InDel	Insertion-Deletion
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria-Bertani Broth
LHA	Left Homology Arm
Mb	Megabase
MIC	Minimum Inhibitory Concentration
MLS_B	Macrolide-lincosamide-streptograminB
mRNA	Messenger RNA
MRSA	Methicillin-resistant *Staphylococcus aureus*
MurNAc	N-acetylmuramic acid
NICE	National Institute for Health and Care Excellence
N.S.	Not Significant
OD	Optical Density
ORF	Open Reading Frame
P30	Passage 30
PaLoc	Pathogenicity Locus
PBS	Phosphate-buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PLG	Phase Lock Gel
PNPase	Polynucleotide Phosphorylases

qRT-PCR	Quantitative Reverse Transcription PCR
RHA	Right Homology Arm
RNA	Ribonucleic Acid
RNAseq	RNA Sequencing
rpm	Revolutions Per Minute
rRNA	Ribosomal RNA
SNP	Single Nucleotide Polymorphism
SOC	Super Optimal Broth
sRNA	Small RNA
T_gen	Generation Time
TBS	Tris Buffered Saline
TCS	Two Component System
TEM	Transmission Electron Microscopy
TNF	Tumour Necrosis Factor
Tris	Tris(hydroxymethyl)aminomethane
TY	Tryptone Yeast
UDP-glucose	Uridine Diphosphate Glucose
UTR	Untranslated Region
v/v	Volume/Volume
Vcf	Variant Call Format
w/v	Weight/Volume
WGS	Whole Genome Sequence
WT	Wildtype

Contents

List of Figures

List of Tables

1 Introduction

1.1 Clostridioides difficile

Clostridioides difficile is a gram-positive obligate anaerobe, capable of causing disease through the faecal-oral transmission of robust endospores (Deakin et al., 2012). These metabolically dormant spores are able to persist in a range of environments, being resistant to oxygen, heat, and many common disinfectants – contributing to both the organism's success as a pathogen, and the associated healthcare costs and difficulty of treating infection (Adams et al., 2013). *C. difficile* is responsible for over 120,000 infections per year in Europe alone, and is the leading cause of hospital-associated diarrhoea (ECDPC, 2023). As well as being an important nosocomial pathogen, a recent paradigm-shift has seen increasing reports of community-acquired *C. difficile* infection (CDI) (ECDPC, 2023). Although often less severe, community-acquired CDI is responsible for an estimated 20-27% of all cases, resulting in a significant burden (Liao et al., 2018). Clinical presentation of CDI covers a large spectrum of disease symptoms, with diarrhoea and colitis being the most common. The significant mortality associated with *C. difficile* typically arises from more severe manifestations, including pseudomembranous colitis, fulminant colitis and toxic megacolon (Rupnik et al., 2009). Infection recurrence, characterised by the reappearance of symptoms after treatment completion, is also common, largely due to the nature of available CDI treatments (Song and Kim, 2019). This results in complex treatment plans and worsened prognosis (Cole and Stahl, 2015; Napolitano and Edmiston, 2017).

Paradoxically, antibiotics constitute both the main treatment and major risk factor for *C. difficile* infection. Administration of broad-spectrum antimicrobials, either prophylactically or to treat another infection, lead to disruption of the gut microbiota, resulting in a dysbiotic state in which *C.*

difficile thrives (Lawley et al., 2012). As well as being associated with the broad-spectrum antimicrobials cephalosporins, clindamycin and fluoroquinolones; antibiotics commonly used to treat *C. difficile* itself can also contribute to CDI, by exacerbating dysbiosis and leaving the patient acutely sensitive to reinfection or relapse (Owens et al., 2008). This recurrence is the most common complication of CDI, arising in up to 30% of patients (Song and Kim, 2019). Collectively, this combination of factors warrants the recent classification of *C. difficile* as an "urgent threat" (CDC, 2024).

1.1.1 CDI

Clinical manifestations of CDI range from self-limiting diarrhoea to often fatal inflammatory complications. Typical presentations of CDI involve mild diarrhoea, which resolves with or without antibiotic treatment. However, notable complications include severe diarrhoea and dehydration, colon perforation, kidney failure, and septicaemia (Czepiel et al., 2019). Severe presentations of CDI may involve pseudomembranous colitis, a colonic inflammatory disease resulting in yellow plaques which form pseudo membranes on the intestinal mucosa; or toxic megacolon, the widening of the colon due to inflammation (Farooq et al., 2015; Sayedy et al., 2010). In such instances, prognosis is poor – patients with toxic megacolon arising from CDI require aggressive treatments, and have a 38-80% mortality rate (Sayedy et al., 2010).

1.1.2 Epidemiology

The phylogenetic diversity of *C. difficile* has allowed for the emergence of several epidemic strains in recent years. In particular, the ribotype 027 lineage was responsible for a 2001 North American epidemic, which spread to the UK, peaking in 2004-2007 (He et al., 2013). This hypervirulent lineage is associated with increased transmission and mortality, although the underlying reasons for the

apparent increase in pathogenicity are far from clear. Ribotype 027 strains display increased expression of toxins, a deletion in *tcdC* (putatively encoding a negative regulator of toxin expression) (Curry et al., 2007; Spigaglia and Mastrantonio, 2002; Warny et al., 2005), and production of an additional binary toxin (Gerding et al., 2014). These strains also displayed high-level resistance to fluoroquinolone antibiotics, including the commonly used ciprofloxacin (Razavi et al., 2007). In the UK, improved hospital management strategies have resulted in an 80% reduction in cases from the height of the 2004-2007 outbreak, with current annual infection rates lying at 13,000 (Mansfield et al., 2020; PHE, 2024). *C. difficile* is responsible for an estimated 29,000 and 1,800 deaths per year in the USA and UK respectively (Lessa et al., 2015; UKHSA, 2022). Case fatality is approximately 15%, however this increases with each subsequent infection recurrence (Cole and Stahl, 2015). Despite strategies to reduce CDI, the costs associated with CDI have increased, with infections costing between $436 million to $3 billion per year in the USA, with total CDI-attributable costs excelling $6.3 billion (Heimann et al., 2018). In England, CDI costs £5,000-£15,000 per case (Heimann et al., 2018). This burden is not solely economic – with the average UK patient stay being 37 days, CDI puts huge pressures on healthcare facilities (Reigadas Ramírez and Bouza, 2018). Despite less emphasis being placed on the burden of CDI in lower-income countries, it is clear that the lack of diagnosis and prevention has led to a severe underestimation of CDI. In many African countries, due to reduced regulation of antibiotics and high HIV prevalence, CDI burden is likely to be high (Roldan et al., 2018). Similarly, in South Africa, CDI incidence was shown to be 9.2%, a third of which was community-acquired (Rajabally and Afzal, 2019).

1.1.3 Toxins

Clinical presentation of CDI is influenced by a range of *C. difficile* virulence factors, including production of various toxins and surface proteins. Primarily, pathogenesis is driven by the activity of toxins A and B, encoded within the pathogenicity locus (*PaLoc*).

1.1.3.1 PaLoc

The *C. difficile PaLoc* spans a 19.6 kb region, with a typically highly conserved genomic localisation and organisation, and encodes 5 proteins involved in toxin-mediated pathogenesis (Figure 1.1a). The five genes in the *PaLoc* include *tcdA* and *tcdB,* encoding toxins A and B respectively; as well as *tcdR, tcdE* and *tcdC.* TcdR is an alternative sigma factor and likely positive regulator of toxin production, since purified *C. difficile* RNA polymerase was unable to bind to the *tcd* promoter regions in the absence of TcdR, and interaction of TcdR with the RNA polymerase holoenzyme allowed transcriptional activation (Govind et al., 2015; Mani and Dupuy, 2001). TcdR also activates its own promoter, in a positive feedback loop, allowing regulation of the entire *PaLoc*. TcdC was thought to be an anti-sigma factor involved in modulating toxin expression through sequestration of TcdR (Matamouros et al., 2007). Indeed, the 18 bp *tcdC* deletion characteristic of ribotype 027 has been attributed to increased toxin expression in these strains. However, the role of TcdC has recently been disputed after identification of low-toxin ribotype 027 clinical isolates, which still possess the archetypal *tcdC* deletion (Anwar et al., 2022). Additionally, correction of the 18 bp *tcdC* deletion through allele exchange had no effect on toxin production (Cartman et al., 2012), suggesting the function of TcdC requires further classification. The exact function of TcdE has also previously been controversial, however it seems likely that this holin-like protein is involved in toxin secretion, as recently demonstrated in clinical strains (Majumdar and Govind, 2022; Mani and Dupuy, 2001). Holins are membrane proteins, commonly encoded by double-stranded DNA phage, which are required for host cell lysis. The *tcdE* open reading frame contains three translational start sites resulting in TcdE isoforms of three different sizes. The involvement of combinations of these isoforms in both toxin release and cell death was demonstrated in the hypervirulent strain R20291 (Govind et al., 2015).

Toxins A (TcdA) and B (TcdB) consist of a broadly similar four-domain structure and are highly similar, with 47% amino acid identity, suggestive of an originating gene duplication event (Figure 1.1a). The N-terminal consists of a glucosyltransferase domain (GTD), next to which is a small cysteine protease domain, involved in auto processing for the release of the GTD (Pruitt et al., 2010). The next domain, often called the Delivery and Receptor Binding Domain (DRBD), contains a hydrophobic region and is thought to be involved in translocation of the GTD from the lumen of endocytic vesicles into the host cell cytoplasm. The C-terminal receptor binding domain (also known as C-terminal combined repetitive oligopeptides (CROPs) domain) can bind to a range of carbohydrates, likely facilitating toxin binding to the cell surface (Hartley-Tassell et al., 2019).

The proposed mode of action, known as the ABCD model (activity (A), binding (B), cutting (C), delivery (D)), is similar between both TcdA and TcdB (Figure 1.1b) (Jank and Aktories, 2008). Here, TcdA and TcdB bind to cellular receptors. Once bound, the toxins undergo endocytosis, through a clathrin-dynamin-dependent pathway (Papatheodorou et al., 2010). Subsequent acidification of the endosome results in a conformational change of the toxin, allowing membrane insertion and formation of a channel through which the glucosyl transferase domain passes (Barth et al., 2001; Qa'Dan et al., 2000). In the cytosol, the toxins undergo a further change, induced by the host cofactor inositol hexakisphosphate. This allows activation of the toxin cysteine protease domain, and results in toxin autocleavage at a position between the cysteine protease and glucosyltransferase domains, releasing the glucosyltransferase domain into the cytosol (Egerer et al., 2009; Oezguen et al., 2012). Despite TcdB being able to induce cellular toxicity independent of the GTD, recent evidence suggests glucosyltransferase activity is still key for disease pathogenesis (Bilverstone et al., 2020). In the cytosol, the GTD is able to inactivate members of the Rho guanosine triphosphatase (GTPase) family, including Rho, Rac and Cdc42, via transfer of glucose from uridine diphosphate glucose (UDP-glucose) to these proteins at a conserved threonine residue (Jank et al., 2007; Just et

al., 1995). Since Rho GTPases control pleiotropic signal transduction pathways, disruption to the host cell is widespread. Most notably, dysregulation of actin depolymerisation leads to disruption of the cytoskeleton, resulting in cell rounding, apoptosis, disruption of tight junctions, and loss of intestinal barrier function (Gerhard et al., 2008; Janoir, 2016). In fact, TcdA and TcdB are capable of causing both type I (apoptosis) and type III (necrosis) programmed cell death (Voth and Ballard, 2005). Changes in Rho GTPase function also evoke changes in proinflammatory signalling pathways, resulting in the production of proinflammatory cytokines interleukin (IL)-1β, tumour necrosis factor (TNF)-α, and IL-8. This, coupled with the subsequent influx of neutrophils, leads to further host tissue damage characteristic of CDI (Madan and Petri, 2012).

1.1.3.2 CDT

Characterisation of the hypervirulent ribotype 027 epidemic strain, first reported at the start of the millennium, showed a combination of factors putatively involved in increased virulence: high-level fluoroquinolone resistance, an 18 bp *tcdC* mutation, and possession of a further toxin – *C. difficile* binary toxin (CDT) (Loo et al., 2005). CDT is an ADP-ribosylating toxin, composed of 2 proteins, the crystal structures of which have been reported recently (Anderson et al., 2020; Xu et al., 2020). CDTa is an ADP-ribosyltransferase, the enzymatic component involved in modifying host cell actin; while CDTb is involved in binding to host cells and translocating CDTa to the host cytosol. CDT first binds to host cells via the lipolysis-stimulated lipoprotein receptor, present in host cells in the liver, kidney, small intestine and colon (Papatheodorou et al., 2011). This binding is followed by accumulation of lipid rafts, oligomerisation and induction of endocytosis (Hemmasi et al., 2015; Papatheodorou et al., 2010). In the resulting endosome, acidification triggers membrane insertion and pore formation by CDTb, allowing translocation of CDTa into the cytosol. Refolding of CDTa after translocation is mediated by host chaperones, including Hsp90 and Cyp40 (Ernst et al., 2018). CDTa then ADP-ribosylates cellular actin at Arg177, producing ADP-ribose and nicotinamide as biproducts. The modified actin is prevented from further polymerisation due to the ADP-ribose moiety. Eventually, this leads to complete depolymerisation of the actin cytoskeleton, resulting in phenotypes typical for

toxins affecting the cytoskeleton, including loss of barrier function and disruption of tight junctions (Aktories et al., 2011; Schwan et al., 2009). However, CDT displays a multifaceted approach to host cell toxicity, since actin polymerisation results in redistribution of the microtubule network. Microtubules are involved in a range of cellular processes, including intracellular transport, cell division and cilia formation (Akhmanova and Steinmetz, 2015). CDT hijacks this network, resulting in formation of long cellular protrusions which increase *C. difficile* adherence to host cells, both *in vitro* and in mouse models (Schwan et al., 2014, 2009). Details of the mechanism of CDT are reviewed in detail elsewhere (Aktories et al., 2011; Gerding et al., 2014).

The regulation of CDT is distinct from, but entwined with, the *PaLoc*; since *cdtA* and *cdtB* are located on a separate 6.2 kb chromosomal region of the genome, known as *CdtLoc. CdtLoc* contains the two genes encoding CDT, and *cdtR* – a LytTR family orphan response regulator. CdtR is a positive regulator of both CDT and the *PaLoc* in hypervirulent strains (Lyon et al., 2016). Non-CDT producing *C. difficile* strains contain either a truncated version of the *CdtLoc*, or a 68 bp insertion sequence at this site (Carter et al., 2007).

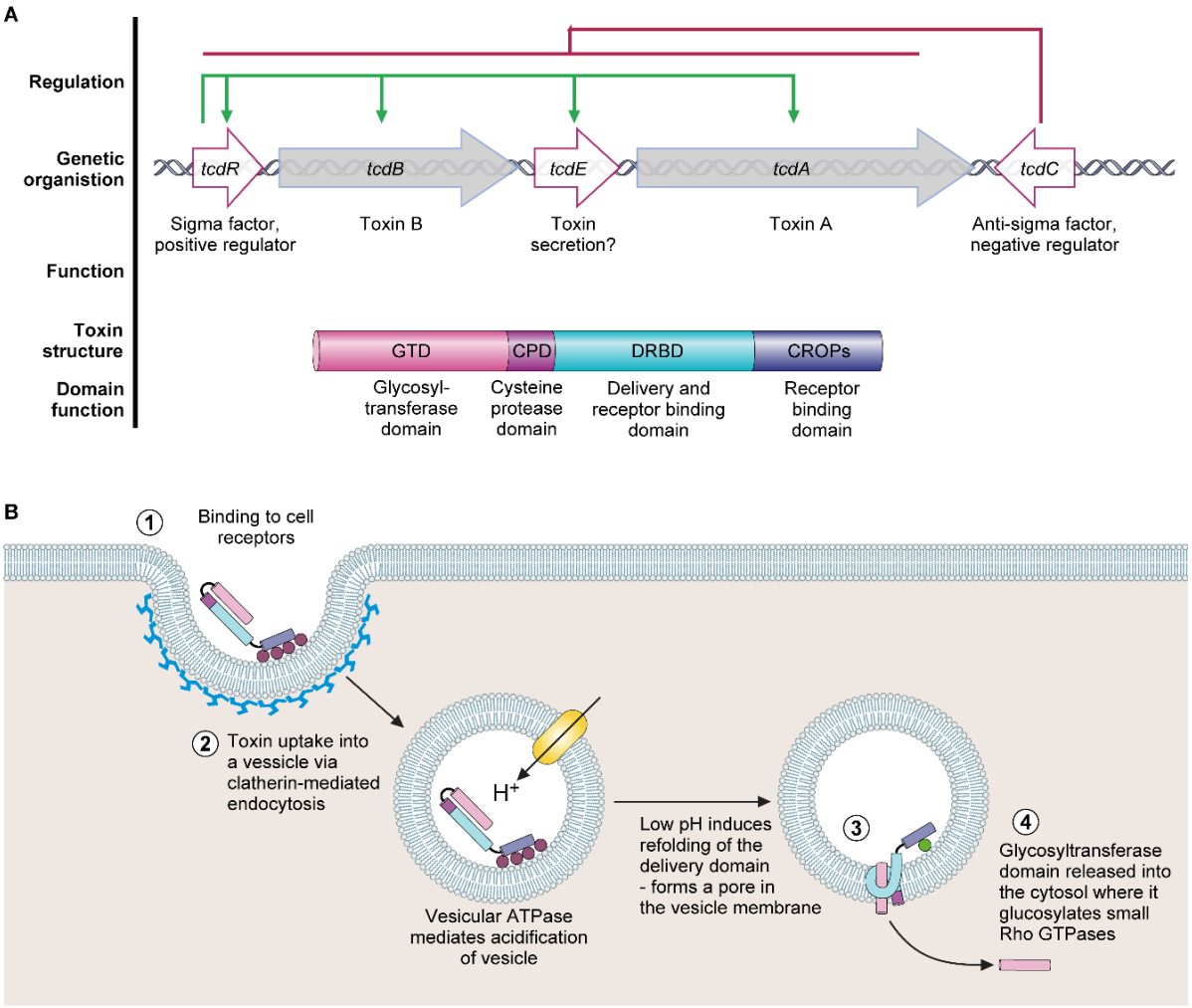


Figure 1.1 - The Pathogenicity Locus and Toxin Mode of Action

(A) The pathogenicity locus (*PaLoc*) is comprised of 5 genes: *tcdA* and *tcdB*, encoding toxins A and B respectively; *tcdR*, encoding an alternative sigma factor and likely positive regulator of the *PaLoc* (regulation shown above in green); *tcdE*, encoding a holin-like protein putatively involved in toxin secretion; and *tcdC*, a putative anti-sigma factor and negative regulator of the *PaLoc* genes (regulation shown above in red). Toxins A and B both consist of a broadly similar four-domain structure. At the N-terminal, the glucosyltransferase domain (GTD) is the active toxin moiety which inactivates members of the Rho GTPase family. A cysteine protease domain is next to the GTD, and is involved in auto-processing and release of the GTD. The next domain, often called the Delivery and Receptor Binding Domain (DRBD), contains a hydrophobic region and is thought to be involved in translocation of the GTD from the lumen of endocytic vesicles into the host cell cytoplasm. The final C-terminal receptor-binding domain (also known as C-terminal combined repetitive oligopeptides (CROPs) domain) binds to a range of cellular receptors. (B) Toxin mode of action. The toxins bind to various cellular receptors via the C-terminal CROPs domain, triggering clathrin-dependent endocytosis (1) followed by acidification of the resulting vesicle (2). The drop in pH triggers a conformational change in the delivery domain which inserts into, and forms a pore in, the vesicle membrane, through which the GTD transits into the host cytoplasm (3). The GTD is then released by a cleavage event mediated by the cysteine protease domain, in a process that is dependent on host inositol hexakisphosphate (4). The GTD is then able to glucosylate and inactive members of the small Rho GTPase family, including Rho, Rac, and Cdc42. Inactivation of Rho GTPases results in multi-level cellular disruption, including dysregulation of actin depolymerisation, which causes disruption of

tight junctions and loss of intestinal barrier function, induction of proinflammatory cytokines and activation of programmed cell death. Reproduced from Buddle and Fagan, 2023 under a CC-BY licence.

1.1.4 Cell Wall

1.1.4.1 Structure

The bacterial cell wall surrounds the membrane and is necessary for growth, shape, maintaining cell integrity, and mediating interactions with the environment. The broadly conserved structure, and pleiotropic functions, of the cell wall make it an attractive antibiotic target. The cell wall is comprised primarily of peptidoglycan – a heteropolymer of glycan chains that are cross-linked by peptides. The glycan chains consist of alternating β-1,4 linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues, making the polysaccharide backbone. These glycan chains can vary in levels of glycan N-deacetylation or O-acetylation between species. *C. difficile* displays high levels of N-deacetylation in GlcNAc residues, which aids in lysozyme resistance (Vollmer and Tomasz, 2000).

The peptide stem is linked to the MurNAc, and consists of L-Alanine, D-Glutamate, meso diaminopimelic acid (A2pm), and two D-Alanines (D-Ala) in *C. difficile* (Coullon and Candela, 2022). The peptide stem is synthesised as a pentapeptide, which may be trimmed during assembly – the majority of peptidoglycan peptides are tetrapeptides, followed by tripeptides, in *C. difficile* (Peltier et al., 2011). These peptide stems are cross-linked to neighbouring peptide stems, to strengthen the cell wall, creating the mature peptidoglycan mesh structure (Figure 1.2a). The process of crosslinking nascent peptidoglycan is essential, and inhibition of crosslinking is a primary mechanism of multiple antibiotics. Crosslinking of peptidoglycan in bacteria most commonly occurs via D-Ala → $A2pm^3$ (4,3), however *C. difficile* exhibits a high abundance of $A2pm^3$ → $A2pm^3$ (3,3) crosslinks. In fact, 3,3 crosslinking accounts for the majority (75%) of crosslinks in *C. difficile* (Coullon et al., 2020).

1.1.4.2 Assembly

The initial stages of peptidoglycan synthesis occur in the cytoplasm (Figure 1.2b). MurNAc is synthesised from GlcNAc precursors (Egan et al., 2020, 2015; Galinier et al., 2023). Further modification of MurNAc occurs via the addition of a pentapeptide stem, catalysed by a series of amino acid ligases, primarily encoded by the *mur* gene cluster. This MurNAc-pentapeptide precursor is bound to UndP in the inner face of the cytoplasmic membrane via MraY, forming lipid I. Addition of a GlcNAc to lipid I, catalysed by MurG, forms lipid II (UndPP-GlcNAc-MurNAc-pentapeptide). Lipid II is then flipped across the cytoplasmic membrane via flippase MurJ, and the nascent peptidoglycan is polymerised via glycosyltransferases. The glycan chains must then be cross-linked via transpeptidases. 4,3 crosslinks are catalysed by D,D-transpeptidases (penicillin-binding proteins (PBPs)), whereby the donor pentapeptide loses its terminal D-Ala, with the crosslink formed between the 4^{th} position D-Ala and the 3^{rd} position of a neighbouring peptide. 3,3 crosslinks are catalysed by L,D-transpeptidases, which use tetrapeptides as donors. In *C. difficile,* there are 3 canonical L,D-transpeptidases, which can be deleted simultaneously with minimal impact on peptidoglycan structure, indicating the presence of novel, non-canonical L,D-transpeptidases which are as yet unidentified (Galley et al., 2024). Carboxypeptidases are highly redundant, and are involved in producing tetrapeptide donors for L,D-transpeptidases, as well as peptidoglycan maintenance and turnover (Egan et al., 2020).

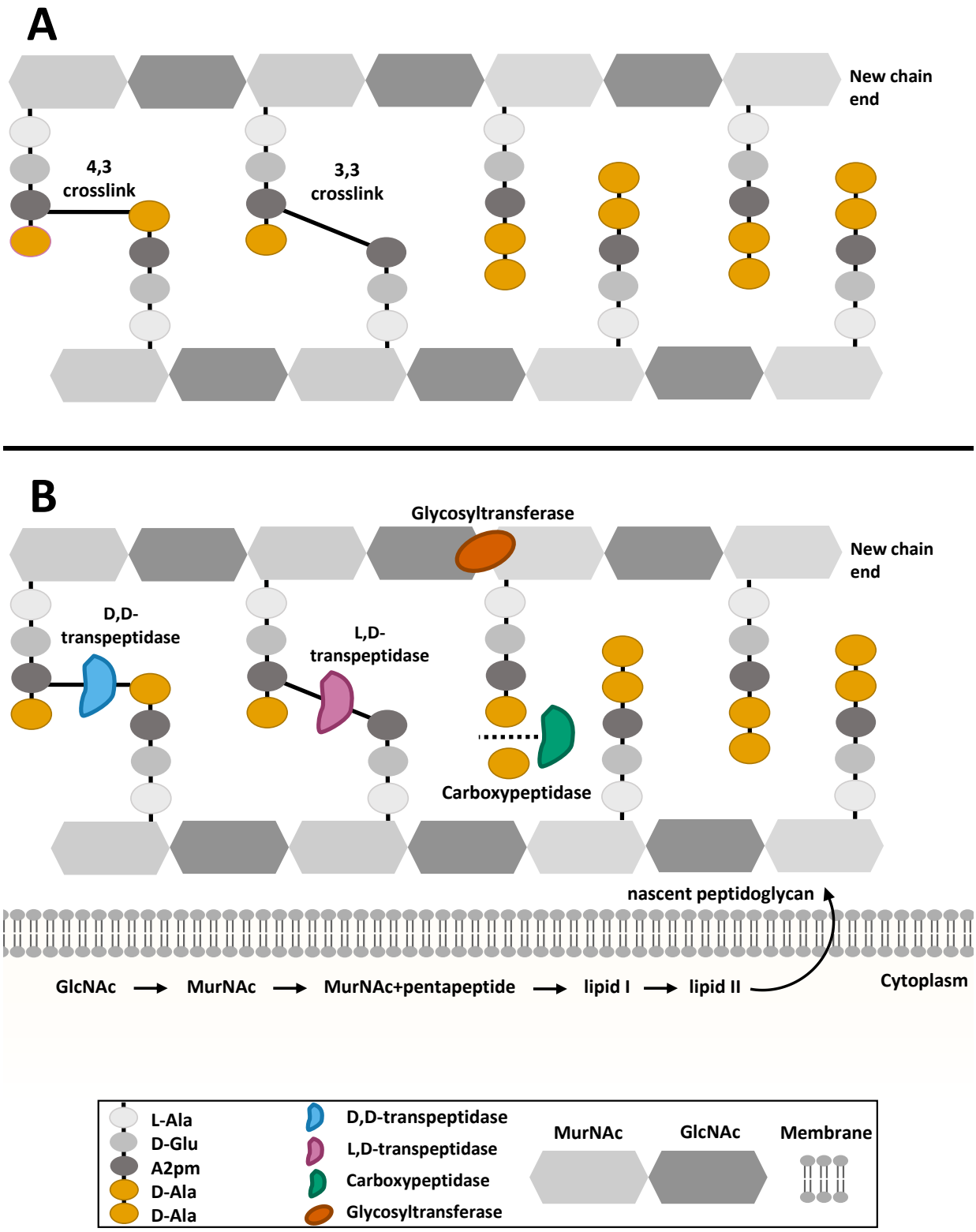


Figure 1.2 - Cell Wall Structure and Assembly

(A) Cell wall structure – the peptidoglycan cell wall is composed of alternating GlcNAc-MurNAc residues. These may be 4,3 (D-Ala → A2pm3) or 3,3 (A2pm3 → A2pm3) crosslinked. (B) Cell wall assembly – MurNAc is synthesised from GlcNAc, and the pentapeptide stem is added to MurNAc through a series of amino acid ligase reactions. MurNAc is converted to lipid I, and GlcNAc is added to produce lipid II. This is flipped across the cytoplasmic membrane, and incorporated by

glycosyltransferase enzymes to produce peptidoglycan chains. These are cross-linked by D,D- and L,D-transpeptidases.

1.2 *C. difficile* Treatments

Antibiotic chemotherapy often results in further dysbiosis and increased risk of recurrent infection. Despite this, treatment for CDI typically relies on metronidazole, fidaxomicin or vancomycin. Non-antibiotic alternatives, such as faecal microbiota transplantation (FMT) or defined microbiome reconstitution, are however becoming more common.

1.2.1 Metronidazole

Metronidazole is a 5-nitroimidazole prodrug, which, although synthetic, derives from 2-nitroimidazole, a compound isolated from *Streptomyces sp.* in the 1950s (Leitsch, 2019). Since metronidazole is a prodrug, it is inactive until taken up and reduced. The transfer of electrons activates metronidazole intracellularly, generating a variety of nitro radicals. These reactive intermediates have pleiotropic effects on the cell, from damaging DNA and proteins, to reducing protective thiol pools. Such reactions occur primarily in low oxygen conditions, since the presence of oxygen re-oxidises metronidazole into its stable prodrug form, meaning metronidazole is primarily used for anaerobic and microaerophilic organisms.

Metronidazole was established as a treatment for CDI in the 1990s, since it allowed comparable cure rates to vancomycin at a cheaper price (Olaitan et al., 2023). Additionally, although noted, resistance was reported to be transient, highly unstable and heterogeneous (Martin et al., 2008; Peláez et al., 2008). More recently however, multiple mechanisms of metronidazole resistance in *C. difficile* have been reported, resulting in metronidazole no longer being recommended as the front-line antibiotic for treating CDI (NICE, 2021).

1.2.2 Fidaxomicin

Fidaxomicin is a macrocyclic lactone antibiotic, a semisynthetic product derived from *Actinomycetes* (Dorst et al., 2020). Fidaxomicin targets the bacterial RNA-polymerase – an attractive antibiotic target, since it is both essential and conserved, but also displays variation among species, allowing the potential to be targeted by both broad- and narrow-spectrum antibiotics (Artsimovitch et al., 2012). Fidaxomicin is a narrow-spectrum antibiotic, which acts to inhibit RNA-polymerase at a site distinct from rifamycin through binding to the DNA-template-RNA-polymerase complex prior to transcription initiation. This traps the complex in an "open clamp" position, preventing interaction with the -35 and -10 sequence (Lin et al., 2018; Venugopal and Johnson, 2012). The specificity of fidaxomicin was uncovered recently using cryo-electron microscopy, since the *C. difficile* RNA polymerase was shown to possess a fidaxomicin-binding determinant, absent from other members of the gut microbiome (Cao et al., 2022).

Fidaxomicin presents many virtues – it is highly specific, and provides high faecal, and low serum, concentrations (Zhanel et al., 2015). It has minimal adverse effects on the gut flora, meaning it is highly effective (and even superior to vancomycin) for treating recurrent infections, allowing sustained clinical responses. It is comparable to the current front-line antibiotic, vancomycin, in relation to clinical responses and safety. The hesitancy surrounding uptake for mainstream use of fidaxomicin is therefore largely centred round cost – a 10-day course of fidaxomicin is estimated to cost 3845.44 USD, vs 23.28 USD for vancomycin (Patel et al., 2023). Despite multiple studies suggesting fidaxomicin is actually cost-saving, by reducing hospital stay times and recurrence rates, the higher upfront cost has remained a barrier to widespread use of this antibiotic (Gallagher et al., 2015; Patel et al., 2023).

1.2.3 Vancomycin

Vancomycin is a glycopeptide broad-spectrum antibiotic used to treat gram positive pathogens. It is a natural product antibiotic, initially isolated in the 1950s, from *Streptomyces orientalis* (Rubinstein and Keynan, 2014). Vancomycin side effects are broad-ranging, from kidney and urinary probems to hearing loss (Peng et al., 2020), meaning although potent, vancomycin was initially abandoned in favour of alternatives with less severe side effects. In later years, vancomycin was revisited as an important treatment for gastrointestinal (GI) infections, since it is poorly absorbed by the GI tract, meaning it reaches high concentrations in the gut lumen. Unlike β-lactam antimicrobials, vancomycin impacts cell wall biogenesis at multiple levels – by binding to the terminal D-Ala of the lipid II pentapeptide chain in nascent peptidoglycan, vancomycin inhibits transpeptidation reactions, preventing peptidoglycan crosslinking. It also impedes glycosyltransferase activity, inhibiting glycan chain elongation (Stogios and Savchenko, 2020). These actions achieve bactericidal activity by compromising the envelope, leading to increased osmotic pressure and lysis (Stogios and Savchenko, 2020).

Despite metronidazole initially being the recommended antibiotic for severe CDI, the results of multiple clinical trials found it to be inferior to vancomycin, leading to a change in practice (Czepiel et al., 2019). As of 2021, vancomycin was declared the current front-line antibiotic for treatment of mild, moderate and severe CDI, as well as for recurrent infections (NICE, 2021). It is highly efficacious, achieving over 85% clinical cure, however recurrence rates from vancomycin treatment are around 25% (Louie Thomas J. et al., 2011).

1.2.4 Faecal Microbiota Transplantation (FMT)

FMT involves administration of faeces from a healthy individual (heterologous), or from one's own previously-healthy microbiome (autologous) to restore the natural gut flora. This has gained popularity as a treatment for CDI over the last decade, however the procedure has yet to be standardised, and there have been reports of adverse events post-transplantation (Lee et al., 2015). Typically, faeces can be delivered via colonoscopy, enema, nasogastric tube or oral capsules (Basson et al., 2020). The virtues of FMT are well established, as both a standalone and combination therapy: one trial suggested clinical resolution following FMT was 92% (Quraishi et al., 2017), while another found FMT with vancomycin provided an 81% clinical resolution of CDI, significantly higher than treatment with vancomycin alone (Nood et al., 2013). Moreover, the potential of FMT to treat the major complication, recurrent CDI, should not be forgotten – a recent study found a 68% success rate across complex patients with recurrent CDI alongside multiple co-morbidities and extended antibiotic use (Nowak et al., 2019). Despite intense effort in recent years, the underlying mechanism of FMT-mediated restoration of colonisation resistance is still disputed but likely involves a combination of competition for resources, immune modulation and production of inhibitory metabolites. Intriguingly however, a very small trial of only 5 patients demonstrated a high rate of CDI resolution using a sterile faecal filtrate, hinting that resident bacteriophage could also be a contributory factor in the effectiveness of FMT (Ott et al., 2017).

Despite the clear effectiveness of FMT, its unconventional nature has limited public acceptance, and the lack of process standardisation poses a worry to clinicians. Further, upon progression to pseudomembranous colitis, FMT has reduced efficacy and often requires repeat treatment (Sbahi and Palma, 2016). There is also a question mark over manipulation of the microbiome – despite huge advancement in metagenomics, a complete understanding of the gut microbiome, and essential constituents, is lacking (Gupta et al., 2016). Most importantly, larger, randomised-

controlled clinical trials are required to fully understand the efficacy and safety of FMT, since the nature of the therapy holds the risk of transferring pathogens to already-vulnerable patients (Lee et al., 2015). Taken together, FMT provides a feasible alternative therapy for CDI, however there are many challenges to overcome before it becomes mainstream. A more refined approach to FMT is clearly desirable, and this is reflected in the array of new microbiome-based therapeutics in clinical development or already undergoing clinical trials for treatment of CDI. Among these are those derived from donor faeces, such as SER-109 from Seres Therapeutics, consisting of spores of mixed Firmicute species purified following ethanol treatment (McGovern et al., 2021), and suspensions of defined bacterial communities grown in pure culture such as the 8-species VE303 from Vedanta Biosciences (Dsouza et al., 2022).

1.3 Resistance to Antibiotics in *C. difficile*

One of the most important factors in *C. difficile* colonisation is antibiotic resistance (Figure 1.3). Prior exposure to antibiotics has long since been accepted as the primary risk factor for CDI, since increased abundance of *C. difficile* in the colon correlates with dysbiosis, most commonly caused through antibiotic exposure (Lawley et al., 2009). Being intrinsically highly resistant to a multitude of antibiotics further increases virulence, and significantly reduces treatment options. The major complication of CDI, recurrence, is also attributed to exacerbation of gut dysbiosis due to antibiotic treatment. Thus, antibiotic resistance allows colonisation, avoidance of clearance, persistence and recurrence – impacting all aspects of infection.

1.3.1 Contribution of Antimicrobial Resistance to *C. difficile* pathogenicity

A multitude of evidence supports antibiotics as the major risk factors for CDI. The mechanism of microbiome-associated colonisation resistance is far from clear, but is likely a multi-faceted

phenomenon involving competition for nutrients, immune modulation, and production of harmful metabolites (Ducarmon et al., 2019). The best understood factor is the impact on bile acid metabolism, and in particular, the conversion of deconjugated primary bile acids to deoxycholate and lithocholate by members of the microbiome with 7α-dehydroxylase activity (Buffie et al., 2015). Treatment with antibiotics, either prophylactically or for another infection, causes severe and unpredictable disruption to the resident microbiome. Changes in diversity and relative abundance of species within the microbiome reduces colonisation resistance in the colon, allowing *C. difficile* to colonise and flourish (Lawley et al., 2009). Supporting this, a wealth of clinical evidence links antibiotic exposure to CDI: retrospective cohort studies have implicated dose, number of antibiotics used, and days of antibiotic exposure with CDI, with risk increasing in a dose-dependent manner (Stevens et al., 2011). Further, a striking recent study suggested that odds of infection increased by 12.8% with every day of antibiotic therapy – however this was dependent on both the antibiotic used, and route of administration (Webb et al., 2020). Although many antimicrobials are associated with CDI, risk is most highly associated with broad-spectrum antibiotics, including cephalosporins, carbapenems, fluoroquinolones, and clindamycin (Webb et al., 2020). In mouse models, clindamycin reduced microbiome diversity by 90% for 28 days, which increased CDI mortality and led to colonic inflammation, even in recovering mice – suggesting antibiotic exposure increases not only the risk, but severity of CDI (Buffie et al., 2012). Of course, recurrence – either through relapse or reinfection – is also highly associated with antibiotic use (Gómez et al., 2017). In paediatric recurrent CDI patients, antibiotic exposure and recent surgery were significant recurrence risk factors (Nicholson et al., 2015). Other clinical studies report similar outcomes, with antibiotics, and previous use of fluoroquinolones, being independent risk factors for recurrence (Song and Kim, 2019). Therefore, antibiotic use undoubtedly has a large impact on the ability of *C. difficile* to act as an opportunistic pathogen.

The success of *C. difficile* as a pathogen is inherently linked to its ability to resist antibiotics. The 4.29 Mb genome of *C. difficile* has demonstrated an extraordinary ability to accumulate resistance determinants to a multitude of antibiotics, including aminoglycosides, tetracyclines, erythromycin, clindamycin, beta-lactams, and cephalosporins (Peng et al., 2017; Spigaglia, 2016). This multidrug resistance was the driving force of the CDI epidemic at the start of the millennium, in addition to emergence of novel epidemic lineages, highlighting the importance of such factors in pathogenesis. Resistance to the macrolide-lincosamide-streptograminB (MLS_B) family of antibiotics, encompassing erythromycin and clindamycin, is achieved through ribosomal methylation, and is gained via acquisition of transposons, such as Tn*5398*, containing *erm* genes (Farrow et al., 2000; Peng et al., 2017). *erm* encodes a 23S rRNA methylase, which modifies the 23S rRNA of the 50S ribosomal subunit, reducing drug binding affinity (Dzyubak and Yap, 2016). However, several *C. difficile* erythromycin resistant strains have been identified which lack *erm* genes – suggesting the presence of yet uncharacterised alternative resistance mechanisms (Spigaglia et al., 2011). Tetracycline resistance is less widespread in *C. difficile,* however conjugative transposons have allowed transfer of *tetM* to certain strains, providing a mechanism of ribosome protection against tetracycline (Dong et al., 2014). Perhaps the most intriguing capability is that of fluoroquinolone resistance. Not unusually, resistance occurs via alterations to the DNA gyrase subunits, typically GyrA (Figure 1.3) (Dridi et al., 2002). However, the emergence of ribotype 027 was associated with widespread fluoroquinolone use, and the epidemic strains possessed recently-acquired high-level fluoroquinolone resistance (Wasels et al., 2015). Since antibiotics target essential cellular processes, resistance is often associated with large fitness burdens. However, competition analysis of strains carrying mutations seen in *C. difficile* 027 clinical isolates found fluoroquinolone resistance did not lead to fitness costs *in vitro* (Wasels et al., 2015).

1.3.2 Resistance to Antibiotics Used to Treat *C. difficile*

Of course, being resistant to a wealth of antibiotics poses two challenges: (i) the extensive resistance displayed greatly reduces treatment options for CDI, warranting the status of *C. difficile* as an urgent threat; and (ii) such treatment options are likely to be further limited through the high degree of adaptation and flexibility in the *C. difficile* genome. As discussed above, until recently, three antibiotics were commonplace for the treatment of CDI. Metronidazole was typically the antibiotic of choice for mild-to-moderate CDI in first instance of infection, while vancomycin was reserved for severe and severe-complicated disease. Fidaxomicin was often overlooked due to higher cost, being significantly more expensive than metronidazole (Cruz, 2012; Nelson et al., 2017). In 2021, vancomycin became the national institute for health and care excellence (NICE)-recommended front-line antibiotic for CDI, replacing metronidazole as the first-instance treatment (NICE, 2021). This move reflects both high metronidazole-related recurrence rates, and increasing reports of metronidazole resistance, but poses risks of its own in terms of increasing vancomycin selection pressures.

1.3.3 Metronidazole Resistance

Alongside the recent emergence of various epidemic lineages, there has been an increase in metronidazole treatment failure (Banawas, 2018). Resistance in *C. difficile* was previously thought to be transient, however a recent explosion in research focussed on characterising metronidazole resistance has led to the discovery of multiple heritable pathways to reduced susceptibility (Figure 1.3). One such mechanism involved a 7 kb plasmid, dubbed pCD-METRO, which increased resistance 25-fold, and conferred stable resistance to metronidazole (Boekhoud et al., 2020). Worryingly, pCD-METRO is thought to be horizontally transferrable, and is already internationally disseminated. However, the lack of universality of this mechanism implies the existence of multiple pathways to metronidazole resistance. Further clinical isolate studies found multiple single nucleotide

polymorphisms (SNPs) in genes affecting iron utilisation and electron transport – hinting at the molecular mechanism of resistance (Lynch et al., 2013). This was further uncovered using an evolutionary approach, which demonstrated the involvement of redox and iron homeostasis genes, in a deterministic route to resistance (Deshpande et al., 2020). Most recently, mutations in the *nimB* promoter, resulting in constitutive expression of NimB, has been implicated in resistance (Olaitan et al., 2023). NimB, a heme-dependent flavin enzyme, promotes resistance through degradation of metronidazole reactive intermediates, explaining the heme-dependancy often observed with metronidazole resistance (Boekhoud et al., 2021). Interestingly, this mutation co-occurred with mutations in *gyrA*, which confer fluoroquinolone resistance in *C. difficile* epidemic strains. This presents a worrying outlook for the future of metronidazole usage. The existence of multiple routes of resistance to what was once the first-instance treatment for CDI demonstrates how even antibiotics used to treat *C. difficile* can further contribute to pathogenesis through treatment failure and recurrence.

1.3.4 Fidaxomicin Resistance

Despite the use of fidaxomicin being limited due to cost, it displays clear benefits to CDI treatment – it prevents spore recovery in *in vitro* gut models (Chilton et al., 2016), and its narrower-spectrum of activity results in reduced rates of recurrence compared to alternative treatments (Burke and Lamont, 2014). It is worrying, therefore, that resistance has already been described (Figure 1.3). Clinical isolate Goe-91 was found to have mutations in *rpoB,* seen previously in laboratory studies (Leeds et al., 2014). Multiple other mutations in *rpoB* resulting in fidaxomicin resistance have also been described (Marchandin et al., 2023). Despite some of these isolates displaying fitness costs in terms of growth, sporulation, and toxin production, Goe-91 displayed no apparent fitness burden (Schwanbeck et al., 2019). Since fidaxomicin is already rarely used, emerging resistance casts doubts over the longevity of this CDI treatment.

1.3.5 Vancomycin Resistance

Despite being well-characterised in other species, and now predominantly used for CDI, vancomycin resistance in *C. difficile* has been poorly defined. That said, vancomycin resistance rates have increased substantially since 2012, correlating with an increased usage worldwide (Saha et al., 2019). The whole genome sequence (WGS) of *C. difficile* published in 2006 revealed a *vanG*-like cluster, proposed to confer resistance through changing the terminal D-Ala-D-Ala residues to D-Ala-D-Ser, reducing vancomycin binding affinity (Sassi et al., 2018; Sebaihia et al., 2006). Using an evolutionary approach, mutations in the two-component system *vanSR*, responsible for regulating the *vanG* operon, were shown to result in constitutive expression of *vanG* in isolates with reduced vancomycin susceptibility (Figure 1.3) (Shen et al., 2020). Further, a VanR Thr115Ala substitution, reported in clinical isolates from Florida, was also associated with reduced vancomycin susceptibility (Wickramage et al., 2023). The mechanism for this substitution-promoting resistance was proposed to involve enhancing the stability of the VanR-DNA interaction. The same mutation was also reported in multiple vancomycin-resistant clinical isolates from Texas and Kenya, suggesting widespread prevalence of this mechanism (Darkoh et al., 2021). 2021 also marked the first documentation of plasmid-mediated vancomycin resistance in *C. difficile,* through a broad-host-range and highly transferable plasmid. Plasmid pX18–498 was associated with reduced vancomycin susceptibility *in vitro,* and more severe CDI *in vivo* in mouse models (Pu et al., 2021). Despite the reductions in vancomycin susceptibility described being relatively small, from 4x the wildtype minimum inhibitory concentration (MIC), these changes potentially have significant effects in clinic – a recent study reported reduced vancomycin susceptibility led to reduced clinical cure rates and decreased sustained clinical response, showing even minor changes can have direct impacts on patients (Eubank et al., 2024).

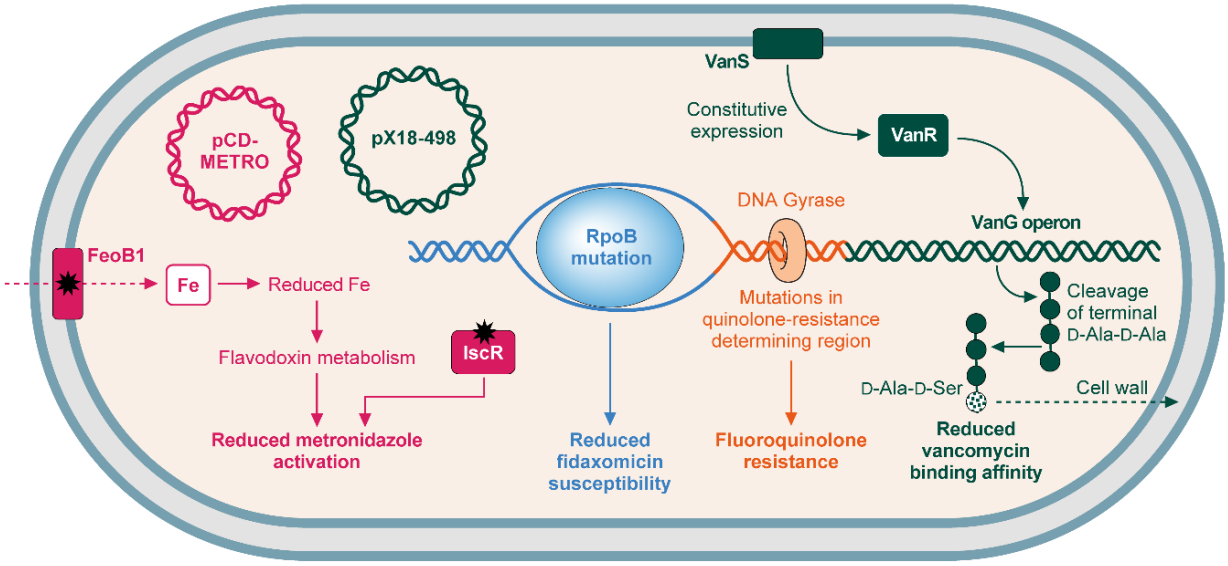


Figure 1.3 - Mechanisms of Resistance to Commonly Used Antibiotics

Mechanisms of *C. difficile* resistance to antibiotics commonly used to treat CDI (vancomycin, fidaxomicin, and metronidazole) and fluoroquinolones. (i) metronidazole (pink): resistance can be gained via the plasmid pCD-METRO. Metronidazole resistance may also be gained through mutation of either FeoB1, which reduces intracellular iron, reducing flavodoxin metabolism and metronidazole activation; or IscR, which also reduces metronidazole activation. (ii) fidaxomicin (blue): mutations in RpoB reduce fidaxomicin susceptibility. (iii) vancomycin (green): Mutations in the *vanSR* two-component system allow constitutive expression of the *vanG*-like operon, which aids vancomycin resistance through replacement of the terminal D-Ala in peptidoglycan pentapeptides with D-Ser, reducing vancomycin binding affinity. Plasmid pX18-498 has also recently been associated with resistance, although the mechanism is not understood. (iv) fluoroquinolones (orange): mutations in the genes encoding DNA gyrase, particularly *gyrA* results in fluoroquinolone resistance. Reproduced from Buddle and Fagan, 2023 under a CC-BY licence.

1.4 Experimental Evolution

Experimental evolution, defined as exploring the evolutionary dynamics and processes occurring in a population evolved under controlled conditions, is a powerful tool with a range of potential uses (McDonald, 2019). Experimental evolution may be used to study bacterial adaptation in response to a multitude of environmental stressors, to study co-evolution and competition, plasmid carriage, laboratory adaptation, or antibiotic resistance. In respect to antibiotic resistance, experimental evolution allows a highly controlled, real-time picture of the evolutionary dynamics and routes to resistance, as well as hinting at precise genes responsible, and potential mechanisms (McDonald, 2019). This principle is straightforward – genes causing a resistance phenotype are likely to be clustered around the biological pathway or event targeted by the antimicrobial. Experimental evolution is a highly valuable approach, since the alternative – isolating and studying clinical isolates – relies on waiting for resistance to emerge; and unpicking the complexities of strain differences, long-term mutation acquisition, and effects of multiple environmental stressors. Furthermore, such clinical studies tend to rely on a candidate gene approach, searching for previously identified or homologous routes to resistance. Experimental evolution, on the other hand, allows a hypothesis-free approach, increasing the potential for discovering novel resistance mechanisms. It should be noted, however, these two approaches are not mutually exclusive – experimental evolution often recapitulates mutations observed in clinical isolates. A recent investigation of known *C. difficile* resistance mutations in response to metronidazole, fidaxomicin, and vancomycin found overlap between mutations discovered by experimental evolution and those present in clinical isolate genomes, including *feoB1*-g.117delA involved in metronidazole resistance, which was present in 11 clinical isolates (Deshpande et al., 2020; Kolte and Nübel, 2024).

1.4.1 The benefits to understanding resistance

There are multiple positive outcomes arising from understanding the routes and mechanisms of resistance. Understanding how resistance evolves can inform combination therapies, which increase the lifespan of antimicrobials – the prototypical example being β-lactam combination therapy (Drawz and Bonomo, 2010). Recognising β-lactamases as a mechanism of resistance informed the development of clavulanic acid, a β-lactamase inhibitor, which increased the scope of β-lactam antimicrobials long after widespread resistance became common. From a drug development perspective, experimental evolution can inform resistance rates of a given therapeutic, and investigating mutation rates can inform which antimicrobials under development to take forward to minimise the risk of the drug becoming obsolete (Palmer and Kishony, 2013). Understanding the development of resistance in some antimicrobials can also enhance their appeal. Resistance to Avidocins, a potential *C. difficile* therapeutic, resulted in S-layer-negative strains through mutation of *slpA*, which were non-pathogenic in hamster models (Kirk et al., 2017). If the pathway to resistance results in loss of virulence, this heightens the potential of developing a potent antimicrobial with a long lifespan.

Understanding resistance mechanisms, in combination with sequencing, can also aid genomic surveillance, helping to inform on the state of resistance to a particular antimicrobial across a bacterial species (Djordjevic et al., 2024). The need for genomic surveillance, to provide a high-resolution characterisation of resistance, multidrug resistance, and transfer of resistance is ever increasing (Djordjevic et al., 2024). This would have multiple benefits – including aiding treatment decisions and enabling forecasting. This idea has recently been trialled – using a metagenomics approach to sequence hospital wastewater, evidence of numerous carbapenemase genes were detected (Markkanen et al., 2023). Although not yet mainstream, genomic surveillance holds great promise for understanding the prevalence and dissemination of resistance.

1.4.2 Methods to Study Experimental Evolution of Resistance

On a fundamental level, experimentally evolving resistance to an antibiotic involves growing the bacteria on increasing concentrations of the desired antibiotic. Broadly, there are three approaches to achieve this: gradient evolution, incremental (stepwise) exposure, and morbidostat (Jahn et al., 2017) (Figure 1.4). Gradient methods allow rapid selection of resistant mutants via exerting maximum selective pressure. Here, an antibiotic dilution gradient is inoculated, and the dilution of the highest concentration permitting growth is passaged to a new gradient of increased concentration. Increment methods benefit from using fewer resources, so can be implemented in vastly parallel experiments – samples are simply transferred to a new plate with a defined incremental increase in antimicrobial concentration. Extinction is common using this approach, so careful consideration of increment on a by-study basis is required (Jahn et al., 2017). As opposed to traditional methods, the morbidostat dynamically adjusts antibiotic concentration to the rate at which resistance evolves. This generates predictable and repeatable data in parallel populations, and has the advantage of continuous culture with growth monitoring (Toprak et al., 2012).

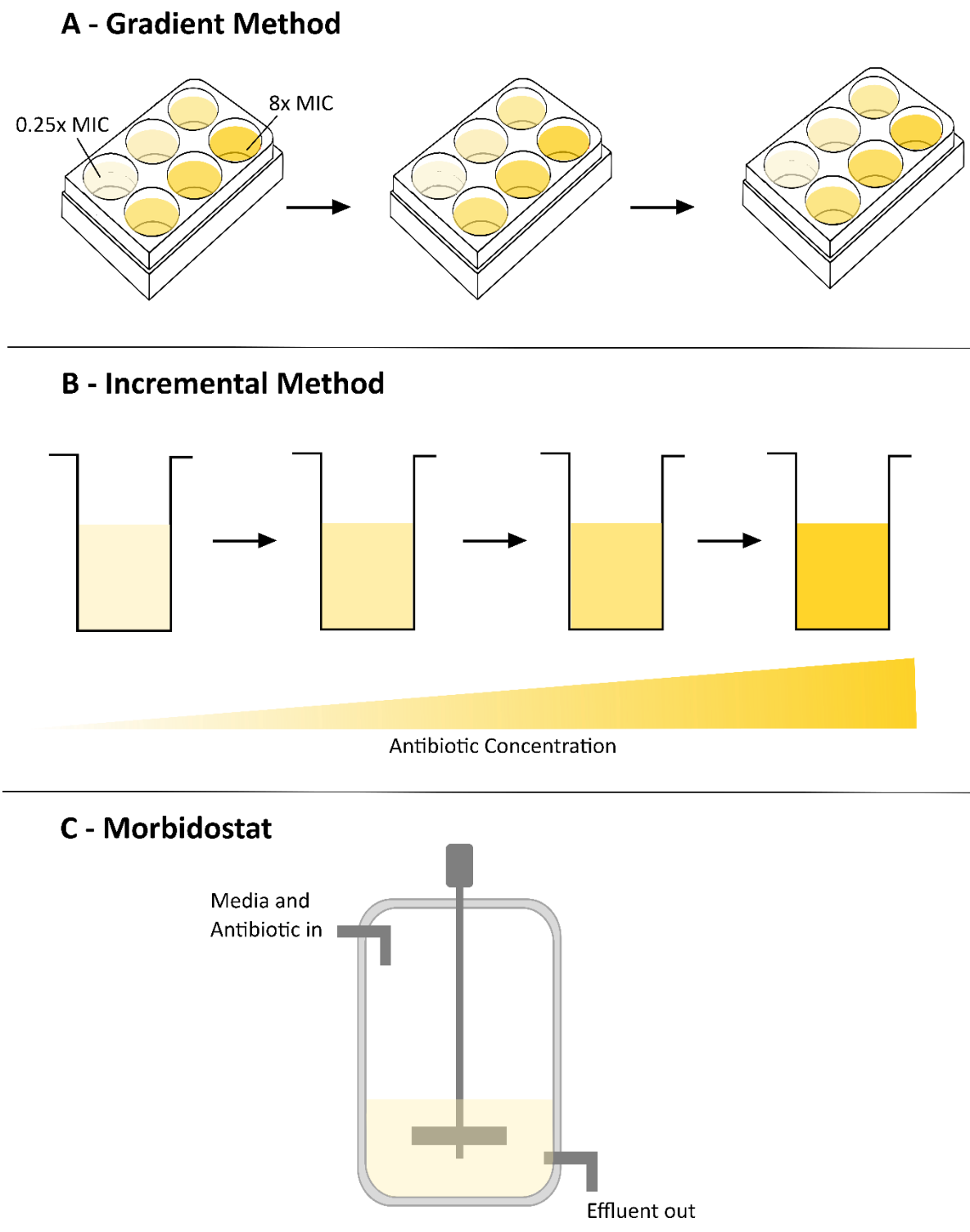


Figure 1.4 - Methods to Study Evolution

Experimental evolution may be studied through multiple methods. (A) Gradient method involves an antibiotic dilution gradient. The sample at highest antibiotic concentration permitting growth is passaged to a new gradient. (B) Incremental method – samples are transferred to new media with incremental increases in antibiotic. (C) Morbidostats allow continuous culture and dynamic adjustment of antibiotic concentration.

1.4.3 The Path to Antibiotic Resistance

1.4.3.1 Multiple Routes to Resistance

Evolution of antibiotic resistance involves a highly complex set of trade-offs, and is influenced by a multitude of external factors. Resistance may occur through acquisition of mutations, or horizontal transmission of genes into new genetic backgrounds, meaning there is often no single pathway to resistance for a given antibiotic – evolutionary trajectories differ depending on the level of resistance gained, the associated fitness cost, and external selection pressures (Santos-Lopez et al., 2019). This was demonstrated *in vivo* through *Escherichia coli* clinical isolate sequencing, as colistin resistance occurred either through plasmid-mediated gain of *mcr-1* (mobile colistin resistance 1), encoding a phosphoethanolamine transferase, or multiple modifications of the *pmrA* and *pmrB* two-component system – two distinct evolutionary pathways (Bourrel et al., 2019). Both the environment and bacterial lifestyle impact the pathway to resistance – *Acinetobacter baumannii* populations were evolved in the presence of ciprofloxacin in either planktonic or biofilm culture. The two populations exhibited vastly different resistance mechanisms, with planktonic cultures demonstrating rapid selection and fixation of few mutations; and biofilm populations being more diverse, generally selecting for increased expression of efflux pumps and showing increased fitness in absence of ciprofloxacin (Santos-Lopez et al., 2019).

1.4.3.2 Evolutionary Trade-offs and Fitness Costs

As is true for all evolution, positive changes which allow adaptation are not always clear-cut – evolving antimicrobial resistance, although beneficial in certain niches, often results in evolutionary trade-offs which impact other aspects of fitness. Since the very nature of antimicrobials is to target fundamental and essential cellular processes, it is not inconceivable that resistance can lead to disruption of such processes, or result in substantial energetic burdens. Evolutionary trade-offs have been widely observed across multiple organisms and antibiotics – ciprofloxacin resistant *E. coli* display reductions in growth rate, and aminoglycoside resistant *E. coli* result in altered ribosome

structure and cellular function (Bagel et al., 1999; Holberger and Hayes, 2009). A meta-analysis of antimicrobial resistant bacteria similarly found high fitness costs to resistance, and reduced competitive fitness compared to their respective wildtypes (Melnyk et al., 2015). Evolutionary trade-offs become especially pertinent when relating to virulence. Multiple studies have reported partial or complete attenuation of virulence following acquisition of resistance – in *Pseudomonas aeruginosa*, overexpression of the multidrug efflux pump MexEF-OprN resulted in reduced production of multiple virulence factors, including pyocyanin and rhamnolipids (Köhler et al., 2001). In *Acinetobacter baumannii*, a nosocomial pathogen notorious for persistence and biofilm production, over production of multidrug efflux pumps resulted in reduction in biofilm formation, and altered membrane composition (Yoon et al., 2015). Similar observations have been reported *in vivo* – *Streptococcus pneumoniae* β-lactam resistance was associated with virulence attenuation in infection models (Beceiro et al., 2013; Rieux et al., 2001). Ultimately, the potential for pleiotropic fitness effects to result from resistance acquisition is an important consideration to fully understand the consequences of antibiotic resistance.

1.4.3.3 Evolution of Cross-Resistance

Nevertheless, gaining resistance need not be solely detrimental. There are multiple examples of mutation accumulation altering sensitivity to other antibiotics – i.e. resistance acquisition leading to cross-resistance (Cherny et al., 2021; Gostev et al., 2023; Lázár et al., 2014). This sometimes, but not always, occurs when antibiotics have similar modes of action, and therefore shared mechanisms of resistance. The first large-scale systematic investigation of this phenomenon occurred in *E. coli*, whereby laboratory evolution to 12 antibiotics was used to establish a cross-resistance network (Lázár et al., 2014). This network evidenced the highly frequent nature of cross-resistance, worryingly even in antibiotics with vastly different modes of action. In fact, with the exception of aminoglycoside antibiotics, all tested antibiotic groups exhibited cross-resistance with multiple other groups. This cross-resistance was not always modest – increases in MIC ranged from 2 to 128-fold. Similar findings were observed in the already difficult-to-treat bacteria methicillin-resistant

Staphylococcus aureus (MRSA) – when evolved with vancomycin and daptomycin, cross resistance between these two antibiotics, as well as multiple semi-synthetic glycopeptides such as oritavancin, was observed (Gostev et al., 2023). This was potentially due to mutations in *walK* (part of the two-component system controlling cell wall metabolism) and multiple peptide resistance factor *mprF,* seen in resistant isolates for both antibiotics. Since vancomycin and daptomycin are the front-line antibiotics for treating MRSA, this cross-resistance poses an alarming threat. Modelling *in vivo* data from hospital patients shows similar trends in cross-resistance – in multiple species, such as *E. coli, Klebsiella pneumoniae* and *P. aeruginosa*, cross resistance between different antibiotics is abundant. In *E. coli*, cross-resistance between gentamicin, ciprofloxacin and trimethoprim was apparent (Cherny et al., 2021). Understanding the wider repercussions of acquired resistance to a particular antibiotic therefore has important clinical consequences.

1.4.3.4 Synergistic Contributions to Resistance

Synergistic mutations are those where the combined effect is greater than the additive effect of the individual mutations (Pérez-Pérez et al., 2009). Synergistic interactions enhance antimicrobial resistance by providing a resistance level higher than either mutation is able to achieve separately. In *C. difficile*, loss of *iscR* worked synergistically with defective *feoB1* to promote metronidazole resistance, by further reducing the metabolism of metronidazole into its active state (Deshpande et al., 2020). The presence of synergistic mutations was also hypothesised to explain why recapitulation of a *tolC* deletion did not fully increase resistance to the same degree as its evolved counterpart in *Flavobacterium johnsoniae* (Chodkowski and Shade, 2023). Alternatively, synergistic mutations may enhance antimicrobial resistance by allowing a specific mutation, which does not provide resistance alone, to have resistance potential. In *Salmonella,* substitutions in the AcrB efflux pump only permitted resistance in the presence of synergistic mutations in either *ramR* or *envZ* (Trampari et al., 2023).

1.4.4 Studying Resistance Mechanisms Through Experimental Evolution

Experimental evolution, coupled with WGS and downstream genetic validation, is a robust approach to identify both the mode of action and resistance mechanisms to a given antibiotic. This approach has been used countless times across multiple organisms and for multiple antibiotics, especially in the post-genomic era (Hartkoorn et al., 2012; Palmer et al., 2011). For example, experimental evolution was used to identify the target of pyridomycin in *Mycobacterium tuberculosis*, as a single non-synonymous mutation in *inhA* was pinpointed in evolved strains (Hartkoorn et al., 2012). Deletion of *inhA* resulted in increased pyridomycin sensitivity, and overexpression increased the MIC 15-fold. InhA was found to be involved in mycolic acid synthesis, and pyridomycin-mediated inhibition of this protein was determined as the mode of action of this antimicrobial. Serial passaging and WGS also allows for novel resistance pathway identification – in *Enterococcus faecalis,* multiple pathways to daptomycin resistance were reported, including alterations to cardiolipin synthase genes, a mechanism distinct from those described in *S. aureus* (Palmer et al., 2011). Validation of identified candidates also enables understanding of the molecular mechanisms of resistance. For example, in vancomycin resistant *S. aureus*, the *walK* histidine kinase was implicated in reduced expression of cell wall metabolism genes, and thickened cell walls (Hu et al., 2015).

A more reserved approach to utilising experimental evolution can be seen in the *C. difficile* field, perhaps reflecting the constraints in tools available prior to the last two decades. Early *C. difficile* evolutionary studies identified mutations in strains evolved in the presence of vancomycin (*rpoC, sdaB*) and fidaxomicin (*rpoB*) respectively (Leeds et al., 2014). This small-scale evolution used a single population, as opposed to multiple parallel lines, and provided no additional genetic characterisation of the mutations or pathway. Nevertheless, this was further supported by a more recent characterisation of *C. difficile* clinical isolates with reduced fidaxomicin susceptibility. Single-molecule real-time sequencing revealed a mutation in *rpoB*, and introducing this mutation into *C.*

difficile 630 resulted in reduced fidaxomicin susceptibility, at the cost of severe defects in toxin production, spore formation and growth (Schwanbeck et al., 2019). Since this fitness cost was not observed in the original clinical isolate, this neatly demonstrates both the evolutionary trade-offs between resistance and fitness, and the possibility of *in vivo* compensatory mechanisms occurring during clinical selection.

A slightly altered approach to experimental evolution, using laboratory-derived hypermutators, was used to characterise the accumulation of mutations leading to metronidazole resistance in *C. difficile* (Deshpande et al., 2020). The DNA repair operon *mutSL* was deleted in *C. difficile* ATCC 700057 – a non-pathogenic, toxin-negative strain – to allow accelerated evolution. By sequencing the genomes of three endpoint colonies, and focussing on proteins involved in redox and iron homeostasis previously reported to be involved in metronidazole resistance, mutations in four genes were identified: *PFOR, feoBI, xdh* and *iscR*. The order of occurrence of these mutations was further investigated by sequencing colonies from mid-evolution time points, demonstrating a deterministic path of mutation accumulation. This time point-driven insight into the evolution of resistance in *C. difficile* provides a novel perspective on the pathways to low- and high-level resistance. However, evolving resistance in a non-pathogenic strain limits the clinical relevance of this work – a fact acknowledged by the authors themselves when commenting on low likelihood of *feoBI* being involved in resistance *in vivo*, due to its essentiality for colonisation and virulence. Additionally, the trend in evolutionary studies to sequence multiple isolates from the same population, to understand parallel evolution and resistance pathways, limits the potential to identify truly alternative routes to resistance. A caveat must also be issued for the hypothesis-driven approach to sequence analysis used here, since narrowing down genes of interest to those previously associated with metronidazole resistance limits discoveries to known genes. Experimental evolution, both in *C.*

difficile and beyond, would benefit from a hypothesis-free approach to pave the way to novel discoveries.

The genetic basis for vancomycin resistance is largely uncharacterised in *C. difficile*. Recently, evolution of first-step resistance in clinically-relevant ribotype 027 strains, coupled with WGS, identified mutations in the *vanSR* two-component system of the *vanG*-like cluster, which were supported by identical mutations from clinical isolates (Shen et al., 2020). These mutations resulted in constitutive expression of *vanG*, conferring reduced susceptibility, presumably via modification of the peptidoglycan from D-Ala D-Ala to D-Ala D-Ser, resulting in reduced vancomycin binding. Although a fantastic example of the virtues of experimental evolution for elucidating resistance mechanisms, this really only scratches the surface regarding the possibilities for evolutionary biology. This work captures first-step resistance, early mutations conferring large fitness benefits, leaving the potential of mutation accumulation, and routes to high-level resistance uncharacterised. Again, the choice to evolve a single population, with no parallel control, limits the scope to understand alternative routes to vancomycin resistance and the evolutionary dynamics at play. Additionally, the fitness costs of constitutive expression of the *van* operon have yet to be explored. In *Enterococci*, the high fitness cost of the *van* operon is well documented, and inducible expression of these genes help to alleviate this burden (Foucault et al., 2010). Further investigation into the costs of constitutive expression of the *C. difficile van* operon therefore merits further work.

1.5 Concluding Remarks

With antimicrobial resistance on the rise, and the array of available treatments diminishing, focus is needed to understand the clinically important pathogen *C. difficile.* Rapid technological advancements have increased the capacity, ease, and scale of microbial genomics, allowing

researchers to push the boundaries of experimental evolution to enable new insights into the evolutionary dynamics of antimicrobial resistance. Despite such victories, there are still many research gaps to address in order to fully comprehend resistance in *C. difficile*. Most urgently, an in-depth characterisation of resistance to vancomycin, the current front-line antibiotic, in clinically relevant *C. difficile* strains is necessary; through a large-scale, highly parallel experimental evolution approach. This would allow insights into the population dynamics involved in the evolution of resistance, as well as potential alternative routes to achieving resistance. Understanding the contributors to high-level resistance, and the presence of synergistic interactions would be of particular clinical relevance, along with characterisation of cross-resistance. Since many examples of antibiotic resistance are associated with fitness costs, a fuller appreciation for the fitness costs associated with vancomycin resistance in *C. difficile* would be useful. Overarchingly, completing such a study in a hypothesis-free manner is paramount for embracing novel discoveries.

1.6 Aims

The *C. difficile* genome contains a complete *van* cluster, the constitutive expression of which results in reduced vancomycin susceptibility (Shen et al., 2020). Beyond these regulatory changes, the molecular mechanisms underpinning the evolution of vancomycin resistance in *C. difficile* remain unknown. Since convergent evolution, whereby an organism evolves the same solution via a genetically different route, is a common hallmark of antibiotic resistance, a complete understanding of the routes to vancomycin resistance in *C. difficile* is lacking (Keshri et al., 2019). As such, it is unclear whether aspects other than altered regulation of the *van* operon, or indeed other loci entirely, contribute to the increasing vancomycin resistance observed clinically. Moreover, although pleiotropic fitness costs to acquiring resistance have been widely described in other organisms, the phenotypic effects of vancomycin resistance acquisition in *C. difficile* have not been characterised. What's more, the traditional, hypothesis-driven approach to experimental evolution for investigating

antibiotic resistance involves narrowing searches to well-characterised genes, including those homologous to resistance genes in other organisms. This severely limits the scope for identification of novel resistance pathways.

The primary aim of this project was to use highly parallel experimental evolution to investigate vancomycin resistance in *C. difficile.* A robust approach to sequencing both evolved isolates, and whole populations, could then be used to assess the mutations involved in vancomycin resistance acquisition on both an individual and population level. This would allow insight into population dynamics, and alternative routes to resistance. Secondly, this project aimed to provide a thorough phenotypic assessment of evolved isolates, to understand whether vancomycin resistance is associated with fitness burdens in *C. difficile,* and if these burdens vary depending on the route to resistance. The final key aim of this project was to ascertain the molecular mechanisms of vancomycin resistance in *C. difficile*.

2 Materials and Methods

All methods relevant to this thesis are described in this chapter, including those used across multiple chapters, and those developed and optimised within this project.

2.1 Bacterial strains and Growth conditions

Bacterial strains made and used are described in Appendix I. The media routinely used throughout this work is described in Table 2.1, and antibiotics used in this work are described in Table 2.2.

All *C. difficile* media was pre-reduced in the anaerobic cabinet (2 h per 5 mL liquid culture, 30 min for agar plates). *C. difficile* strains were cultured in tryptone yeast (TY) broth or on pre-reduced brain heart infusion (BHI) agar, unless otherwise stated. *C. difficile* was grown at 37°C in an anaerobic cabinet (Don Whitley Scientific), with an atmosphere composed of 80% N_2, 10% CO_2 and 10% H_2. Cultures were supplemented with antibiotics as appropriate.

E. coli cultures were grown in Luria-Bertani (LB) broth (37°C, 200 rpm shaking), or on LB agar at 37°C. Cultures were supplemented with antibiotics as appropriate.

Bacterial strains were stored in 20% glycerol at -80°C.

Table 2.1 - Media Used in This Study

Media used in this work, components per 400 ml and manufacturer.

Media Name	Components per 400ml	Manufacturer	Notes
BHI agar	20.8 g BHI agar powder	Thermo Scientific	Used for *C. difficile* growth
BHIS+Taurocholate	20.8 g BHI agar powder 2 g Yeast Extract 0.4 g L-Cysteine, non-animal source 0.4 g Taurocholate	Thermo Scientific Bacto Sigma Sigma	Sporulation efficiency determination
TY broth	12 g Bacto Tryptose 8 g Yeast Extract	Bacto Bacto	Used for overnight growth of *C. difficile*
LB Agar	14 g LB agar powder	Thermo Scientific	Used for *E. coli* growth
LB Broth	8 g LB Broth	Thermo Scientific	Used for overnight growth of *E. coli*
CDMM	See Below*	See Below	For mutagenesis using *codA* allele exchange vector
SOC	NEB Super Optimal broth with Catabolic repression (SOC) (500 µL per 25 µL transformation)	NEB	For increased *E. coli* transformation efficiency using NEB5α

**C. difficile* minimal medium (CDMM) composition. Components in the below table were added to 15 g/L agar

Component	Stock concentration (mg/mL)	Final concentration (mg/mL)
Amino acids (5x)		
Casamino Acids	50	10
L-Tryptophan	2.5	0.5
L-Cysteine	2.5	0.5
Salts (10x)		
Na_2HPO_4	50	5
$NaHCO_3$	50	5
KH_2PO_4	9	0.9
NaCl	9	0.9
Glucose (20x)		
D-Glucose	200	10
Trace Salts (50x)		
$(NH_4)_2SO_4$	2	0.04
$CaCl_2\_2H_2O$	1.3	0.026
$MgCl_2\_6H_2O$	1	0.02
$MnCl_2\_4H_2O$	0.5	0.01
$CoCl_2\_6H_2O$	0.05	0.001

Iron (100x)		
$FeSO_4$ _$7H_2O$	0.4	0.004
Vitamins (1000x)		
D-Biotin	1	0.001
Calcium-D-pantothenate	1	0.001
Pyridoxine	1	0.001
5-fluorocytosine	10	0.05

Table 2.2 - Antibiotics Used in This Study

Antibiotics used to supplement broth/agar (stock and working concentrations).

Antibiotic	**Stock Concentration**	**Working Concentration**	**Solvent**	**Brand**
Chloramphenicol	30 mg/mL	15 µg/mL	70% ethanol	Acros Organics
Thiamphenicol	15 mg/mL	15 µg/mL	100% Methanol	Sigma
Colistin	50 mg/mL	50 µg/mL	H_2O	Sigma
Kanamycin	50 mg/mL	50 µg/mL	H_2O	Sigma
Anhydrotetracycline (aTc)	2 mg/mL	60 ng/mL	100% ethanol	Sigma
Vancomycin	100 mg/mL	N/A	H_2O	Sigma
Rifampicin	32 µg/mL	N/A	Methanol	Sigma

2.2 DNA Manipulation

2.2.1 Genomic DNA Isolation

High quality genomic DNA (gDNA) was obtained from overnight cultures of *C. difficile* through a modified version of the phenol-chloroform method as described previously (Wren and Tabaqchali, 1987). To lyse the sample, 1.5 mL *C. difficile* overnight culture was harvested via centrifugation (2 min at 12000 x g) and resuspended in 200 µL lysis buffer (200 mM NaCl, 50 mM ethylenediaminetetraacetic acid (EDTA), 20 mM Tris-HCl pH 8.0) with 10 µL of purified bacteriophage phiCD27 endolysin (CD27L catalytic domain), and incubated for 1 hour at 37°C. The sample was then incubated for 1 hour at 55°C, following the addition of 10 µL pronase (final concentration 1 mg/mL). 80 µL (final concentration 2%) of N-lauroylsarcosine was then added, and

the sample was incubated for a further 1 hour at 37°C. Finally, 200 µL RNase (final concentration 0.2 mg/mL) was added, before incubation for 1 hour at 37°C.

For gDNA extraction, 700 µL of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the sample and mixed by gentle inversion, before centrifugation in a heavy phase lock gel (PLG) tube at 13,000 x g for 2 min. This facilitated easy separation of the organic and DNA-containing aqueous phases. The aqueous phase was recovered, mixed with phenol:chloroform:isoamyl alcohol, and centrifuged in a second PLG tube as before. This process was repeated a further 2 times using 500 µL chloroform:isoamyl alcohol (24:1) to remove the phenol.

Precipitation of the sample commenced with the addition of 500 µL of ice-cold isopropanol to the recovered aqueous phase, followed by incubation at -20°C overnight. The gDNA was harvested via centrifugation (15 min at 4,000 x g at 4°C), before washing in 500 µL 70% ethanol and harvesting again (10 min, 4,000 x g). The pellet was air-dried to remove residual ethanol, and resuspended in 50 µL of either nuclease free water (NEB) or elution buffer (EB) (Thermo Scientific).

The quality of the gDNA was determined through agarose gel electrophoresis, the quantity was measured via Qubit, and the purity was determined via A_{260} spectrophotometry.

Crude gDNA extraction, for use in screening PCRs, was performed using Chelex 100 resin (Sigma). Single colonies of *C. difficile* were resuspended in 100 µL of nuclease free water (NEB) containing a small amount of Chelex. Samples were incubated at 100°C for 10 min and briefly centrifuged. The supernatant was used in PCR reactions (2.5 µL of DNA per 20 µL PCR reaction). Chelex-based gDNA extraction was performed on the day for PCR.

2.2.2 Polymerase Chain Reaction (PCR)

PCR was used to amplify sections of genomic or plasmid DNA. PCR was performed using a T100 thermocycler (Bio-Rad), and using primers synthesised by Eurofins.

Phusion high-fidelity polymerase mastermix (NEB) was used in instances where high fidelity was necessary. A 20 µL reaction comprised 10 µL Phusion 2x mastermix, 100 ng gDNA (or 1 ng plasmid DNA) template, 10 µM forward and reverse primers, and nuclease free water (NEB). Dimethyl sulfoxide (DMSO) was added at 5 to 10% (v/v) for primer lengths over 35 bp to help with DNA melting. The reaction was performed using an initial denaturation of 98°C for 30 s, followed by 35 cycles of 98°C denaturation for 30 s, 50-65°C annealing for 30 s, 72°C extension for 15-30 s per kb DNA, and a final extension of 72°C for 5 min, to complete partial products.

Taq red polymerase (PCR biosystems) was used for colony PCR screening, where fidelity was less important. A 20 µL reaction contained 10 µL Taq red mastermix, 10 µM forward and reverse primers, 8 µL nuclease free water (NEB) and 1 colony. The reaction was performed as follows: initial denaturation (and *E. coli* cell lysis) at 95°C for 3 min, 35 cycles of 95°C denaturation for 30 s, 58°C annealing for 30 s, 72°C extension for 1 min per kb, and a final extension at 72°C for 5 min.

2.2.3 Agarose Gel Electrophoresis

Agarose (VWR) was made to 0.8-3% (w/v) in Tris-acetate-EDTA buffer (Thermo Scientific) and dissolved by microwaving. 0.8% (w/v) was used as standard, however higher percentage gels were used to resolve smaller DNA fragments. 55°C molten agarose was cast with SybrSafe (1/10,000 final volume, Invitrogen). The set gel was submerged in 1 x Tris-acetate-EDTA buffer (Thermo Scientific) in the electrophoresis tank, and 5 µL of 1 kb plus DNA ladder (NEB) was added to the first well. Purple loading dye (NEB) or UView (Bio-Rad, for gel extraction only) was added to the samples prior to

loading. Electrophoresis was generally performed at 110 V for 30 min, with minor alterations dependant on gel composition. Gels were imaged using a ChemiDoc MP imager (Bio-Rad).

2.2.4 PCR Purification

PCR fragments were purified using the GeneJET PCR Purification kit (Thermo Scientific), as per the manufacturer's instructions. DNA was eluted in 20 µL of nuclease free water (NEB), and quantified using A_{260} spectrophotometry.

2.2.5 Gel Extraction of DNA

Samples were resolved as above (2.2.3), and visualised using an ultraviolet transilluminator. Bands were excised with a scalpel and placed in a microfuge tube with 1:1 concentration of binding buffer and 2.5 µL of 3 M sodium acetate. DNA was extracted from the gel using the GeneJET Gel Extraction kit (Thermo Scientific), as per the manufacturer's instructions. The DNA was eluted using 20 µL of 50°C nuclease free water (NEB), and quantified using A_{260} spectrophotometry.

2.2.6 Restriction Endonuclease Digestion of DNA

Digestion of DNA using restriction endonucleases (NEB) was performed as per the manufacturer's instructions. A 20 µL reaction was assembled using 1-2 µg DNA, 2 µL of the appropriate 10x buffer (NEB), and 1 µL of enzyme. The reaction was incubated at the enzyme's optimum temperature (usually 37°C) for 1 h, unless otherwise specified.

For PCR screening by digest, which was used in identifying SNP mutants, digestion of a large number of samples was performed directly in the PCR tube. Here, 1 µL of enzyme was added to the 20 µL

PCR sample after the PCR had finished and cooled. Reactions were incubated at the enzyme's optimum temperature for 1 h.

To remove DNA template from completed Phusion PCR reactions, 2 µL of DpnI (NEB) was added to the 20 µL PCR sample, after the PCR had finished and cooled, and incubated at 37°C for 2 h. The sample was then purified as above (2.2.5).

2.2.7 Ligation of DNA

DNA fragments were ligated using T4 DNA ligase (NEB), and were assembled as follows: 1 µL T4 ligase buffer, 0.5 µL T4 ligase, 25 ng vector DNA, insert DNA at a 1:3 molar ratio vector:insert, and nuclease free water (NEB) up to a final volume of 10 µL. Reactions were incubated for 1 h at room temperature, before being used for *E. coli* transformations (2.2.10).

2.2.8 Gibson Assembly of DNA Fragments

Gibson assembly was used to ligate multiple overlapping fragments of DNA in a single step (Gibson et al., 2009). NEBBuilder was used to design primers to amplify fragments with a 30 bp overlap to the adjacent fragment.

The vector DNA was linearised by digestion (2.2.6) and purified by gel extraction (2.2.5). Insert fragments were amplified by PCR (2.2.2) and purified (2.2.4). 3-fragment Gibson assembly was performed in a final volume of 20 µL, composed as follows: 10 µL of 2x isothermal assembly mastermix (Gibson et al., 2009), 50 ng of linearised vector, insert fragments at a 1:2:2 molar ratio of vector:insert:insert, and nuclease free water (NEB) up to 20 µL. Reactions were incubated at 50°C for 4 h, before being used for *E. coli* transformations (2.2.10).

2.2.9 Production of Chemically Competent *E. coli*

Overnight cultures of *E. coli* CA434 were sub-cultured (1:100) and grown to exponential phase (optical density $(OD)_{600nm}$ 0.4 - 0.6). Cells were harvested by centrifugation (10 min, 4,000 x g, 4°C) and resuspended in 5 mL ice-cold $CaCl_2$ (100 mM), before being incubated on ice for 15 min. Cells were harvested as before, and resuspended in 1 mL of a solution containing 100 mM $CaCl_2$ and 15% (v/v) glycerol. Cells were incubated for 2 h on ice, before being aliquoted (50 µL), snap frozen in liquid nitrogen, and stored at -80°C.

2.2.10 Transformation of *E. coli*

NEB5α competent *E. coli* cells (NEB) were used for cloning and plasmid propagation, and CA434 were used as the conjugation donor, to transfer plasmids into *C. difficile*. Both *E. coli* strains were transformed using the same heat shock method, except using 2 µL of ligation product or Gibson assembly (NEB5α), or 0.5 µL plasmid miniprep (CA434).

A 50 µL aliquot of competent cells was thawed on ice for 10 min, and split into 25 µL volumes in pre-chilled microfuge tubes. DNA was added as above, and the samples were incubated on ice for 30 min. Samples were heat shocked at 42°C for 30 s, and incubated on ice for 2 min, before the addition of 500 µL SOC media (NEB). The samples were incubated with shaking at 37°C for 1 h, before 100 µL was spread on LB agar plates supplemented with appropriate antibiotics.

2.2.11 Plasmid DNA Isolation

5 mL of overnight *E. coli* culture was harvested via centrifugation (10 min, 4000 x g). Plasmid DNA was extracted via the GeneJET Plasmid Miniprep kit (Thermo Scientific), through selective binding to the silica spin column, according to the manufacturer's instructions. DNA was eluted in 50 µL

nuclease free water (NEB), and quantified using the A_{260} spectrophotometry. Plasmids used in this project can be found in Appendix II.

2.2.12 Sanger Sequencing of DNA

Sanger sequencing of PCR fragments and plasmid DNA was carried out by Genewiz. Sequencing alignments and analyses were performed using Geneious (Biomatters).

2.2.13 Conjugative Transfer of Plasmid DNA into *C. difficile*

Transfer of plasmid DNA into *C. difficile* was performed as previously described (Kirk and Fagan, 2016). In brief, 200 µL of overnight *C. difficile* culture was heat-treated at 50°C for 15 min. 1 ml of overnight *E. coli* CA434 culture was harvested via centrifugation (3000 x g, 2 min), and resuspended using the heat-treated *C. difficile*. The cell mixture was spotted onto BHI agar, and incubated in the anaerobic cabinet for 24 h. Growth was subsequently harvested from the plate using 1 mL TY broth. The harvested material (both neat and diluted 1:5) was spread on a series of BHI plates supplemented with thiamphenicol and colistin. Thiamphenicol was used for plasmid selection, whilst colistin was used to kill *E. coli.* Single transconjugant colonies were restreaked to purity on the above selective plates, before being frozen at -80°C.

2.2.14 Mutagenesis using Allele Exchange in *C. difficile*

Allele-exchange mutagenesis, utilising the *codA* heterologous counterselection system, was performed as described previously (Cartman et al., 2012). Vector insert fragments were either synthesised by Genewiz, or manually generated by PCR (2.2.2). In the latter case, primers designed for Gibson assembly (2.2.8) were used to amplify 1200 bp regions of the genome upstream and downstream of the region of interest (homology arms). Amplified fragments were purified as above

(2.2.4). The vector backbone was digested (2.2.6) and purified (2.2.5). The new plasmid, comprising the vector and homology arm inserts, was assembled via Gibson assembly (2.2.8). Alternatively, for synthesis fragments, the vector backbone was digested (2.2.6) and purified (2.2.5), and the new plasmid was assembled via ligation (2.2.7). In either case, the resultant plasmid was transformed (2.2.10) into *E. coli* NEB5α, sequenced (2.2.12), and transformed (2.2.10) into *E. coli* CA434, before being conjugated into *C. difficile* (2.2.13).

Single recombination events, whereby the plasmid had integrated into the *C. difficile* genome, were visualised as larger colonies on selective BHI agar, since the growth rate of such colonies was faster, as plasmid replication rate was no longer limiting. Putative single crossovers were confirmed via PCR using combinations of primers situated within the vector and flanking the homology arms. Single recombinants were restreaked onto non-selective BHI agar, and incubated for 2-3 days, to allow a second recombination event to occur, and for the plasmid to be lost. Growth from the non-selective BHI plates was harvested using 1 mL of sterile phosphate-buffered saline (PBS), serially diluted, and plated onto CDMM with 5-fluorocytosine. In the presence of *codA*, 5-fluorocytosine is lethal, thus selecting against plasmid carriage. After 48 h, colonies were screened by PCR using primers flanking the homology arms. Positive colonies were restreaked to purity and tested for growth on BHI supplemented with thiamphenicol to ensure the plasmid had been lost. Mutants were confirmed with Sanger sequencing and frozen in 80% glycerol at -80°C.

2.2.15 Overexpression in *C. difficile*

Overexpression was achieved using the tet-inducible expression system, as described previously (Fagan and Fairweather, 2011). The gene of interest was amplified from the genome using PCR (2.2.2), digested (2.2.6), and purified as above (2.2.4). The vector backbone was digested (2.2.6) and purified (2.2.5), and the new plasmid was assembled via ligation (2.2.7). The new plasmid was then

transformed (2.2.10) into *E. coli* NEB5α, sequenced (2.2.12), and transformed (2.2.10) into *E. coli* CA434, before being conjugated into *C. difficile* (2.2.13).

2.2.16 Bioinformatics Relating to Mutagenesis

Geneious software v7.1.9 (Biomatters) was routinely used for primer design, *in silico* digestion, and Sanger sequencing alignments. Primers for Gibson assembly were designed using NEBuilder. Compatible endonuclease enzymes for use in diagnostic restriction digests were determined using NEBcutter.

2.3 Evolution Methods

2.3.1 Directed Evolution of *C. difficile*

Directed evolution of *C. difficile* was performed using a broth-based gradient approach, whereby 10 individually barcoded parallel lines were evolved for a period of 30 passages (60 days). For each passage, a 6-well plate was assembled for each parallel line, containing 4 mL of TY broth with a gradient spanning 0.25 to 8x the current vancomycin MIC (determined from the previous passage, or from the ancestral MIC, in the case of the first passage). This allowed the gradient to rise with increasing levels of vancomycin resistance.

Overnight *C. difficile* cultures, derived from single colonies, were adjusted to OD_{600nm} 1.0. The evolution was initiated by adding 10 µL of adjusted culture to each well of the assembled 6-well plates, before incubating for 48 h at 37°C. Plates were visually inspected after 48 h, and fresh plates were assembled as above. 10 µL of the well with the highest antibiotic concentration supporting growth was used to inoculate the wells of the subsequent passage. For each parallel line, a control well was passaged without antibiotic. 1 mL of each parallel population, and the corresponding

control, was frozen at -80°C in 15% glycerol whenever the MIC increased; and after passages 10, 20 and 30.

2.3.2 Genome Sequencing of *C. difficile*

For short-read sequencing of *C. difficile* isolates, library prep (Nextera XT Library Prep Kit (Illumina, San Diego, USA)) and 30x illumina sequencing (NovaSeq 6000, 250 bp paired end protocol) was performed at MicrobesNG (Birmingham, UK). Reads were trimmed at MicrobesNG using Trimmomatic (v0.30) with a sliding window quality cut-off of Q15.

For long-read sequencing of *C. difficile* isolates, library prep (Oxford Nanopore Technologies SQK-RBK114.96 kit (ONT, UK)) and nanopore sequencing (GridION, FLO-MIN114 (R.10.4.1) flow cell) was performed at MicrobesNG (Birmingham, UK).

For sequencing of bacterial populations, library prep (Nextera DNA Flex Library Prep Kit (Illumina, San Diego, USA)) and 250x Illumina sequencing (NovaSeq 6000, 150bp paired end protocol) was performed at SNPsaurus (Oregon, USA). Reads were trimmed using Trimmomatic (v0.39) with the following criteria: leading:3; trailing:3; slidingwindow:4:15; minlen:36.

2.3.3 Genome Sequencing Analysis of *C. difficile* Isolates

To ensure sufficient quality for analysis, trimmed reads were checked using FastQC (v0.11.9) (Simon Andrews, 2019).

A custom script, inspired by a previously described mutant analysis pipeline (Wright et al., 2019), was used to analyse isolates. First, reads were aligned to the *C. difficile* reference (R20291, accession number: FN545816) using BWA-mem (v0.7.17) (Li and Durbin, 2009) and sorted using SAMtools

(v1.43) (Li et al., 2009). Coverage across the genome was inferred using Bedtools (v2.30.0) (Quinlan and Hall, 2010) genomecov and map functions. PCR duplicates were removed via Picard (v2.25.2) (http://broadinstitute.github.io/picard/). SAMtools (v1.43) mpileup was used to generate the required mpileup file for Varscan. Variants were then called using Varscan (v2.4.3-1) (Koboldt et al., 2009) mpileup2cns (calling SNPs, insertions and deletions) using the following parameters: min-coverage 4; min-reads2 4; min-var-freq 0.80; p-value 0.05; variants 1; output-vcf 1. This required a minimum of 4 reads, and a minimum frequency of 80%, to support a variant. Variant call format (Vcf) files were annotated using snpEff (v5.0) (Cingolani et al., 2012).

The Breseq (v0.35.5) (Deatherage and Barrick, 2014) pipeline was also used to call variants using default parameters, and putative variants were retained if detected in both analysis pipelines. Variants identified were manually verified using the integrative genomics viewer (IGV) (v2.8.6) (Robinson et al., 2011).

2.3.4 Closing Isolate Genomes Using Nanopore

Genome assembly of long-read sequencing data was performed at MicrobesNG (Birmingham, UK): reads were assembled using Flye (v2.9.2-b1786) (Kolmogorov et al., 2019), polished using Medaka (v1.8.0) (ONT, 2024), and annotated using Bakta (v1.8.1) (Schwengers et al., 2021).

For a given isolate, Nanopore and Illumina data were amalgamated using Polypolish (v0.6.0) (Wick and Holt, 2022), to generate complete closed genomes and to reduce erroneous base calls. Short read sequences were first aligned to the Nanopore assembly using BWA-mem (v0.7.17) (Li and Durbin, 2009), using the -a parameter. Alignments were filtered by insert size using Polypolish, with default parameters, to remove spurious repeat alignments. The Nanopore assembly was then polished with the Illumina data using Polypolish, generating a polished fasta. This was annotated

with Prokka (v1.14.6) (Seemann, 2014). Variants were then called using Snippy (v4.6.0) (Seemann, 2015) with default parameters. Identified variants were manually verified using IGV (v2.8.6) (Robinson et al., 2011).

2.3.5 Genome Sequencing Analysis of *C. difficile* Populations

C. difficile populations were analysed using a pipeline that followed the same custom script as for isolates (2.3.3), with modifications to SNP calling parameters and filtering. Realistic sequencing data was simulated using InSilicoSeq (v1.5.4) (Gourlé et al., 2019) at multiple coverage depths (80x, 100x, 300x), with SNPs seeded at 5% frequency. Simulated sequences were analysed using the custom pipeline, and a range of Varscan parameters (P-values, min-coverage, min-reads2) were trialled to generate a set of parameters to accurately call variants without false positives.

Different varscan parameters were used depending on the coverage depth of a particular population, to allow calling of low frequency variants with a strong evidence base. For samples with an average coverage across the genome of 100x or higher, the following Varscan (v2.4.3-1) mpileup2cns parameters were used: min-coverage 80; min-var-freq 0.05; p-value 0.05; variants 1; output-vcf 1. A minimum of 4 reads were required to support a variant, and the cut-off for variant calling was 5%. Samples with an average coverage across the genome of below 100x used the following Varscan (v2.4.3-1) mpileup2cns parameters: min-coverage 4; min-reads2 4; min-var-freq 0.05; p-value 0.05; variants 1; output-vcf 1. Here, a minimum of 4 reads were required to support a variant (as with high coverage samples), meaning the minimum variant frequency possible to call was inversely scaled (4/coverage at position). Variants were filtered using Varscan (v2.4.3-1) compare to discard variants that never reached above 10% frequency by the end of the evolution. Variants were also manually inspected to ensure no variants remained the same frequency across all time points, indicating a false call.

Mutations occurring within genes were visualised in the Kyoto encyclopedia of genes and genomes (KEGG) colour mapper (Kanehisa et al., 2022) to view KEGG pathways.

2.3.6 Principal Component Analysis (PCA)

Multiple results can be visualised in a single graph using principal component analysis. PCAs were computed using the prcomp() function in Base R (http://www.rproject.org/), and visualised using the factoextra package.

2.4 Phenotypic Assays

2.4.1 Minimum Inhibitory Concentration (MIC)

MICs were performed using standard agar dilution methods (CLSI, 2024). Overnight cultures of *C. difficile* were adjusted to OD_{600nm} 0.1. 2.5 µL of sample was spotted onto BHI plates with ranging antibiotic concentrations, in biological triplicate and technical duplicate. Plates were incubated for 48 h before the MIC was determined. Plates were then imaged using a Scan 4000 colony counter (Interscience). In the case of overexpression strains, the appropriate inducing agent, and antibiotics to select for the plasmid, were included.

2.4.2 Growth Analysis

Growth curves were performed anaerobically in 96 well plates using a Stratus microplate reader (Cerillo). Overnight cultures of *C. difficile* were adjusted to OD_{600nm} 0.05 and incubated at 37°C for 1 h to equilibrate the cultures. The 96 well plate lid was treated with 0.05% Triton X-100 + 20% ethanol to prevent condensation. 200 µL of the equilibrated samples were used to assemble the 96 well plate, and measurements were taken at minimum in biological and technical triplicate. The OD_{600nm}

was taken every 3 minutes over a 24 h period. Values for the blank (media only) wells were removed from all well measurements, and the data was plotted in Graphpad Prism (v9.0.2), before being analysed in RStudio (v4.1.0) using the GrowthCurver package (v0.3.0) (Sprouffske and Wagner, 2016).

2.4.3 Sporulation Efficiency Determination

Sporulation efficiency was determined based on methods previously described (Dembek et al., 2015). Briefly, *C. difficile* overnight cultures were subcultured in TY to OD_{600nm} 0.01 and grown for 8 h. Cultures were adjusted again to OD_{600nm} 0.01, subcultured 1:100 immediately into 10 mL TY broth, and grown overnight to obtain spore-free cultures in early stationary phase (T = 0). At T = 0, and the following 5 days, total viable counts were enumerated by spotting 10-fold serial dilutions, made in PBS, in technical triplicate onto BHIS agar with 0.1% sodium taurocholate. Colonies were counted after 24 h incubation. Spore counts were enumerated using the above method following heat treatment (65°C for 30 min), to kill vegetative cells in the sample.

2.4.4 Mutation Rate Assay

The mutation rate of mutants compared to their respective wildtypes was determined using rifampicin. 10 mL *C. difficile* overnight cultures were concentrated 10-fold. 200 µL of concentrated cells were spread onto BHI plates supplemented with 0.015 µg/mL rifampicin (8x MIC). The remaining culture was used to determine total viable counts by spotting 10-fold serial dilutions, made in PBS, in technical triplicate onto non-selective BHI plates. After 24 h incubation, total counts and mutation counts were calculated, and mutation rates were determined. A minimum of 5 biological replicates were used to derive the mutation rate for a particular strain.

2.5 Microscopy

2.5.1 Phase Contrast Microscopy

2.5.1.1 Cell Fixation

1 mL of sample was harvested via centrifugation (2 min at 4000 x g), and washed 3 times in PBS, before addition of 4% paraformaldehyde. Samples were incubated at room temperature for 15 min, washed twice more in PBS, and resuspended in dH_2O. Samples were stored at 4°C.

2.5.1.2 Phase Contrast Imaging

For phase contrast microscopy, samples were fixed as above and resuspended in sterile dH_2O to an appropriate cell density before being dried onto glass slides. Slides were then rinsed with distilled water and dried. Cover slips were then mounted using 80% glycerol, and samples were imaged using a 100x Phase Contrast objective on the Nikon Ti eclipse widefield imaging microscope, using NIS elements software, at the University of Sheffield Wolfson Light Microscopy Facility.

2.5.1.3 Cell Length Analysis Using MicrobeJ

Phase contrast images were analysed to determine cell length in Fiji (v2.9.0) using MicrobeJ (v5.13l) (Ducret et al., 2016). MicrobeJ recognises bacteria and takes cell length measurements. These were manually verified for each image. At least 185 cells were used to average cell length for each sample.

2.5.2 Fluorescence Microscopy

2.5.2.1 Cell Fixation

Samples were grown to an OD_{600nm} of 0.3 before staining with 3 µg/mL Vancomycin BODIPY FL Conjugate (ThermoFisher) for 30 min. Samples were fixed using a fixation cocktail consisting of 20 µL 1 M Na_3PO_4 (pH 7.4), 100 µL 16% (w/v) paraformaldehyde and 4 µL 25% (w/v) glutaraldehyde (Ransom et al., 2015). 500 µL of sample was added to the fixation cocktail and incubated for 15 min

at 37°C. Cells were harvested via centrifugation (2 min at 6000 x g) and washed five times in PBS, before resuspending in 30 µL PBS.

2.5.2.2 Fluorescence Microscopy Imaging

For fluorescence microscopy, samples were fixed as above mounted immediately on an agarose pad (2% w/v) in a 35 mm glass bottom dish (Ibidi). Samples were imaged using a 100x Plan Apochromat phase contrast objective on a Nikon Ti-E microscope with Perfect Focus system, using NIS elements software. Samples were imaged in phase contrast and with Green Fluorescent Protein (GFP) filter. Fluorescence imaging was performed by Anne Williams in the William Durham Laboratory.

2.5.2.3 Fluorescence Intensity Analysis

Fluorescence microscopy analysis was performed by Anne Williams in the William Durham Laboratory. Images were segmented using Ilastik (Berg et al., 2019) and analysed using FAST (Meacock and Durham, 2023). A custom Matlab (vR2023a) script was used to extract fluorescence intensity profiles across the central 70% of the cell along the long axis (Figure 2.1). The fluorescence intensity profile for each cell was then normalised by its average background fluorescence, derived from the average fluorescence intensity of the outer 15% of the cell at each pole. The maximum intensity along the fluorescence profile was determined for each cell and used to generate a swarm plot comparing maximum fluorescence across samples.

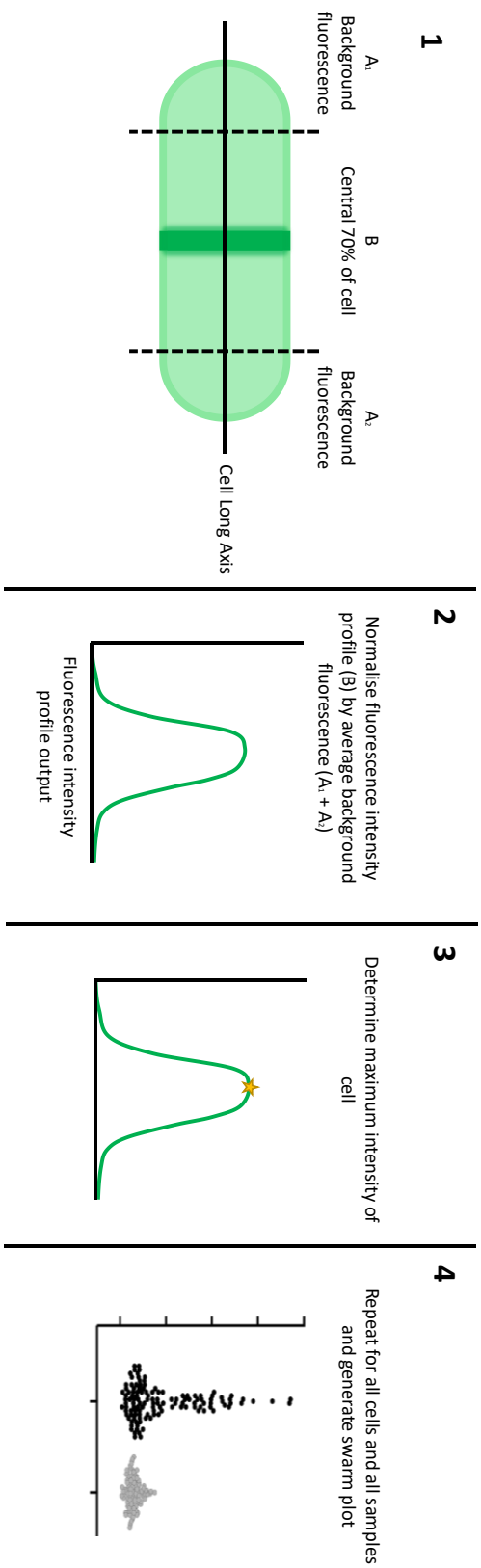


Figure 2.1 - Fluorescence Intensity Analysis

Fluorescence intensity was determined by a custom Matlab script. Intensity profiles were normalised to the cell's average background fluorescence. The maximum intensity for each cell was used to generate a swarm plot. Schematic adapted with permission from Anne Williams.

2.5.3 Transmission Electron Microscopy

2.5.3.1 Cell Fixation

For electron microscopy, 500 µL of sample was added to the fixation cocktail described above (2.5.2.1) and incubated for 30 min at 37°C, before a further 15 min incubation on ice. Fixed cells were harvested via centrifugation (2 min at 6000 x g) and washed 3 times in tris-buffered saline (TBS). Pellets were resuspended in 30 µL TBS and stored at 4°C.

2.5.3.2 Electron Microscopy

Electron microscopy imaging and analysis was performed by Christopher Hill at the electron microscopy unit at the University of Sheffield. Glutaraldehyde-fixed samples were washed in 0.1 M sodium cacodylate buffer, before a secondary fixation using 1% aqueous osmium tetroxide. Samples were washed again using 0.1 M sodium cacodylate buffer, before being dehydrated with ethanol, embedded in resin, and thin sectioned using a Reichert Ultramicrotome. Sections were then transferred onto copper coated grids and stained with uranyl acetate and Reynold's lead citrate. Sections were visualised using the FEI Tecnai T12 with a tungsten source operated at 80 kV, and imaged using a Gatan Orius SC1000B camera.

2.6 RNA and Quantitative Reverse Transcription PCR (qRT-PCR)

2.6.1 RNA Isolation

RNA was extracted from 5 mL of log-phase cells. On reaching OD_{600nm} 0.5, 10 mL of RNA Protect (Qiagen) was added to samples. Samples were incubated for 5 min at 37°C, before being harvested by centrifugation (10 min at 4,000 x g at 4°C). Pellets were resuspended in 1 mL of RNA Pro solution from the FastRNA Pro Blue Kit (MP Biosciences), and added to a matrix tube containing 0.1 mm

silica beads. Cells were lysed using the FastPrep-24 5G instrument (MP Biomedicals) using 2 cycles of 20s (6 m/s), with 2 min on ice between cycles. Matrix beads were removed via centrifugation (16000 x g for 10 min at 4°C), and liquid was transferred to a microfuge tube and incubated at room temperature for 5 min. 300 µL of chloroform was then added, and samples were vortexed for 10 s. Samples were centrifuged (16000 x g for 10 min at 4°C) to separate the phases, and the aqueous phase was transferred to a new microfuge tube containing 500 µL 100% ethanol. Samples were incubated at -20°C overnight, and then centrifuged (16000 x g for 15 min at 4°C). The pellets were washed in 500 µL 70% ethanol, centrifuged again (16000 x g for 5 min at 4°C), and air dried for 5 min. RNA was eluted in 45 µL of nuclease free water (NEB).

DNAse treatment was performed using the Turbo DNase kit (Ambion), as per the manufacturer's instructions. The samples were then cleaned up using the RNeasy MinElute cleanup kit (Qiagen), according to the manufacturer's protocol.

RNA quantity and purity was measured via A_{260} spectrophotometry, and samples were stored at -80°C.

2.6.2 Generation of cDNA

First strand complementary DNA (cDNA) was synthesised from 5 µg RNA using SuperScript III Reverse Transcriptase (ThermoFisher). 26 µL (5 µg) RNA was mixed with 2µL deoxyribonucleotide triphosphate (dNTP) mix and 1 µL of 10 µM random hexaoligos, heated to 65°C (5 min), and cooled on ice (2 min). A mastermix containing 8 µL 5x reverse transcriptase buffer, 2 µL 0.1 M dithiothreitol (DTT), 1 µL RiboLock RNase inhibitor (Thermo), and 2 µL SuperScript III Reverse Transcriptase was added to samples. Reverse transcription was performed using a temperature cycle of 25°C for 5 min,

50°C for 30 min, and 70°C for 15 min. An additional 86 µL of nuclease free water (NEB) was added to samples, so that all concentrations were equal to 40 ng/µL, and 5 µL was equal to 200 ng of cDNA.

Reverse Transcriptase negative samples were also prepared, as above, but without the addition of SuperScript III, to monitor DNA contamination.

2.6.3 Plasmid for qRT-PCR

Expression of genes using qRT-PCR was measured as described in Turner et al., 2009, whereby exact copy number of target gene was measured using a plasmid, containing a single copy of each target gene, to standardise the assay. This plasmid was designed to include each target transcript back-to-back, plus the housekeeping gene *rpoA* (most commonly used reference in *C. difficile* quantification studies). The fragment was synthesised by Genewiz and cloned into a pUC-GW-Kan backbone.

The plasmid was transformed into NEB5α (2.2.10), purified, and linearised using an enzyme chosen (depending on the plasmid sequence) to cut only once. This was then diluted to $2x10^7$ copies per µL.

2.6.4 Melt Curve Analysis

Melt curve analysis was used prior to qRT-PCR to assess whether the reactions produced single specific products, and to determine the optimal temperature to measure fluorescence. Target fragments were amplified using standard PCR (2.2.2) using the qRT-PCR plasmid (2.6.3) as a template, and purified (2.2.4). The melt curve reaction was assembled by combining 20 µL of SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich) and 5 µL of purified PCR product. The reaction was run on the BioRad CFX Connect Real Time System, using the melt curve template in the BioRad CFX manager (v3.1).

2.6.5 Primer Matrix

qRT-PCR reactions were optimised using a primer matrix. Reactions were set up as below, using the qRT-PCR plasmid (2.6.3) as a template. 4 dilutions of plasmid ($2x10^7$ - $2x10^4$ copies per µL) were measured, using 3 primer dilutions (1/5, 1/10, 1/20). A matrix was completed for all target genes. Optimal conditions were chosen based on efficiency (80-90%, within 5% for all targets).

2.6.6 qRT-PCR

qRT-PCR was performed using SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich). Each reaction comprised 25 µL SYBR Green JumpStart Taq ReadyMix, 7 µL $MgCl_2$, 11 µL nuclease free water (NEB), 2 µL of primer pair (appropriate dilution determined by primer matrix), and 5 µL cDNA (concentration of 40 ng/µL, giving 200 ng). Standard curves were made for each experiment using the plasmid, serially diluted from $2x10^7$ - $2x10^1$ copies per µL in lambda DNA (Promega) to mimic DNA-rich sample conditions. Negative (water) and reverse transcriptase negative samples were run in parallel to monitor DNA contamination. Experiments were run using the BioRad CFX Connect Real Time System. Amplification was initiated via a denaturation (95°C, 3 min), which was followed by 40 cycles of 95°C for 10 s, 58°C for 10 s, 72°C for 15 s, and ~ 70°C for 10 s to measure fluorescence (depending on optimal temperature identified in the melt curve). After cycle completion, a melt curve was performed as standard. qRT-PCR was performed in biological and technical triplicate. Copy number was calculated using the BioRad CFX manager (v3.1), and results were expressed as copies per 1000 copies of *rpoA*.

3 Experimental Evolution of Vancomycin Resistance in *C. difficile*

3.1 Introduction

Large-scale, highly parallel experimental evolution has been accomplished in multiple bacterial species, the most famous of these being the long-term *E. coli* evolution performed by Lenski and colleagues (Lenski, 2017). Evolving multiple populations in parallel is invaluable for studying antibiotic resistance, allowing appreciation of alternative routes to resistance and population dynamics. This was demonstrated again in *E. coli*, whereby a 6-population evolution of triclosan resistance revealed how resistance can occur through multiple different mutations in the *fabI* gene, the antibiotic's molecular target (Leyn et al., 2021). This convergent evolution not only suggested that FabI is the sole target of triclosan, but showed how multiple mutations could lead to the same outcome. Beyond this, sequencing isolates from multiple time points during the evolution provided a detailed picture of initial mutations which were later outcompeted by variants providing higher-level resistance.

Multiple variations of classical experimental evolution have been used to advance understanding of antimicrobial resistance. A notable example, the intentional inclusion of hypermutators to drive accelerated evolution and mutation accumulation, was recently used to explore metronidazole resistance in *C. difficile* (Deshpande et al., 2020). A hypermutator phenotype can be induced by

deletion of *mutSL*, encoding a DNA damage repair system (Prunier and Leclercq, 2005). Since stable metronidazole resistant isolates have been historically difficult to obtain, evolving resistance in a *mutSL* knockout was used to investigate the accumulation of mutations leading to resistance. This led to identification of multiple genes involved in both low- and high-level metronidazole resistance. Vancomycin resistance has been slow to evolve in *C. difficile* in clinic, with relatively few examples of resistant clinical isolates (Greentree et al., 2022). Experimental evolution of vancomycin resistance in *C. difficile* may therefore benefit from a similar approach, to enable investigation beyond first-step resistance.

However, the lack of reported vancomycin resistant clinical isolates could potentially be explained by a high fitness cost of resistance carriage. Since a full phenotypic characterisation of laboratory-derived vancomycin resistant *C. difficile* isolates has not been reported, an investigation into important fitness parameters, such as growth and sporulation efficiency, would be beneficial. Similarly, as cross-resistance is a common phenomenon, often reported in other species, understanding the relationships between vancomycin resistance and resistance to other antibiotics in *C. difficile* would be of particular clinical importance.

3.2 Aims and Outcomes

The work presented in this chapter aimed to address these shortfalls, to provide a new standard for highly parallel laboratory evolution in *C. difficile*. Mutagenesis was used to generate a clinically-relevant, but non-pathogenic, strain of *C. difficile* for experimental evolution of vancomycin resistance. This strain was further modified by the deletion of *mutSL,* and by genetic barcoding for future characterisation. A highly parallel evolution was performed, producing vancomycin resistant populations. Phenotypic characterisations, including growth and sporulation efficiency analyses,

were performed to assess potential fitness costs of resistance acquisition. Sequencing of both isolates and populations was then performed to assess vancomycin resistance on both an individual and population level, to allow understanding of the routes to resistance, and the population dynamics involved in high-level resistance.

3.3 Results

3.3.1 Large-scale, Parallel Evolution in *C. difficile*

3.3.1.1 Production of Parental Isolates

C. difficile with an MIC >16 µg/mL may be described as highly resistant to vancomycin (Greentree et al., 2022). In order to evolve high-level vancomycin resistance in the clinically relevant ribotype 027 *C. difficile* strain R20291, it was first prudent to generate a non-pathogenic background. This was achieved previously by the Fagan lab group (data not published) by the deletion of 18 kb, spanning the entire *PaLoc*, including the genes encoding both major toxins and associated regulatory proteins, resulting in the strain R20291Δ*PaLoc*.

To accelerate vancomycin resistance evolution, a subsequent deletion, removing the *mutSL* genes encoding a DNA-damage repair system, generated the hypermutable variant R20291Δ*PaLoc*Δ*mutSL*. Mutagenesis of *C. difficile* was performed by allelic exchange. Specifically, regions approximately 1,200 bp either side of the *mutSL* cluster were amplified from the genome via PCR, using the primer pairs RF2066/RF2067 and RF2068/RF2069. The resulting fragments, the left and right homology arms, were then joined together and inserted into pJAK112 (linearised by BamHI/SacI restriction digest) via Gibson assembly (Figure 3.1). The new construct, pJEB002, was transformed into *E. coli* NEB5α. Once the pJEB002 sequence was confirmed to be correct, the plasmid was transformed into *E. coli* CA434, before conjugation into *C. difficile* R20291Δ*PaLoc.* Colonies were screened by PCR to confirm single recombination – i.e. insertion of the plasmid into the *C. difficile* genome at the desired site. Colonies where single recombination was confirmed were grown on BHI, before harvesting to, and incubation on, CDMM agar with 5-fluorocytosine. The plasmid-encoded *codA* gene selects against plasmid carriage in the presence of 5-fluorocytosine, since *codA* encodes a cytosine

deaminase, which converts 5-fluorocytosine to cytotoxic 5-fluorouracil. The resulting colonies were screened for the deletion of *mutSL* by PCR, and the deletion was confirmed via Sanger sequencing.

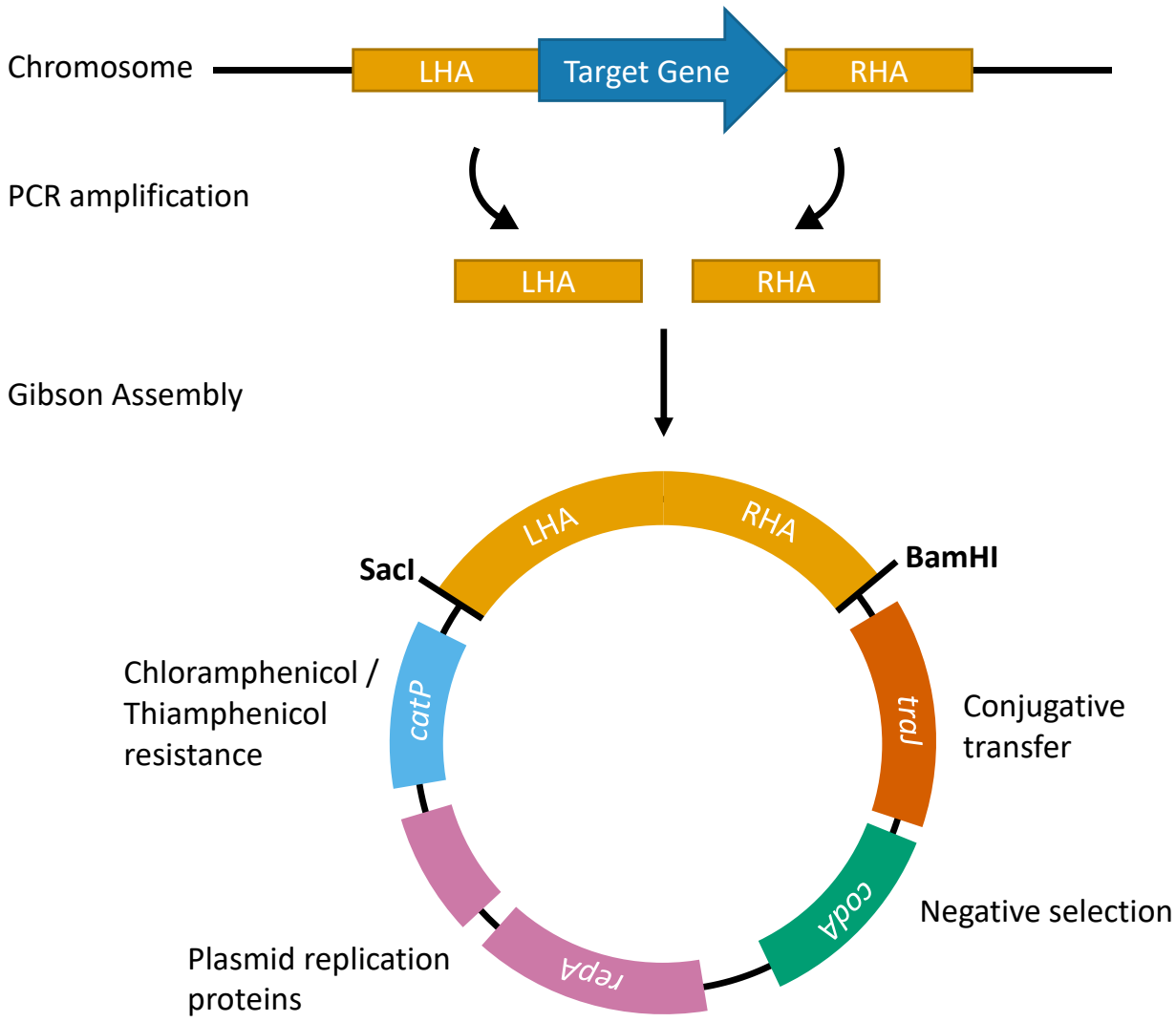


Figure 3.1 - Generating a Mutagenesis Vector

Fragments approximately 1200 bp either side of the target gene(s) – known as the left homology arm (LHA) and right homology arm (RHA) – were amplified. Gibson assembly was used to incorporate the LHA and RHA into the linearised pJAK112 vector backbone to produce the mutagenesis vector construct. The pJAK112 vector contains *traJ* for conjugative transfer, *codA* for negative selection, plasmid replication genes for both *E. coli* and *C. difficile*, and *catP* for chloramphenicol/thiamphenicol resistance (plasmid selection).

To differentiate the 10 replicate lines used in the experimental evolution, 9-nucleotide barcode (Bc) sequences were inserted downstream of the *pyrE* gene in five R20291Δ*PaLoc*, and five R20291Δ*PaLoc*Δ*mutSL* isolates. This resulted in 10 individually barcoded replicate lines (R20291Δ*PaLoc pyrE*::barcode 1-5; R20291Δ*PaLoc*Δ*mutSL pyrE*::barcode 7-11). Barcodes were designed to be differentiable even in the presence of mutations, accounting for potential disruption in hypermutators. Table 3.1 summarises the construction of plasmids for strains relevant to this chapter.

Table 3.1 - Strains and Constructs Relevant to This Chapter

Strain	**Construction**	**Plasmid**
R20291Δ*PaLoc*	Previously constructed.	pJAK143
R20291ΔPaLocΔmutSL	Homology arm amplification using RF2066/RF2067 and RF2068/RF2069. Gibson assembly into linearised pJAK112.	pJEB002
R20291Δ*PaLoc* *pyrE::*barcode 1	Plasmid previously constructed via Genewiz synthesis, followed by ligation into pJAK081 linearised by RF1810/RF1811.	pJAK201
R20291Δ*PaLoc* *pyrE::*barcode 2	Plasmid previously constructed via Genewiz synthesis, followed by ligation into pJAK081 linearised by RF1810/RF1811.	pJAK202
R20291Δ*PaLoc* *pyrE::*barcode 3	Plasmid previously constructed via inverse PCR of pJAK201 with RF1902/RF1903	pJAK203
R20291Δ*PaLoc* *pyrE::*barcode 4	Plasmid previously constructed via inverse PCR of pJAK201 with RF1904/RF1905	pJAK204
R20291Δ*PaLoc* *pyrE::*barcode 5	Plasmid previously constructed via inverse PCR of pJAK201 with RF1906/RF1907	pJAK205
R20291ΔPaLocΔmutSL pyrE::barcode 7	Plasmid previously constructed via inverse PCR of pJAK201 with RF1912/RF1913	pJAK207
R20291Δ*PaLoc*Δ*mutSL* *pyrE::*barcode 8	Plasmid previously constructed via inverse PCR of pJAK201 with RF1914/RF1915	pJAK208
R20291Δ*PaLoc*Δ*mutSL* *pyrE::*barcode 9	Plasmid previously constructed via inverse PCR of pJAK201 with RF1916/RF1917	pJAK209
R20291Δ*PaLoc*Δ*mutSL* *pyrE::*barcode 10	Plasmid previously constructed via inverse PCR of pJAK201 with RF1918/RF1919	pJAK210
R20291Δ*PaLoc*Δ*mutSL* *pyrE::*barcode 11	Plasmid previously constructed via inverse PCR of pJAK201 with RF1920/RF1921	pJAK211

3.3.1.2 Δ*mutSL* Increases Mutation Rate 20-fold

To assess the change in mutation rate following the deletion of DNA-damage repair system genes *mutSL*, a rifampicin resistance assay was performed. Rifampicin is commonly used for mutation rate assays due to resistance being conferred by single base pair substitutions in *rpoB* (Guérillot et al., 2018). Overnight cultures of R20291Δ*PaLoc* and R20291Δ*PaLoc*Δ*mutSL* were concentrated 10-fold and used to calculate colony forming units (CFU). The remainder of the sample was spread on BHI plates containing 8x MIC rifampicin. The mutation rate was calculated by dividing mutants per mL by CFU per mL. The deletion of *mutSL* from R20291Δ*PaLoc* yielded an approximate 20-fold increase in mutation rate.

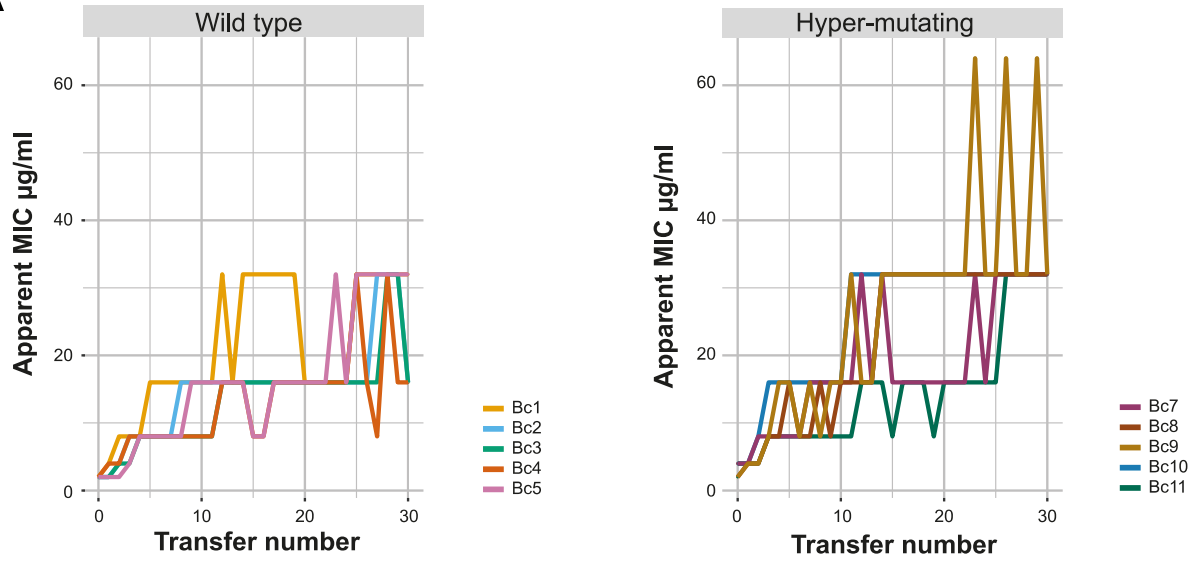
3.3.1.3 Vancomycin Resistance Evolves Rapidly in *C. difficile* During *in vitro* Experimental Evolution

The experimental evolution was performed using a broth-based gradient approach. This method was chosen since it allows the exertion of high selective pressures, whilst minimising the risk of extinction, since sub-MIC growth is captured at each passage. The evolution was initiated by inoculating 6-well plates filled with TY media (4 mL per well) with the 10 barcoded strains. Each well was supplemented with vancomycin at either 0.25x, 0.5x, 1x, 2x, 4x, or 8x the initial MIC of 1 µg/mL. The populations were then passaged every 48 h. After a visual inspection, cells from the well with the highest antibiotic concentration permitting growth were propagated in a 1:400 dilution to a fresh 6-well plate. A vancomycin gradient across the plate was maintained at 0.25x – 8x MIC each passage, by using the MIC from the previous passage to calculate subsequent vancomycin concentrations. This process was repeated for a total of 30 serial transfers per replicate line. Populations underwent approximately 8.64 generations per transfer, yielding approximately 259 generations throughout the course of the experiment. Alongside the vancomycin evolution, ten matched control populations were propagated under equivalent conditions in the absence of vancomycin, to control for mutations arising due to laboratory adaptation.

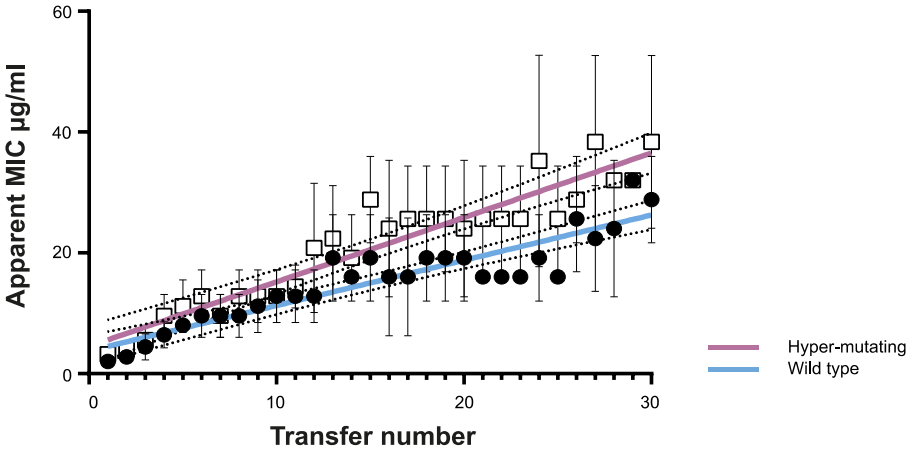
Resistance evolved rapidly in all ten replicate populations propagated with vancomycin selection (Figure 3.2a). Nine of the ten replicate lines evolved to grow in 2 µg/mL vancomycin by the end of the second passage (P2). Although not measured by a classical MIC assay, growth in 2 µg/mL vancomycin suggests an MIC of 4 µg/mL (referred to hereafter as the apparent MIC), meaning nine of the ten evolved populations reached the European committee on antimicrobial susceptibility testing (EUCAST) breakpoint within 144 h. An apparent MIC of 16 µg/mL was relatively stable in many replicate populations, being maintained for long periods. In contrast, an apparent MIC of 32 µg/mL was often unstable, being lost in subsequent passages at least once in eight of the ten replicate populations, however this was usually regained (six of the ten replicate populations). An apparent MIC of 64 µg/mL was only observed in one replicate line (Bc9), but evolved three separate times across the evolution. This was a highly unstable occurrence, which was immediately lost in the subsequent passage each time. All ten replicate lines had an apparent MIC of 16 (Bc2, 3, 4) or 32 (Bc1, 5, 7, 8, 9, 10, 11) µg/mL vancomycin by passage 30 (P30).

To investigate whether R20291Δ*PaLoc*Δ*mutSL* promoted accelerated vancomycin resistance evolution, the apparent MICs at each time point for wild type (WT) (R20291Δ*PaLoc pyrE*::barcode 1-5) and hyper-mutating (R20291Δ*PaLoc*Δ*mutSL pyrE*::barcode 7-11) replicate evolved populations were averaged and subjected to linear regression analysis. The hyper-mutating replicate lines showed significantly accelerated vancomycin resistance evolution compared to WT replicate lines (Figure 3.2b).

A



B



C

Evolved Isolate	Van Evo endpoint	MIC 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Colony for sequencing
P30 Bc1	32	16	16	16	16	32	32	32	16	32	16	16	16	16	**32**	P30 Bc1 Col14 (MIC32)
P30 Bc2	16	8	**16**	16	8	16	16									P30 Bc2 Col2 (MIC16)
P30 Bc3	16	8	8	8	16	16	4	16	**16**	16						P30 Bc3 Col8 (MIC16)
P30 Bc4	16	16	16	16	16	**16**	16									P30 Bc4 Col5 (MIC16)
P30 Bc5	32	16	8	16	16	16	**32**									P30 Bc5 Col6 (MIC32)
P30 Bc7	32	16	16	16	**32**	32	16									P30 Bc7 Col4 (MIC32)
P30 Bc8	32	16	16	16	16	16	16	32	32	**32**						P30 Bc8 Col9 (MIC32)
P30 Bc9	32	16	16	16	16	16	16	16	16	16	16	16	**32**	16	16	P30 Bc9 Col12 (MIC32)
P30 Bc10	32	32	32	32	32	32	**32**									P30 Bc10 Col6 (MIC32)
P30 Bc11	32	32	32	16	32	**32**	32									P30 Bc11 Col5 (MIC32)

Figure 3.2 - Evolution of Vancomycin Resistance

(A) Changes in apparent vancomycin MIC over the course of a 30 transfer experimental evolution. Apparent MIC was determined as the well with the lowest vancomycin concentration showing no clear growth. (B) Shown are the means and standard deviations of the apparent MIC for five wild type (open squares) and five hyper-mutating (closed circles) populations. Linear regressions fitted to each data set, blue and pink respectively, are significantly different by Analysis of Covariance (ANCOVA), P=0.0008. (C) Agar dilution MICs of clones isolated from the P30 populations. Dark pink shows an MIC consistent with the apparent MIC of the population from which it was isolated. Medium pink shows a 2-fold MIC reduction, and light pink shows a 4-fold MIC reduction, compared to the apparent MIC of the respective populations. The colony chosen for sequencing (and subsequent phenotypic characterisations) is shown in the final column.

To assess the heterogeneity of the replicate evolved populations at the end of the evolution, multiple individual clones (minimum 6 per replicate line) were isolated from each evolved population at P30. The MIC of each individual clone was measured by agar dilution. Of the 82 clones tested, 38 (46%) had an MIC consistent with the apparent MIC of the population from which they were isolated (Figure 3.2c). In fact, only two evolved populations showed complete homogeneity in MIC – Bc4 and Bc10 reliably displayed an MIC consistent with the apparent MIC of their respective populations. The remaining replicate populations produced variable MICs: 42 clones displayed MICs 2-fold lower than the apparent MIC of their respective populations, while two clones displayed MICs 4-fold lower. The most heterogeneous populations in terms of range of clonal MICs were Bc3 and Bc5. Overall, this demonstrates significant variation within evolved populations by the end of the evolution.

3.3.2 Phenotypic Analysis of Resistant Isolates

Clones isolated from the P30 evolved populations, referred to hereafter as evolved isolates, were phenotypically assayed in terms of growth, sporulation efficiency, and morphology. The evolved isolate assayed for each replicate barcode was the same as the isolate chosen for sequencing (shown in Figure 3.2c).

3.3.2.1 Vancomycin Resistant Clones Display Reduced Growth

To assess the consequences of vancomycin resistance on evolved isolate fitness, growth was measured *in vitro* in rich liquid media (TY broth) (Figure 3.3a). Each evolved isolate was compared to its respective parental strain. All 10 evolved isolates displayed growth defects, with some exhibiting more severe impairments. In particular, Bc1, 7 and 10 showed slower growth and a reduced maximum optical density (OD). Bc11 showed the most extreme growth phenotype, with very slow growth and the lowest OD after 8 h.

To statistically compare growth differences between evolved isolates and their matched parental strains, growth curves were analysed using the R package Growthcurver (Sprouffske and Wagner, 2016). Growthcurver summarises plate reader outputs, providing multiple parameters for growth measurement, such as generation time (t_gen) and area under the curve (AUC). To assess which growth measurement was the most representative, a cross correlation of Growthcurver outputs was performed using the R package Hmisc (Harrell and Dupont, 2019). Cross correlations were visualised using the R package Corrplot to produce a correlation matrix, highlighting the most associated variables (Wei and Simko, 2021) (Figure 3.3b). The best measurement of growth was determined to be AUC, which was subsequently used to statistically compare growth of each evolved isolate with its matched parental strain, using Student's t-tests with Welch's correction. All evolved isolates displayed significantly different growth profiles compared with their matched parental strain, showing vancomycin resistance negatively impacts growth in rich media.

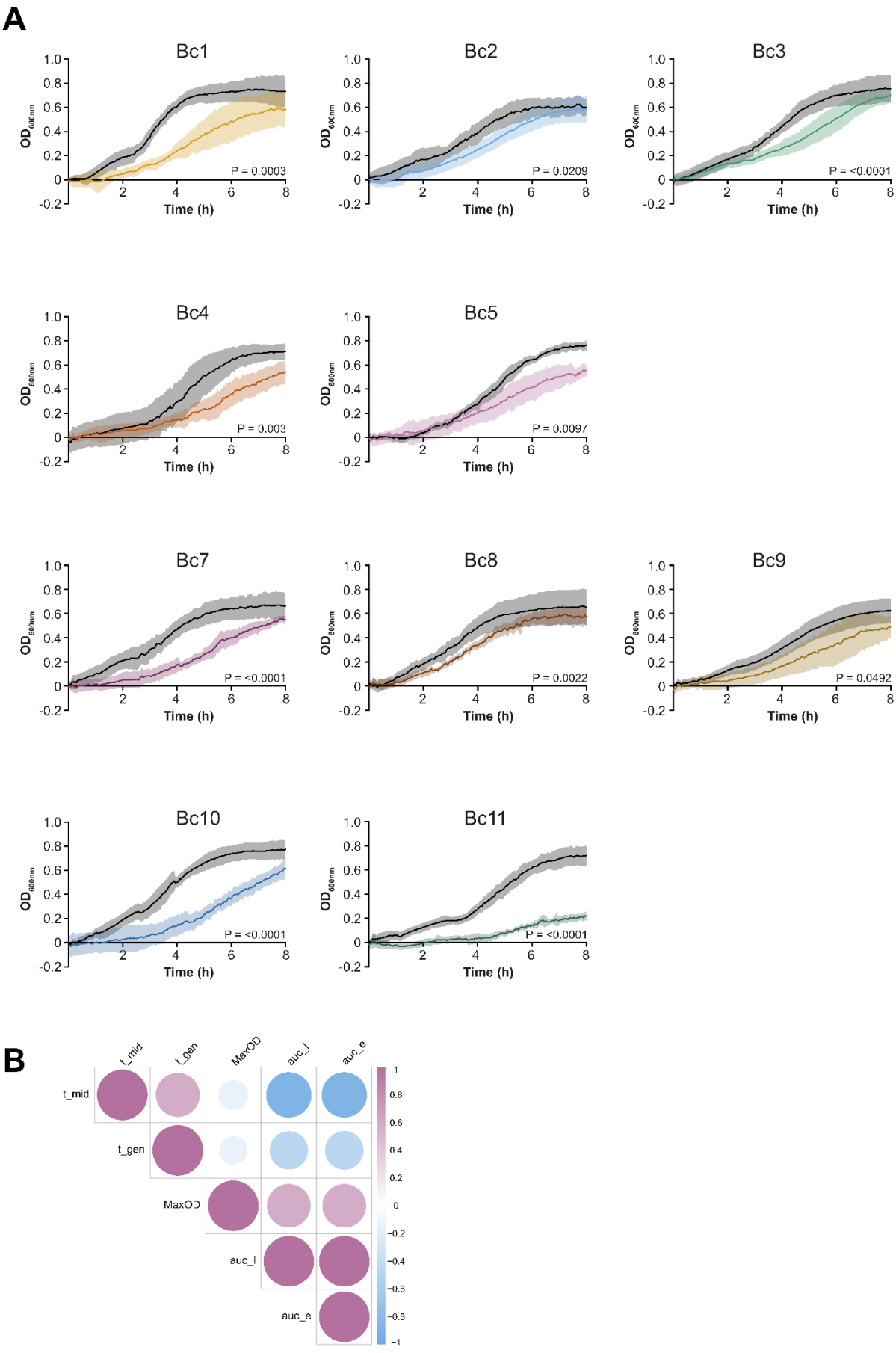


Figure 3.3 - Growth of Evolved Isolates

(A) Growth over time in rich media (TY broth) was evaluated by measuring OD at 600 nm in a 96 well microplate spectrometer. Growth of each evolved isolate (coloured lines) was compared to its matched parental strain (black lines). Shown are the mean and standard deviation of repeats, assayed at minimum in biological and technical triplicate. For each strain, area under the curve was determined using the GrowthCurver R package and these were compared using Student's t-tests with Welch's correction, with the P value shown on each graph. All pairwise differences are significant. (B) Cross-correlation of Growthcurver outputs to determine the best measure of growth to compare samples.

3.3.2.2 Vancomycin Resistance Results in a Range of Sporulation Phenotypes

Since sporulation is a key hallmark of *C. difficile* virulence and pathogenicity, the sporulation (and germination) efficiency of evolved isolates was assessed. Sporulation efficiency was determined using heat resistance by comparing total cell counts with heat-treated (65°C, 30 min) spore counts over a period of 5 days. Each evolved isolate was compared to R20291Δ*PaLoc* in biological duplicate and technical triplicate. Sporulation efficiency was statistically assessed by comparing the spore count AUC of each evolved isolate to the WT spore count AUC, as this provided a representative measure of sporulation rate and efficiency. From this, a range of sporulation phenotypes were apparent – Bc3 and Bc5 displayed no difference in sporulation efficiency, with a profile almost identical to the WT. Several strains did, however, display impaired sporulation phenotypes of ranging severities (Figure 3.4). Bc4 showed mild defects in sporulation, with a similar apparent rate to the WT, but with reduced sporulation efficiency. Bc8 showed a delayed sporulation phenotype, whereby efficiency at time point 5 was comparable to the WT, but earlier time points demonstrated a slower population transition from vegetative cells to spores. Bc7, 9 and 10 had severe sporulation defects, with both reduced sporulation rates and efficiencies. The most severe sporulation phenotype, a total lack of sporulation, was found in Bc11. Taken together, it is clear vancomycin resistance can result in pleiotropic sporulation defects.

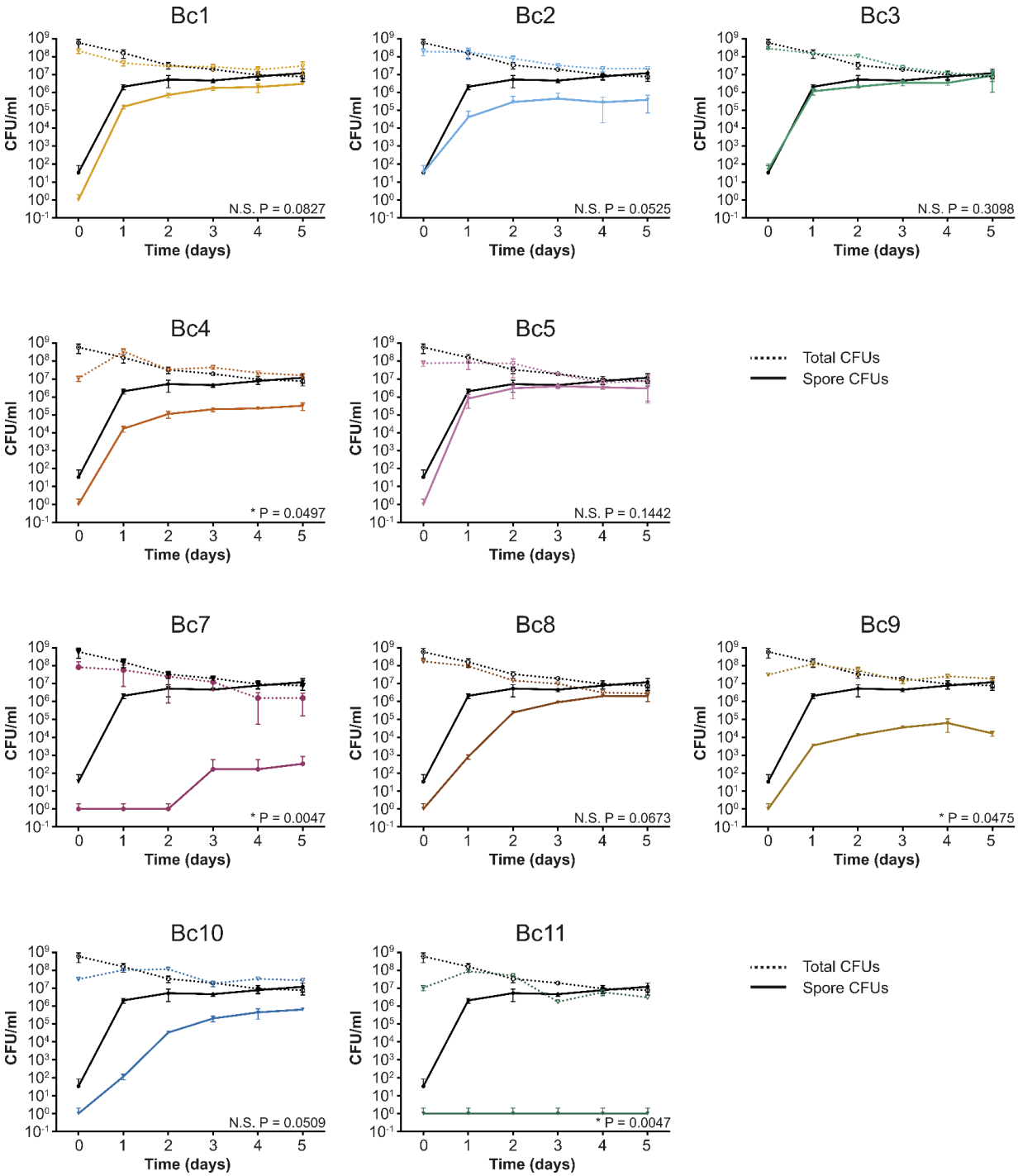


Figure 3.4 - Sporulation of Evolved Isolates

Sporulation efficiencies of each evolved isolate (coloured lines) were compared to the same parental R20291Δ*PaLoc* (black lines). All evolved isolates were compared to R20291Δ*PaLoc*, regardless of their parental mutational background (Δ*mutSL*, barcode). Stationary phase cultures were incubated anaerobically for 5 days with samples taken daily to enumerate total CFUs (dotted lines) and spores (solid lines), following incubation at 65°C for 30 min to kill vegetative cells. Shown are the mean and standard deviations of biological duplicates assayed in triplicate. For each strain, spore CFU area under the curve was determined using GraphPad Prism, and these were compared using Dunnett's T3 multiple comparisons test with the adjusted P value shown on each graph. * = significant difference, N.S. = not significant.

3.3.2.3 Altered Morphological Changes at the Cellular Level

Since vancomycin impacts the cell wall, it was not unfeasible to consider the potential morphological ramifications of resistance acquisition. On account of this, evolved isolates and the R20291 WT were imaged via phase contrast microscopy to determine whether there were morphological differences on a cellular level. Despite cell morphology being broadly similar, large variations in cell length were observed between evolved isolates (Figure 3.5a). To investigate this further, cell lengths of evolved isolates and the R20291Δ*PaLoc* WT were analysed using microbeJ (Ducret et al., 2016). Cells were first segmented, before length measurements of all cells in-frame were performed. A minimum of 185 cells, across multiple images and multiple biological replicates, were measured. The length measurements were visualised using a violin plot, and statistically analysed using a one-way analysis of variance (ANOVA) with Dunnett's T3 multiple comparisons (Figure 3.5b). Bc 2, 7, 8 and 9 showed no significant difference in cell length compared to the WT. Of the evolved isolates with a significantly different median cell length, Bc4 was the only replicate to be significantly smaller than the WT. Bc 1, 3, 5, 10 and 11 were all significantly longer than the WT, with Bc5 being the largest.

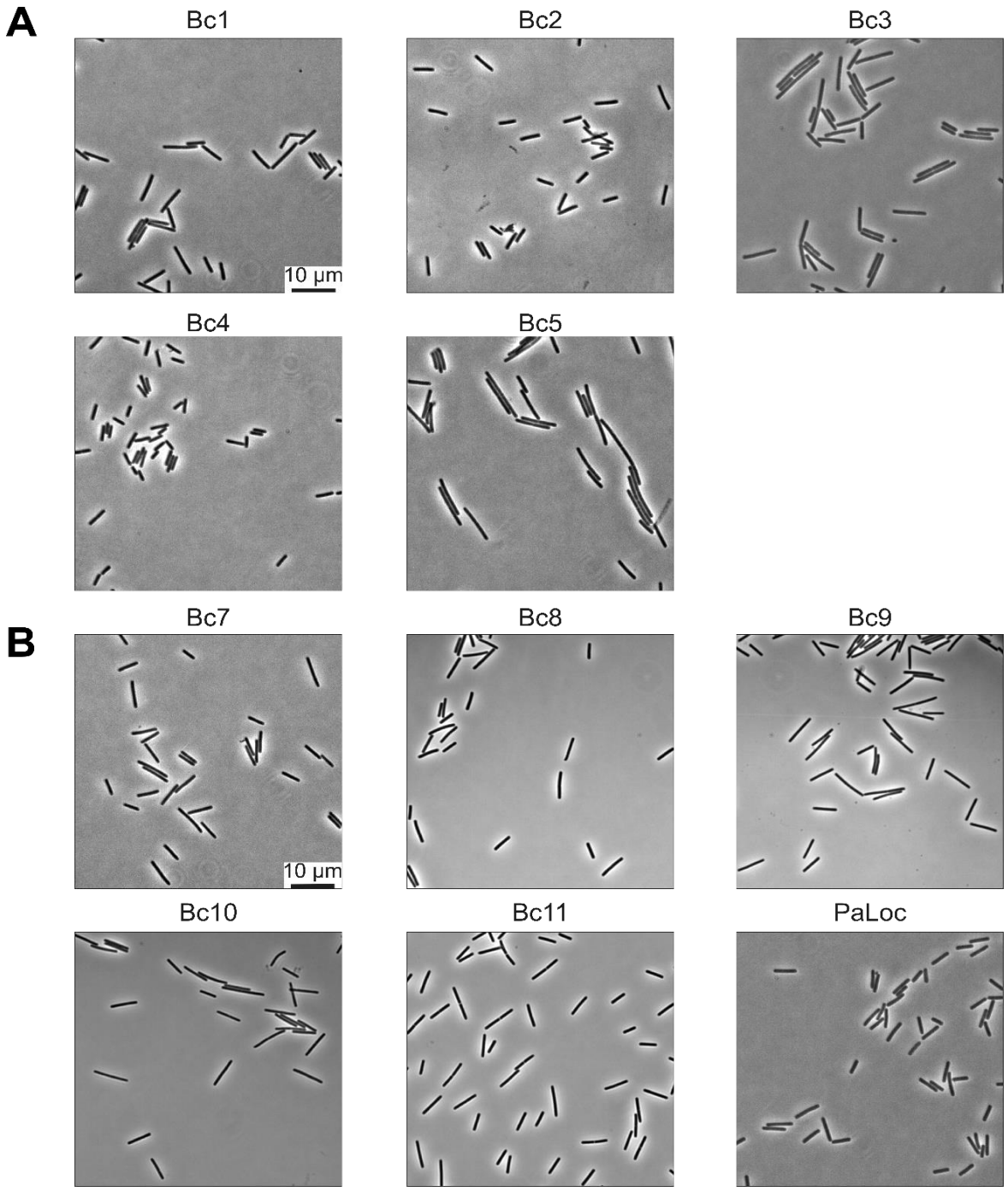

A
Bc1
Bc2
Bc3
10 µm
Bc4
Bc5
B
Bc7
Bc8
Bc9
10 µm
Bc10
Bc11
PaLoc
C
Cell length (µm)
20
15
10
5
0
** N.S. ** ** ** N.S. N.S. N.S. ** **
ΔPaLoc
Bc1
Bc2
Bc3
Bc4
Bc5
Bc7
Bc8
Bc9
Bc10
Bc11
Strain

Figure 3.5 - Cell Morphology of Evolved Isolates

Phase contrast light microscopy of mid-log cultures of each wild type (A) and hyper-mutating (B) evolved isolate, with R20291Δ*PaLoc* for comparison. Shown is a representative field of view for each strain. (C) Images were analysed using MicrobeJ to determine lengths of at least 185 individual cells for each strain. Shown is an all-point violin plot with the median indicated by a solid horizontal line. Statistical significance of evolved isolates against the R20291Δ*PaLoc* control was calculated using a one-way ANOVA with Dunnett's T3 multiple comparisons test, ** = P<0.0001, N.S. = not significant.

3.3.2.4 Electron Microscopy of Evolved Isolates

Since vancomycin resistance resulted in varied sporulation and cell length phenotypes, changes in spore ultrastructure as a consequence of vancomycin resistance were investigated via transmission electron microscopy (TEM). Evolved isolates, and the R20291Δ*PaLoc* WT control, were grown in TY broth for 5 days before fixation, processing and imaging. Electron microscopy imaging was performed by Christopher Hill at the University of Sheffield Electron Microscopy Unit.

C. difficile spores are complex, multi-structured entities, the basic components of which are labelled in Figure 3.6. The dense, darkly stained spore core houses the DNA, and is well protected by the outer layers (Buddle and Fagan, 2023). These layers comprise an inner membrane, primordial cell wall, a flexible peptidoglycan cortex, outer membrane, proteinaceous spore coat and exosporuim. Most of the evolved isolates had typical spore morphologies with clearly identifiable spore structures, consistent with those of the WT (Figure 3.6). The exosporuim was clearly discernible on Bc1, 2, 3, 5 and 8. Bc9 displayed altered spore morphologies, with a large core, but thin cortex and spore coat layers. No obvious examples of spores could be found in Bc11 samples, consistent with the lack of sporulation observed in the sporulation efficiency assay.

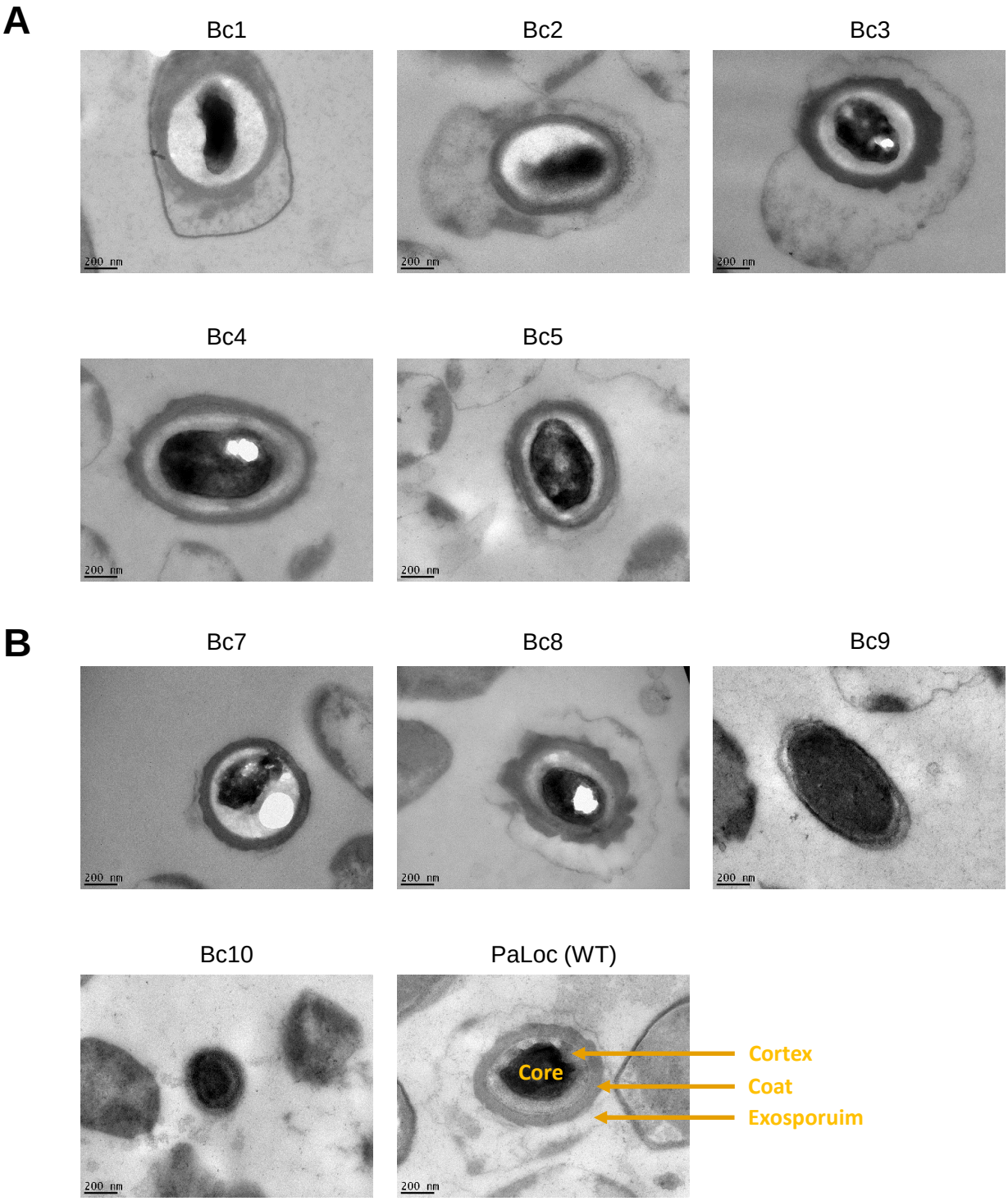


Figure 3.6 - Thin Sections of Evolved Isolate Spores

Spore ultrastructure of *C. difficile* WT (A) and hyper-mutating (B) evolved isolates, with R20291Δ*PaLoc* for comparison. Basic spore features are labelled on R20291Δ*PaLoc.* Images taken after 5 days of growth in TY medium. Shown is a representative spore structure for each replicate line.

3.3.2.5 Evolved Isolates Have Similar Evolutionary Trajectories

Pulling together the above phenotypic observations allows visualisation of the evolutionary trajectories of the evolved isolates. There are multiple approaches to determining patterns of phenotypic evolution, with principal component analysis – plotting phenotypic observations in a phenotypic change vector – being the gold-standard (Adams and Collyer, 2009). PCA allows consolidation of multiple variables into a single plot, reducing the dimensionality of datasets whilst preserving information (Jolliffe and Cadima, 2016). This provides an assessment of trajectories, their magnitude, and parallelism between replicate lines. For simplicity, only the five WT-derived evolved isolates, and the R20291Δ*PaLoc* ancestor, were assessed. To generate the PCA, phenotypic observations were first condensed into single representative measurements. For growth analysis, this was determined to be AUC (section 3.3.2.1). Similarly, sporulation AUC was used to represent sporulation variation (section 3.3.2.2). Cell length and MIC were also incorporated, and all measurements were made relative to the WT. The PCA was computed using the prcomp() function in Base R (http://www.rproject.org/), and visualised using the factoextra package (Kassambara and Mundt, 2020). The first 2 principal components were plotted, since together these explained 93% of the variance (Figure 3.7).

Visualising the evolved isolates in multivariate phenotype space showed all five evolved isolates followed similar evolutionary trajectories away from the parental WT ancestor. These trajectories were generally associated with lower sporulation efficiency and growth defects, albeit with divergence in relation to cell size, and extent of the fitness burdens. Bc3 was closest in fitness to the parental WT in terms of phenotypes measured. Bc2 and Bc4 appeared to have the worst fitness/resistance trade off, as both showed reductions in growth and sporulation, without the benefit of the high MICs displayed by Bc1 and Bc5. There was also no sub-clustering by resistance

mechanism, suggesting fitness is not solely determined by mutations relating to resistance, but also by accumulation of non-beneficial mutations throughout the course of the evolution.

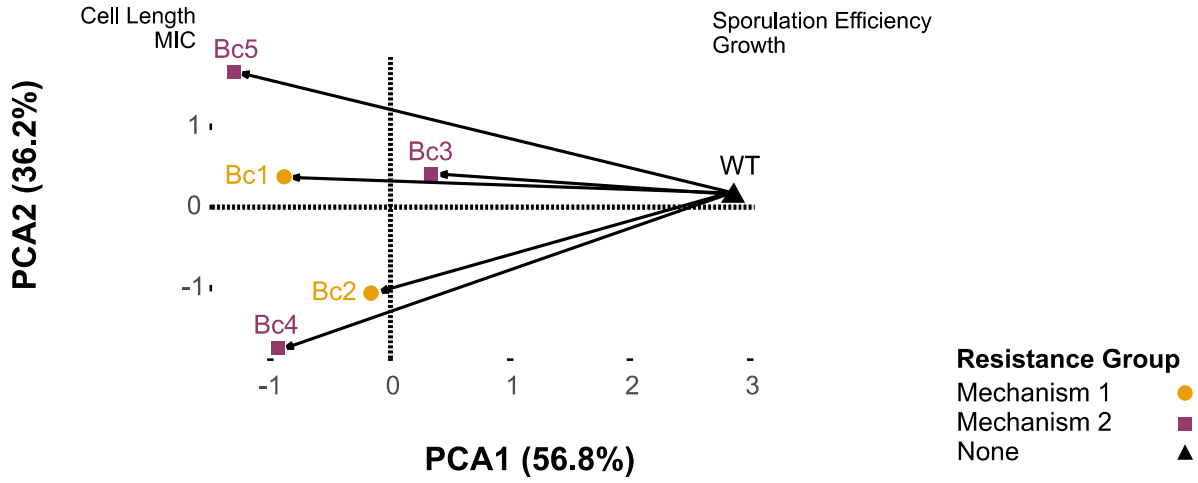


Figure 3.7 - Principal Component Analysis of Evolved Isolates Bc1-5

Principal Component Analysis (PCA) of P30 isolates from populations Bc1-5 (coloured points) vs the ancestral strain (black triangle), with PC1 versus PC2 plotted, accounting for 93% variance. The loadings (sporulation efficiency, growth, MIC, cell length) are shown in respective locations. Arrows show the evolutionary trajectories of wild-type replicate lines from their ancestor in multivariate phenotype space.

3.3.2.6 Evolved Isolates Display Teicoplanin Cross-Resistance

To investigate whether evolution of vancomycin resistance is associated with cross-resistance to other antibiotics, the MICs of evolved isolates and the R20291Δ*PaLoc* WT were measured for teicoplanin. Teicoplanin is a semisynthetic glycopeptide antibiotic, acting in a similar manor to vancomycin to inhibit cell wall synthesis. On the grounds of this, it is not out of the question that one or more of the resistance mechanisms displayed by the evolved isolates may also provide protection against this antibiotic.

All of the ten evolved isolates displayed elevated teicoplanin MICs relative to the WT (WT MIC 0.25 μg/mL) (Figure 3.8). Increases were, however, relatively small – a 2-fold increase was observed for six of the ten evolved isolates (Bc2, 3, 4, 5, 7, and 11), and a 4-fold increase was observed for the remaining four evolved isolates (Bc1, 8, 9, and 10). 4-fold increases were most often observed in hyper-mutating replicate lines. There was no correlation between vancomycin and teicoplanin resistance amongst the evolved isolates (P=0.2, data not shown). Despite being highly resistant to vancomycin, the low-level cross resistance observed suggests there are alternative teicoplanin resistance mechanisms which do not overlap with the captured vancomycin resistance mechanisms.

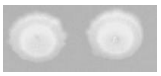
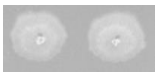
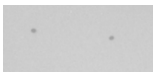
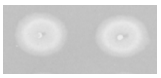
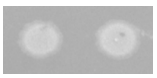
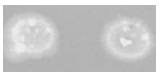
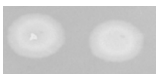
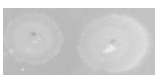
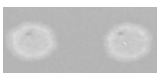
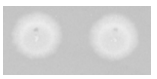
Isolate	Teicoplanin (µg/mL)					
	0	0.12	0.25	0.5	1	MIC
R20291Δ *Paloc*						0.25
P30 Bc1						1
P30 Bc2						0.5
P30 Bc3						0.5
P30 Bc4						0.5
P30 Bc5						0.5
P30 Bc7						0.5
P30 Bc8						1
P30 Bc9						1
P30 Bc10						1
P30 Bc11						0.5

Figure 3.8 - Teicoplanin Cross-resistance of Evolved Isolates

Teicoplanin MIC of evolved isolates. MIC measured by agar dilution on BHI, with R20291Δ*PaLoc* as the WT control. Assays were performed in biological triplicate and technical duplicate. Shown is a single representative biological replicate.

3.3.3 Genetic Basis of Vancomycin Resistance (Isolates)

In order to understand the genetic basis of vancomycin resistance on an individual level, evolved isolates were sequenced. A single isolate from each of the ten replicate populations, which had an MIC consistent with the apparent MIC of the population from which it was isolated (section 3.3.1.3),

was selected for whole genome sequencing. One random isolate from each of the ten control populations was also sequenced as a matched control. Short-read 30x whole genome sequencing was performed using the Illumina NovaSeq 6000 at MicrobesNG.

3.3.3.1 Creating an Optimised Isolate Sequencing Analysis Pipeline

An optimised sequencing analysis pipeline was developed as part of this project, based on a resistant mutant analysis pipeline described previously (Wright et al., 2019). Illumina reads were trimmed by Trimmomatic (v0.30) at MicrobesNG. Trimmed reads were checked using FastQC (v0.11.9) (S Andrews, 2019), and aligned to the R20291 reference genome using BWA-mem (v0.7.17) (Li and Durbin, 2009). PCR duplicates, which occur when an error is amplified through PCR (increasing its likelihood of being falsely called as a variant), were removed using Picard (v2.25.2) (http://broadinstitute.github.io/picard/). The mpileup utility within SAMtools (v1.43) was subsequently used to generate the mpileup file required for Varscan. Mpileup summarises each base in a sequence, generating a table comprising the reference base, read bases, coverage and quality. Varscan (v2.4.3-1) (Koboldt et al., 2009) was chosen for variant detection due to its high sensitivity and specificity, its potential for use in both isolate and pooled sequencing samples, and since it provides increased freedom to optimise various parameters. Variants were called using Varscan parameters optimised for the genome sequence coverage, to reduce false positive and false negative calls. A minimum of 80% frequency was required for variant calls, a standard for isolate variant calling (Kaewprasert et al., 2022). To validate variants, samples were re-analysed using Breseq (v0.35.5) (Deatherage and Barrick, 2014), with variants being retained if detected in both analysis pipelines. The validated variants were also manually checked in IGV (v2.8.6) (Robinson et al., 2011).

3.3.3.2 SNPs, Frameshifts and Deletions Contribute to Vancomycin Resistance

Analysis was focussed on variants which were observed in isolates evolved with vancomycin, but not in ancestral or matched-control isolates – so-called vancomycin-unique variants – since these are the

most likely to have evolved in response to vancomycin selection. Overall, 114 vancomycin-unique variants were observed across the ten evolved isolates (Appendix V - Isolate Variants). Wild type-derived evolved isolates contained between one and three vancomycin-unique mutations per genome. In hyper-mutating evolved isolates, between 14 and 26 vancomycin-unique mutations per genome were observed. Of the 114 vancomycin-unique mutations observed, within-gene SNPs accounted for 43% of the variants, of which 67.4% were nonsynonymous and 32.6% were synonymous. Frameshifts accounted for a further 31% of variants, and in-frame deletions accounted for 1.8%.

3.3.3.3 Parallel Evolution Observed Across all Ten Replicate Lines

Parallel evolution is defined as the repeated evolution of a genotype in independent populations – for example, variants occurring in the same genetic locus in multiple replicate populations. Given the size of the *C. difficile* genome, parallel evolution should be improbable – even with a strong selective pressure such as vancomycin, there ought to be countless ways to evolve resistance, given the many genes involved in peptidoglycan biosynthesis and maintenance, and the many more genes involved in regulation. The presence of parallel evolution in these circumstances therefore suggests strong selection, and a potential role of these mutations in vancomycin resistance.

In the five WT-derived evolved isolates, parallel evolution was observed at three genomic loci: *vanT* in 3 isolates, *CDR20291_3437* (simplified to *CD3437* hereafter) in 2 isolates and *comR* in 2 isolates (Figure 3.9). What's more, mutations in *comR* always co-occurred with mutations in *vanT* (Bc3 and Bc4) (although conversely, *vanT* occurred without *comR* in Bc5). VanT is a putative serine racemase encoded within a VanG-type cluster (referred to hereafter as the *van* cluster) that was previously implicated in decreased vancomycin susceptibility in *C. difficile* (Shen et al., 2020). *comR* encodes a homologue of the RNA degradosome component polynucleotide phosphorylase (PNPase), suggesting that RNA stability may play a role in the *vanT*-associated mechanism of vancomycin

resistance. Consistent with this possibility, the coexisting mutations in Bc5, the other WT-derived evolved isolate carrying a *vanT* mutation, included *maa*, which encodes a putative maltose O-acetyltransferase; and a 75 bp deletion that completely removed an intergenic region downstream of *rpmH* and before *rnpA*, a locus encoding another predicted component of the RNA degradosome. Mutations affecting RNA stability were not observed in the other two WT-derived evolved isolates.

CD3437, mutated in Bc1 and Bc2, encodes a predicted two-component system histidine kinase, with its cognate response regulator encoded by *CD3438*. The nearby locus *CD3439* encodes a putative D,D-carboxypeptidase that likely plays a role in peptidoglycan modification through the removal of the terminal D-Ala from the stem peptide. Based on their predicted functions, these genes were renamed to *dacS* (*CD3437*, histidine kinase), *dacR* (*CD3438*, response regulator) and *dacJ* (*CD3439*, D,D-carboxypeptidase). Despite not being implicated in vancomycin resistance previously, the *dacJRS* cluster displayed strong evidence of selection. Consistent with this, all five hyper-mutating evolved isolates had mutations in either the *van* cluster (1 had a nonsynonymous mutation in *vanT* and 1 in *vanS*) or the *dacJRS* cluster, meaning all ten evolved isolates carried mutations in at least one of these two gene clusters. Mutations in *vanT* and *dacS* appeared to be mutually exclusive in the WT-derived evolved isolates, however in the hyper-mutating isolate Bc8, mutations in both *dacS* and *CD1523* (*vanS*), encoding a two-component system sensor histidine kinase that is thought to regulate the *van* operon in response to vancomycin, were observed. This suggests the two pathways to resistance are not entirely mutually exclusive.

dacS displayed no obvious co-occurrence with other mutations. In fact, illumina sequence data suggested there were no other unique mutations in Bc1. Conversely, in Bc2, nonsynonymous mutations in *bclA3* and *CD3124* were observed. *bclA3* encodes a spore surface protein with no known function in vegetative cells, while *CD3124* encodes an orphan histidine kinase of unknown

function. *CD3124* showed striking parallelism at the SNP-level across the evolved isolates, being mutated in four of the ten isolates (Bc2, 7, 9, and 10), with an identical frameshift mutation in three (Bc7, 9, and 10) (Figure 3.9).

Beyond those already mentioned, there were multiple occurrences of parallel evolution across the five hyper-mutating evolved isolates. One notable example was *mreB2*, which was mutated in 3 hyper-mutating evolved isolates (Bc8, 9, and 10). MreB2, a bacterial actin homologue, is a rod shape determining protein and a component of the *C. difficile* elongasome, which directs localisation of peptidoglycan synthesis. Two *mreB* genes exist in *C. difficile, mreB* and *mreB2,* both of which are essential for growth *in vitro* (Dembek et al., 2015). *mreB2* lies within the core elongsasome operon alongside *mreCD* and *min* genes, displaying the closest homology to *B. subtilis mreB*, and is the primary MreB-encoding gene in *C. difficile* (Enany, 2017). Interestingly, a further gene in this cluster, *CD0985*, displayed mutations in 2 hyper-mutating evolved isolates (Bc7 and Bc8), suggesting a role of the elongasome in vancomycin resistance. Mutations in elongasome-encoding genes were only observed in hyper-mutating evolved isolates, demonstrating the interesting evolutionary concept that hypermutators possess a larger evolutionary potential, allowing them to potentiate novel evolutionary paths not explored by the WT-derived evolved isolates (Figure 3.9). Furthermore, *sdaB*, which encodes an L-serine dehydratase, displayed mutations in two of the five hyper-mutating evolved isolates (Bc9 and Bc11). *sdaB* has previously been implicated in vancomycin resistance in early *C. difficile* evolution studies (Leeds et al., 2014). Although not subject to parallel evolution in this study, *asnB1*, encoding an asparagine synthetase, was mutated in Bc11. AsnB1 has previously been associated with vancomycin sensitivity in *C. difficile,* being responsible for peptidoglycan amidation in the presence of vancomycin (Ammam et al., 2020). The targeting of this gene therefore demonstrates a wider, more general, parallelism beyond this experimental evolution.

The remainder of this project focusses on the two key pathways to vancomycin resistance observed in the WT-derived evolved isolates – *dacJRS* and the *van* cluster. These two pathways were present in all ten of the evolved isolates, and therefore likely represent significant drivers of vancomycin resistance in *C. difficile*.

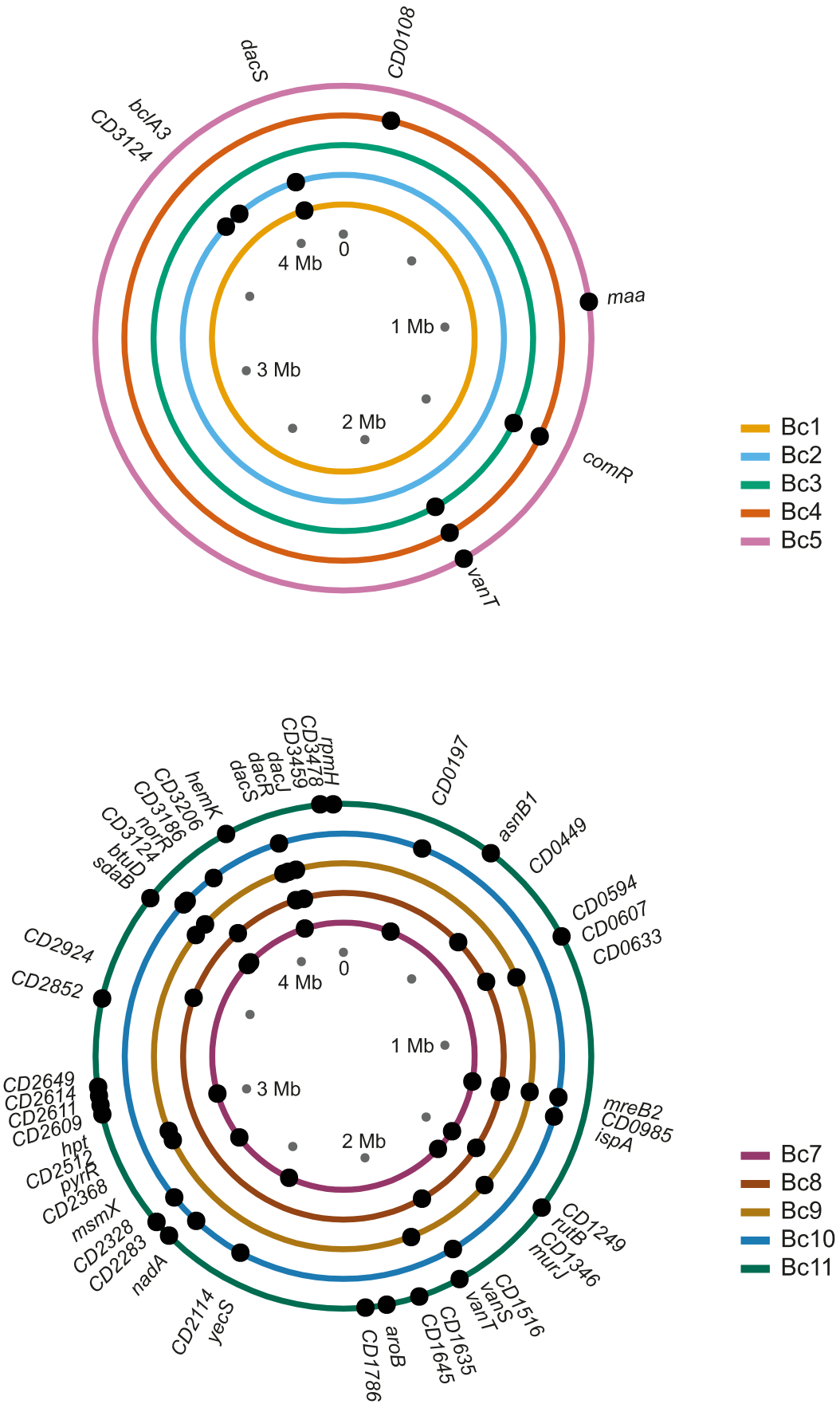


Figure 3.9 - Genetic Basis of Isolate Resistance

The chromosomal locations of non-synonymous, vancomycin-unique variants in WT (top) and hyper-mutating (bottom) derived *C. difficile* evolved isolates. Each concentric circle represents a single *C. difficile* genome, colour coded according to population as indicated in the key on the right. A full list of all variants shown here and including synonymous and intergenic mutations is included in Appendix V.

3.3.4 Population Dynamics in Evolving Populations

Having identified the presence of two major routes to resistance on an individual level, the contributions of these pathways to vancomycin resistance at the population level were assessed through pooled population sequencing. To understand the evolutionary dynamics driving the evolving populations, pooled population sequencing was performed at multiple time points. Whole population samples from all ten replicate lines were sequenced at three time points: passage 10, 20 and 30, to capture changes across the entire experimental evolution. Control populations, from matched-control replicate lines (passaged in absence of vancomycin) were also sequenced at passage 10, 20 and 30, to control for changes occurring due to laboratory adaptation. Short-read pooled genome sequencing was performed using the Illumina NovaSeq 6000 at SNPsaurus. Sequencing was performed at 250x coverage to enable identification of low-frequency mutations.

3.3.4.1 Creating an Optimised Pooled Sequencing Analysis Pipeline

Population sequencing data differs from isolate sequence data, since the major aim of population sequencing is to discern both the variants present, and their relative frequencies in the population. To get the richest possible population data, low frequency variants must therefore be identifiable. For evolutionary study, 5% frequency is the gold-standard aim. This means adjustments are necessary to enable calling of low frequency SNPs, without false positive (erroneous calls) or false negative (not calling a true positive) errors. The result of this is a frequency/coverage trade-off: the variant frequency it is possible to call is inversely scaled with coverage. Low frequency variant calling thus requires high levels of coverage across the genome.

A sequence analysis pipeline, optimised for pooled sequencing samples and adjustable for coverage level, was developed as part of this project, based on the isolate sequencing pipeline described above (section 3.3.3.1). The initial stages of the pooled sequencing analysis pipeline were identical to the isolate pipeline, with divergence arising after generation of mpileup files. Coverage across the genome was plotted using Bedtools (v2.30.0) (Quinlan and Hall, 2010) genomecov and map functions. Average genome-wide coverage across all samples was 322x, however average coverage varied from 22x to 641x in vancomycin evolved populations. This disparity in coverage created additional complexity for analysis, which was solved using sequencing simulations.

Sequencing data based on the R20291 genome was simulated at multiple coverage depths (80x, 100x, 300x), using InSilicoSeq (v1.5.4) (Gourlé et al., 2019), with defined SNPs seeded at 5% frequency. Simulated sequences were evaluated using the analysis pipeline, and a range of Varscan parameters were trialled to generate a set of parameters to accurately call variants at all coverage levels. In all cases, a minimum coverage setting of 80% of the average coverage across the genome resulted in identification of all seeded SNPs. As is the gold-standard for variant calling (Warden et al., 2014), a minimum of four reads were required to support a variant call. In order to achieve 5% frequency calling, the evolutionary gold-standard, with a minimum of four supporting reads and a minimum coverage of 80% of the average coverage, 100x coverage was required. This resulted in two sets of varscan parameters, dependent on average coverage across the genome (with 100x as the cut-off), to allow the lowest frequency calls possible with a strong evidence base. For samples with an average coverage across the genome of 100x or higher, the following Varscan (v2.4.3-1) mpileup2cns parameters were used: min-coverage 80; min-var-freq 0.05; p-value 0.05; variants 1; output-vcf 1 – allowing variant calls as low as 5% frequency. Conversely, samples with an average coverage across the genome of below 100x used the following Varscan (v2.4.3-1) mpileup2cns parameters: min-coverage 4; min-reads2 4; min-var-freq 0.05; p-value 0.05; variants 1; output-vcf 1.

As with high-coverage samples, a minimum of four supporting reads were required to call a variant, however the minimum variant frequency call possible was inversely scaled (4/coverage at position), allowing low frequency calls only with good evidence. Despite potentially losing some richness, this allowed the quality of variant calls to remain consistent across samples.

Due to the large number of variants across the populations, variants were filtered to focus on those which were most likely to have evolved in response to vancomycin selection. As with evolved isolate sequences, analysis focussed on variants which were observed in populations evolved with vancomycin, but not in matched-control populations (vancomycin-unique variants). Further filtering discarded variants that never reached above 10% frequency across the entire evolution, since these were unlikely to have large impacts on population resistance. Variants were also manually inspected ensure no variants remained the same frequency across all time points, indicating a false call.

3.3.4.2 Multiple KEGG Pathways are Affected by Vancomycin Resistance

Nonsynonymous vancomycin-unique mutations were visualised using KEGG colour mapper, to understand the distribution of mutations within functional classes. Mutations were clustered within 17 functional classes (Figure 3.10). The best represented classes were two component systems and ATP Binding Cassette (ABC) transporters, with six and four matched genes respectively. Genes involved in resistance and peptidoglycan biosynthesis were next best represented.

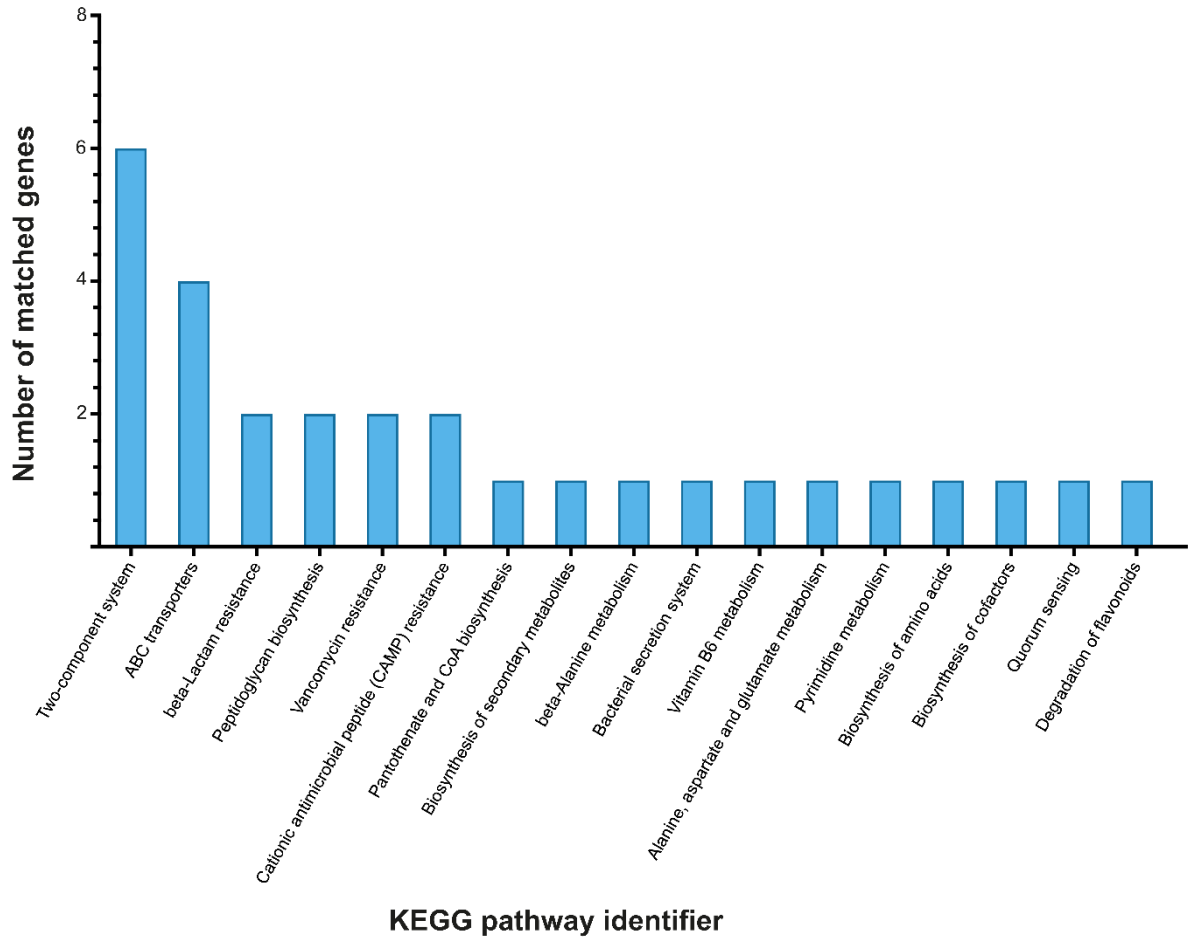


Figure 3.10 - KEGG Pathway Analysis of Genes Identified in Evolved Populations

Genes impacted by mutations during evolution (except those in the transiently hyper-mutating Bc1 P20) were visualised in KEGG colour mapper and assigned to cellular pathways. Two component systems and ABC transporters were the best-represented functional classes.

3.3.4.3 Alternate Evolutionary Pathways Display Different Dynamics

In total, discounting variants found in Bc1 P20, 535 vancomycin-unique variants were identified across all ten replicate populations throughout all time points in the evolution. Bc1 P20 was removed from the analysis, as this sample contained 520 additional mutations, which likely reflect random mutation due to the emergence of a spontaneous hyper-mutator phenotype (Figure 3.11a). This phenotype was transient – only 12 mutations were present in Bc1 P10, and only 4 were present in Bc1 P30 – however no high frequency mutations in DNA repair genes were identified as potentiators.

Variants were defined as fixed if they were present in more than 95% of the population, and remained so throughout the remaining time points of the evolution. P30 variants present in over 95% of the population were also included in analysis. Mutations which sweep through a population to reach fixation may be indicative of positive selection, especially if similar variants evolve in parallel across multiple populations. The population dynamics and fixation rates of the evolved populations were investigated, with particular focus on the two key pathways to resistance identified in the evolved isolates (Figure 3.11 and Figure 3.12). The *dacS* and *vanT* pathways showed highly contrasting selection dynamics: *dacS* mutations rapidly rose to high frequency, reaching fixation by P10. Interestingly, the same pattern was observed for *dacR* mutations in Bc7 and Bc10. At P10, the apparent population MICs ranged from 8-16 µg/mL, suggesting such variants are important for providing early, albeit lower-level, vancomycin resistance.

By contrast, mutations in *vanT* arose later and only reached fixation by P20 or P30. These mutations were preceded by other, high frequency, mutations at P10 that did not survive, being supplanted by *vanT* variants which presumably conferred higher levels of vancomycin resistance and/or fitness. In Bc3, the two preceding high frequency mutations (both T>TA) were very close together, separated

by only 7 bp, in an intergenic region downstream of *CD0482*, encoding a uridine kinase, and upstream of *glsA*, encoding a glutaminase. These mutations are outside of the likely *glsA* promoter region, but could affect its regulation. Interestingly, changes in glutamine metabolism have previously been linked to vancomycin resistance in *Staphylococcus aureus* (Cui et al., 2000). The single high frequency mutation in Bc5 at P10 was a nonsynonymous substitution in *CD3034*, which introduced a Gly255Asp mutation in the encoded D-hydantoinase, an enzyme which may play a role in D-amino acid synthesis. Taken together, this suggests *vanT* is unlikely to be required for first-step resistance. *vanT* mutations may provide higher-level vancomycin resistance, allowing supplantation of earlier mutations, or may require potentiating prerequisite mutations to arise first, however no secondary mutations common to all populations with *vanT* mutations were identified.

Beyond the two key pathways, interesting population dynamics were observed for other putative vancomycin resistance mechanisms. Like *dacS, CD3124* mutations arose early, reaching fixation by P10 in four of the ten evolved populations, again hinting this may be important for first-step, lower level vancomycin resistance. Mutations in *mreB2*, however, displayed no consistency in population dynamics, becoming fixed at P10 in Bc9, and P20 in the other populations. In *CD0985*, another gene in the elongasome cluster mutated in Bc7 and Bc8, mutations were slower to arise. Bc8 displayed an interesting selection dynamic here, since both *mreB2* and *CD0985* contained high frequency mutations at P20, however by P30, the *mreB2* variant was no longer present, leaving only the fixed mutation in *CD0985.* This suggests mutations in *CD0985* may provide increased resistance, or perhaps increased fitness, compared to *mreB2.* Intriguingly, despite arising in multiple populations, mutations in *vanS* were consistently low frequency in the populations they arose in, and never reached fixation – either remaining low frequency throughout the course of the evolution, or being lost from the populations (Figure 3.11f). This perhaps suggests *vanS* does not provide a large increase in vancomycin resistance, allowing selection to favour alternative mutations.

Overall, sequencing populations over time presents a valuable method to understand both the drivers of vancomycin resistance, and the comparative dynamics of alternative resistance pathways. Sequencing in this way also provides an explanation for the lack of homogeneity in MIC observed in isolates from the populations at the end of the evolution (Figure 3.2c), as multiple low frequency P30 mutations can be observed for each population (Figure 3.11 and Figure 3.12). This lack of homogeneity in the population suggests the potential to evolve resistance which is higher-level, more stable, or less costly.

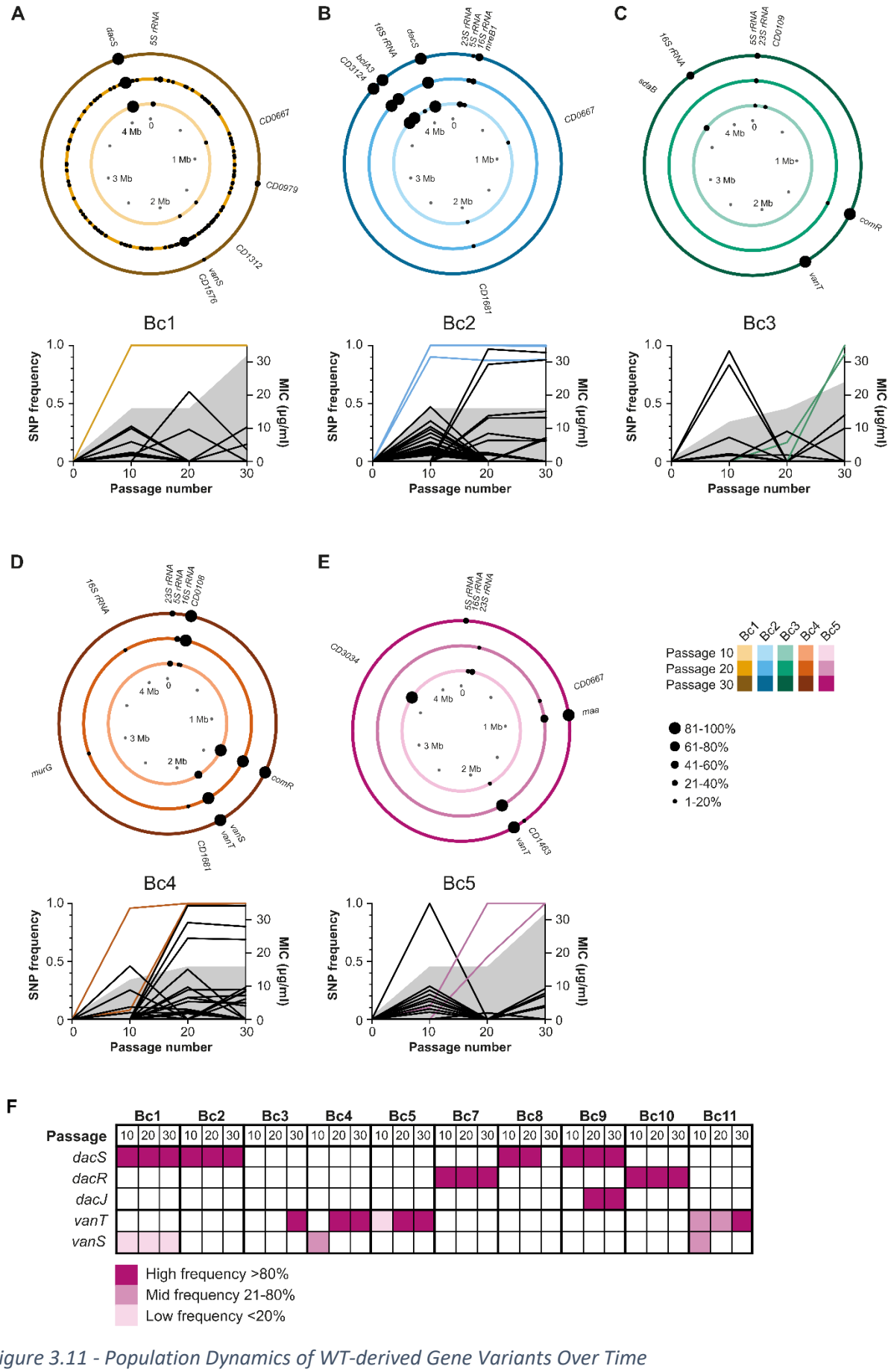


Figure 3.11 - Population Dynamics of WT-derived Gene Variants Over Time

Accumulation of variants in the WT-derived *C. difficile* lineages Bc1 (A), Bc2 (B), Bc3 (C), Bc4 (D) and Bc5 (E). Each circle plot represents the 4.2 Mb genome of a single evolving population after 10 (inner ring), 20 (middle ring) and 30 passages (outer ring), with the locations of non-synonymous within gene variants indicated with black circles and the penetrance of each mutation in the population indicated by the size of the circle. The line graphs show the frequency of all variants (intergenic, synonymous, non-synonymous, frameshifts and nonsense) in each population. The vancomycin MIC for each population is also indicated by the shaded region. Mutations also identified in the respective end point clone (Figure 3.9) are highlighted by the coloured lines. Note population Bc1 evolved an apparent hypermutator phenotype prior to P20, with 520 variants identified at that time point. For simplicity only variants present in P10 and P30 are labelled. A full list of all variants shown here, including those in Bc1 P20, is included in Appendix VI. (F) Relative frequencies and time of emergence of mutations in genes *dacS*, *dacR*, *dacJ*, *vanT* and *vanS* across all ten evolving populations.

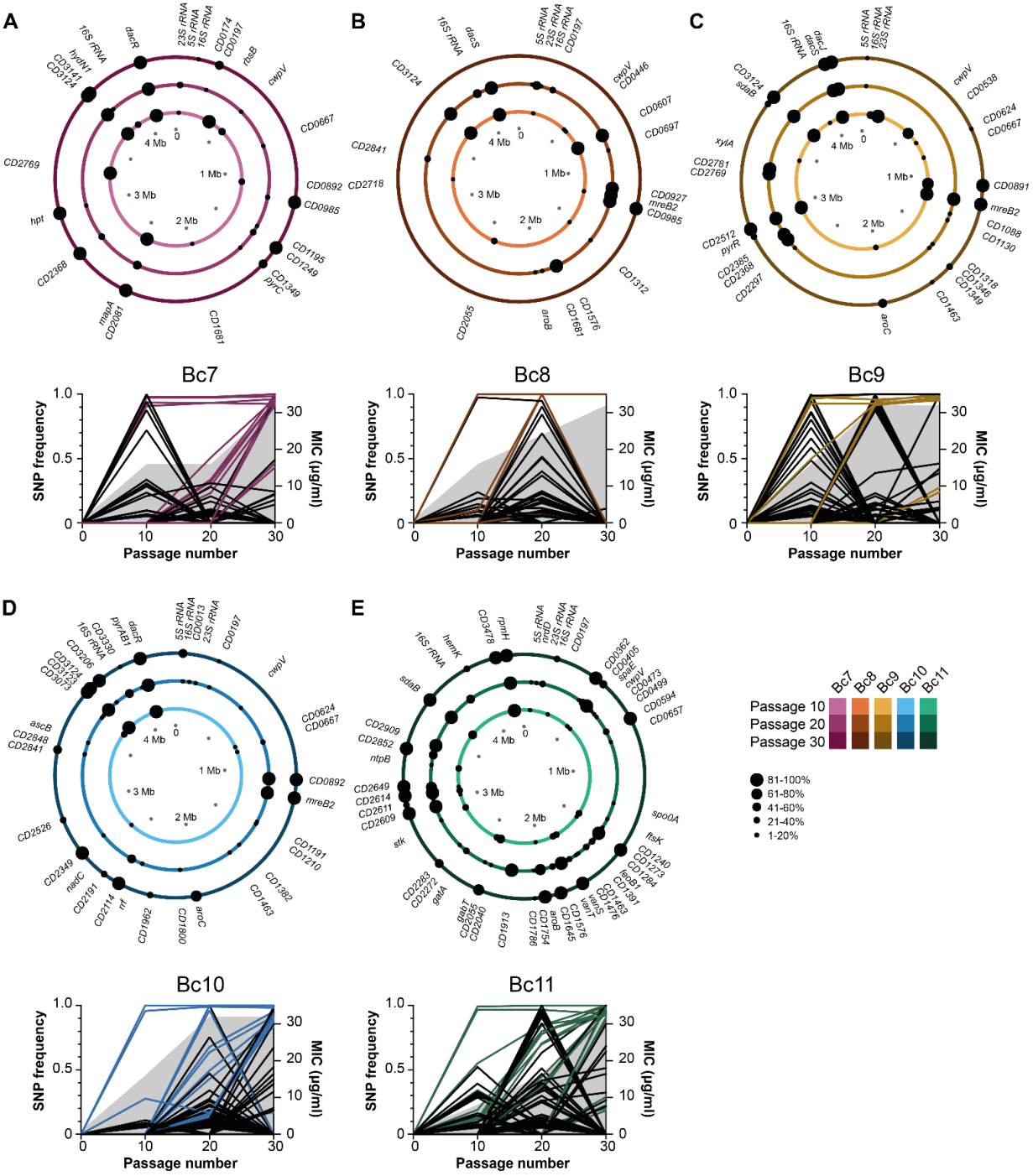


Figure 3.12 - Population Dynamics of Hyper-mutating Gene Variants Over Time

Accumulation of variants in the hyper-mutating *C. difficile* lineages Bc7 (A), Bc8 (B), Bc9 (C), Bc10 (D) and Bc11 (E). Each circle plot represents the 4.2 Mb genome of a single evolving population after 10 (inner ring), 20 (middle ring) and 30 passages (outer ring), with the locations of non-synonymous within gene variants indicated with black circles and the penetrance of each mutation in the population indicated by the size of the circle. The line graphs show the frequency of all variants (intergenic, synonymous, non-synonymous and nonsense) in each population. The vancomycin MIC for each population is also indicated by the shaded region. Mutations also identified in the respective end point clone (Figure 3.9) are highlighted by the coloured lines. A full list of all variants shown here is included in Appendix VI.

3.4 Discussion

Vancomycin is the current front line antibiotic for treatment of CDI. Despite vancomycin resistance being well established in other organisms (Ahmed and Baptiste, 2018; Cong et al., 2019; Dyrkell et al., 2021), relatively little evidence of vancomycin resistance in *C. difficile* has come to light. Whether this reflects an underlying constraint on resistance, such as a large associated fitness burden, or simply a lack of routine clinical monitoring, is unknown. Studying vancomycin resistance through experimental evolution presents a unique opportunity to unpick the routes and dynamics of resistance acquisition, without the complexities of strain divergence and long-term genetic drift which accompanies clinical isolate study. In this chapter, an extensive exploration of vancomycin resistance was attained through experimental evolution, phenotypic characterisation and whole genome sequencing. *C. difficile* was shown to evolve high-level resistance rapidly *in vitro*, primarily by two key mechanisms. Resistance was, however, accompanied by pleiotropic growth and sporulation defects.

3.4.1 Extending the Bounds of the Current Evolution

Vancomycin resistance evolved rapidly over 30 passages in all 10 replicate populations, reaching an MIC of 16-32x that of the ancestral WT. Sequencing these populations at multiple time points allowed visualisation of the population dynamics. An interesting observation, arising from population sequencing but also reflected in evolved isolate MIC measurements, was the striking heterogeneity of the populations at P30, the end of the evolution. Multiple sub-populations, with low- to mid-frequency mutations, were observed at P30 across all 10 replicates. This was mirrored in the evolved isolate MIC measurements, whereby isolates displayed variation in MICs, ranging from the apparent MIC of their respective populations, to 2- or 4-fold lower. This variation, coined

heteroresistance, is common in antibiotic resistant populations (Band and Weiss, 2019). Heteroresistance, the phenomenon describing the co-existence of subpopulations with a range of antibiotic susceptibilities, is however poorly understood. The presence of subpopulations with multiple degrees of vancomycin resistance, perhaps also with multiple levels of fitness, may be indicative of population adaptation, preserving both high-level and less-costly resistance. The continued evolution of these heterogeneous populations beyond P30, selecting for higher-level resistance to further understand mutation accumulation would make interesting further work. Higher-level resistance is evidently possible to achieve, having evolved unstably on three occasions in Bc9. Although, the unstable nature of high-level resistance observed during this evolution may reflect the extreme fitness costs required for such resistance levels. An alternative approach, therefore, would involve continued evolution to instead select for increased fitness, through accumulation of refining mutations. Compensatory evolution, the process whereby fitness costs may be ameliorated by additional mutations elsewhere in the genome (Kimura, 1985), would provide useful insight into the ability of *C. difficile* to evolve high-level, less-costly resistance to vancomycin. Indeed, examples of refinement and succession were already observed here, with early high frequency mutations conferring moderate increases in MIC being completely supplanted by later variants. Notably, compensation can occur in the presence or absence of the antibiotic, and confer a fitness advantage in either environment (Schulz zur Wiesch et al., 2010). This has been widely demonstrated across multiple species – for example, compensatory evolution of rifampin resistant *E. coli* in the absence of rifampin produced isolates with increased fitness, without depletion of rifampin resistance (Reynolds, 2000). The dynamics of such compensation in *C. difficile* thus merits future work.

3.4.2 Further Considerations of Fitness

Vancomycin resistance was accompanied by severe, albeit varied, fitness defects. A reduction in fitness was observed across all evolved isolates in one such measure of fitness, growth in rich media. Taken in conjunction with sequence data, it is clear that evolved isolates display reduced growth regardless of their route to resistance, suggesting either that all captured vancomycin resistance mechanisms negatively impact fitness; or that the observed reduction in fitness is due to accumulation of additional mutations, for example those which are non-vancomycin-unique, or those which affect cell metabolism and stress responses. Importantly, significant sporulation defects were also observed in four of the ten evolved isolates, with a complete absence of sporulation in Bc11. The basis of these genetic defects remains unclear, as sporulation is a complex and poorly understood process. Some of the evolved isolates contained frameshift mutations in genes found to be essential for sporulation, such as *pyrR* in Bc9 (Dembek et al., 2015). However this pattern cannot explain all of the defects observed. Interestingly, although the WT-derived evolved isolates Bc3, 4 and 5 possess the same vancomycin resistance mechanism, mutations in *vanT*, variation in sporulation efficiency was observed. The Bc4 evolved isolate displayed a reduction in sporulation efficiency, whilst Bc3 and Bc5 showed sporulation comparable to the WT control. The Bc4 evolved isolate did not display any increase in resistance which may account for the increased fitness burden, suggesting the presence of hitchhiking mutations giving rise to reduced fitness. Genetic hitchhiking describes the process in which neutral, or even deleterious, mutations rise to high frequency through linkage to beneficial mutations (Smith and Haigh, 1974). This means a mutation promoting resistance would be selected in the population, even if it arose in a clone containing other mutations, deleterious for fitness. In the case of the Bc4 evolved isolate, the observed mutation in *CD0108* may be the causative deleterious mutation, or perhaps a more general accumulation of non-vancomycin-unique mutations. Regardless of culpable mutations, the observed sporulation defects have wider implications for the survival of vancomycin resistant *C. difficile* isolates. Since sporulation

is a key part of the *C. difficile* lifecycle, and is essential for disease transmission, such defects would likely be an evolutionary dead-end if similar variants were to emerge in clinic (Deakin et al., 2012).

This chapter explored some of the key phenotypes arising from experimental evolution, to understand the costs of vancomycin resistance. Nevertheless, multiple aspects of fitness remain unexplored. Further investigations of fitness to include germination efficiencies, and toxin production, two more key features of *C. difficile* virulence, would be beneficial. A caveat of evolving a non-toxigenic strain, R20291Δ*PaLoc*, is the inability to directly measure the latter, however performing toxin assays in recapitulated strains would be helpful to understand whether acquisition of vancomycin resistance impacts the pathogenicity of *C. difficile*. Exploring the clinical pertinence of the observed fitness defects, through competition studies, would allow better conclusions to be made regarding the implications of these defects. Competition studies, in mice treated with vancomycin, would be the gold-standard approach to understand the survival of vancomycin resistant strains *in vivo* (Collins et al., 2015).

3.4.3 Wider Implications of Cross-resistance

Low-level teicoplanin cross-resistance, ranging from 2 to 4x the WT MIC, was observed across all ten evolved isolates. Teicoplanin, like vancomycin, has a mode of action involving binding the terminal D-Ala on the pentapeptide stem of nascent peptidoglycan. However unlike vancomycin, the lipoglycopeptide teicoplanin contains a hydrophobic moiety, allowing interaction with the membrane (Zeng et al., 2016). The comparatively low reduction in susceptibility, compared to that of vancomycin, suggests the existence of teicoplanin resistance pathways alternative to the vancomycin resistance mechanisms observed here. Interestingly, this low-level cross resistance was observed in the inverse direction in teicoplanin resistant *S. aureus,* which only displayed low-level

vancomycin resistance (Vaudaux et al., 2001). Although it is well documented that certain *van* operons can induce teicoplanin resistance phenotypes, explanations beyond this are lacking (Selim, 2022). Regardless, even low levels of cross-resistance have the potential to be clinically significant – reports on both vancomycin and metronidazole resistance in *C. difficile* suggest that even low reductions in susceptibility can significantly alter clinical cure rate (Eubank et al., 2024; Gonzales-Luna et al., 2021). Although teicoplanin is not routinely used for CDI, evidence suggests it is a strong candidate for treatment of severe CDI, displaying good efficacy and low recurrence rates (Popovic et al., 2018). The observed cross resistance therefore has significant implications regarding reducing the clinical repertoire for treatment of CDI.

Beyond teicoplanin, cross-resistance to other antibiotics remains unexplored. In *S. aureus*, evidence of cross-resistance between daptomycin and vancomycin has been observed (Thitiananpakorn et al., 2020). A broader exploration of vancomycin cross-resistance, with specific focus on antibiotics which are potent against *C. difficile*, would be beneficial. Generating a *C. difficile* vancomycin cross-resistance network, in a similar approach to that of Lázár and colleagues, could be instrumental for advising treatment regimes in clinical settings (Lázár et al., 2014).

Alternatively, collateral sensitivity – whereby resistance to one antibiotic results in sensitivity to another – would be another useful research avenue to further understand *C. difficile* vancomycin resistance (Roemhild and Andersson, 2021). This has been widely studied in other organisms – for example, tigecycline resistant *E. coli* display collateral sensitivity to nitrofurantoin (Roemhild et al., 2020). Collateral sensitivity resulting from vancomycin resistance has also been reported – in *Enterococcus faecium*, vancomycin resistance resulted in sensitivity to pleuromutilin antibiotics (Li et al., 2022). In mouse models, use of pleuromutilin antibiotics improved survival against vancomycin-resistant *E. faecium*, suggesting collateral sensitivity could be used to expand the arsenal of potential

treatments. As some, but not all, pleuromutilin antibiotics display activity against *C. difficile*, investigation of this phenomena would be interesting.

3.4.4 Pathways to Vancomycin Resistance

Sequencing the evolved isolates revealed two predominant pathways to vancomycin resistance, observed in all ten replicate lines. These pathways were centred around mutations in the *van* cluster (mainly mutations in *vanT*, encoding the serine/alanine racemase component of the cluster), or in the *dacJRS* gene cluster (encoding a two component system and a D,D-carboxypeptidase). Although around 85% of *C. difficile* strains are thought to carry the *van* cluster (Ammam et al., 2012), the majority of isolates are vancomycin-sensitive, which has historically sparked controversy regarding whether the *van* cluster in *C. difficile* is involved in resistance. Early reports suggested the *van* cluster, although active on a transcriptional level, does not affect cell wall composition (Ammam et al., 2013; Peltier et al., 2013). However, recent evolutionary work found mutations in the *vanSR* two component system led to de-repression of the *van* cluster, allowing constitutive expression of the *van* operon and reduced vancomycin susceptibility (Shen et al., 2020). These mutations were mirrored in vancomycin resistant clinical isolates, together confirming a role of this cluster in vancomycin resistance. Interestingly, mutations in *vanS* were only transiently present in the experimental evolution presented here, consistently remaining at low frequency across the three populations in which they were observed. This suggests only a minor role of *vanS* in the vancomycin resistance observed here. In contrast, mutations to *vanT* were common, fixing in four of ten replicate lines. The *vanT*-encoded serine/alanine racemase is comprised of two domains, an N-terminal membrane-bound domain, likely involved in L-Ser uptake, and a C-terminal racemase domain, which converts serine or alanine from L- to D- form (Meziane-Cherif et al., 2015). VanT likely contributes to vancomycin resistance by converting L-Ser to D-Ser, required for the production of D-Ser terminating pentapeptides in nascent peptidoglycan. Mutations in *vanT* alone have not before

been reported in relation to vancomycin resistance. The mechanism by which these mutations alter VanT function to confer vancomycin resistance in *C. difficile* therefore merits further study.

Mutations in the *dacJRS* cluster have not been previously associated with vancomycin resistance, but were observed in six of the ten evolved populations. *dacJRS* encodes a two component system (*dacSR*) and a putative D,D-carboxypeptidase (*dacJ*), suggesting a plausible mechanism of vancomycin resistance through removal of the terminal D-Ala residue in nascent peptidoglycan (Ghosh et al., 2008). This mechanism may provide "first-step" vancomycin resistance, since fixation of variant *dacS* alleles occurred by P10 in three populations (Bc1, 2 and 9), as did the identical *dacR*c.532A>G variant in two populations (Bc7 and 10). Although this parallelism suggests strong evolutionary selection under vancomycin pressures, such mutations have not yet been reported in clinical isolates (Kolte and Nübel, 2024). It would be interesting, however, to quantify the strength of selection imposed on this cluster, compared to other areas of the *C. difficile* genome, both in the evolved isolates presented here and in clinical isolates. This approach is common in evolutionary biology, and can be calculated in multiple ways (Cadzow et al., 2014).

Aside from the two key pathways identified, parallel evolution was observed across genes encoded within the elongasome gene cluster, *mreB2* and *CD0985.* Mutations in these two genes were observed across four of the ten replicate populations (Bc7, 8, 9 and 10). Mutations in genes responsible for directing peptidoglycan synthesis and maintenance are perhaps unsurprising, however, because of their function, both *mreB2* and *CD0985* are essential in *C. difficile*. This essentiality limits the further investigation of these genes by classical molecular methods. No evidence of the involvement of either gene in vancomycin resistance has been noted previously, however *mreB* was found to be mutated in daptomycin-resistant *Bacillus subtilis* (Hachmann et al., 2011; Tran et al., 2015). Importantly, recapitulation of *mreB* mutations had little effect on

daptomycin resistance, suggesting these mutations were compensatory, rather than resistance-determining. It is not out of the question, therefore, that the missense mutations observed in this evolution are also compensatory.

A mutation in *asnB1,* encoding an asparagine synthetase, was observed in Bc11. Population sequencing showed this mutation gradually rose to fixation throughout the evolution, suggesting it may be a refining mutation, not required for first-step resistance. AsnB1 is responsible for peptidoglycan amidation of A2pm, which controls the activity of an L,D-carboxypeptidase (Bernard et al., 2011). This enzyme has actually been implicated in vancomycin sensitivity in *C. difficile* (Ammam et al., 2020). AsnB1 was shown to be responsible for peptidoglycan amidation in the presence of vancomycin, which resulted in increased sensitivity, potentially via interfering with vancomycin resistance mechanisms. The mutation observed in Bc11 may therefore reduce the activity of AsnB1, leading to a reduction of peptidoglycan amidation in the presence of vancomycin to increase resistance. This putative mechanism of resistance could easily be tested through recapitulation of the observed *asnB1* mutation, along with measurement of MIC, enzyme activity assays, and peptidoglycan structural analysis in the presence and absence of vancomycin.

As with most experimental evolution studies, the focus of this project has primarily been on nonsynonymous, vancomycin-unique mutations. However, it is important not to forget the putative impacts of synonymous mutations on resistance. Synonymous mutations, by their very nature, do not lead to change in the encoded protein, however can affect gene expression through codon usage bias, mRNA structure modification, or mRNA stability alterations (Bailey et al., 2021). Codon usage bias, the non-uniform use of synonymous codons, has long since been appreciated, and the observation that essential genes contain higher frequencies of optimal codons is not new (Gouy and Gautier, 1982; Grantham et al., 1980). However, research into synonymous mutations as drivers of

adaptation is in its infancy. Recently, multiple experimental evolution studies have reported synonymous mutations as drivers of adaptive evolution, and even drivers of antibiotic resistance – a synonymous mutation repeatedly identified in the *ompK36* gene of carbapenem resistant *Klebsiella pneumoniae* was shown to reduce OmpK36 translation through alterations to the mRNA secondary structure. As OmpK36 promotes carbapenem influx, this synonymous mutation-driven gene knockdown was able to promote carbapenem resistance (Wong et al., 2022). Future investigations may therefore involve a deeper examination of genetic parallelism of synonymous mutations observed during the evolution presented here.

3.4.5 Clinical Significance of Observed Mutations

The clinical significance of the mutations observed during this experimental evolution remains an open question. So far, *dacS* mutations identical to those observed here have not been reported in publically available *C. difficile* isolate sequences (Kolte and Nübel, 2024). However, vancomycin has only been used as the front-line antibiotic for CDI since 2021 (NICE, 2021), meaning comparatively few genomes are available in which vancomycin is known to have been a significant selection pressure. Therefore, longitudinal observations would be needed over the coming years to monitor new *C. difficile* clinical isolates. This would uncover whether the mutations observed during this evolution arise in clinical isolates, allowing a true appreciation of the clinical significance of this work.

4 Exploring the Role of *dacJRS* in Vancomycin Resistance

4.1 Introduction

Two component systems (TCS) are ubiquitous across bacteria, and are an essential regulatory mechanism which modulate a multitude of fundamental cellular processes (Hirakawa et al., 2020). TCS allow bacteria to sense and respond to environmental stimuli, and regulate metabolism, virulence, and antibiotic resistance. Although not completely understood, it is clear that TCS are also prominent regulators of sporulation in *C. difficile* (Edwards and McBride, 2023, 2014). In their simplest form, TCS comprise of two proteins, a histidine kinase and their cognate response regulator. The histidine kinase is typically a homodimeric transmembrane protein, which consists of a sensor domain, typically extracellular, and a cytoplasmic signalling domain. The response regulator consists of a receiver domain and DNA-binding domain (Hirakawa et al., 2020; Yan et al., 2019). Signal transduction occurs when the histidine kinase senses a stimulus and auto-phosphorylates at a conserved histidine residue. This phosphoryl group is transferred to the response regulator, altering its conformation and thus DNA-binding affinity. Response regulators may be activators or repressors, modifying the expression of their regulons as such (Rajeev et al., 2020). Tens of thousands of TCS have now been identified in bacteria, including many permutations of the canonical TCS (Ulrich and Zhulin, 2010). For example, a single histidine kinase can regulate multiple response regulators, and multiple histidine kinases can regulate a single response regulator (Schaller et al., 2011). Many complex multistep phosphorelays have also been characterised – for example, the regulation of sporulation in *B. subtilis* involves five histidine kinases acting on a single response regulator, Spo0F,

which phosphorylates Spo0B, which in turn phosphorylates Spo0A, the global sporulation regulator (Piggot and Hilbert, 2004).

Since two component systems promote bacterial survival and adaptation by allowing them to respond to environmental stressors, their large contribution to antibiotic resistance, observed across bacterial species, is unsurprising. A huge number of antibiotic resistance-promoting TCS are now well characterised – for example, increased expression of the VraSR TCS, involved in the regulation of cell wall stress, was associated with daptomycin resistance in *S. aureus* (Mehta et al., 2012). TCS may also contribute to resistance by regulating antibiotic efflux – PhoPQ, a TCS responsible for regulation of an ABC transporter capable of tetracycline efflux, was associated with tetracycline resistance in *P. aeruginosa* (Chen and Duan, 2016). Additionally, TCS may control more targeted resistance mechanisms – in *P. aeruginosa*, the CreBC TCS was associated with high-level β-lactam resistance through induction of β-lactamase AmpC (Zamorano et al., 2014). Similarly, in vancomycin resistant *Enterococcus* species, the VanSR TCS promotes resistance through increased expression of the *van* operon genes, resulting in alteration of the vancomycin target (Evers and Courvalin, 1996). As discussed previously, *C. difficile* also possesses a homologous VanSR TCS, which can promote vancomycin resistance when constitutively expressed (Shen et al., 2020).

C. difficile encodes around 50 TCS, the majority of which remain uncharacterised, although recently, a few examples of antibiotic resistance-associated TCS have been identified. One such TCS, DraRS, was found to promote daptomycin resistance, likely mediated by induction of genes involved in cell envelope biogenesis and stress responses (Pannullo et al., 2023b). The HexRK TCS also promotes daptomycin resistance in *C. difficile*, through regulation of HexSDF proteins which alter the cell membrane composition (Pannullo et al., 2023a). Additionally, the WalRK TCS, which has been well-characterised in other species, was recently found to be essential for viability, and responsible for

cell envelope biogenesis, the downregulation of which resulted in increased vancomycin sensitivity in *C. difficile* (Müh et al., 2022).

In the previous chapter, experimental evolution coupled with whole genome sequencing was used to understand the pathways to vancomycin resistance in *C. difficile*. Exploring the population dynamics and genetic parallelism across replicate populations allowed identification of the previously uncharacterised TCS *dacRS,* and upstream gene *dacJ*, that were mutated in six of the ten evolved isolates, showing clear evidence of selection. In the evolved populations, mutations in *dacRS* always rose rapidly to fixation, suggesting a role in first-step vancomycin resistance. However, the individual contribution of this TCS to resistance, and its mechanism, was unknown.

4.2 Aims and Outcomes

This chapter describes the full characterisation the *dacJRS* gene cluster, including its contribution to vancomycin resistance, the fitness costs associated with this pathway, and the mechanism by which mutations in *dacJRS* result in increased resistance. To achieve this, *dacS* mutations observed in evolved isolates were recapitulated in the parental R20291Δ*PaLoc* background. The contribution of these mutations to resistance were tested using MIC assays, and the phenotypic effects of SNP carriage were assessed using growth curves and sporulation efficiency assays. The mechanism of *dacJRS*-mediated resistance was determined through qRT-PCR and fluorescence microscopy.

4.3 Results

4.3.1 The Contribution of *dacJRS* to Vancomycin Resistance

To assess the contribution of *dacJRS* to vancomycin resistance, mutations observed in the evolved isolates were precisely recapitulated in the parental R20291Δ*PaLoc* background. Recapitulation, as opposed to classical gene knockout, was used for these investigations, since the missense mutations observed were more subtle than crude deletions, calling for a more nuanced approach. *dacJRS* mutations were observed in six of the ten evolved isolates (Bc1,2,7,8,9 and 10). Mutations in *dacS*, the histidine kinase, were chosen for further study. Of these, the Bc1 and Bc8/9 mutations were chosen for recapitulation: illumina data suggested that the only nonsynonymous, vancomycin-unique mutation present in the Bc1 isolate occurred in *dacS* (Appendix V), and both Bc8 and Bc9 evolved isolates displayed an identical *dacS* SNP, showing parallel evolution to the SNP level, and thus strong selection for this allele. Table 4.1 summarises the *dacJRS* mutations present in evolved isolates.

Table 4.1 - dacJRS Mutations in Evolved Isolates

Barcode	**Gene**	**Mutation**
Bc1	*dacS*	*dacS*c.714G>T
Bc2	*dacS*	*dacS*c.798A>T
Bc7	*dacR*	*dacR*c.532A>G
Bc8	*dacS*	*dacS*c.548T>C
Bc9	*dacS*	*dacS*c.548T>C
	dacJ	*dacJ*c.478A>G
Bc10	*dacR*	*dacR*c.532A>G

4.3.1.1 Recapitulation of *dacS*c.714G>T and *dacS*c.548T>C

Mutagenesis of *C. difficile* was performed by allelic exchange. Specifically, 2 kb fragments comprising approximately 1 kb either side of the SNP, with the altered allele in the middle, were designed. A

further synonymous mutation was introduced near to the SNP, to remove a restriction enzyme site and simplify screening. The fragments, flanked by BamHI and SacI sites, were synthesised by Genewiz and inserted into pUC-GW-Kan vectors. The fragments were then subcloned into pJAK112 vectors (linearised via BamHI/SacI restriction digest), resulting in the constructs pJEB026 and pJEB019 (Bc1 and Bc8/9, respectively). The new constructs were transformed and conjugated as previously described. Colonies were screened for the new SNP using PCR, followed by diagnostic digest. Removal of a restriction site in the mutants resulted in an altered fragment profile that was readily distinguishable from the WT. Recapitulation of the SNPs was confirmed using Sanger sequencing, generating strains R20291Δ*PaLoc dacS*c.714G>T (Bc1) and *R20291*Δ*PaLoc dacS*c.548T>C (Bc8/9). Table 4.2 summarises the construction of plasmids for strains relevant to this chapter.

Table 4.2 - Strains and Constructs Relevant to This Chapter

Strain	**Construction**	**Plasmid**
R20291Δ*PaLoc dacS*c.548T>C	Fragment designed to contain the *dacS*c.548T>C point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB019
R20291Δ*PaLoc dacS*c.714G>T	Fragment designed to contain the *dacS*c.714G>T point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB026
R20291Δ*PaLoc dacS*c.548T>C *vanT*c.347G>A	Fragments containing point mutations synthesised by Genewiz, followed by ligation into linearised pJAK112. Double mutant generated by addition of pJEB008 to R20291Δ*PaLoc dacS*c.548T>C.	pJEB019, pJEB008
Bc1Δ*dacJ*	Fragment designed to contain 1 kb regions either side of *dacJ*. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112. Strain and plasmid made by Lucy Thompson.	pLMT003
R20291Δ*PaLoc* Δ*dacRS*	Fragment designed to contain 1 kb regions either side of *dacRS*. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112. Strain and plasmid made by Lucy Thompson.	pLMT002

4.3.1.2 *dacS* SNPs Contribute to Vancomycin Resistance

To assess whether mutations in *dacS* alone are sufficient to confer vancomycin resistance, vancomycin MICs were performed for recapitulated strains using standard agar dilution. Both R20291Δ*PaLoc dacS*c.714G>T (Bc1) and R20291Δ*PaLoc dacS*c.548T>C (Bc8/9) showed an MIC of 4 µg/mL, 4-fold higher than the parental R20291Δ*PaLoc* (Figure 4.1). This confirmed that *dacS* plays a significant role in vancomycin resistance in *C. difficile*, and doesn't require the presence of additional mutations or prerequisites. Together with the population sequencing data, this confirms the involvement of the *dacS* pathway in first-step resistance.

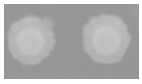
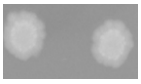
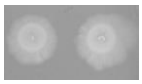
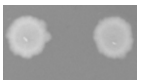
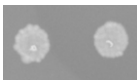
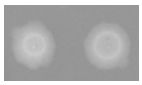
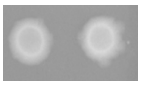
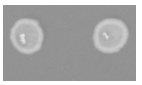
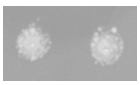
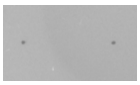
Isolate	Vancomycin (µg/mL)					
	0	0.5	1	2	4	MIC
R20291Δ*Paloc*						1
R20291Δ*PaLoc dacS*c.714G>T						4
*R20291ΔPaLoc dacS*c.548T>C						4

Figure 4.1 - dacS Mutations Result in Vancomycin Resistance

Vancomycin MICs of R20291Δ*PaLoc*, R20291Δ*PaLoc dacS*c.714G>T and R20291Δ*PaLoc dacS*c.548T>C. MICs determined by agar dilution on BHI. Assays were performed in biological triplicate and technical duplicate. Shown is a single representative biological replicate.

4.3.2 Phenotypic Analysis of Recapitulated Strains

To understand how the *dacJRS* pathway contributes to the loss of fitness observed in evolved isolates, phenotypic assessments of growth and sporulation were performed.

4.3.2.1 *dacS* Mutations Result in Growth Defects

To evaluate growth of the *dacS* recapitulated strains, R20291Δ*PaLoc dacS*c.714G>T and R20291Δ*PaLoc dacS*c.548T>C were assayed in rich media as described in the previous chapter, and compared with the parental R20291Δ*PaLoc* control (Figure 4.2a). Again, AUC was used to statistically compare growth of each of each strain with the control. Both *dacS* SNP strains displayed significant growth defects, with longer generation times and reduced AUC. R20291Δ*PaLoc dacS*c.714G>T showed a growth profile similar to that of the Bc1 evolved isolate, suggesting the *dacS* SNP is a large contributor to the Bc1 phenotype. Interestingly, the growth defects in evolved isolates Bc8 and Bc9, although significant, were much less prominent than those displayed by R20291Δ*PaLoc dacS*c.548T>C, indicating the action of compensatory evolution in these evolved isolates to reduce the fitness burden of *dacS* SNP carriage. This phenomena happened in both Bc8 and Bc9 evolved isolates, but not Bc1, providing further evidence of accelerated evolution in these hyper-mutating lines.

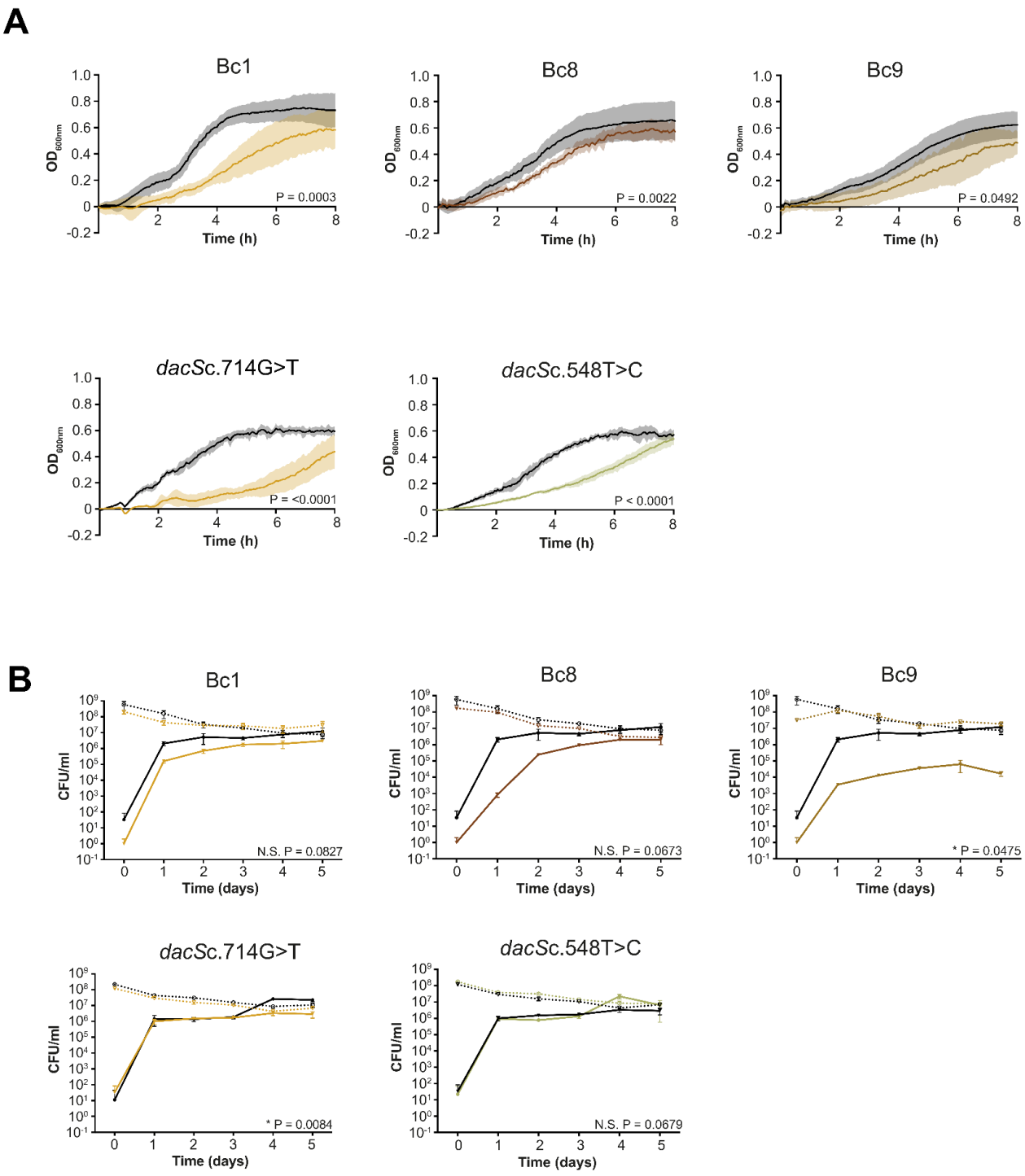


Figure 4.2 - Growth and Sporulation of dacS Recapitulated Strains

(A) Growth over time in rich media (TY broth) was evaluated by measuring OD at 600 nm in a 96 well microplate spectrometer. Growth of R20291Δ*PaLoc dacS*c.714G>T and R20291Δ*PaLoc dacS*c.548T>C shown below the evolved isolates (from Figure 3.3) from which the mutations were derived. Each recapitulated strain is compared to R20291Δ*PaLoc* (black lines). The mean and standard deviation of repeats, assayed in biological and technical triplicate, are presented. For each strain, area under the curve was determined using the GrowthCurver R package. AUC was compared using Student's t-tests with Welch's correction, with the P value shown on each graph. (B) Sporulation efficiencies of R20291Δ*PaLoc dacS*c.714G>T and R20291Δ*PaLoc dacS*c.548T>C (coloured lines) compared to the parental R20291Δ*PaLoc* (black lines). Sporulation efficiencies of evolved isolates from Figure 3.4 are

also shown. Stationary phase cultures were incubated for 5 days with samples taken daily to enumerate total colony forming units (CFUs, dotted lines) and spores (solid lines), following incubation at 65°C for 30 min to kill vegetative cells. Shown are the mean and standard deviations of samples assayed in biological and technical triplicate. For each strain, spore CFU area under the curve was determined using Graphpad Prism and these were compared using Dunnett's T3 multiple comparisons test with the adjusted P value shown on each graph. * = significant difference, N.S. = not significant.

4.3.2.2 *dacS* Mutations Do Not Alter Sporulation Efficiencies

The impact of *dacS* SNPs on sporulation efficiency was assessed as described in the previous chapter, using total and sporulation counts over a period of 5 days. The sporulation efficiencies of R20291Δ*PaLoc dacS*c.714G>T and R20291Δ*PaLoc dacS*c.548T>C were compared to R20291Δ*PaLoc* (Figure 4.2b). Neither *dacS* SNP-containing strain resulted in visual differences in sporulation rate or efficiency. Although the difference in sporulation efficiency for R20291Δ*PaLoc dacS*c.714G>T was statistically significant, it was clear there was no biologically significant difference in sporulation, highlighting some of the shortfalls in statistical methods for this data type. This data suggests the reduction in sporulation rate (Bc8) and efficiency (Bc9) observed in the evolved isolates was due to other mutations, and not *dacS*.

Taken together, these assays show that *dacS* mutations partially explain the phenotypes observed in evolved isolates, since they do contribute to fitness defects in terms of growth, but do not affect sporulation efficiency.

4.3.2.3 *dacS* Mutations Do Not Contribute to Teicoplanin Cross-Resistance

As a 4-fold increase in teicoplanin resistance was observed in the evolved isolates from which the *dacS* recapitulated SNPs derive, the teicoplanin MICs for R20291Δ*PaLoc dacS*c.714G>T and R20291Δ*PaLoc dacS*c.548T>C were measured, to investigate whether *dacS* mutations result in antibiotic cross-resistance. Both *dacS* recapitulated SNP strains showed an MIC of 0.25 µg/mL, identical to that of the R20291Δ*PaLoc* control. This suggests not only that the *dacS* pathway to

vancomycin resistance does not contribute to teicoplanin cross-resistance, but also that the cross-resistance observed in the evolved isolates is due to a further, thus far unknown, pathway.

4.3.3 dacS Does Not Act on the *van* Cluster

Through experimental evolution and genome sequencing, two key pathways to vancomycin resistance, the *van* cluster and *dacJRS*, were identified. Analysis of population dynamics showed clearly that although the two pathways evolved across all ten populations, they were largely mutually exclusive – although mutations in *dacS* and *vanS* coexisted in a small number of cases, there was no evidence of *dacJRS*/*vanT* co-occurrence across the evolved populations. Since both pathways displayed high levels of parallel evolution, this exclusivity raised questions regarding the potential cross-actions of the two mechanisms. The hypothesis that this mutual exclusivity arose from downstream convergence of the *dacS*/*van* pathways was therefore investigated.

4.3.3.1 No Synergy Observed for *dacS* and *vanT* Mutations

To investigate whether the *dacS* and *van* cluster pathways were cross-acting, a strain containing both *dacS* (Bc1) and *vanT* (Bc5) recapitulated SNPs (R20291Δ*PaLoc dacS*c.548T>C *vanT*c.347G>A) was generated, the construction of which is summarised in Table 4.2. If mutations in *dacS* and *vanT* were indeed involved in different pathways, targeting separate mechanisms, then additive (combined effects are summed), or even synergistic (combined effects are greater than the sum of both), interactions would be expected. Alternatively, if *dacS* and *vanT* pathways had precisely the same target, the effects of both would be unlikely to differ from that of either alone. It was hypothesised either of these eventualities would be reflected in the vancomycin MIC of R20291Δ*PaLoc dacS*c.548T>C *vanT*c.347G>A. No positive interaction, additive or synergistic, was observed in the R20291Δ*PaLoc dacS*c.548T>C *vanT*c.347G>A strain. The MIC was the same as that of R20291Δ*PaLoc dacS*c.548T>C, at 4 μg/mL. This lack of interaction may suggest the two pathways do indeed overlap, however, since the MIC was identical to that of the *dacS* SNP alone, the possibility

that *vanT* was simply not contributing to resistance in this setting could not be ruled out. Another approach to investigate the potential cross-action of the two pathways was therefore explored.

4.3.3.2 *dacS* Mutations Do Not Alter Expression of the *van* Cluster

Since the *van* cluster is well-characterised, and has been confirmed as a causative mechanism of vancomycin resistance in *C. difficile* (Shen et al., 2020), the *dacS* and *vanT* pathways may be regarded as truly separate mechanisms if DacS does not target the *van* cluster. To investigate whether mutations in *dacS* result in increased expression of genes in the *van* cluster, qRT-PCR was performed.

qRT-PCR was used to compare R20291Δ*PaLoc* and R20291Δ*PaLoc dacS*c.714G>T, in the presence (0.5x MIC) and absence of vancomycin. mRNA copy number was assessed against an absolute copy number control using serial dilution of the plasmid pJEB032, which contained a single copy of each target gene fragment of the *van* cluster (*vanR, vanS, vanG, vanXY, vanT*) and housekeeping gene *rpoA*. The copy numbers were standardised using the *rpoA* housekeeping gene, and results were expressed as copies per 1000 copies of *rpoA*.

No difference in expression between R20291Δ*PaLoc* and R20291Δ*PaLoc dacS*c.714G>T was observed for any of the *van* cluster genes, in either vancomycin condition (Figure 4.3). This showed that mutations in *dacS* do not alter expression of the *van* cluster, therefore confirming that the *dacS* pathway is novel, acting independently to the *van* cluster. Although the reason for the mutual exclusivity of these two mechanisms is still unidentified, and the downstream mechanisms of the *dacS* and *vanT* pathways may overlap, this work demonstrates the two key pathways to vancomycin resistance observed in the experimental evolution are fundamentally different.

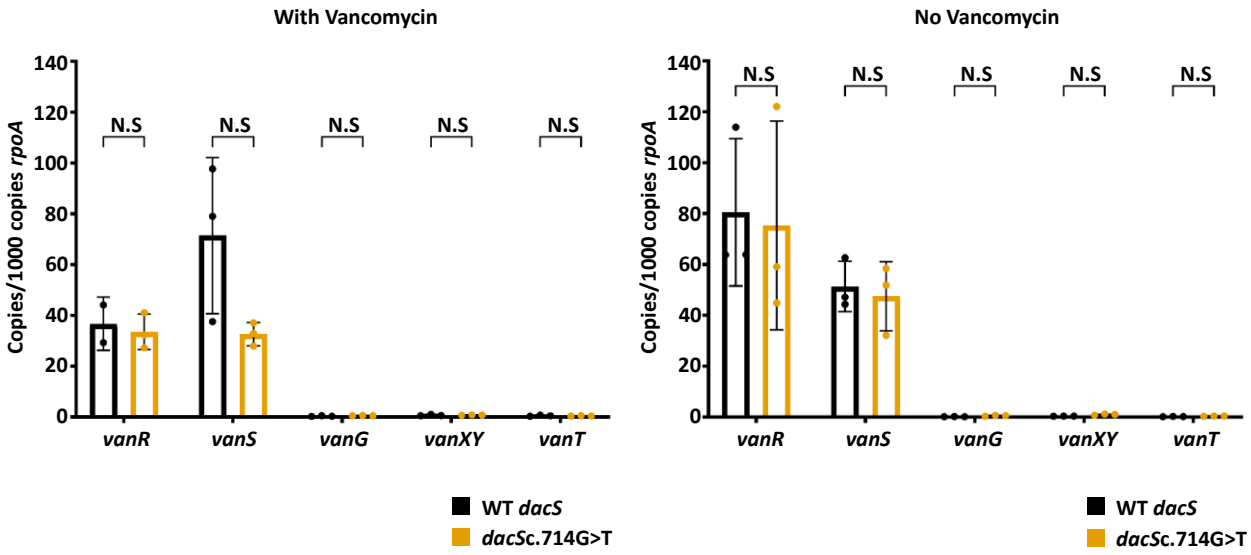


Figure 4.3 - dacS Does Not Alter Expression of the van Cluster

qRT-PCR analysis of the *van* operon genes, and TCS *vanSR*, in R20291Δ*PaLoc* (black) and R20291Δ*PaLoc dacS*c.714G>T (yellow). Exact copy numbers, normalised relative to the house-keeping gene *rpoA,* are shown. qRT-PCR assays were performed in biological and technical triplicate. Statistical significance was calculated using a two-way ANOVA with the Tukey post-hoc test. No differences were significant.

4.3.4 *dacS* Mutations Result in Overexpression of *dacJ*

After determining that mutations in *dacS* do not result in vancomycin resistance by acting on the *van* cluster, the mechanism of resistance associated with *dacS* was examined.

4.3.4.1 The *dacJRS* Gene Cluster

The genomic organisation of the *dacJRS* cluster was visualised using Geneious (v7.1.9 http://www.geneious.com/). Promoters were also predicted using previous global transcription site mapping (Fuchs et al., 2021). The *dacJRS* cluster consists of *dacRS*, the TCS histidine kinase and response regulator, and the upstream *dacJ,* controlled by a separate promoter (Figure 4.4a). As *dacJ* encodes a putative D,D-carboxypeptidase, which is likely involved in peptidoglycan modification through removal of the terminal D-Ala in nascent peptidoglycan, it was hypothesised that DacRS acts on DacJ.

4.3.4.2 The DacS Histidine Kinase

Structural prediction of the *dacS*-encoded histidine kinase was performed using AlphaFold to understand potential impacts of the missense mutations observed (Jumper et al., 2021). DacS was modelled as a homodimer, as histidine kinases often exist in this formation, yielding a plausible structural model (Figure 4.4b). Both R20291Δ*PaLoc dacS*c.714G>T (Bc1) and *R20291*Δ*PaLoc dacS*c.548T>C (Bc8/9) mutations alter the DacS cytoplasmic domain - Glu238Asp (Bc1) within the predicted catalytic ATPase domain, and Val183Ala (Bc8/9) within the dimerization and histidine phosphorylation domain, suggesting the mutations impact signal transduction as opposed to sensing. Although the direct impacts of these mutations on DacS function were not clear, the possibility that the mutations in *dacS* led to altered expression of *dacJ* was examined.

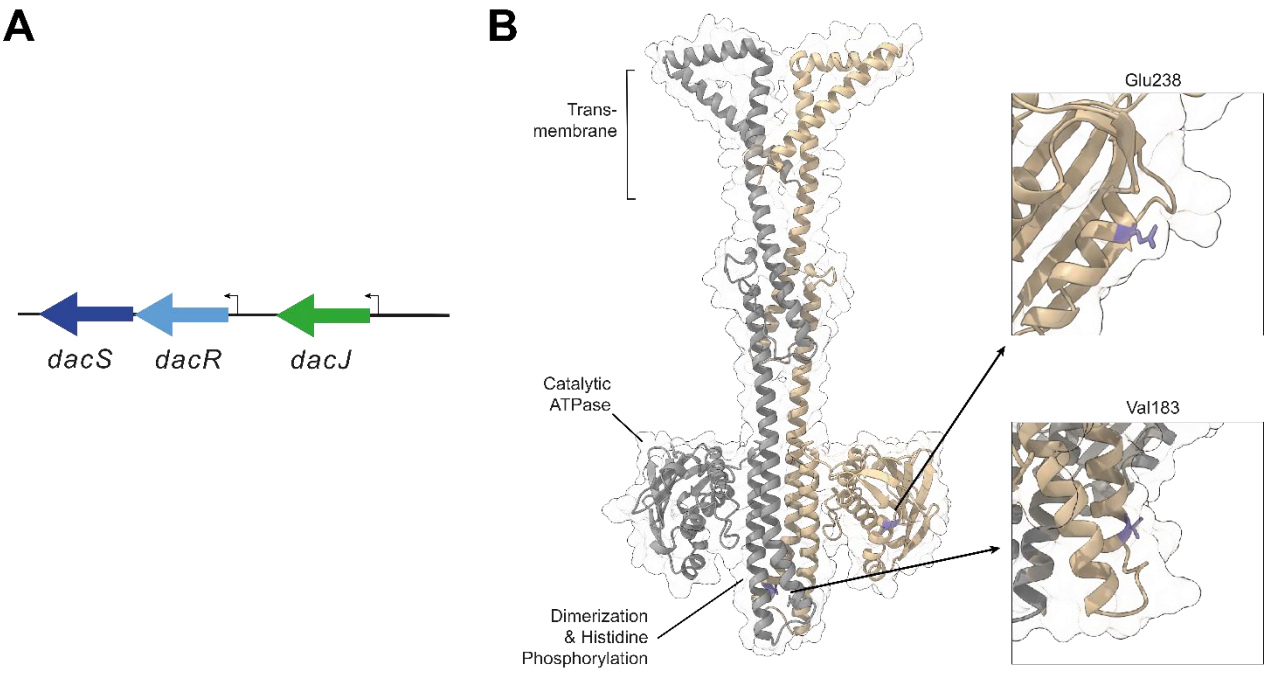


Figure 4.4 - dacJRS Gene Cluster Organisation and the DacS Histidine Kinase

(A) Genomic organisation of the *dacJRS* cluster. (B) AlphaFold model of DacS as a dimer (Jumper et al., 2021). The transmembrane domains were identified using DeepTMHMM (Hallgren et al., 2022) and the Catalytic ATPase and Dimerization and Histidine Phosphorylation domains were predicted

using InterProScan (Jones et al., 2014). The locations of Val183 and Glu238 are highlighted in purple on one chain.

4.3.4.3 *dacS* SNPs Result in *dacJ* Overexpression

qRT-PCR was used to investigate whether the observed mutations in *dacS* resulted in altered expression of the *dacJRS* gene cluster. RNA from R20291Δ*PaLoc dacS*c.714G>T (Bc1), *R20291*Δ*PaLoc dacS*c.548T>C (Bc8/9), and the R20291Δ*PaLoc* control in the presence (0.5x MIC) and absence of vancomycin was extracted, and expression of *dacJRS* and the housekeeping gene *rpoA* were measured. Expression was quantified as described before (4.3.3.2), whereby a qRT-PCR control plasmid, pJEB029, containing a single copy of each target gene fragment was used to generate a standard curve from which the exact copy number of target transcripts could be determined. The copy numbers were again standardised using the *rpoA* housekeeping gene, and results were expressed as copies per 1000 copies of *rpoA*. Differences in expression were then statistically analysed using a 2-way ANOVA.

Both *dacS* SNPs resulted in increased expression of all genes in the *dacJRS* gene cluster, showing the *dacRS* TCS acts on both *dacRS* and *dacJ* promoters, to regulate both itself and *dacJ* (Figure 4.5a). This increase was prominent in both the presence and absence of vancomycin, suggesting either that *dacS* SNPs result in constitutive expression of *dacJRS,* or that *dacRS* responds to an unknown signal. Although there was a significant increase across all genes for both *dacS* SNP strains, the magnitude of change varied from 4.8 fold (for *dacS* in the presence of vancomycin in the R20291Δ*PaLoc dacS*c.714G>T strain) to 94.6 fold (for *dacJ* in the presence of vancomycin in the R20291Δ*PaLoc dacS*c.548T>C strain).

This data suggests mutations in *dacS* result in vancomycin resistance through overexpression of *dacJ*. Overexpression of the putative DacJ D,D-carboxypeptidase would likely reduce vancomycin sensitivity via removal of the terminal D-Ala from nascent peptidoglycan, the vancomycin target.

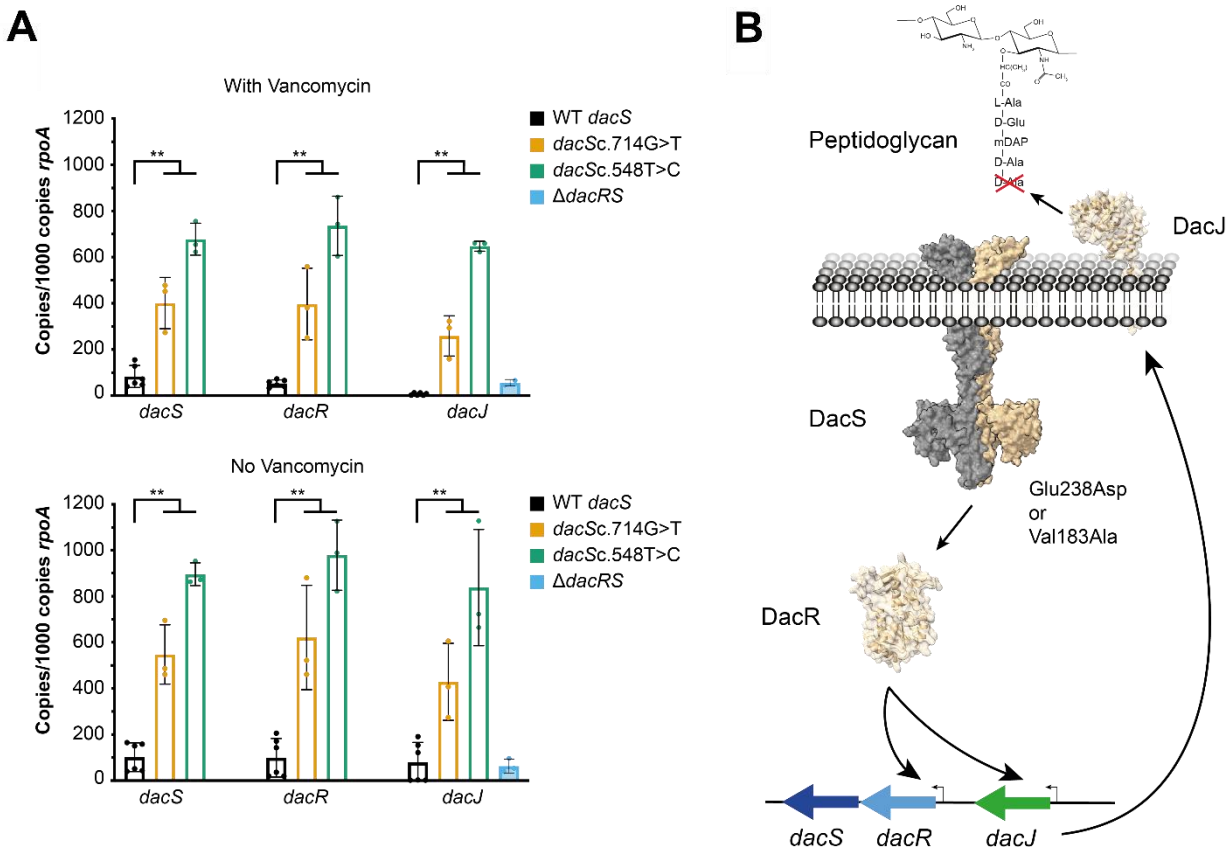


Figure 4.5 - dacS Mutations Lead to Increased Expression of dacJRS

(A) qRT-PCR analysis of *dacJRS* expression in R20291Δ*PaLoc* (black bars), R20291Δ*PaLoc dacS*c.714G>T (yellow bars), R20291Δ*PaLoc dacS*c.548T>C (green bars) and R20291Δ*PaLoc*Δ*dacRS* (blue bars), in the presence (0.5x MIC) and absence of vancomycin. Exact copy numbers, normalised relative to the house-keeping gene *rpoA,* are shown. Assays were performed in biological and technical triplicate. Statistical significance was calculated using a two-way ANOVA with the Tukey-Kramer test, ** = P<0.001. (B) Schematic of the consequences of *dacS* SNPs. Lack of *dacJ* upregulation in R20291Δ*PaLoc*Δ*dacRS* suggests that phosphorylated-DacR acts as an activator of the two promoters in the *dacJRS* cluster, with DacS Glu238Asp or Val183Ala substitutions constitutively activating the function of the TCS respectively. The consequence is over-expression of DacJ which is then translocated to the cell surface where it can cleave the terminal D-Ala residue in nascent peptidoglycan, thereby preventing vancomycin binding.

4.3.4.4 *dacS* is a Positive Regulator of *dacJ*

Through qRT-PCR, the DacRS TCS was shown to regulate itself and *dacJ*. However, whether DacRS was an activator or repressor of expression – and whether the observed *dacS* mutations resulted in increased *dacJRS* expression through activation of an activator, or deactivation of a repressor – was unknown. Therefore, to investigate how *dacS* SNPs result in increased expression of *dacJRS*, a *dacRS* knockout, R20291Δ*PaLoc*Δ*dacRS,* was constructed by allelic exchange as summarised in Table 4.2. qRT-PCR was used to measure expression of *dacJ* in R20291Δ*PaLoc*Δ*dacRS* and R20291Δ*PaLoc* as described above (section 4.3.4.3). If DacR were a repressor, Δ*dacRS* would have the same effect as the *dacS* SNPs, removing the repression to increase expression of *dacJ*. However, if DacR were an activator, Δ*dacRS* would remove the activation, resulting in little *dacJ* expression. No increase in *dacJ* expression was observed in the *dacRS* knockout, confirming the role of DacRS as a positive regulator of *dacJ* (and likely *dacRS*) transcription (Figure 4.5a). This suggests the *dacS* SNPs increase the activity of the DacRS TCS, leading to increased *dacJRS* expression (Figure 4.5b). Work relating to Δ*dacRS* was performed by Lucy Thompson (unpublished master's dissertation).

4.3.4.5 *dacS* SNPs Result in Reduced Vancomycin Binding

The role of DacJ as a D,D-carboxypeptidase could not be directly assayed, since overexpression of *dacJ* was lethal in *E. coli*, possibly due to the inhibition of cell wall cross-linking from DacJ carboxypeptidase activity (data not shown). To confirm that vancomycin resistance in the *dacS* recapitulated SNP mutants was due to overexpression of *dacJ*, resulting in removal of terminal D-Ala residues from peptidoglycan precursors and thus reduced abundance of D-Ala D-Ala vancomycin binding sites, vancomycin binding was visualised. It was hypothesised that *dacS* recapitulated SNP mutants would display reduced vancomycin binding due to *dacJ* overexpression. Vancomycin-BODIPY is a vancomycin analogue containing a single BODIPY fluorescent dye molecule per vancomycin molecule, meaning vancomycin binding to the cell wall can be visualised using fluorescence microscopy. Vancomycin-BODIPY binding is typically observed at the mid-cell, the site of maximal peptidoglycan synthesis, due to the abundance of pentapeptide precursors at this site.

Log phase cells were incubated with vancomycin-BODIPY, fixed, and washed before imaging using phase contrast microscopy with a GFP filter. The vancomycin labelling of R20291Δ*PaLoc dacS*c.714G>T (Bc1) and R20291Δ*PaLoc dacS*c.548T>C (Bc8/9) was compared to the R20291Δ*PaLoc* control. An unlabelled R20291Δ*PaLoc* control was also included.

The parental R20291Δ*PaLoc* strain displayed clear mid-cell vancomycin-BODIPY labelling, consistent with the presence of pentapeptide precursors (Figure 4.6a). Conversely, the *dacS* recapitulated SNP mutants R20291Δ*PaLoc dacS*c.714G>T and R20291Δ*PaLoc dacS*c.548T>C showed no evidence of vancomycin binding, demonstrating that overexpression of *dacJ* in these strains results in a depletion of D-Ala D-Ala vancomycin binding sites.

To quantify vancomycin binding in the R20291Δ*PaLoc* and *dacS* recapitulated SNP strains, fluorescence intensity profiles were collected for each cell across the long axis. After normalisation to the cell's background fluorescence, the maximal intensity of all imaged cells was plotted for each strain. The quantification mirrored the observations from the raw images – a clear increase in maximum intensity for the R20291Δ*PaLoc*, and profiles similar to the unlabelled control for the *dacS* recapitulated strains (Figure 4.6b). Fluorescence quantification was performed by Anne Williams.

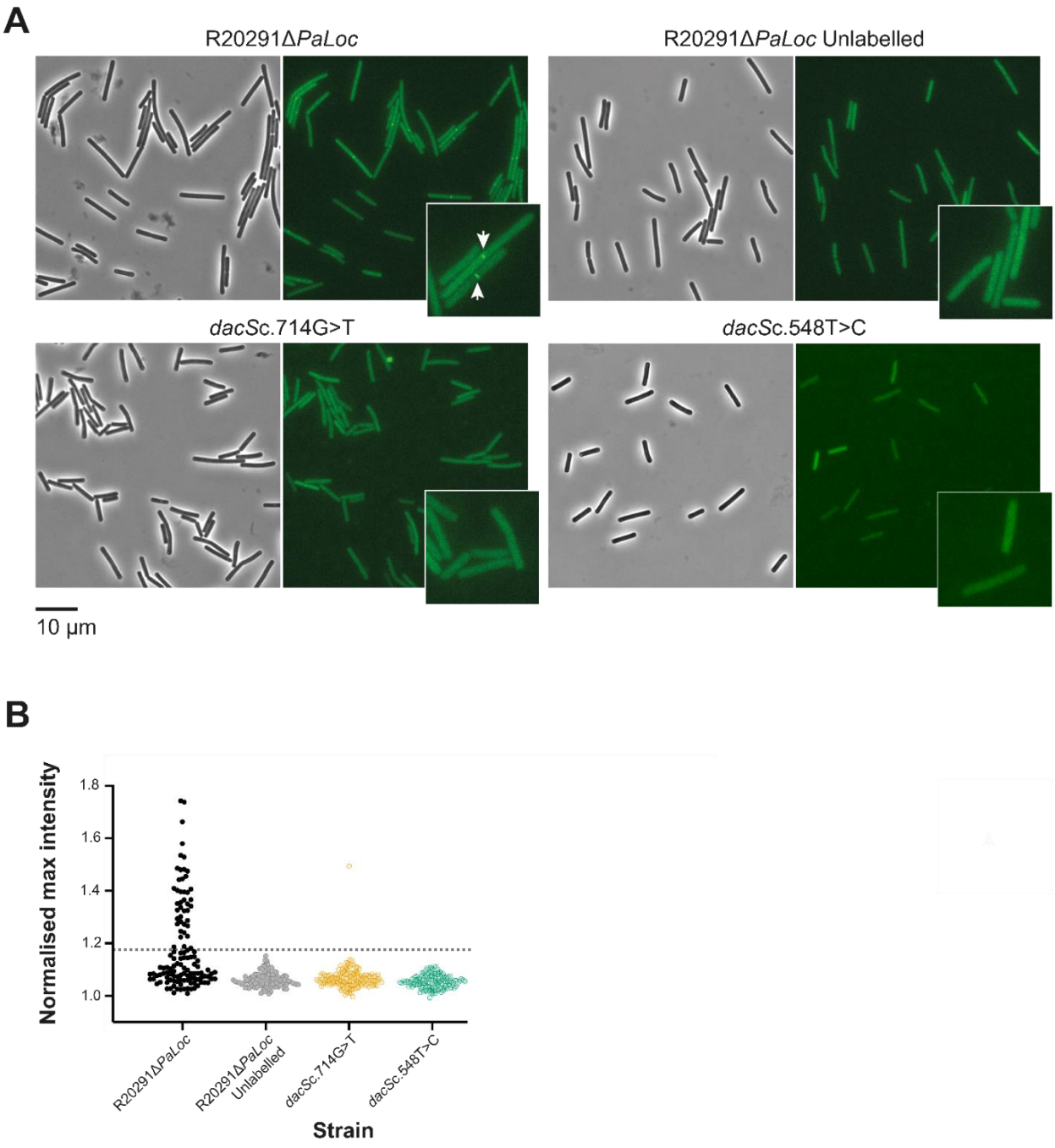


Figure 4.6 - DacJ Activity Reduces Vancomycin Binding Sites in the Cell Wall

(A) Representative images of R20291Δ*PaLoc*, R20291Δ*PaLoc dacS*c.714G>T and *R20291*Δ*PaLoc dacS*c.548T>C labelled with vancomycin-BODIPY. An unlabelled R20291Δ*PaLoc* control is also shown. Mid-cell staining is apparent in R20291Δ*PaLoc* but not in the two *dacS* recapitulated strains. (B) Normalised maximum fluorescence intensity of each cell, gained from the maximum value across its intensity profile, was plotted for each strain to quantify differences in vancomycin binding across strains.

4.3.5 The Contribution of *dacJRS* to Bc1 Resistance

Work in this chapter has shown that evolved *dacS* SNPs result in vancomycin resistance through overexpression of the *dacJ*-encoded D,D-carboxypeptidase, which reduces vancomycin binding at the mid-cell. However, how the *dacS* pathway, which by itself led to a 4-fold increase in vancomycin MIC, fits into the 32-fold increase in MIC observed in the evolved isolate Bc1 was not determined. Further, although *dacS* was shown to increase *dacJ* expression, and mutants with recapitulated *dacS* SNPs displayed reduced vancomycin binding, the role of DacJ was not directly tested to confirm its causative effect on resistance and vancomycin binding reduction. To address these points, *dacJ* was deleted from the evolved isolate Bc1, resulting in the strain Bc1Δ*dacJ*. Construction of this strain is described in Table 4.2. Investigations relating to Bc1Δ*dacJ* were performed by Lucy Thompson (unpublished master's dissertation).

4.3.5.1 Bc1Δ*dacJ* Displays Reduced Vancomycin Resistance

The vancomycin MIC of Bc1Δ*dacJ* was tested by standard agar dilution, and compared to the evolved isolate (Bc1) and the R20291Δ*PaLoc* control. Bc1Δ*dacJ* exhibited an MIC of 2 µg/mL, an 8-fold reduction from the Bc1 MIC observed in the same assay (Figure 4.7a). This huge reduction in resistance from the removal of *dacJ* directly confirms the major role of the *dacJRS* pathway in the vancomycin resistance observed in Bc1, and confirms the causative role of *dacJ*.

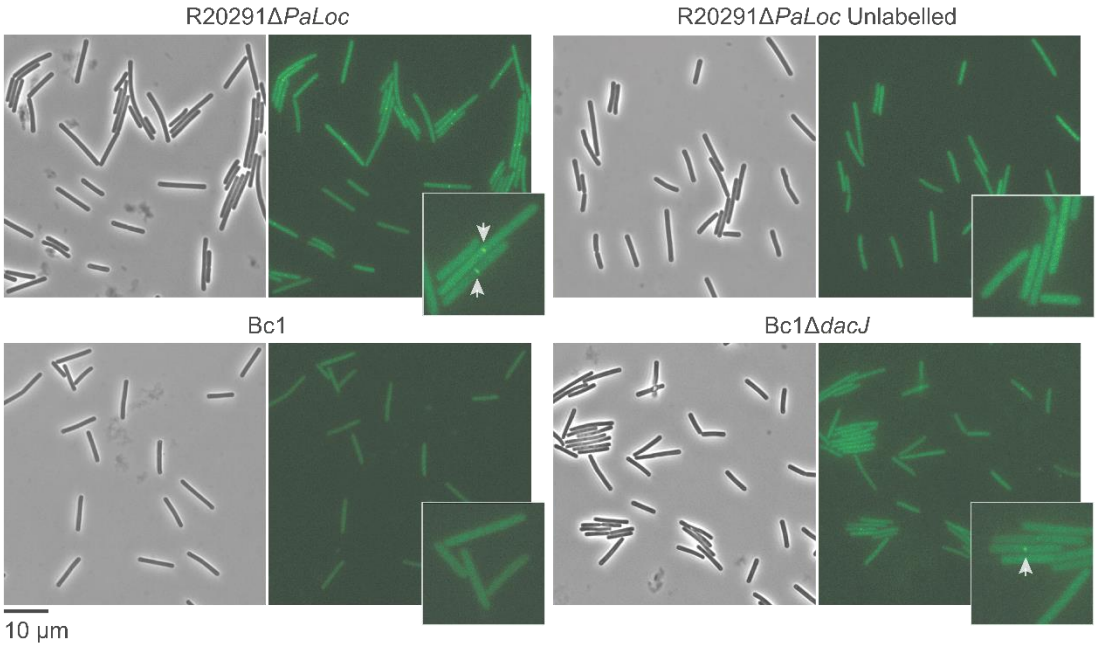
Interestingly, removal of *dacJ* from Bc1 did not reduce its MIC to WT-level, as the MIC of Bc1Δ*dacJ* was 2-fold higher than the R20291Δ*PaLoc* control. Additionally, the presence of *dacJ* in the Bc1 evolved isolate increased its MIC from 2 to 16 µg/mL (8-fold), however the overexpression of *dacJ* from the recapitulated *dacS* SNP resulted in just a 4-fold increase in MIC. Taken together, this suggests the presence of a second, minor mechanism of vancomycin resistance in Bc1, providing an MIC of 2 µg/mL. The disparity between resistance gains of Bc1Δ*dacJ* and R20291Δ*PaLoc dacS*c.714G>T could potentially be explained by the synergistic effects of these two mechanisms.

However, no second mechanism could be identified from the illumina data (discussed in more detail in Chapter 5).

A

Isolate	Vancomycin (µg/mL)							
	0	0.5	1	2	4	8	16	MIC
R20291Δ*Paloc*								1
Bc1Δ*dacJ*								2
Bc1								16*

B



C

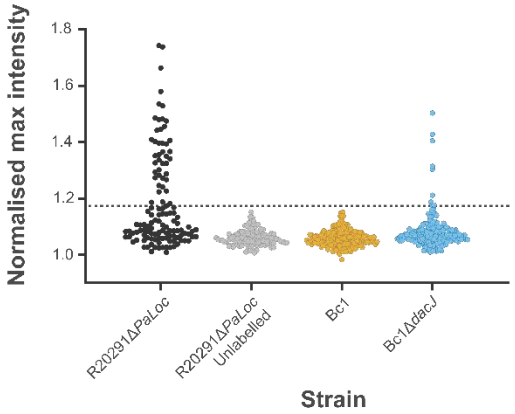


Figure 4.7 - ΔdacJ Partially Restores Vancomycin Binding in Bc1

(A) Vancomycin MICs of R20291Δ*PaLoc*, Bc1Δ*dacJ* and Bc1. MICs determined by agar dilution on BHI. Assays were performed in biological triplicate and technical duplicate. Shown is a single representative biological replicate. (B) Representative images of R20291Δ*PaLoc*, Bc1Δ*dacJ* and Bc1 labelled with vancomycin-BODIPY. An unlabelled R20291Δ*PaLoc* control is also shown. Mid-cell staining is apparent in R20291Δ*PaLoc* but not in the Bc1 evolved isolate. Deletion of *dacJ* from Bc1 partially restores vancomycin binding, indicating the direct role of DacJ in depletion of vancomycin binding sites. (C) Normalised maximum fluorescence intensity of each cell, gained from the maximum value across its intensity profile, was plotted for each strain to quantify differences in vancomycin binding across strains.

4.3.5.2 Partial Restoration of Vancomycin Binding was Observed in Bc1Δ*dacJ*

To confirm the causative effect of *dacJ* on vancomycin binding reduction, vancomycin binding of Bc1Δ*dacJ* was compared to the evolved isolate Bc1 and the R20291Δ*PaLoc* control. To assess vancomycin binding, samples were incubated with vancomycin-BODIPY and imaged as above. An unlabelled R20291Δ*PaLoc* control was also included. It was hypothesised that if *dacJ* were involved in depletion of vancomycin binding sites, deletion of *dacJ* in Bc1 would restore vancomycin binding.

As expected, the Bc1 evolved isolate displayed no evidence of vancomycin binding (Figure 4.7b). Bc1Δ*dacJ* displayed partial restoration of vancomycin binding – mid-cell binding was clearly present, but in a smaller proportion of cells than the R20291Δ*PaLoc* control. This observation was also reflected in the fluorescence quantification (Figure 4.7c). This data confirms the direct role of DacJ in vancomycin resistance through depletion of vancomycin binding sites; and suggests that DacJ is the major, but not sole, contributing factor to vancomycin resistance in Bc1.

4.4 Discussion

dacJRS emerged as an important locus during the *in vitro* evolution presented in this thesis, with mutations in the *dacJRS* gene cluster occurring in six of the ten replicate populations, suggesting a prominent role in vancomycin resistance. Detailed dissection of the *dacJRS* mechanism of resistance described in this chapter found mutations in *dacS* are capable of invoking vancomycin resistance alone, without the presence of potentiating prerequisite mutations. Together with the observed *dacS* population dynamics, this suggests *dacJRS* is an important mechanism of first-step vancomycin resistance. Through qRT-PCR, the *dacS* SNPs were shown to elicit resistance through overexpression of putative D,D-carboxypeptidase DacJ, via increasing the activity of the DacRS TCS, a positive regulator of *dacJRS.* Increased expression of DacJ was associated with reduced vancomycin binding to the cell wall, and the direct involvement of DacJ was confirmed through construction of a Bc1Δ*dacJ* strain, which resulted in reduced resistance and partial restoration of vancomycin binding.

4.4.1 Compensable Fitness Costs of *dacJRS*

Although the *dacJRS* pathway was associated with fitness burdens in terms of growth, no defects were observed relating to sporulation efficiency. Since sporulation is an inherent part of the *C. difficile* life cycle, it is quite alarming that, given the right circumstances, routes to vancomycin resistance can exist which do not pose detrimental effects on transmission. Furthermore, the reduction in growth in the evolved isolates Bc8 and Bc9 was less prominent than in R20291Δ*PaLoc dacS*c.548T>C, indicating amelioration of the *dacS*-associated growth burden through accumulation of refining mutations. Investigating the mechanisms of this compensatory evolution, either through a more extensive resequencing of Bc8 and Bc9, or via further evolution of the Bc1 isolate (which did not display evidence of amelioration), would therefore be pertinent; to provide a crucial

complement to directed evolution, allowing identification of the routes to high-fitness vancomycin resistance.

4.4.2 Remarking on *vanT*/*dacS* Mutual Exclusivity

Through qRT-PCR, *dacS* mutations were shown to have no impact on expression of the *van* cluster, confirming the two observed pathways to vancomycin resistance were separate – *dacS* does not elicit vancomycin resistance through alternative activation of the *van* cluster. However, there are multiple other circumstances in which the *vanT*/*dacS* pathways may overlap, accounting for the lack of *vanT*/*dacS* co-occurrence observed amongst evolved isolates. For example, cross-action of these mechanisms cannot be completely ruled out – the complete regulon of the *vanRS* TCS has not been described, meaning activation of *dacJRS* via *vanRS* may still occur. Additionally, despite *dacJRS* and *van* pathways being seemingly separate, acting through independent mechanisms, both have a similar end point – alteration of the cell wall vancomycin binding site to promote resistance. Indeed, in this chapter and previous work, reduction in vancomycin mid-cell binding can be observed for either pathway (Shen et al., 2020), which could be suggestive of downstream crossover between pathways. The *van* cluster in *C. difficile* is complex, and its role in vancomycin resistance has previously been controversial (Ammam et al., 2013; Peltier et al., 2013). The contribution of *vanT* mutations to vancomycin resistance is not understood (explored in Chapter 6). Until this *vanT*-mediated mechanism of resistance is clear, the reason for the apparent *vanT*/*dacS* mutual exclusivity cannot be fully appreciated.

4.4.3 Further Intricacies of the *dacJRS* System

More immediately, many molecular intricacies of the *dacJRS* system remain unknown. Using a *dacRS* knockout, the role of the *dacRS* TCS as an activator of *dacJRS* expression was revealed. However,

how the mutations in the *dacS* histidine kinase result in increased activity of the TCS remains undetermined. Structure-function predictions, linking mutations with downstream effects, were used to discern the likely mechanism of *vanR*-mediated resistance in *C. difficile*, as the mutation in *vanR* was predicted to stabilise the response regulator to allow DNA engagement (Shen et al., 2020). Thus, structural insights into the *dacS* histidine kinase would provide a better understanding of how the Bc1 and Bc8/9 mutations result in increased DacRS activity. These structural predictions are more widely relevant, as it is clear from the experimental evolution that the same outcome (increased DacRS activity) can be achieved via multiple different mutations (Bc1, Bc8/9). Understanding the relationship between mutations and functional effects may allow prediction of other mutations which would give rise to the same outcome, which would be essential for monitoring *dacS* in clinical isolates.

Similarly to *dacS, dacR* displayed SNP-level parallelism, with an identical mutation observed in Bc7 and Bc10, suggesting strong selection of this allele. Further exploration of *dacR,* to understand how this mutation alters the DacR structure to increase activation of the *dacJRS* cluster would therefore be beneficial. The Bc7/10 *dacR* mutation could not be recapitulated during this project (data not shown), however a concerted effort to generate this mutation to understand its contribution to resistance would be important for a fuller understanding of the *dacJRS* resistance pathway.

Visualisation of the genomic organisation of the *dacJRS* cluster showed the *dacRS* TCS was downstream of the *dacJ* D,D-carboxypeptidase. This, along with knowledge of peptidoglycan biogenesis and the mechanism of action of vancomycin, led to the hypothesis that *dacRS* acted on *dacJ*, which was confirmed through qRT-PCR. However, TCS regulon size can vary hugely, both by TCS and by species (Müh et al., 2022; Pettersen et al., 2024). To further understand the *dacRS* TCS, it may be useful to construct a Bc1Δ*dacRS* strain – comparing the phenotype of this strain with

Bc1Δ*dacJ* would indicate whether *dacRS* regulates other vancomycin resistance determinants, as Bc1Δ*dacRS* would be more sensitive than Bc1Δ*dacJ*. Alternatively, since this study did not investigate DacR-promoter binding, a broader view of the *dacRS* regulon may be achieved using ChIP-seq to identify further DacR-DNA binding sites (Myers et al., 2015). An exhaustive investigation of the *dacRS* regulon may also be achieved through RNA seq.

This chapter identified DacJ as a key mediator of vancomycin resistance, since overexpression of the DacJ putative D,D-carboxypeptidase, through increased DacRS activation, resulted in reduced vancomycin binding ability. Intriguingly, Bc9 displayed mutations in both *dacS* and *dacJ*. The *dacJ* mutation could not be recapitulated during this project, despite multiple attempts (data not shown), however understanding the contribution of this mutation to resistance is imperative. Since mutations in *dacS* result in overexpression of *dacJ,* the observed co-occurrence of *dacS* and *dacJ* in Bc9 may suggest the additional *dacJ* mutation enhances overexpression further, perhaps through alteration of the mRNA secondary structure. Alternatively, the mutation in *dacJ* may increase DacJ activity, acting as a refining mutation to provide a small resistance increase. To unpick this, *dacJ* and *dacJS* recapitulated strains would be necessary. Comparison of *dacS, dacJ* and *dacJS* via qRT-PCR would indicate whether the *dacJ* SNP further increases *dacJ* expression, and MIC comparisons of these strains would enable assessment of vancomycin resistance gain. Together, these additional investigations would enable understanding of an alternative secondary mechanism of the *dacJRS* pathway, and also how the accumulation of mutations can promote vancomycin resistance.

Although it is clear DacJ overexpression results in vancomycin resistance through reducing availability of vancomycin binding sites, presumably via cleavage of the terminal D-Ala in nascent peptidoglycan, the downstream mechanisms of this pathway are unknown. For example, in a mechanism similar to that of the *vanG* operon, the terminal D-Ala may be replaced with D-Ser, to

maintain the pentapeptide for peptidoglycan biogenesis (Reynolds and Courvalin, 2005). Alternatively, as is common in *C. difficile*, the resultant tetrapeptide may act as a substrate for LD-transpeptidases, further increasing the abundance of 3,3 crosslinks (Coullon et al., 2020). Peptidoglycan analysis of the *dacS* recapitulated SNP mutants may therefore help to uncover the downstream mechanisms of the *dacJRS* pathway.

4.4.4 Translating *dacJRS* Mechanistic Insights

Although mutations in *dacS* led to overexpression of *dacJRS* transcripts, this work does not quantify the resultant protein-level changes. Indeed, transcript levels are not always predictive of protein levels (Liu et al., 2016). Measuring differences in DacJ expression between the WT and strains possessing *dacS* mutations would therefore provide further evidence in support of the mechanism proposed in this chapter.

4.4.5 *dacJRS* and Bc1

Overexpression of *dacJ,* arising from mutations in *dacS* in evolved isolates, resulted in a 4-fold increase in vancomycin resistance. However, the difference in MIC between Bc1 and Bc1Δ*dacJ* suggested overexpression of *dacJ* was responsible for an 8-fold increase in resistance. Although this confirmed *dacJRS* as the major mechanism of vancomycin resistance in Bc1, this disparity, coupled with the observation that Bc1Δ*dacJ* had an MIC 2-fold higher than the R20291Δ*PaLoc* parental strain, suggested the presence of a second, minor mechanism of resistance in Bc1 acting synergistically with *dacJRS*. However, no other vancomycin-unique mutations were observed in the Bc1 Illumina data. This could therefore be explained through either the existence of epigenetic resistance mechanisms, or by the presence of further mutations, not captured in illumina data. Recently, long read sequencing has been used for high-quality variant calling in complex or difficult

to sequence regions (Goenka et al., 2022). Further exploration of vancomycin resistance in Bc1, through use of Nanopore sequencing to close the Bc1 genome, is detailed in the subsequent chapter.

5 Understanding Multifactorial Vancomycin Resistance in Bc1

5.1 Introduction

Through experimental evolution and short-read sequencing, *dacS*c.714G>T was identified as the sole vancomycin-unique mutation in the Bc1 evolved isolate (referred to in this chapter as Bc1), and extensive exploration of the *dacJRS* cluster found that *dacS*c.714G>T did in fact constitute the major mechanism of resistance in Bc1. However, the resistance observed in Bc1 could not be fully recapitulated by introduction of this single SNP to the parental strain. Additionally, although construction of Bc1Δ*dacJ* confirmed the major role of *dacJ* in resistance, multiple discrepancies in resistance gains were observed: first, the deletion of *dacJ* resulted in an MIC which was still 2-fold higher than the parental R20291Δ*PaLoc.* Moreover, the presence of *dacJ* resulted in an 8-fold increase in resistance between Bc1Δ*dacJ* and Bc1, yet overexpression of *dacJ* via *dacS*c.714G>T resulted in only a 4-fold increase, suggesting the presence of a second mechanism working synergistically with *dacJRS.* Synergistic interactions, whereby the combined effect is greater than the sum of individual effects, are widely recognised as enhancers of antibiotic resistance. Few examples of resistance-promoting synergy have been described in *C. difficile*, however high-level lysozyme resistance was found to result from the combined expression of peptidoglycan deacetylases PgdA and PdaV (Ho et al., 2014). More recently, the accumulation of mutations leading to metronidazole resistance in *C. difficile* was examined, observing that loss of *iscR* worked synergistically with defective *feoB1* to reduce metabolism of metronidazole into its active state (Deshpande et al., 2020). Since accumulation of mutations leading to vancomycin resistance has not been previously

studied in *C. difficile*, no evidence of synergistic mutations which heighten vancomycin resistance have thus far been reported.

Despite evidence suggesting the presence of a second, possibly synergistic, mechanism in Bc1, this was not identified with the available data. A possible approach to address this uncertainty would be to perform additional sequencing to identify interactions promoting high-level resistance in Bc1. This approach would enable assembly of a complete, closed bacterial genome through long-read sequencing with short-read error correction (coined hybrid assembly) (Moss et al., 2020; Wick et al., 2023). The addition of long-read sequencing enables variant calling in complex or difficult to sequence regions, and has been used increasingly used in recent years (Goenka et al., 2022). Multiple comparison studies, evaluating short-read, long-read and hybrid assemblies have concluded the hybrid approach is the most comprehensive for variant calling, since long-read sequencing is (although improving) error-prone, and short-read sequencing suffers from incomplete coverage and often fails to detect duplications (Ashton et al., 2015; Juraschek et al., 2021; Ruan et al., 2020). The power of hybrid sequencing was neatly demonstrated during an investigation of high-level methicillin resistance in *S. aureus*. Illumina sequencing suggested overrepresentation of reads from the Staphylococcal cassette chromosome *mec* element, however the nature of this overrepresentation could not be discerned from the short-read sequences. Hybrid assembly of the genome showed the presence of 10 tandem repeats of the element, representing a new mechanism of adaptation and resistance in *S. aureus* (Gallagher et al., 2017). Overall, hybrid sequencing represents a valuable method for detection of novel mechanisms of resistance, especially those arising from duplication.

5.2 Aims and Outcomes

The aim of the work presented in this chapter was to fully comprehend the mechanisms of vancomycin resistance occurring in the evolved isolate Bc1, including the identification and characterisation of additional resistance mechanisms – how they interact with the previously identified *dacJRS* pathway, whether they are able to fully recapitulate the resistance observed in Bc1, and how they achieve resistance on a molecular level. To achieve this, Bc1 was re-sequenced using Oxford Nanopore long-read technology to close the genome. The additional mutations identified from hybrid assembly variant calling were recapitulated in all possible combinations, and their individual contributions to resistance were assessed. The identified secondary mechanism, a *vanS* insertion, was examined using qRT-PCR.

5.3 Results

5.3.1 Closing the Bc1 Genome

Previous characterisation of the *dacJRS* cluster hinted at the existence of a secondary mechanism of resistance in Bc1. However, no additional vancomycin-unique mutations were observed in the Bc1 Illumina data. To address this, long-read sequencing was performed on the Bc1 evolved isolate, to provide coverage across difficult-to-sequence areas to close the genome. Long-read 50x whole genome sequencing was performed using the Oxford Nanopore Technologies GridION flow cell at MicrobesNG.

5.3.1.1 Improving Long-Read Sequencing with Short-Read Error Correction

Although long-read assembly was performed by MicrobesNG, additional steps were taken to polish this data using the matched Bc1 short-read data, to improve accuracy and variant calling. The Bc1 Illumina reads were first aligned to the Nanopore assembly using BWA-MEM (v0.7.17) (Li and Durbin, 2009). Reads were filtered to remove spurious alignments, and the Nanopore assembly was polished using Polypolish (v0.6.0) (Wick and Holt, 2022). The resulting polished fasta file was annotated using Prokka (v1.14.6) (Seemann, 2014), and variants were called using Snippy (v4.6.0) (Seemann, 2015). Identified variants were manually verified using IGV (v2.8.6) (Robinson et al., 2011).

5.3.1.2 Additional Variants Identified Through Long-read Sequencing

Long-read variant analysis displayed a high degree of similarity to Illumina data, highlighting the high quality of both the Illumina reads and variant calling parameters. However, two additional variants were observed in the Bc1 long-read data (Table 5.1, full dataset in Appendix VII). Both of these additional variants were large insertion-deletions (InDels) – a 44 bp deletion downstream of *CD0978* and upstream of *CD0980* (1,197,357_1,197,400del) and a 30 bp duplication within the open reading frame (ORF) of *vanS* (*vanS*c.367_396dup). In the R20291 genome, a 168 bp ORF, denoted *CD0979*, is

annotated between *CD0978* and *CD0980.* The observed Bc1 deletion occurs within this ORF, removing the stop codon and extending the putative *CD0979* into *CD0980* (encoded on the opposite strand), increasing the ORF size to 207 bp. However, as other *C. difficile* strains, including the well-characterised 630, do not have an equivalent ORF annotated in this region, and since RNA seq analysis did not detect any transcription associated with *CD0979,* it is likely that *CD0979* is a misannotation, meaning the 44 bp deletion observed affects an intergenic region (Fuchs et al., 2021; Sebaihia et al., 2006). The *vanS*c.367_396dup occurs within *vanS*, which encodes the histidine kinase of the *van* cluster. Multiple mutations in *vanS* have previously been associated with vancomycin resistance in *C. difficile* (Kolte and Nübel, 2024), however all of these were SNP variants. The mutation observed here was a 30 bp duplication, adding 10 amino acids to the VanS protein.

Table 5.1 - Vancomycin-Unique Bc1 Mutations from Long-read Data

Position	Reference	Alternative	Locus Tag	Gene Name	Gene Function
1197338	ATATGGAATAATTTCCATATCTCTAATGAGTACATAAAACAGAGG	A	*CDR20291_0979*	*CD0979**	
1795507	T	TGAACAACGTAAATTTGAGACACAAATGGCA	*CDR20291_1523*	*vanS*	Two-component sensor histidine kinase
4085308	C	A	*CDR20291_3437*	*dacS*	Two-component sensor histidine kinase

5.3.2 Fully Recapitulating Vancomycin Resistance in Bc1

To assess contributions of the individual mutations to vancomycin resistance, and to examine potential additive or synergistic interactions between mechanisms, the mutations detected from long-read sequencing analysis of Bc1 were precisely recapitulated in the parental R20291Δ*PaLoc* background. A collection of strains was constructed containing all possible combinations of single,

double and triple mutations. Recapitulating all mutation combinations was performed to provide an exhaustive insight into the exact combination of mutations required to completely recapitulate the level of resistance observed in the evolved isolate Bc1, as well as to gauge the possibility of compensatory evolution or hitchhiking.

5.3.2.1 Construction of Recapitulated Strains

Recapitulation of mutations in *C. difficile* was performed via allelic exchange as previously described. Since the recapitulated mutations were InDels, the 30/44 bp differences were screened via PCR, followed by gel electrophoresis on high percentage (3%) agarose. Recapitulation of the SNPs was confirmed using Sanger sequencing, generating strains R20291Δ*PaLoc vanS*c.367_396dup, R20291Δ*PaLoc* 1,197,357_1,197,400del, R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup, R20291Δ*PaLoc vanS*c.367_396dup 1,197,357_1,197,400del, R20291Δ*PaLoc dacS*c.714G>T 1,197,357_1,197,400del, and R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup 1,197,357_1,197,400del. The strains relevant to this chapter, and their respective plasmids, are summarised in Table 5.2.

Table 5.2 - Strains and Constructs Relevant to This Chapter

Strain	Construction	Plasmid
R20291Δ*PaLoc* *dacS*c.714G>T	Fragment designed to contain the *dacS*c.714G>T point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB026
R20291Δ*PaLoc* *dacS*c.714G>T *vanS*c.367_396dup	Fragments containing mutations synthesised by Genewiz, followed by ligation into linearised pJAK112. Double mutant generated by addition of pJEB034 to R20291Δ*PaLoc dacS*c.714G>T.	pJEB026, pJEB034
R20291Δ*PaLoc* *dacS*c.714G>T *vanS*c.367_396dup 1,197,357_1,197,400del	Fragments containing mutations synthesised by Genewiz, followed by ligation into linearised pJAK112. Triple mutant generated by addition of pJEB034 to R20291Δ*PaLoc dacS*c.714G>T 1,197,357_1,197,400del.	pJEB026, pJEB033, pJEB034
R20291Δ*PaLoc* *dacS*c.714G>T 1,197,357_1,197,400del	Fragments containing mutations synthesised by Genewiz, followed by ligation into linearised pJAK112. Double mutant generated by addition of pJEB033 to R20291Δ*PaLoc dacS*c.714G>T.	pJEB026, pJEB033
R20291Δ*PaLoc* *vanS*c.367_396dup	Fragment designed to contain the *vanS*c.367_396dup 30 bp duplication with approximately 1 kb either side of the mutation. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB034
R20291Δ*PaLoc* *vanS*c.367_396dup 1,197,357_1,197,400del	Fragments containing mutations synthesised by Genewiz, followed by ligation into linearised pJAK112. Double mutant generated by addition of pJEB034 to R20291Δ*PaLoc* 1,197,357_1,197,400del.	pJEB033, pJEB034
R20291Δ*PaLoc* 1,197,357_1,197,400del	Fragment designed to contain 1,197,357_1,197,400del 44 bp deletion with approximately 1 kb either side of the mutation. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB033

5.3.2.2 *dacS* and *vanS* Mutations Fully Recapitulate Bc1 Resistance

To evaluate the contribution of each mutation and combination to vancomycin resistance, vancomycin MIC assays were performed for recapitulated strains using standard agar dilution. The parental R20291Δ*PaLoc,* and the evolved isolate Bc1, were also included in the same assay for direct comparison. R20291Δ*PaLoc dacS*c.714G>T presented an MIC of 4 µg/mL (as shown previously), and the *vanS* duplication alone showed only a small (2-fold) increase in MIC (Figure 5.1). The 44 bp intergenic deletion (1,197,357_1,197,400del) provided no contribution to resistance, either alone or in combination with any other mutation, suggesting it was not involved in vancomycin resistance.

This deletion may therefore have persisted through genetic hitchhiking, or may act as a refining mutation for either resistance or fitness.

Crucially, the R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup strain had an MIC of 16 µg/mL, fully recapitulating the Bc1 MIC observed in the same assay. As the combination of *dacS* and *vanS* mutations gave rise to an MIC much higher than the moderate increases observed in either mutation alone, these mutations were concluded to act synergistically to achieve the level of resistance observed in Bc1. This provides the first evidence of synergistic interactions promoting vancomycin resistance in *C. difficile*, and demonstrates that only two mutations are required for high-level vancomycin resistance.

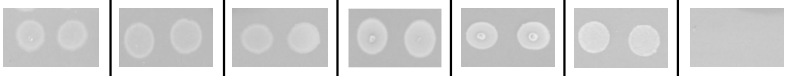
Isolate	Vancomycin (µg/mL)							
	0	0.5	1	2	4	8	16	MIC
R20291Δ*Paloc*								1
Bc1								16
R20291Δ*PaLoc* *dacS*c.714G>T								4
R20291Δ*PaLoc* 1,197,357_1,197,400del								1
R20291Δ*PaLoc* *vanS*c.367_396dup								2
R20291Δ*PaLoc* *dacS*c.714G>T *vanS*c.367_396dup								16
R20291Δ*PaLoc* *vanS*c.367_396dup 1,197,357_1,197,400del								2
R20291Δ*PaLoc* *dacS*c.714G>T 1,197,357_1,197,400del								4
R20291Δ*PaLoc* *dacS*c.714G>T *vanS*c.367_396dup 1,197,357_1,197,400del								16

Figure 5.1 - vanS+dacS Mutations Fully Recapitulate Bc1 Resistance

Vancomycin MICs of R20291Δ*PaLoc*, Bc1, R20291Δ*PaLoc dacS*c.714G>T, R20291Δ*PaLoc* 1,197,357_1,197,400del, R20291Δ*PaLoc vanS*c.367_396dup, R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup, R20291Δ*PaLoc vanS*c.367_396dup 1,197,357_1,197,400del, R20291Δ*PaLoc*

*dacS*c.714G>T 1,197,357_1,197,400del, and R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup 1,197,357_1,197,400del. MICs determined by agar dilution on BHI. Assays were performed in biological triplicate and technical duplicate. Shown is a single representative biological replicate.

5.3.3 Phenotypic Analysis of Recapitulated Strains

Growth and sporulation efficiency assays were performed on recapitulated mutant combinations to understand how each mutation contributes to the Bc1 phenotype.

5.3.3.1 *dacS*c.714G>T is Responsible for Bc1 Growth Defects

The Bc1 *dacS*c.714G>T mutation was previously shown to significantly reduce growth, displaying defects similar to those displayed by the Bc1 evolved isolate. To identify whether *dacS*c.714G>T was the sole contributor to the Bc1 phenotype, growth of all recapitulated mutant combinations was assayed in rich media as described previously, and compared with the parental R20291Δ*PaLoc* control. Severe growth defects were only observed for strains containing the *dacS*c.714G>T mutation, suggesting this mutation was responsible for the growth defect observed in Bc1 (Figure 5.2). Conversely, the growth profile of the *vanS* insertion mutation was not significantly different from that of the R20291Δ*PaLoc* control, showing it did not adversely affect fitness. One possible explanation of why the 44 bp intergenic deletion (1,197,357_1,197,400del) did not contribute to vancomycin resistance was that it may act as a refining mutation, providing a fitness advantage. However, the 1,197,357_1,197,400del strain displayed a lower maximum OD_{600nm} than the control. Additionally, growth was not restored in the R20291Δ*PaLoc dacS*c.714G>T 1,197,357_1,197,400del strain, suggesting the persistence of the 44 bp deletion in Bc1 may be a product of genetic hitchhiking.

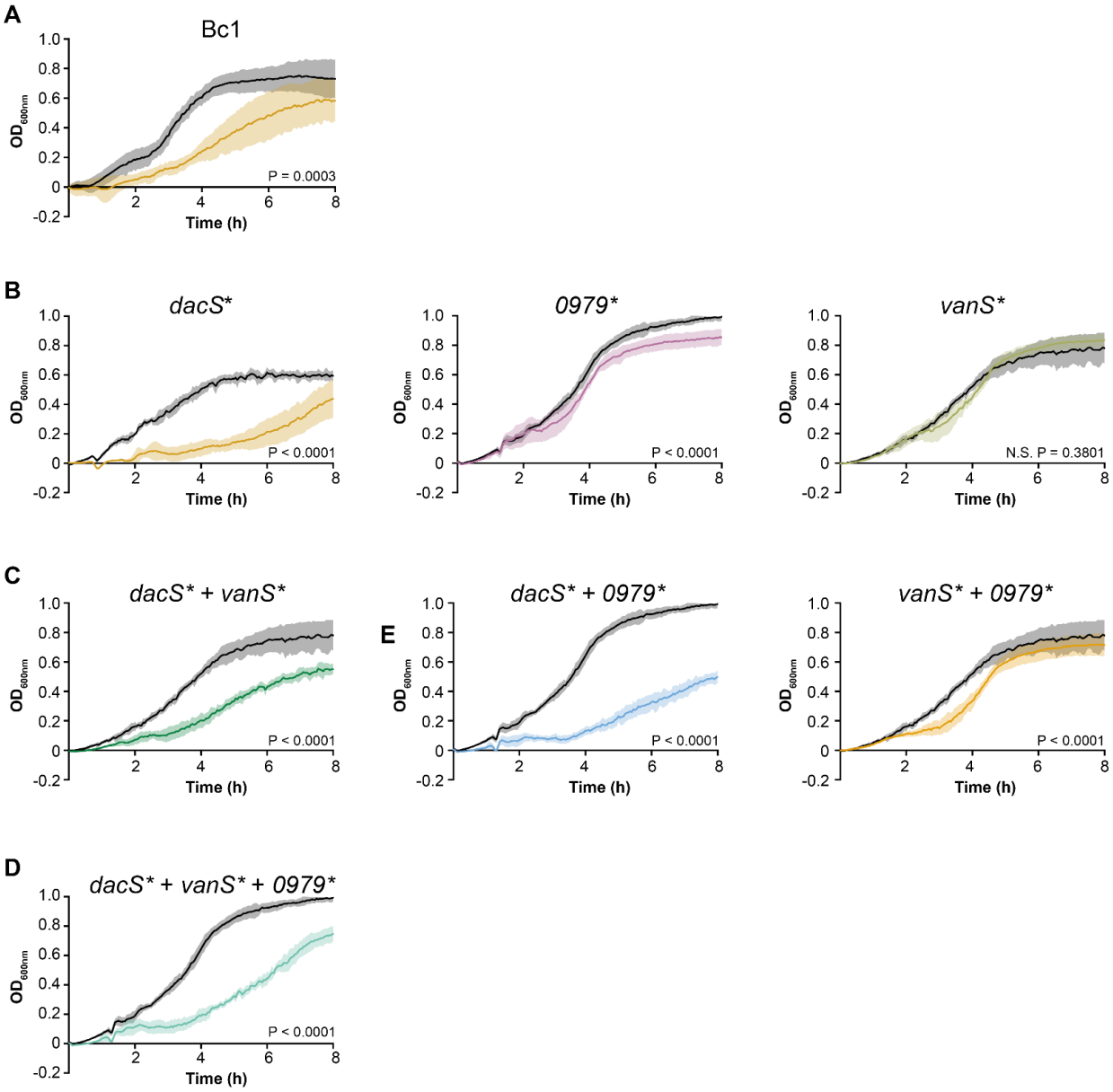


Figure 5.2 - Growth of Bc1 Recapitulated Strains

Growth over time in rich media (TY broth) was evaluated by measuring OD at 600 nm in a 96 well microplate spectrometer. Growth of evolved isolate Bc1 (A), R20291Δ*PaLoc dacS*c.714G>T (*dacS**), 1,197,357_1,197,400del (*0979**) and *vanS*c.367_396dup (*vanS**) single (B), double (C) and triple (D) mutants (coloured lines) were compared to R20291Δ*PaLoc* (black lines). Shown are the mean and standard deviation of repeats, assayed at minimum in biological and technical triplicate. For each strain, AUC was determined using the GrowthCurver R package. AUC was compared using Student's t-tests with Welch's correction, with the P value shown on each graph. N.S. indicates differences were not significant.

5.3.3.2 Bc1 Mutations Do Not Alter Sporulation Efficiencies

The sporulation efficiency of R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup 1,197,357_1,197,400del (referred to hereafter as the Bc1 triple mutant) was compared to the parental R20291Δ*PaLoc* control. Similar to the evolved isolate Bc1, the Bc1 triple mutant displayed a sporulation efficiency which was not significantly different to R20291Δ*PaLoc* (Figure 5.3). Together, this indicates that not only can high-level vancomycin resistance arise from just two synergistically-acting mutations, but these mutations do not impact fitness in terms of sporulation.

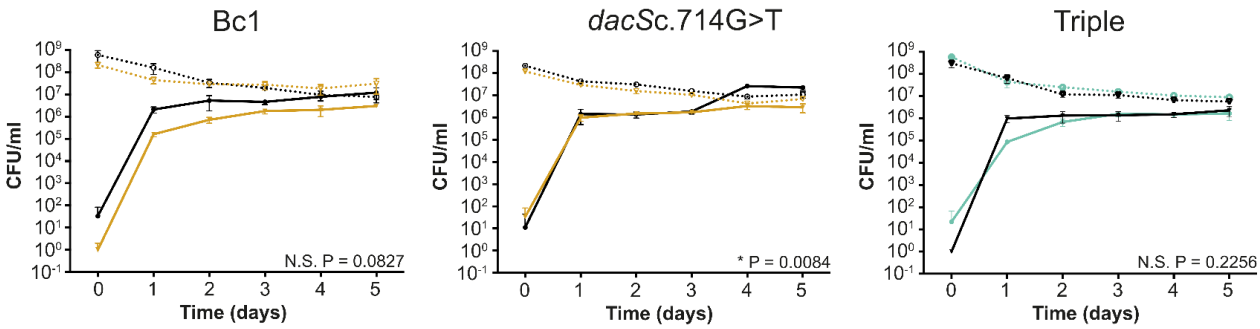


Figure 5.3 - Sporulation of Bc1 Recapitulated Strains

Sporulation efficiencies of the evolved isolate Bc1, R20291Δ*PaLoc dacS*c.714G>T and the R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup 1,197,357_1,197,400del triple mutant (coloured lines) were compared to R20291Δ*PaLoc* (black lines). Stationary phase cultures were incubated anaerobically for 5 days with samples taken daily to enumerate total colony forming units (CFUs, dotted lines) and spores (solid lines), following incubation at 65°C for 30 min to kill vegetative cells. Shown are the mean and standard deviations of biological triplicates assayed in triplicate. For each strain, spore CFU area under the curve was determined using GraphPad Prism and these were compared using Dunnett's T3 multiple comparisons test with the adjusted P value shown on each graph. * = significant difference, N.S. = not significant.

5.3.4 *vanS* Mutations Result in Overexpression of the *van* Operon

Through hybrid sequencing and recapitulation of mutant combinations, the synergistic actions of *dacS*c.714G>T and *vanS*c.367_396dup were determined to be the causative components of the 16-fold increase in vancomycin resistance observed in Bc1. The mechanism of resistance associated with *dacS*c.714G>T was explored in the previous chapter. Since the single *vanS*c.367_396dup mutant

strain showed a 2-fold increase in MIC, suggesting the *vanS* insertion alone provides low-level vancomycin resistance, the remainder of this chapter explores the mechanism of resistance associated with *vanS*.

5.3.4.1 The *van* Cluster

The genomic organisation of the *van* cluster was visualised using Geneious (v7.1.9 http://www.geneious.com/). Promoters were predicted using global transcription site mapping (Fuchs et al., 2021). The *van* cluster in *C. difficile* consists of *vanRS,* which encodes the TCS responsible for *van* gene regulation (Figure 5.4a). The remaining genes, *vanG*, *vanXY*, and *vanT*, are controlled by a separate promoter. These genes encode a D-Ala D-Ser ligase, a D-Ala D-Ala carboxypeptidase, and a serine racemase respectively, which promote vancomycin resistance through replacement of the terminal D-Ala with D-Ser in nascent peptidoglycan (Reynolds and Courvalin, 2005).

5.3.4.2 The *vanS* Histidine Kinase

To visualise the effect of the observed 30 bp duplication in *vanS*, structural prediction of the WT and mutated VanS proteins was performed using AlphaFold (Jumper et al., 2021). VanS was modelled as a homodimer (Figure 5.4b). Intriguingly, the ten amino acid insertion resulted in a predicted 90° rotation of the transmembrane domains relative to the catalytic ATPases. An understanding of the mechanistic dynamics of VanS in *C. difficile* is lacking, meaning the direct impact of this mutation could not be discerned. However, since previous works have characterised *vanS*-mediated resistance, and as signal transduction in histidine kinases commonly occurs via helical rotation (Gushchin et al., 2020; Neiditch et al., 2006), the *vanS*c.367_396dup mutation was hypothesised to result in constitutive activation of VanS, resulting in overexpression of the *van* cluster.

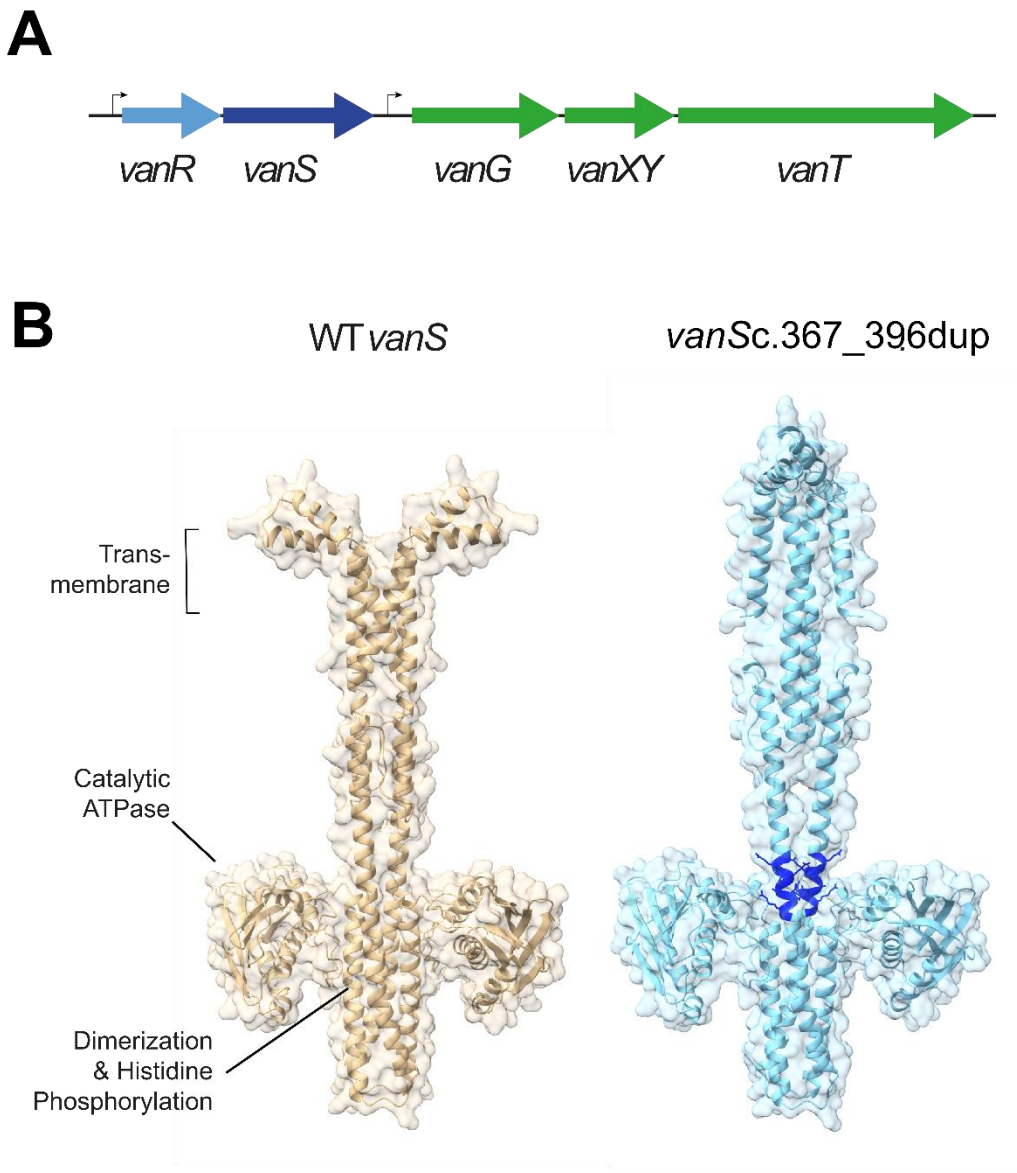


Figure 5.4 - Visualising vanS at the Gene and Protein Level

(A) Genomic organisation of the *van* cluster. (B) AlphaFold model of VanS as a dimer with (right) and without (left) the *vanS*c.367_396dup-encoded 10 amino acid duplication (Jumper et al., 2021). The transmembrane domains were identified using DeepTMHMM (Hallgren et al., 2022) and the Catalytic ATPase and Dimerization and Histidine Phosphorylation domains were predicted using InterProScan (Jones et al., 2014). The location of the 10 amino acid duplication is highlighted on both chains in dark blue. The insertion results in an approximately 90° rotation of the transmembrane domains relative to the catalytic ATPases.

5.3.4.3 The *vanS* Variant Results in Overexpression of *van* Genes

qRT-PCR was performed to investigate the impact of the *vanS*c.367_396dup on the *van* cluster. The *vanS*c.367_396dup strain was compared to the R20291Δ*PaLoc* control, and expression of the *van* genes (*vanR, vanS, vanG, vanXY, vanT*), as well as housekeeping gene *rpoA*, was measured in the presence (0.5x MIC) and absence of vancomycin. Expression was quantified as defined earlier, using the previously described pJEB032 plasmid to determine exact copy numbers. The copy numbers were again standardised using the *rpoA* housekeeping gene, and results were expressed as copies per 1000 copies of *rpoA*. Differences in expression were then statistically analysed using a 2-way ANOVA.

The *vanS* mutation had no effect on *vanR* expression in either the presence or absence of vancomycin (Figure 5.5). Intriguingly, a significant decrease (4.5- to 4.8-fold) in *vanS* expression was observed in the *vanS*c.367_396dup strain. Since the *vanRS* genes are thought to comprise a bicistronic operon (Fuchs et al., 2021), it is likely that the observed decrease in *vanS* expression is due to altered 3' stability, or transcriptional efficiency, of the *vanR-vanS*c.367_396dup transcript. Despite this decrease, expression of the *van* operon genes *vanG, vanXY,* and *vanT* was significantly increased in the *vanS*c.367_396dup strain, from a 10-fold increase in transcription of *vanG* in the presence of vancomycin to a 136-fold increase in *vanG* without vancomycin. Importantly, as these increases were observed in both the presence and absence of vancomycin, this suggested that the *vanS* insertion results in constitutive expression of *vanG, vanXY,* and *vanT.*

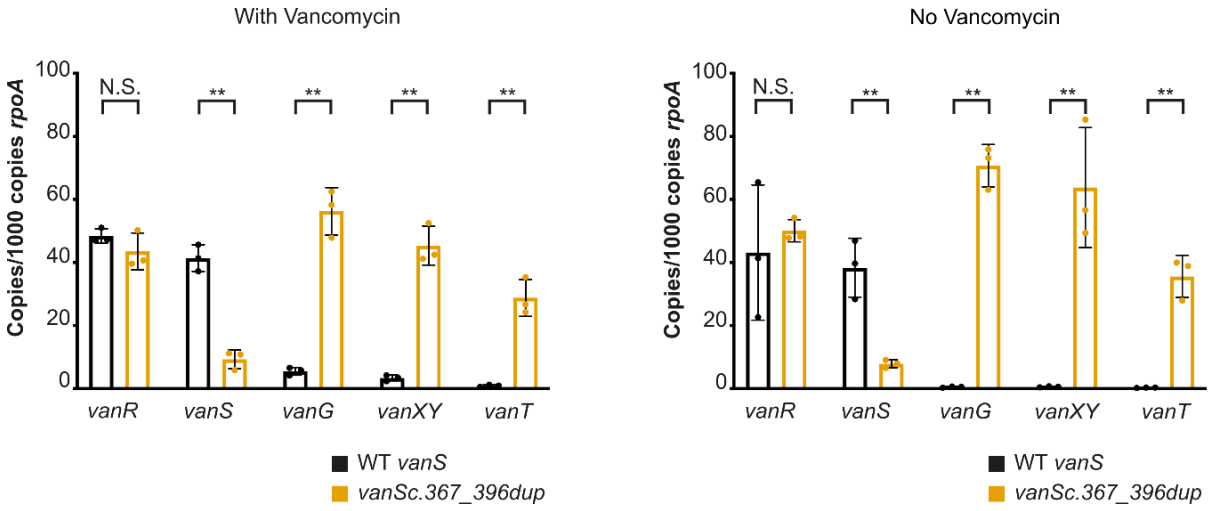


Figure 5.5 - vanSc.367_396dup Alters Expression of van Genes

qRT-PCR analysis of the *van* operon genes, and TCS *vanSR*, in R20291Δ*PaLoc* (black) and R20291Δ*PaLoc vanSc.*367_396dup (yellow). Expression was quantified against a standard curve and normalised relative to the house-keeping gene *rpoA*. Assays were performed in biological and technical triplicate. Statistical significance was calculated using a two-way ANOVA with the Tukey-Kramer test, ** = $P < 0.001$.

5.4 Discussion

The *dacSc.*714G>T mutation was found to be the major mechanism of vancomycin resistance in Bc1. However, the level of vancomycin resistance exhibited by Bc1 could not be fully recapitulated with this mutation alone, and despite the presence of a second mechanism being evident, no such mechanism could be identified from Illumina data. This chapter provides an extensive insight into mechanisms of vancomycin resistance present in the individual isolate Bc1. Re-sequencing of Bc1, coupled with hybrid assembly, identified an additional two mutations present in the Bc1 genome, both of which were large InDels. Through recapitulation of all Bc1 mutations, in all combinations, the *vanSc.*367_396dup was shown to act synergistically with the *dacSc.*714G>T SNP to provide high-level vancomycin resistance, fully recapitulating the Bc1 phenotype. The mechanism of resistance associated with the *vanSc.*367_396dup – overexpression of *vanG, vanXY,* and *vanT* – was determined using qRT-PCR.

5.4.1 The Wider Implications of Discovering InDel-mediated Resistance

Hybrid assembly has previously been shown to be invaluable for determining resistance mechanisms not discernible from Illumina data (Gallagher et al., 2017). The two additional large InDels described here, *vanS*c.367_396dup and 1,197,357_1,197,400del, were only identified through hybrid sequencing of Bc1. This enabled identification of all vancomycin resistance-associated mutations in Bc1, allowing full recapitulation of the Bc1 phenotype. The clear value of hybrid sequencing demonstrated here may set the precedence for future experimental evolution works – a more focussed, but comprehensive, approach should be considered to gain the most expansive insights.

Since InDels are often missed when performing short-read sequencing, yet can clearly be important in *C. difficile* vancomycin resistance, this presents important implications for clinical resistance monitoring. *C. difficile* clinical isolates are often routinely monitored through Illumina sequencing (Miles-Jay et al., 2023), however this data may be missing important InDels or rearrangements. Interestingly, although *vanS* mutations are widely reported in vancomycin resistant *C. difficile* isolates, all reported mutations are SNPs (Kolte and Nübel, 2024), leaving open the possibility that InDels are also occurring, but are not captured by current monitoring efforts.

5.4.2 Clinical Considerations Regarding *dacS/vanS* Synergy

Through combining recapitulated mutations, the 16-fold increase in vancomycin resistance observed in Bc1 was shown to be due to just two synergistically-acting mutations – *dacS*c.714G>T and *vanS*c.367_396dup. To best knowledge, this is the first report of synergistic interactions promoting vancomycin resistance in *C. difficile.* Additionally, the finding that only two mutations are required for high-level resistance has important clinical consequences. Consistent with the R20291Δ*PaLoc* *vanS*c.367_396dup strain assayed here, increasing reports of clinical isolates with reduced

vancomycin susceptibility identify *vanS* as a causative mutation in low-level resistance (Kolte and Nübel, 2024). The work presented here shows a single SNP in *dacS* has the potential to dramatically increase vancomycin resistance in these isolates, suggesting an increased need for rigorous clinical monitoring.

5.4.3 Further Intricacies of the *vanS* Resistance Mechanism

Since the *vanS*c.367_396dup provided low-level vancomycin resistance, the mechanism of resistance relating to this mutation was elucidated. The 30 bp insertion in *vanS* resulted in constitutive expression of *vanG, vanXY,* and *vanT* genes. Similar *vanS*-mediated mechanisms have been observed in both clinical isolates and experimental evolution (Shen et al., 2020). However, precisely how the insertion of ten amino acids results in constitutive activation of *van* genes is unknown. Previous studies have associated helical rotation with histidine kinase activation (Gushchin et al., 2020; Neiditch et al., 2006). The *vanS*c.367_396dup-containing VanS structure predicted by AlphaFold suggested a 90° rotation of the transmembrane domains relative to the catalytic ATPases, suggesting this mutation may be responsible for maintaining VanS in its active conformation. Validation of this predicted structure via X-ray crystallography (Moraes et al., 2014), along with a fuller understanding of the structural basis of VanS activity may therefore be beneficial, to aid a more complete comprehension of the functional consequences of mutations arising in *vanS.*

Despite *dacS* and *vanS* mutations displaying clear synergy, the mechanisms underpinning this are unknown. The mechanism of resistance associated with the *van* cluster, replacement of D-Ala with D-Ser in nascent peptidoglycan, has been widely described (Reynolds and Courvalin, 2005; Shen et al., 2020); and the mechanism of *dacJRS* resistance, cleavage of terminal D-Ala binding sites, was elucidated in this project. Multiple downstream interactions between these pathways may explain

the synergy observed – for example, the VanXY carboxypeptidase may be the bottleneck in *van* cluster-mediated resistance, meaning overexpression of the DacJ D,D-carboxypeptidase would provide relief to this mechanism. Alternatively, the combination of increased 3,3 crosslinks and D-Ser pentapeptide precursor peptides may aid high-level resistance. In any case, further investigation into this relationship, starting with peptidoglycan analysis, would merit interesting future work. Moreover, additional fluorescence microscopy of Bc1Δ*dacJ*Δ*vanS*, to assess whether deletion of both genes fully restores vancomycin binding, would complete the Bc1 story.

6 Investigating the Roles of *vanT* and *comR* in Vancomycin Resistance

6.1 Introduction

In previous chapters, the roles of *dacRS* and *vanRS* in promoting vancomycin resistance were explored. However, aside from TCS, there are multiple mechanisms by which resistance may be regulated. As post-transcriptional regulation is one of the major levels of gene expression control, one important regulatory mechanism is the RNA degradosome. The RNA degradosome is a multi-enzyme complex responsible for bulk RNA decay (Tejada-Arranz et al., 2020). Although well-studied in *E. coli*, the degradosome varies across bacterial species, and has only recently been fully characterised in *S. aureus.* Common to all characterised RNA degradosomes are two core components: an RNase, and a DEAD-box RNA helicase. In many species, however, polynucleotide phosphorylases (PNPases) and multiple additional RNases are typical (Cho, 2017). In fact, the *S. aureus* degradosome contains four RNases (Y, J1, J2, and RnpA), PNPase, enolase, phosphofructokinase, and RNA helicase. These enzymes function together to degrade RNA through unfolding the RNA secondary structure (DEAD-box RNA helicase), endoribonuclease activity (RNase Y, J1, J2, RnpA), 5’ → 3’ exoribonuclease activity (RNase J1 and J2), and 3’ → 5’ exoribonuclease activity (PNPase) (Cho, 2017; Tejada-Arranz et al., 2020).

Multiple examples of the RNA degradosome promoting survival under antibiotic pressures have been reported. In *Mycobacterium tuberculosis*, RNase J was found to be disproportionately mutated across multidrug resistant *M. tuberculosis* isolates, and deletion of RNase J resulted in increased multidrug tolerance (Martini et al., 2022). Multiple studies have implicated PNPase in both persistence and resistance: in *E. coli*, the PNPase gene deletion resulted in reduced persister formation. The association between PNPase and persistence was determined through RNAseq – PNPase was found to negatively regulate *crp*, a global positive regulator of metabolism (Wu et al., 2022). In *P. aeruginosa*, a PNPase mutant which led to downregulation of pyocin biosynthesis genes was directly implicated in ciprofloxacin resistance (Fan et al., 2019). Taken together, it is clear that alterations to PNPase can affect the expression of numerous genes, and have pleiotropic cellular effects.

Alternatively, resistance may be regulated epigenetically. Epigenetics describes the heritable changes which occur without changes to the DNA sequence. In bacteria, the best described epigenetic mechanism is that of DNA methylation, typically of adenine bases (Casadesús and Low, 2006). Following replication, methylated DNA is converted into two hemimethylated duplexes, whereby methylation exists on only one strand. These sites are rapidly re-methylated, ensuring heritability. Methylation occurs via DNA methyltransferases, and can modulate multiple cellular processes, including restriction-modification systems, initiation of replication, and regulation of gene expression (Marinus and Casadesus, 2009). Epigenetic gene expression regulation is a complex, multi-level process – regulation may occur via modifications to promoters or regulatory sequences, which may in turn affect binding of RNA polymerases or transcriptional regulators. Recently, evidence of regulation via intragenic methylation has also been reported (Zhao et al., 2023). Although full-genome methylation analysis can be performed with increasing ease, owing to the ever-reducing costs of single molecule real-time sequencing, the combinatorial use of this data with

transcriptomics, to associate methylation with regulation, is still a relatively new field (Payelleville and Brillard, 2021). That said, the epigenome of multiple *C. difficile* clinical isolates was recently elucidated. Through methylome and transcriptomic analyses, a novel DNA methyltransferase, CamA, was discovered, inactivation of which resulted in pleiotropic effects on sporulation, cell length, biofilm formation and virulence (Oliveira et al., 2020).

The incomplete restoration of resistance through recapitulation of genetic mechanisms is a well-documented observation, supporting the notion that epigenetic changes can modulate antibiotic resistance. Epigenetic mechanisms have been directly associated with antibiotic resistance in *E. coli*, whereby changes in DNA methylation were associated with altered antibiotic resistance-associated gene expression patterns. Interestingly, these changes were antibiotic-dependent, and resulted in modulation of genes commonly mutated in the genomes of resistant isolates (D'Aquila et al., 2023). In *M. tuberculosis*, differential methylation and expression analysis was used to explore the role of epigenetics in para-aminosalicylic acid resistance, reporting global changes in metabolism and ABC transporter expression (Li et al., 2020). However, the epigenetic mechanisms of antibiotic resistance in *C. difficile* are as yet unknown.

Beyond *dacRS* and *vanRS*, a further pathway to vancomycin resistance was observed in the evolved isolates. Mutations in *vanT* occurred in four of the ten evolved isolates, co-occurring with *comR* mutations in two isolates (Bc3 and Bc4). In isolates displaying *vanT*/*comR* co-occurrence, few other vancomycin-unique mutations were observed – in Bc3, *vanT* and *comR* mutations were the only vancomycin-unique mutations, and in Bc4, only one other vancomycin-unique mutation (*CD0108*) was identified – suggesting the combination of *vanT* and *comR* is a major contributor to the high-level resistance observed. *comR* encodes a homologue of the RNA degradosome component PNPase. Consistent with this observed *vanT*-degradosome co-occurrence, Bc5 displayed a SNP in

vanT, and a 75 bp deletion in the intergenic region between *rpmH* (encoding 50S ribosomal protein L34) and *rnpA* (encoding a predicted RNA degradosome endoribonuclease). Together, these observations suggest a link between the RNA degradosome and *vanT* in promoting vancomycin resistance.

6.2 Aims and Outcomes

This chapter describes the initial characterisation of the *vanT-comR*-mediated vancomycin resistance pathway, including its contribution to resistance, the associated fitness costs, initial mechanistic insights, and investigations into its regulation. To achieve this, *vanT* and *comR* mutations were recapitulated in the parental R20291Δ*PaLoc* background. The contributions of these mutations, individually and in combination, were tested using MIC assays, and the phenotypic effects of SNP carriage were assessed using growth curves and sporulation efficiency assays. Early mechanistic insights were achieved through qRT-PCR analysis of gene expression.

6.3 Results

6.3.1 The Contribution of *vanT* and *comR* Mutations to Vancomycin Resistance

To assess the contributions of *vanT* and *comR* to resistance, the mutations observed in the evolved isolates were precisely recapitulated in the parental R20291Δ*PaLoc* background. *vanT* SNPs were observed in four of the ten evolved isolates (Bc3, 4, 5 and 11), and co-existing *comR* mutations were observed in Bc3 and Bc4. All individual *vanT* and *comR* mutations, as well as the *vanT/comR* combinations observed in Bc3 and Bc4, were recapitulated. Although concurrent mutations relating to *vanT* and the RNA degradosome were also observed in Bc5, the 75 bp deletion preceding *rnpA* could not be constructed, as the deletion vector was lethal in *E. coli*, meaning investigations into this pathway were focussed solely on the *comR* PNPase homologue.

6.3.1.1 Construction of Recapitulated Strains

Recapitulation of mutations in *C. difficile* was performed via allelic exchange as previously described, generating strains R20291Δ*PaLoc comR*c.836_856del, R20291Δ*PaLoc comR*c.836_856del *vanT*c.688G>A, R20291Δ*PaLoc comR*c.1337G>T, R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A, R20291Δ*PaLoc vanT*c.271C>A, R20291Δ*PaLoc vanT*c.347G>A, R20291Δ*PaLoc vanT*c.359A>G, and R20291Δ*PaLoc vanT*c.688G>A. Table 6.1 summarises the construction of plasmids for recapitulated mutant strains relevant to this chapter.

Table 6.1 – Recapitulated Mutant Strains Relevant to This Chapter

Strain	Construction	Plasmid
R20291Δ*PaLoc comR*c.836_856del	Fragment designed to contain the *comR*c.836_856del deletion with approximately 1 kb either side of the mutation. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB021
R20291Δ*PaLoc comR*c.836_856del *vanT*c.688G>A	Fragments containing mutations synthesised by Genewiz, followed by ligation into linearised pJAK112. Double mutant generated by addition of pJEB027 to R20291Δ*PaLoc comR*c.836_856del.	pJEB021 pJEB027

R20291Δ*PaLoc* *comR*c.1337G>T	Fragment designed to contain the *comR*c.1337G>T point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB022
R20291Δ*PaLoc* *comR*c.1337G>T *vanT*c.271C>A	Fragments containing mutations synthesised by Genewiz, followed by ligation into linearised pJAK112. Double mutant generated by addition of pJEB031 to R20291Δ*PaLoc comR*c.1337G>T.	pJEB022 pJEB031
R20291Δ*PaLoc* *vanT*c.271C>A	Fragment designed to contain the *vanT*c.271C>A point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB031
R20291Δ*PaLoc* *vanT*c.347G>A	Fragment designed to contain the *vanT*c.347G>A point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB008
R20291Δ*PaLoc* *vanT*c.359A>G	Fragment designed to contain the *vanT*c.359A>G point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB012
R20291Δ*PaLoc* *vanT*c.688G>A	Fragment designed to contain the *vanT*c.688G>A point mutation with approximately 1 kb either side of the SNP. Fragment synthesised by Genewiz, followed by ligation into linearised pJAK112.	pJEB027

6.3.1.2 *comR* and *vanT* Only Partially Recapitulate the Evolved Isolate Resistance

To understand the contribution of each mutation and combination to vancomycin resistance, vancomycin MIC assays were performed for recapitulated strains using standard agar dilution. None of the four single *vanT* recapitulated strains conferred vancomycin resistance, all displaying an MIC of 1 µg/mL, identical to that of the R20291Δ*PaLoc* control (Figure 6.1). Conversely, both single *comR* recapitulated strains displayed low-level vancomycin resistance, with MICs of 2 µg/mL (Bc3 *comR*c.836_856del) or 4 µg/mL (Bc4 *comR*c.1337G>T). No multiple effects, neither additive nor synergistic, were observed in the R20291Δ*PaLoc comR*c.836_856del *vanT*c.688G>A and R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A combination strains – both displayed MICs identical to the respective single *comR* recapitulated strains. This suggests *comR* is not a potentiating mutation required for *vanT*-mediated resistance. Together, these findings suggest *comR* is involved

in providing low-level (up to 4-fold) vancomycin resistance, and is capable of doing so without any additional mutations.

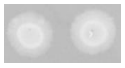
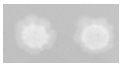
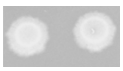
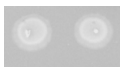
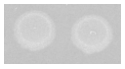
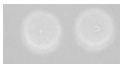
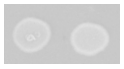
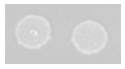
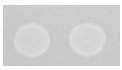
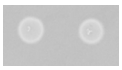
Isolate	Vancomycin (µg/mL)					
	0	0.5	1	2	4	MIC
R20291Δ*Paloc*						1
R20291Δ*PaLoc comR*c.836_856del						2
R20291Δ*PaLoc comR*c.836_856del *vanT*c.688G>A						2
R20291Δ*PaLoc comR*c.1337G>T						4
R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A						4
R20291Δ*PaLoc vanT*c.271C>A						1
R20291Δ*PaLoc vanT*c.347G>A						1
R20291Δ*PaLoc vanT*c.359A>G						1
R20291Δ*PaLoc vanT*c.688G>A						1

Figure 6.1 - comR Mutations Partially Recapitulate Evolved Isolate Resistance

Vancomycin MICs of R20291Δ*PaLoc*, R20291Δ*PaLoc comR*c.836_856del, R20291Δ*PaLoc comR*c.836_856del *vanT*c.688G>A, R20291Δ*PaLoc comR*c.1337G>T, R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A, R20291Δ*PaLoc vanT*c.271C>A, R20291Δ*PaLoc vanT*c.347G>A, R20291Δ*PaLoc vanT*c.359A>G, and R20291Δ*PaLoc vanT*c.688G>A. MICs were determined by agar dilution on BHI. Assays were performed in biological triplicate and technical duplicate. Shown is a single representative biological replicate.

6.3.1.3 *vanT* Overexpression Does Not Increase Vancomycin Resistance

Since the recapitulated *vanT* mutants did not confer any level of vancomycin resistance, further investigations relating to the contributions of these mutations to resistance were performed. All four observed *vanT* mutations were SNPs, meaning loss of VanT function was unlikely. The observed mutations could therefore have multiple possible outcomes. First, the *vanT* mutations may affect *vanT* expression – although the mutations occurred within *vanT,* as opposed to the *vanRS* TCS, *vanT* is the last gene in the *van* cluster transcript (Fuchs et al., 2021). As it is well-documented that expression decreases towards the end of an operon (Lim et al., 2011), it is possible that the observed SNPs alter the mRNA secondary structure, providing increased stability to the end of the *van* cluster transcript, thus increasing *vanT* expression. Importantly, SNPs with this effect would only provide resistance in specific environmental conditions when the *van* cluster is active, for example during experimental evolution, which may explain why the recapitulated *vanT* mutants did not display increased resistance during the MIC assay. Alternatively, it is possible that *vanT* SNPs, which arose later in the evolution, are refining mutations, providing minor enhancements to the VanT racemase to increase *van* cluster-mediated vancomycin resistance. To assess these possibilities, several *vanT* overexpression strains were constructed. If the *vanT* SNPs increase mRNA stability to increase *vanT* expression, then overexpression of the WT *vanT* gene should provide increased vancomycin resistance. However, if the *vanT* SNPs are refining mutations, perhaps providing only a small increase in resistance not discernible in the *vanT* recapitulated strains, then overexpression of the evolved isolate *vanT* genes may help to illuminate this phenotype.

A collection of *vanT* overexpression strains were constructed in the R20291Δ*PaLoc* genetic background, each carrying a plasmid-borne copy of *vanT* (WT *vanT*, *vanT*c.271C>A, *vanT*c.347G>A, *vanT*c.359A>G, or *vanT*c.688G>A), under the control of a tetracycline-inducible promoter. A control strain, harbouring the same plasmid backbone but expressing *gusA* (which was confirmed not to

impact vancomycin resistance) was also constructed. In each case, *vanT* was amplified by PCR using the primer pair RF2364/RF2365. The resulting *vanT* fragments were ligated into pRPF185 (linearised by BamHI/SacI restriction digest), completing the *vanT* overexpression vectors. The pRPF185 vector backbone contains a tet-inducible expression system, consisting of divergent promoters PtetR and Ptet (Fagan and Fairweather, 2011). TetR, expressed from the PtetR promoter, negatively regulates both promoters in the absence of tetracycline. De-repression of TetR in the presence of tetracycline allows expression of the gene of interest (from Ptet) and increased expression of TetR, creating a negative feedback loop to allow dose-dependent induction. The *vanT* overexpression vectors were conjugated into R20291Δ*PaLoc* as described previously, generating the desired *vanT* overexpression strains. Table 6.2 summarises the construction of overexpression plasmids relevant to this chapter.

Table 6.2 - Overexpression Constructs Relevant to This Chapter

Description	**Construction**	**Plasmid**
R20291Δ*PaLoc* + control vector	Previously constructed.	pRPF185
R20291Δ*PaLoc* + WT *vanT* vector	*vanT* amplification from R20291Δ*PaLoc* gDNA using RF2364/RF2365. Ligated into pRPF185 linearised by BamHI/SacI restriction digest.	pJEB013
R20291Δ*PaLoc* + *vanT*c.271C>A vector	*vanT*c.271C>A amplification from Bc4 gDNA using RF2364/RF2365. Ligated into pRPF185 linearised by BamHI/SacI restriction digest.	pJEB014
R20291Δ*PaLoc* + *vanT*c.347G>A vector	*vanT*c.347G>A amplification from Bc5 gDNA using RF2364/RF2365. Ligated into pRPF185 linearised by BamHI/SacI restriction digest.	pJEB015
R20291Δ*PaLoc* + *vanT*c.359A>G vector	*vanT*c.359A>G amplification from Bc11 gDNA using RF2364/RF2365. Ligated into pRPF185 linearised by BamHI/SacI restriction digest.	pJEB016
R20291Δ*PaLoc* + *vanT*c.688G>A vector	*vanT*c.688G>A amplification from Bc3 gDNA using RF2364/RF2365. Ligated into pRPF185 linearised by BamHI/SacI restriction digest.	pJEB017

To assess whether overexpression of *vanT* resulted in vancomycin resistance, MICs of the overexpression strains were evaluated. MIC assays were performed using standard agar dilution, with the addition of 15 µg/mL thiamphenicol to select for the plasmid. MICs were assayed with and without the presence of 60 ng/mL aTc, a non-antibiotic analogue of tetracycline used to induce overexpression. Overexpression of *vanT* did not result in vancomycin resistance in any of the overexpression strains (Table 6.3). Overexpression of both WT and *vanT* SNP genes provided an MIC of 1 µg/mL, identical to that of the *gusA* control, suggesting the *vanT* mutations do not promote mRNA stability to increase *vanT* expression. It is possible, therefore, that the observed *vanT* SNPs are refining mutations, which provide a small enhancement to vancomycin resistance, which (even with overexpression) is not discernible through the standard agar dilution MIC method.

Table 6.3 - Vancomycin MICs of vanT Overexpression Strains

Overexpressed Gene	**Vancomycin MIC (µg/mL)**	
	+ aTc	**- aTc**
gusA control	1	1
WT *vanT*	1	1
*vanT*c.271C>A	1	1
*vanT*c.347G>A	1	1
*vanT*c.359A>G	1	1
*vanT*c.688G>A	1	1

6.3.2 Probing the *comR*-mediated Mechanism of Resistance

As *vanT* mutations may provide only minor contributions to the resistance profiles of evolved isolates, the pathway of low-level resistance associated with *comR* was examined. The *comR* SNP observed in the Bc4 evolved isolate was chosen for further characterisation, as this mutation supplied the largest increase (4-fold) in resistance.

6.3.2.1 *comR* Mutations Result in Severe Growth and Sporulation Defects

The Bc4 evolved isolate displayed significant defects in growth and sporulation efficiency. To investigate whether *comR*c.1337G>T was responsible for the loss of fitness in Bc4, and to examine the fitness costs directly associated with the *comR*-mediated resistance pathway, phenotypic assessments of growth and sporulation were performed. *comR* is a PNPase homologue, meaning mutations in this gene could potentially lead to severe pleiotropic effects. It is possible, therefore, that the co-existing *vanT*c.271C>A mutation actually refines fitness, helping to rescue any severe fitness defects associated with *comR*. To assess this possibility, the *vanT*c.271C>A strain, and the combination mutant *comR*c.1337G>T *vanT*c.271C>A, were also included in analyses.

To evaluate growth of the recapitulated strains, R20291Δ*PaLoc comR*c.1337G>T, R20291Δ*PaLoc vanT*c.271C>A and R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A were assayed in rich media as described in previous chapters, and compared with the parental R20291Δ*PaLoc* control. AUC was used to statistically compare growth of each of each strain with the control. R20291Δ*PaLoc comR*c.1337G>T displayed significant growth defects, with a similar profile to the Bc4 evolved isolate, suggesting *comR*c.1337G>T is a main contributor to the Bc4 phenotype (Figure 6.2a). Interestingly, the growth profile of R20291Δ*PaLoc vanT*c.271C>A was significantly better than R20291Δ*PaLoc*, possessing a comparable generation time but reaching a higher maximum OD_{600nm}. However, this mutation did not rescue the *comR*-mediated growth defects, since the R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A combination strain showed similar growth to the *comR* single mutant. This suggests the *comR* pathway to vancomycin resistance is associated with significant fitness costs in terms of growth.

To determine the contribution of *comR*c.1337G>T to the sporulation defects observed in the Bc4 evolved isolate, sporulation efficiencies of R20291Δ*PaLoc comR*c.1337G>T, R20291Δ*PaLoc*

*vanT*c.271C>A and R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A were assayed as described previously. In each case, sporulation efficiency was compared to the parental R20291Δ*PaLoc* control. Significant sporulation defects were observed in the R20291Δ*PaLoc comR*c.1337G>T strain – although sporulation efficiency was comparable to the control by the end of the experiment, this strain displayed a delayed sporulation phenotype (Figure 6.2b). Despite no sporulation defects being apparent in the *vanT*c.271C>A mutant, this mutation again did not rescue the defects observed in the *comR*c.1337G>T strain, with the combination mutant showing the same sporulation profile as R20291Δ*PaLoc comR*c.1337G>T. Together, this suggests that the *comR* SNP is largely responsible for the growth and sporulation defects observed in the Bc4 evolved isolate, demonstrating that the *comR* pathway provides low-level, but costly, vancomycin resistance.

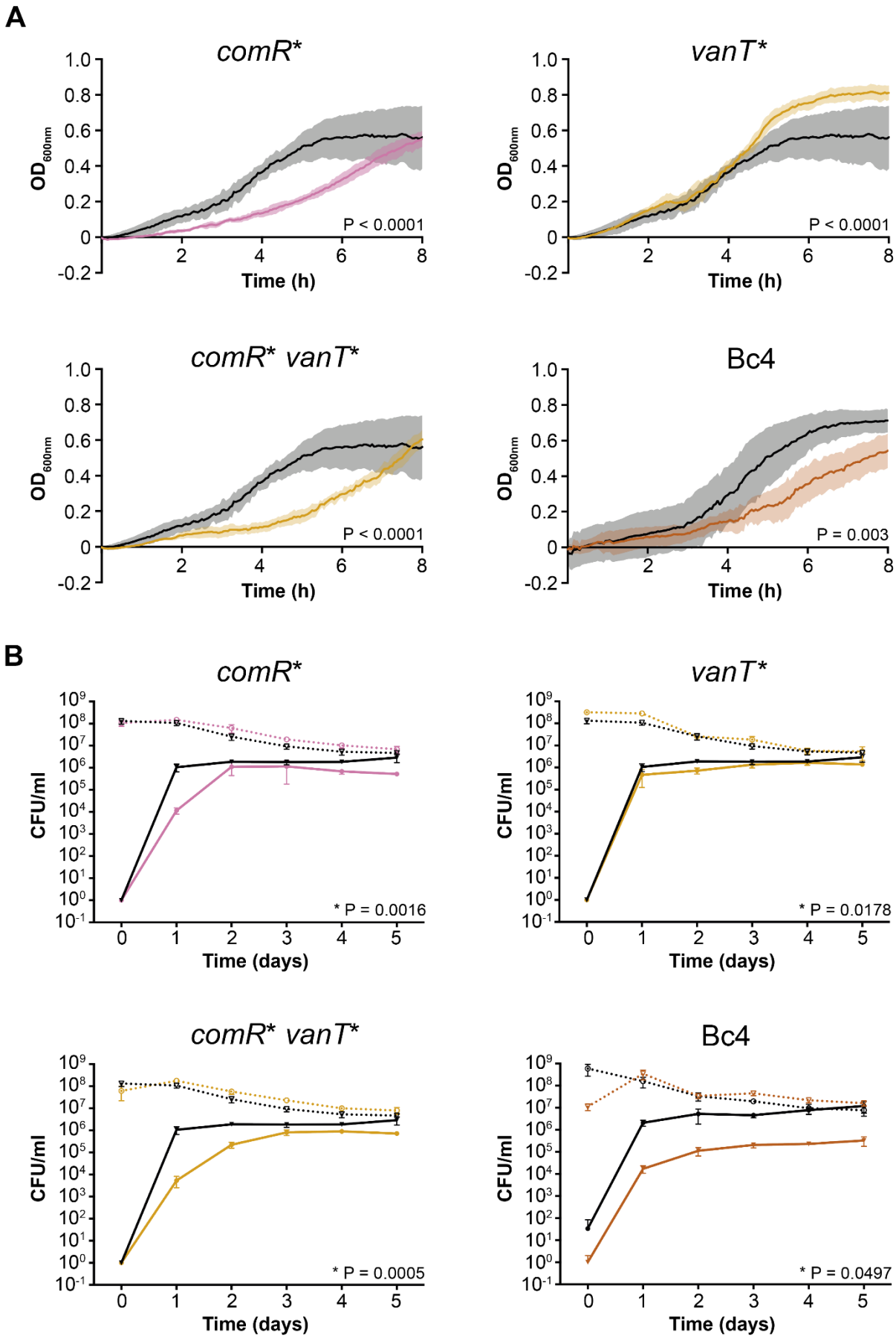


Figure 6.2 - Growth and Sporulation of Bc4 Recapitulated Strains

(A) Growth over time in rich media (TY broth) was evaluated by measuring OD at 600 nm in a 96 well microplate spectrometer. Growth of R20291Δ*PaLoc comR*c.1337G>T (*comR**), R20291Δ*PaLoc vanT*c.271C>A (*vanT**), and R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A (*comR** *vanT**) (coloured lines) was compared to R20291Δ*PaLoc* (black lines). Growth of evolved isolate Bc4 (coloured line) and its matched control (black line) was included for comparison. The mean and standard deviation of repeats, assayed in biological and technical triplicate, are presented. For each strain, area under the curve was determined using the GrowthCurver R package and these were compared using Student's t-tests with Welch's correction, with the P value shown on each graph. (B) Sporulation efficiencies of R20291Δ*PaLoc comR*c.1337G>T (*comR**), R20291Δ*PaLoc vanT*c.271C>A (*vanT**), R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A (*comR** *vanT**) and endpoint clone Bc4 (coloured lines) compared to the parental R20291Δ*PaLoc* (black lines). Stationary phase cultures were incubated anaerobically for 5 days with samples taken daily to enumerate total colony forming units (CFUs, dotted lines) and spores (solid lines), following incubation at 65°C for 30 min to kill vegetative cells. Shown are the mean and standard deviations of biological and technical triplicate. For each strain, spore CFU area under the curve was determined using GraphPad Prism and these were compared using Dunnett's T3 multiple comparisons test with the adjusted P value shown on each graph. * = significant difference, N.S. = not significant.

6.3.2.2 *comR* Mutations Do Not Contribute to Teicoplanin Cross-resistance

The Bc4 evolved isolate displayed a 2-fold increase in teicoplanin resistance. To assess whether this increase was due to the *comR*-mediated resistance pathway, the teicoplanin MIC of R20291Δ*PaLoc comR*c.1337G>T was measured. The MIC of R20291Δ*PaLoc comR*c.1337G>T was 0.25 µg/mL, identical to the R20291Δ*PaLoc* control, suggesting that teicoplanin resistance in Bc4 arises from a separate pathway, and that *comR* does not contribute to cross-resistance to this antibiotic.

6.3.2.3 *comR* Mutations Do Not Act on Genes in Previously Defined Pathways

As PNPases possess exoribonuclease activity to degrade RNA, mutations in *comR* likely have global effects on transcript stability. However, how mutations in this post-transcriptional global regulator directly impact vancomycin resistance are unknown. To attempt to identify vancomycin resistance-promoting targets with increased RNA stability resulting from the *comR*c.1337G>T mutation, the expression of the two major pathways to vancomycin resistance identified in this project were assessed.

Since *comR* and *vanT* mutations co-occurred in two of the ten evolved isolates, expression of the *van* cluster was assessed first, to determine whether the *comR*c.1337G>T mutation led to increased stability of the *van* cluster transcripts. R20291Δ*PaLoc comR*c.1337G>T and R20291Δ*PaLoc* were compared in the presence (0.5x MIC) and absence of vancomycin. *vanR, vanS, vanG, vanXY, vanT* and housekeeping gene *rpoA* were measured using qRT-PCR, and mRNA copy numbers were quantified using serial dilution of the pJEB032 plasmid as described previously. The copy numbers were standardised using the *rpoA* housekeeping gene, and results were expressed as copies per 1000 copies of *rpoA*. No difference in expression was observed between R20291Δ*PaLoc comR*c.1337G>T and R20291Δ*PaLoc* for any of the *van* cluster genes, in either the presence or absence of vancomycin, suggesting mutations in *comR* do not increase transcript stabilities of the *vanRS* or *vanGXYT* operons (Figure 6.3a).

As the *comR*c.1337G>T mutation did not increase stability of the *van* gene transcripts, the possibility that *comR*c.1337G>T elicits resistance through increasing the stability of *dacJRS* transcripts was explored. Again, R20291Δ*PaLoc comR*c.1337G>T and R20291Δ*PaLoc* were compared in the presence (0.5x MIC) and absence of vancomycin, however qRT-PCR was used to measure expression of the *dacJRS* cluster. mRNA copy numbers were quantified using serial dilution of the pJEB029 plasmid as described previously. No difference in *dacJRS* expression was observed in either vancomycin condition, suggesting mutations in *comR* do not result in vancomycin resistance through increasing *dacJRS* transcript stability (Figure 6.3b).

Despite being unable to elucidate the vancomycin resistance-promoting target of R20291Δ*PaLoc comR*c.1337G>T, this work demonstrates the *comR*-mediated vancomycin resistance pathway is separate to the two previously identified pathways (the *van* cluster and *dacJRS*), acting via thus far undefined mechanisms to promote vancomycin resistance. To fully characterise transcripts with

altered stability as a result of the *comR*c.1337G>T mutation, and to associate this mutation with the downstream pathways leading to vancomycin resistance, RNAseq would be required.

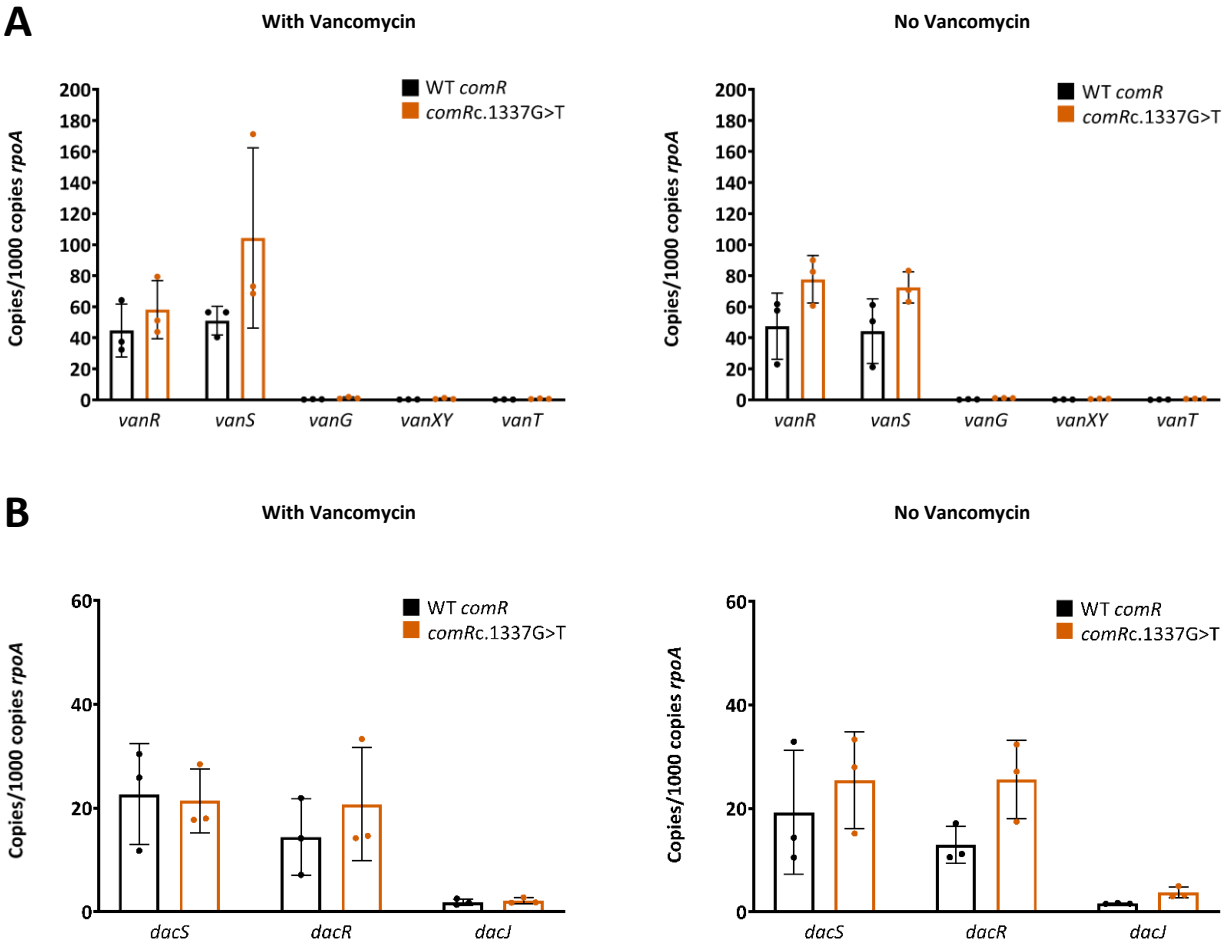


Figure 6.3 - comRc.1337G>T Does Not Alter Expression of van or dacJRS Genes

qRT-PCR analysis of the *van* (A) and *dacJRS* (B) genes, in R20291Δ*PaLoc* (black) and R20291Δ*PaLoc comR*c.1337G>T (orange). Expression was quantified against a standard curve and normalised relative to the house-keeping gene *rpoA*. Assays were performed in biological and technical triplicate. Statistical significance was calculated using a two-way ANOVA with the Tukey-Kramer test. No differences were significant.

6.3.3 Closing the Bc3 and Bc4 Genomes

The 16-fold increase in vancomycin MIC observed in the Bc3 and Bc4 evolved isolates could not be fully recapitulated by engineering individual SNPs in a clean genetic background. Despite recapitulating the only two vancomycin-unique mutations identified in Bc3 from illumina data, *comR*c.836_856del and *vanT*c.688G>A, only a 2-fold increase in resistance was observed. Similarly, despite the Bc4-derived *comR*c.1337G>T SNP providing an MIC of 4µg/mL, this was still 4-fold lower than the MIC of the Bc4 evolved isolate. The previous chapter demonstrated the virtues of hybrid genome sequencing, as additional long-read sequencing of the Bc1 genome led to the identification of two further InDels, which allowed full recapitulation of the level of resistance displayed by Bc1. The same approach was therefore used to determine whether any additional variants were present in the Bc3 and Bc4 evolved isolate genomes, which could explain the incomplete recapitulation of resistance for these isolates. 50x whole genome long-read sequencing was performed using the Oxford Nanopore Technologies GridION flow cell at MicrobesNG. Long-read data was polished with matched Bc3 or Bc4 short-read data, as described in the previous chapter, to generate a complete closed genome and to improve variant calling. The same pipeline was used to call and validate variants.

No additional variants were identified in the Bc3 and Bc4 hybrid assemblies (full dataset in Appendix VII). This highlights the quality of the rigorous analysis previously performed on Illumina data to reduce false negative, and positive, calls. However, as resistance could not be fully recapitulated in these isolates, and no further variants were identified in the closed genomes, this leaves a question regarding the resistance pathways driving high-level vancomycin resistance in these isolates.

6.3.4 Investigating the Epigenetics of Vancomycin Resistance

As the 16-fold increase in resistance observed in the Bc3 and Bc4 evolved isolates could not be fully recapitulated, despite no further mutations being identified in the genome, it is likely the high-level resistance in these isolates arises from an alternative mechanism which does not alter the genome sequence. Epigenetic mechanisms are commonly cited in instances where recapitulation of resistance is incomplete (Muhammad et al., 2022), and have been previously implicated in antibiotic resistance (Li et al., 2020). Full characterisation of epigenetic resistance pathways occurring in the evolved isolates would require a combination of differential epigenomics and transcriptomics (Payelleville and Brillard, 2021). However, to investigate whether epigenetic mechanisms could explain the incomplete recapitulation observed, initial work focussed on exploring the idea that *vanT* mutations may contribute to vancomycin resistance in the right conditions, and that epigenetic changes could explain why recapitulation of this phenotype was not observed in *vanT* SNP mutants.

6.3.4.1 Methylation Sites in the *van* Cluster

The *C. difficile* methyltransferase CamA is involved in regulating a range of cellular processes, from sporulation to virulence (Herbert et al., 2003; Oliveira et al., 2020; van Eijk et al., 2015). CamA methylates adenine (m^6A), and has a target sequence of CAAAA**A** (whereby the final A is methylated). The CamA methylation sites of the *van* cluster were predicted using the target sequence and visualised using Geneious (v7.1.9 http://www.geneious.com/).

No methylation sites were observed in *van* cluster promoters (Figure 6.4). However, 12 potential m^6A sites were observed across the *van* cluster: nine intragenic, and two intergenic sites, and one in the 5' UTR of *vanR,* predicted by global transcription site mapping (Fuchs et al., 2021). Since intragenic methylation has also previously been associated with regulation (Zhao et al., 2023), together this suggests the *van* cluster may be subject to epigenetic regulation. The evolved isolates

may, for example, display altered *van* gene methylation patterns which are not present in the *vanT* recapitulated strains, potentially explaining the observed discrepancies in resistance.

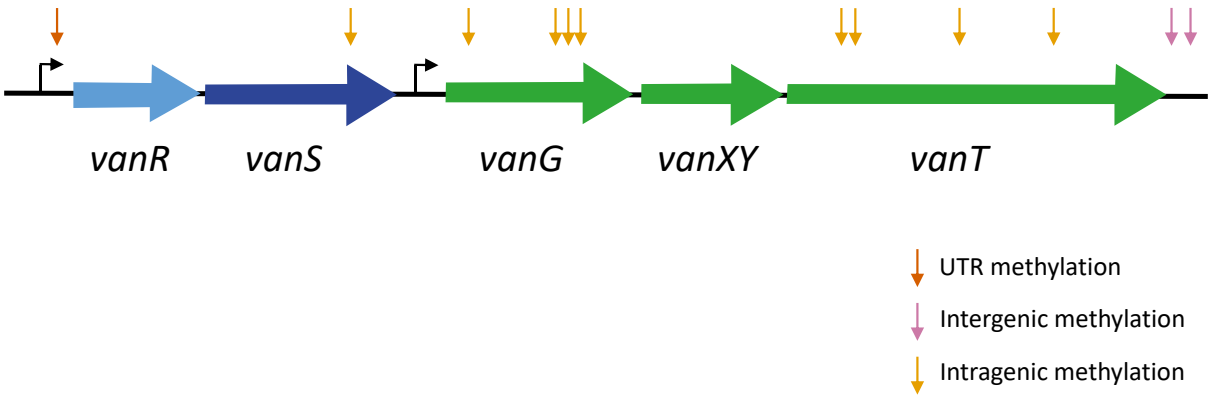


Figure 6.4 - Methylation Sites in the van Cluster

Shown are the intergenic and intragenic recognition sites of CamA (m^6A) in the *van* cluster. The CamA recognition site (CAAAA**A**) was annotated and visualised using Geneious (v7.1.9 http://www.geneious.com/).

6.3.4.2 Differences in *vanR* Expression Between Evolved Isolate and *vanT*c.271C>A

To assess whether the *van* genes are subject to epigenetic regulation, the expression of *van* gene transcripts was compared between R20291Δ*PaLoc,* R20291Δ*PaLoc vanT*c.271C>A, R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A, and Bc4. It was hypothesised that if epigenetic regulation was operational, differences in *van* transcript expression would be observed which would not be explainable through the strains' known genetic backgrounds. R20291Δ*PaLoc,* R20291Δ*PaLoc vanT*c.271C>A, R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A, and Bc4 were compared in the absence of vancomycin, and expression of *rpoA, vanR* and *vanG* (representing the two *van* cluster transcripts predicted by global transcription site mapping (Fuchs et al., 2021)) was measured using

qRT-PCR. mRNA copy numbers were quantified using serial dilution of the pJEB032 plasmid as described previously.

Significant upregulation of *vanR* was observed for both Bc4 and the *comR*c.1337G>T *vanT*c.271C>A combination mutant, compared with R20291Δ*PaLoc* and R20291Δ*PaLoc vanT*c.271C>A (Figure 6.5). Since the only genetic difference between R20291Δ*PaLoc vanT*c.271C>A and the upregulated R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A strain was the *comR* SNP, which was previously shown to have no effect on expression of the *van* cluster, this suggests epigenetic differences between these strains could explain the differences in *vanR* regulation.

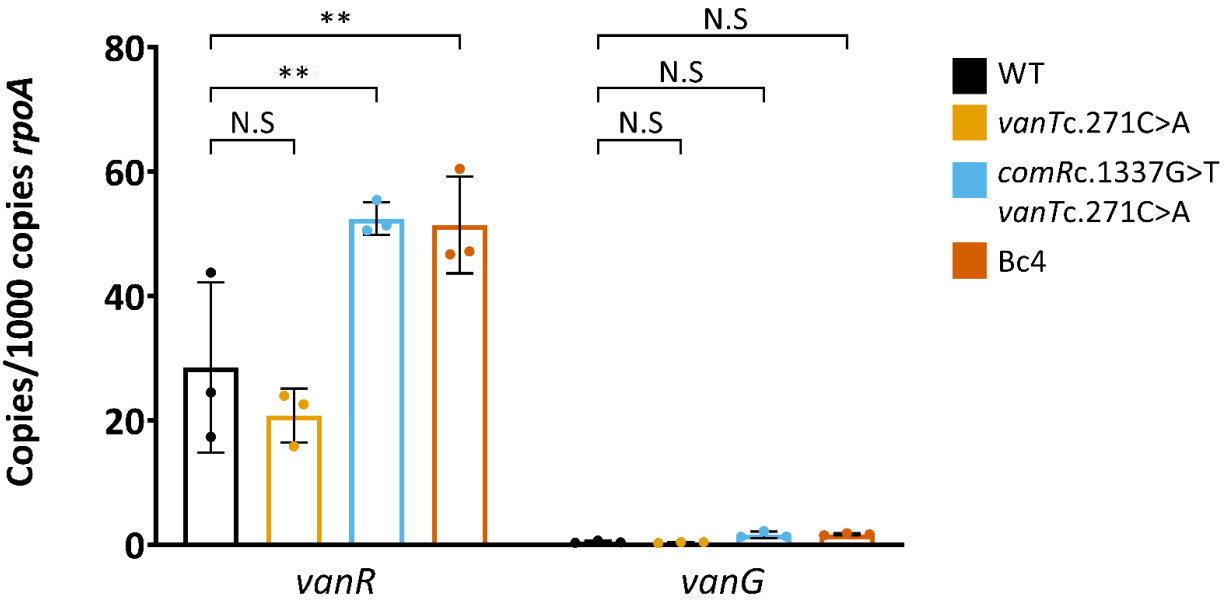


Figure 6.5 - Differences in vanR Expression Between Bc4 and vanTc.271C>A

qRT-PCR analysis of the two *van* cluster transcripts, represented by *van*R and *vanG*, in R20291Δ*PaLoc* (black), *vanT*c.271C>A (yellow), *comR*c.1337G>T *vanT*c.271C>A (blue), and evolved isolate Bc4 (orange). Expression was quantified against a standard curve and normalised relative to the house-keeping gene *rpoA*. Assays were performed in biological and technical triplicate. Statistical significance was calculated using a two-way ANOVA with the Tukey-Kramer test, ** = $P < 0.001$, N.S. indicates differences were not significant.

6.4 Discussion

Aside from *dacJRS,* mutations in *vanT,* which occurred in four of the ten evolved isolates, were identified as the other major resistance pathway observed during experimental evolution. Since *vanT* mutations consistently arose late in the evolution, this pathway is unlikely to be a mechanism of first-step resistance, perhaps instead requiring a prerequisite event or mutation. The striking co-occurrence of *comR* and *vanT* mutations in two evolved isolates inspired further questions relating to the nature of this resistance pathway. This chapter explores resistance mediated by *vanT* and *comR* mutations in Bc3 and Bc4. Through recapitulation of all *vanT* and *comR* mutations, singly and in combination, the contributions of these mutations to vancomycin resistance were determined. None of the *vanT* SNP mutants displayed an increased MIC, suggesting this mutation may be refining, as opposed to resistance-driving. Conversely, R20291Δ*PaLoc comR*c.1337G>T demonstrated increased, albeit costly, vancomycin resistance. However, although this pathway was shown to act separately from previously identified *van* and *dacJRS* pathways, the mechanism could not be identified. Hybrid assembly identified no further mutations in Bc3 and Bc4, so the possibility of epigenetic mechanisms driving high-level resistance was considered.

6.4.1 Bc4 Resistance Beyond *vanT* and *comR*

To understand the roles of *vanT* and *comR* in the resistance displayed by evolved isolates Bc3 and Bc4, *vanT* and *comR* mutations were recapitulated singly and in combination, and their contribution to resistance was assessed via MIC assays. In Bc3, only two mutations, *comR*c.836_856del and *vanT*c.688G>A were identified, meaning the combination strain R20291Δ*PaLoc comR*c.836_856del *vanT*c.688G>A provided a complete genetic recapitulation of the Bc3 evolved isolate. However, the MIC of this strain was just 2 µg/mL, 8-fold lower than the Bc3 MIC. Conversely, Bc4 possessed an additional *CD0108*c.173C>G mutation. Although this gene was annotated as a putative membrane protein, sequence (Altschul et al., 1990) and structural (Kelley et al., 2015) homology suggested *CD0108* encodes the cyclic di-AMP synthase CdaA. The second messenger synthase CdaA has direct

links to the regulation of peptidoglycan synthesis, as cyclic-di-AMP depletion reduces cell wall integrity, and cyclic-di-AMP accumulation leads to a reduction in D-Ala-D-Ala muropeptide precursors (Massa et al., 2020; Schwedt et al., 2023; Zhu et al., 2016). It is possible, therefore, the *CD0108*c.173C>G mutation results in increased cyclic-di-AMP production, resulting in fewer available D-Ala-D-Ala precursors, favouring D-Ala-D-Ser incorporation into the pentapeptide stem. Recapitulating the *CD0108*c.173C>G mutation, singly and in combination with *comR*c.1337G>T and *vanT*c.271C>A, would enable a fuller understanding of the resistance pathways occurring in Bc4, and would highlight any additive or synergistic interactions between this mutation and the *vanT/comR* pathway.

6.4.2 Further Insights into *vanT* Resistance

Despite *vanT* demonstrating remarkable parallel evolution, with SNPs occurring in Bc3, Bc4, Bc5 and Bc11, recapitulation of these mutations did not increase vancomycin resistance. Two initial hypotheses were investigated to assess how *vanT* may contribute to resistance without the recapitulated strains displaying a discernible change in MIC. First, *vanT* SNPs could alter the mRNA secondary structure to increase transcript stability and thus *vanT* expression. However, an observable increase in MIC in this case would rely on expression of the *van* cluster, and as these genes are not constitutively expressed in *C. difficile* (Ammam et al., 2013), this may explain the lack of resistance observed in the MIC assays of recapitulated *vanT* mutants. To test this possibility, WT *vanT* was overexpressed. Overexpression of WT *vanT* did not result in increased resistance, suggesting *vanT* mutations do not elicit resistance through alteration of transcript stability. Second, the *vanT* SNPs may be refining mutations, providing minor improvements to resistance, meaning the incremental resistance gains were not observable in the MICs of recapitulated strains. Overexpression of the *vanT* mutants was performed in an effort to enhance such refining effects, making the benefit to resistance observable. However, the MICs of strains overexpressing *vanT*c.271C>A, *vanT*c.347G>A, *vanT*c.359A>G, and *vanT*c.688G>A were not significantly different

from the control strain. Importantly, although the concentration of aTc used to induce *vanT* expression in the MIC assays was 60 ng/mL, which is within the range typically used for *C. difficile* Ptet induction, dose-dependent induction can be observed up to 500 ng/mL aTc (Fagan and Fairweather, 2011). It may be interesting, therefore, to investigate whether even higher levels of overexpression make these putative refinements observable. Overall, *vanT* mutations did not significantly contribute to vancomycin resistance, and therefore may be refining mutations. This is consistent with the late appearance of these mutations in P20 or P30 of the evolution, since refinement is the last stage in the evolution of a trait (Quandt et al., 2014). To further understand the possible refining nature of the *vanT* SNPs, exploring the functional relevance of the mutations, for example through racemase activity (Noda et al., 2005) or serine transport (Khozov et al., 2023) assays, would be interesting.

6.4.3 The *comR* PNPase Homologue

Since *vanT* mutations did not provide observable enhancements to vancomycin resistance, the mechanism of resistance associated with *comR*c.1337G>T, which provided a 4-fold increase in resistance, was examined. *comR*c.1337G>T was found to be the main contributor to the fitness defects observed in the Bc4 evolved isolate, suggesting *comR*-mediated resistance is costly, likely due to the pleiotropic effects of mutating a global regulator. As qRT-PCR showed *comR*c.1337G>T did not lead to an increase in *van* or *dacJRS* transcripts, the mechanism of vancomycin resistance associated with this mutation was not determined. RNAseq would thus be required to investigate global transcriptional differences between R20291Δ*PaLoc comR*c.1337G>T and R20291Δ*PaLoc*. A full understanding of the *comR* regulon would illuminate the direct connection between *comR*c.1337G>T and vancomycin resistance.

Since three of the evolved isolates (Bc3, Bc4 and Bc5) all displayed links to the RNA degradosome, considerations of PNPase were confined to the assessment of degradosome-associated PNPase

resistance. However, it is important not to forget that PNPase can also function as a 3' → 5' exoribonuclease independently of the degradosome. In fact, evidence suggests PNPase primarily functions independently – global expression profiling of *E. coli* found many transcripts stabilised by mutation of PNPase were not impacted by mutation of the degradosome (Bernstein et al., 2004; Cameron et al., 2018). Notably, in *Listeria monocytogenes*, PNPase was recently reported as a positive regulator of biofilm formation, and deletion of PNPase resulted in reduced biofilm biomass and increased antibiotic susceptibility (Quendera et al., 2023). The *comR*c.1337G>T mutation observed here could therefore increase biofilm formation to promote vancomycin resistance. Beyond gene expression modulation, PNPase functions independently of the degradosome in small RNA (sRNA) regulation (De Lay and Gottesman, 2011). It is widely documented that sRNAs contribute to antibiotic resistance (including vancomycin) and tolerance, meaning it is possible the PNPase-mediated vancomycin resistance mechanism involves altered regulation of sRNAs (Mediati et al., 2021; Sinel et al., 2017). Taken together, the wide-reaching influence of PNPase in modulating expression of RNAs suggests this vancomycin resistance pathway is multifactorial, meaning unpicking the intricacies of PNPase-mediated resistance is unlikely to be straightforward.

6.4.4 Considering Epigenetic Resistance Mechanisms

The incomplete recapitulation of resistance observed in evolved isolates Bc3 and Bc4, despite introduction of both *comR* and *vanT* mutations, suggested the presence of a further vancomycin resistance mechanism in these isolates. However, hybrid assembly using long-read sequencing identified no further vancomycin-unique variants in these isolates, suggesting the further mechanism of resistance was not genetically encoded. Since the CamA methyltransferase is highly conserved, and was shown to influence transcriptional responses in *C. difficile* (Oliveira et al., 2020), m^6A sites were predicted across the *van* cluster. 12 putative methylation sites were revealed, suggesting this cluster may be subject to epigenetic regulation. Importantly, *C. difficile* possesses multiple other methyltransferases (van Eijk et al., 2015), which may also be responsible for

transcriptional modulation. A more holistic insight, encompassing all putative modifications, may therefore be gained through comparison of the *van* cluster in single molecule real-time sequenced genomes.

To explore the possibility of epigenetic vancomycin resistance, initial investigations focussed on comparing expression of *van* gene transcripts between recapitulated mutant strains and the evolved isolate Bc4. Slight, but significant, differences in regulation of *vanR* were observed between R20291Δ*PaLoc vanT*c.271C>A and R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A. As the only genetic difference between these strains, the *comR*c.1337G>T mutation, was shown not to impact *van* gene expression, this is suggestive of epigenetic regulation. Although both the methylomes and transcriptomes of these strains would be required to unequivocally implicate epigenetic alterations in vancomycin resistance, these initial observations suggest a more in-depth epigenetic characterisation would be worthwhile.

7 Discussion

7.1 Overview

Despite vancomycin being one of the few routine treatments for CDI worldwide, and the frontline drug of choice in the UK (NICE, 2021), resistance to this antibiotic in *C. difficile* has not been widely reported. However, the increased use of vancomycin since its promotion to frontline status in 2021, coupled with emerging evidence of resistance in clinical isolates (Darkoh et al., 2021), and anecdotal reports of treatment failure (Molleti et al., 2023), suggest increased priority should be placed on understanding vancomycin resistance in *C. difficile*. Previous evolutionary studies have shown that stable vancomycin resistance in *C. difficile* is possible, at least *in vitro*, however the impacts on fitness and pathogenicity remain unclear. Additionally, beyond regulatory changes to the *van* cluster (Shen et al., 2020), the pathways to vancomycin resistance, including alternative routes to resistance, and the accumulation of mutations leading to high-level resistance, remain unexplored. This project therefore aimed to address these unknowns by (i) using a parallel approach to evolve high-level vancomycin resistance in *C. difficile in vitro,* (ii) providing phenotypic assessments of evolved isolates to understand the fitness costs involved in different routes to resistance, (iii) using whole genome sequencing to assess the mutations involved in resistance acquisition on an individual and population level, and (iv) determining the molecular mechanisms of resistance. High-level resistance evolved rapidly in *C. difficile*, at the cost of severe phenotypic changes. Multiple mutations, including *dacS* and *vanS* variants, were implicated in vancomycin resistance, and the molecular mechanisms associated with these pathways were discerned through qRT-PCR and fluorescence microscopy.

7.2 Large-scale, Parallel Evolution in *C. difficile*

Over the last decade, experimental evolution has begun to be utilised to investigate antibiotic resistance in *C. difficile.* Such studies have determined the role of *rpoB* in fidaxomicin resistance (Leeds et al., 2014), and the accumulation of mutations required for metronidazole resistance (Deshpande et al., 2020). Few works exploring the evolution of vancomycin resistance have been described in *C. difficile*, perhaps reflective of the historically low rate of resistance reported in clinical isolates. Mechanistic insights into vancomycin resistance have largely centred on the *C. difficile van* cluster, which displays homology to the *Enterococci vanG* operon (Reynolds and Courvalin, 2005; Shen et al., 2020). However, recent reports of vancomycin resistant clinical isolates which do not display mutations in the *van* cluster suggest alternative routes to resistance are possible (Eubank et al., 2024). Furthermore, although large-scale experimental evolution is commonplace in *S. aureus* and *E. coli* (Lenski, 2017; Papkou et al., 2020; Sulaiman and Lam, 2021), experimental evolution studies in *C. difficile* have generally been confined to assessment of single populations, meaning convergent evolution cannot be evaluated. Here, a large-scale experimental evolution was performed using 10 individually genetically barcoded replicate lines, evolved in parallel along with matched controls. Under ramping vancomycin selection, resistance emerged rapidly over approximately 250 generations, with all 10 replicate lines reaching an MIC of 16-32x that of the WT – the highest level of vancomycin resistance reported to date through *in vitro* evolution of *C. difficile* – suggesting high-level vancomycin resistance in *C. difficile* is possible, at least under permissive laboratory conditions.

To assess the mutations involved in vancomycin resistance, whole genome sequencing of evolved isolates, whole populations, and their respective matched controls was performed. Isolate sequencing showed striking genetic parallelism, as all evolved isolates displayed mutations in either *dacRS* or *vanT*, suggesting the presence of two major pathways to vancomycin resistance. This

clearly demonstrates the virtues of large-scale evolution, since observing the patterns of parallel evolution allowed previously uncharacterised genes to be implicated in resistance. Population sequencing across multiple time points enabled visualisation of population dynamics, highlighting highly contrasting dynamics between the *dacRS* and *vanT* pathways. *dacRS* mutations always fixed at P10, suggesting involvement of this pathway in first-step resistance; whereas *vanT* mutations only fixed later in the evolution, suggesting either a contribution to higher-level resistance, or a requirement of a prerequisite event to enable resistance. Overall, sequencing in this way was shown to be invaluable for identification of mechanisms beyond first-step resistance, alternative routes to resistance, and the combination of mutations required for higher-level resistance.

Together, large-scale parallel evolution, coupled with isolate and population sequencing, may provide a new gold-standard for experimental evolution in *C. difficile*. This robust method enables a non-candidate gene approach to study antibiotic resistance, since resistance pathways can be determined through observing genetic parallelism, allowing the identification of novel mechanisms not previously associated with resistance. Of course, this work really only scratches the surface regarding the possibilities for studying vancomycin resistance evolution – it may be worthwhile evolving even higher-level resistance to determine further mechanisms and synergistic interactions, or perhaps to perform a long-term *C. difficile* evolution similar to that of *E. coli* (Lenski, 2017), to investigate how evolutionary dynamics play out over extended time periods. Either way, further insights into resistance pathways of this important antibiotic will have useful applications in clinical monitoring. Needless to say, this approach has further applications beyond the study of vancomycin resistance (or indeed the study of *C. difficile*) – future work may therefore assess novel alternative routes to resistance for other commonly used antibiotics.

7.3 Fitness Costs of Vancomycin Resistance

The carriage of antibiotic resistance determinants can have variable effects on bacterial fitness outside of environments where antibiotics constitute a significant selection pressure. Although such effects are typically deleterious (Beceiro et al., 2013), and can have consequences for both pathogenicity and virulence (Cameron et al., 2015), resistance acquisition can also be neutral – fluoroquinolone resistance in ribotype 027 *C. difficile* was not associated with fitness costs *in vitro* (Wasels et al., 2015). Since previous studies of vancomycin resistance in *C. difficile* involved only genotypic characterisations, the phenotypic consequences of vancomycin resistance were assessed.

Growth and sporulation efficiency assays of the evolved isolates showed vancomycin resistance was associated with pleiotropic fitness burdens. The recapitulated strains R20291Δ*PaLoc dacS*c.714G>T, R20291Δ*PaLoc dacS*c.548T>C, and R20291Δ*PaLoc comR*c.1337G>T also showed defects in growth and/or sporulation efficiency, suggesting these mutations were largely responsible for the loss of fitness observed in their respective evolved isolates, directly implicating vancomycin resistance determinants (rather than additional deleterious mutations which do not contribute to resistance) in the observed fitness defects. Interestingly, principal component analysis showed no clustering by resistance pathway – isolates harbouring the same resistance mechanism displayed different fitness profiles. This implies that the resistance determinants are large, but not sole, contributors to the fitness of evolved isolates, suggesting that either mutation accumulation, or compensatory evolution, may also be at play. Indeed, evidence of compensatory evolution was observed when comparing the growth profiles of R20291Δ*PaLoc dacS*c.548T>C and evolved isolates Bc8 and Bc9. Although still significant, growth defects in evolved isolates were much less prominent than in R20291Δ*PaLoc dacS*c.548T>C, indicating the action of compensatory evolution in these isolates to reduce the fitness burden of *dacS* SNP carriage. It is largely accepted that amelioration of fitness costs through compensatory evolution is favoured over loss of resistance, even in antibiotic-free

environments (Schulz zur Wiesch et al., 2010), meaning further investigation of the mechanisms of compensatory evolution would merit interesting further work. Performing a further evolution of the evolved isolates, in an antibiotic-free (or constant concentration) environment would enable insights into how *C. difficile* may circumvent the observed fitness costs to achieve a high-resistance/high-fitness phenotype.

Nevertheless, the pleiotropic fitness defects observed across multiple vancomycin resistance pathways have important implications regarding the virulence of resistant strains. Sporulation defects were observed in four of the ten evolved isolates, with a complete absence of sporulation in Bc11. Since sporulation is essential for *C. difficile* pathogenicity, being an absolute requirement for onward transmission, severe sporulation defects arising in response to vancomycin treatment in patients would likely result in an evolutionary dead-end. Importantly, sporulation efficiencies comparable to the WT were observed for *dacS* recapitulated strains, suggesting pathways to vancomycin resistance can exist which do not result in significantly reduced virulence. Further insights into fitness, and the accompanying effects on virulence, may be gained through competition studies. Competing the evolved isolates with the WT *in vitro* may be the first stage of testing the impacts of growth and sporulation defects, however an *in vitro* gut model (Chilton et al., 2015), or *in vivo* competition in animal models, would be required to test the true clinical significance of the fitness defects observed here.

Despite being associated with varying fitness burdens, vancomycin resistant *C. difficile* isolates did display gains in terms of antibiotic cross-resistance, as all evolved isolates exhibited a 2- to 4-fold increase in teicoplanin resistance. It would be beneficial, therefore, to determine whether vancomycin resistance results in more expansive cross-resistance (or indeed collateral sensitivity) to

other antibiotics. Building such networks in *C. difficile* may become invaluable for advising treatments for complex or severe CDI (Lázár et al., 2014).

7.4 *dacJRS*–mediated Vancomycin Resistance

Perhaps the most important finding of this thesis was the identification of a novel mechanism of vancomycin resistance in *C. difficile*. Previously, *dacJRS* was an uncharacterised gene cluster, which had never before been associated with vancomycin resistance. However, *dacJRS* was subject to remarkable parallel evolution, with mutations occurring in six of the ten evolved isolates. Further, *dacS* was one of only a few genes possessing vancomycin-unique mutations in evolved isolates Bc1 and Bc2, suggesting a prominent involvement of this gene in the resistance displayed by these isolates. Recapitulation of the *dacS* SNPs observed in Bc1 and Bc8/9 showed *dacS* alone was capable of invoking a 4-fold increase in vancomycin resistance. Taken with the early fixation of *dacS* mutations observed through population sequencing, it is likely *dacJRS* constitutes a mechanism of first-step resistance.

Mechanistic insights into this pathway through qRT-PCR showed mutations in *dacS* led to increased activity of the DacRS TCS, a positive regulator of *dacJRS*, resulting in increased expression of the *dacJRS* cluster. Increased expression of DacJ, a putative D,D-carboxypeptidase, was hypothesised to elicit resistance through removal of the terminal D-Ala in nascent peptidoglycan, the binding site of vancomycin (Figure 7.1). Vancomycin-BODIPY microscopy showed a complete loss of vancomycin binding in *dacS* recapitulated strains, confirming the role of *dacJRS* in depletion of D-Ala D-Ala vancomycin binding sites. The direct involvement of DacJ was confirmed through the deletion of *dacJ* in the Bc1 evolved isolate, which resulted in an 8-fold reduction in resistance and partial restoration of vancomycin binding.

Implicating *dacJRS* in vancomycin resistance indicates this cluster could help guide genomic surveillance efforts. However, even across the ten evolved isolates, multiple mutations were observed which culminated in identical outcomes, suggesting any number of further mutations in these genes may also result in resistance. Differentiating random mutations from resistance-determining mutations in this cluster would therefore pose a challenge to genomic surveillance. One method to bridge this gap, and to move this cluster towards becoming a biomarker of resistance, would be to adopt lessons from rational design approaches to protein engineering, to more thoroughly understand the functional relevance of genetic changes (Korendovych, 2018). However, for now, an increased focus on this cluster during analysis of clinical isolates would suffice.

7.5 Synergistic Interactions in Vancomycin Resistance

Although the *dacS*c.714G>T mutation was found to be a major contributor vancomycin resistance in the Bc1 evolved isolate, providing a 4-fold increase in MIC, the level of resistance exhibited by Bc1 could not be fully recapitulated with this mutation alone. Multiple lines of evidence pointed toward the presence of a second mechanism of resistance – the MIC of Bc1Δ*dacJ* was 2 µg/mL, still 2-fold higher than the parental control; and only partial restoration of vancomycin binding was observed in the Bc1Δ*dacJ* strain, suggesting a reduction in vancomycin binding affinity in Bc1 beyond the actions of *dacJRS*. Through hybrid assembly of the Bc1 genome, a further two mutations, *vanS*c.367_396dup and 1,197,357_1,197,400del, were identified. Although failure to detect InDels through short-read sequencing is not a new observation (Juraschek et al., 2021), this has important implications for genomic surveillance, which routinely relies on Illumina sequencing (Miles-Jay et al., 2023), as it

suggests current surveillance efforts may not be capturing all clinically important variants. As hybrid assembly becomes less costly, a shift to this approach may prove beneficial.

Although the individual contribution of *vanSc.*367_396dup to vancomycin resistance was modest, the combination of *dacSc.*714G>T and *vanSc.*367_396dup mutations allowed full recapitulation of the level of resistance observed in Bc1, illustrating one way by which accumulation of mutations can result in high-level resistance (Figure 7.1). As well as being the first report of synergistic interactions promoting vancomycin resistance in *C. difficile*, this observation highlights that only two mutations are required for high-level resistance. Previous works have demonstrated that even minor increases in resistance can have negative effects on patient cure rates (Eubank et al., 2024), suggesting the resistance observed here is of considerable clinical significance. Additionally, as mutations in *vanS* are often reported in clinical isolates with reduced vancomycin susceptibility (Kolte and Nübel, 2024), the alarming notion that only a single SNP in *dacS* may be required for dramatically increased levels of resistance suggests an even greater need for routine clinical monitoring.

7.6 *vanT* and *comR*-mediated Vancomycin Resistance

Although *vanT* mutations were identified from the experimental evolution as the other major pathway to vancomycin resistance, displaying parallel evolution in four of the ten evolved isolates (and co-occurrence with *comR* in two), investigations into this mechanism are largely incomplete. Despite extensive recapitulation of all four *vanT* mutations observed in evolved isolates, the individual contribution of *vanT* to vancomycin resistance could not be discerned. Taken with the consistently late fixation of *vanT* SNPs observed though population sequencing, it is likely that these SNPs act as refining mutations – providing minor enhancements to the *van*-mediated resistance

mechanism without providing a significant contribution to resistance. Further characterisation of *vanT*, for example through the construction of Bc3Δ*vanT,* or through functional comparison of *vanT* mutants in racemase activity assays (Noda et al., 2005), may provide greater insights into the roles of *vanT* mutations in resistance. Conversely, the recapitulation of *comR* mutations, which co-occurred with *vanT* mutations in Bc3 and Bc4, provided a 2- to 4-fold increase in MIC, suggesting an alternative, albeit costly, route to low-level resistance. Beyond observing no effect on expression of *van* and *dacJRS* genes, the mechanism of resistance associated with *comR* was not determined. The complete *comR* regulon, deduced by RNAseq, would therefore be helpful to make headway in understanding this pathway.

Intriguingly, the level of vancomycin resistance observed in Bc3 and Bc4 could not be fully recapitulated, despite no further mutations being identified in these isolates via hybrid assembly. The possibility of epigenetic regulation of vancomycin resistance was therefore touched upon in this thesis, however, although possible (Li et al., 2020), the complete epigenome, and its association with vancomycin resistance, presents a large and challenging gap in the *C. difficile* antimicrobial resistance sphere.

7.7 Concluding Remarks

In summary, this thesis comprehensively examines vancomycin resistance in *C. difficile*, identifying genetic loci worthy of attention in future clinical monitoring, as well as providing a solid foundation for future work. Experimental evolution was shown here to be an invaluable tool for studying antimicrobial resistance, demonstrating high-level resistance can evolve rapidly in *C. difficile* given the right environmental conditions. Evolving at scale, using multiple replicate populations, allowed the identification of multiple alternative routes to vancomycin resistance which had not been

previously identified. Phenotypic assessment of *dacS*, *vanS* and *comR*-mediated resistance highlighted the costly nature of resistance acquisition, and molecular characterisation allowed the mechanisms associated with these pathways to be determined. The first evidence of synergistic interactions promoting vancomycin resistance was also reported here, demonstrating that just two mutations are required for high-level resistance in *C. difficile*. This work has undoubtedly come at the right time, with vancomycin being promoted to the recommended front-line drug for CDI during the course of this project (NICE, 2021). As genomic data for vancomycin resistant clinical isolates becomes more abundant, it may be possible to understand how resistance evolves in *C. difficile in vivo* through genome-wide association studies. For now, this project provides sufficient groundwork to guide better genomic surveillance.

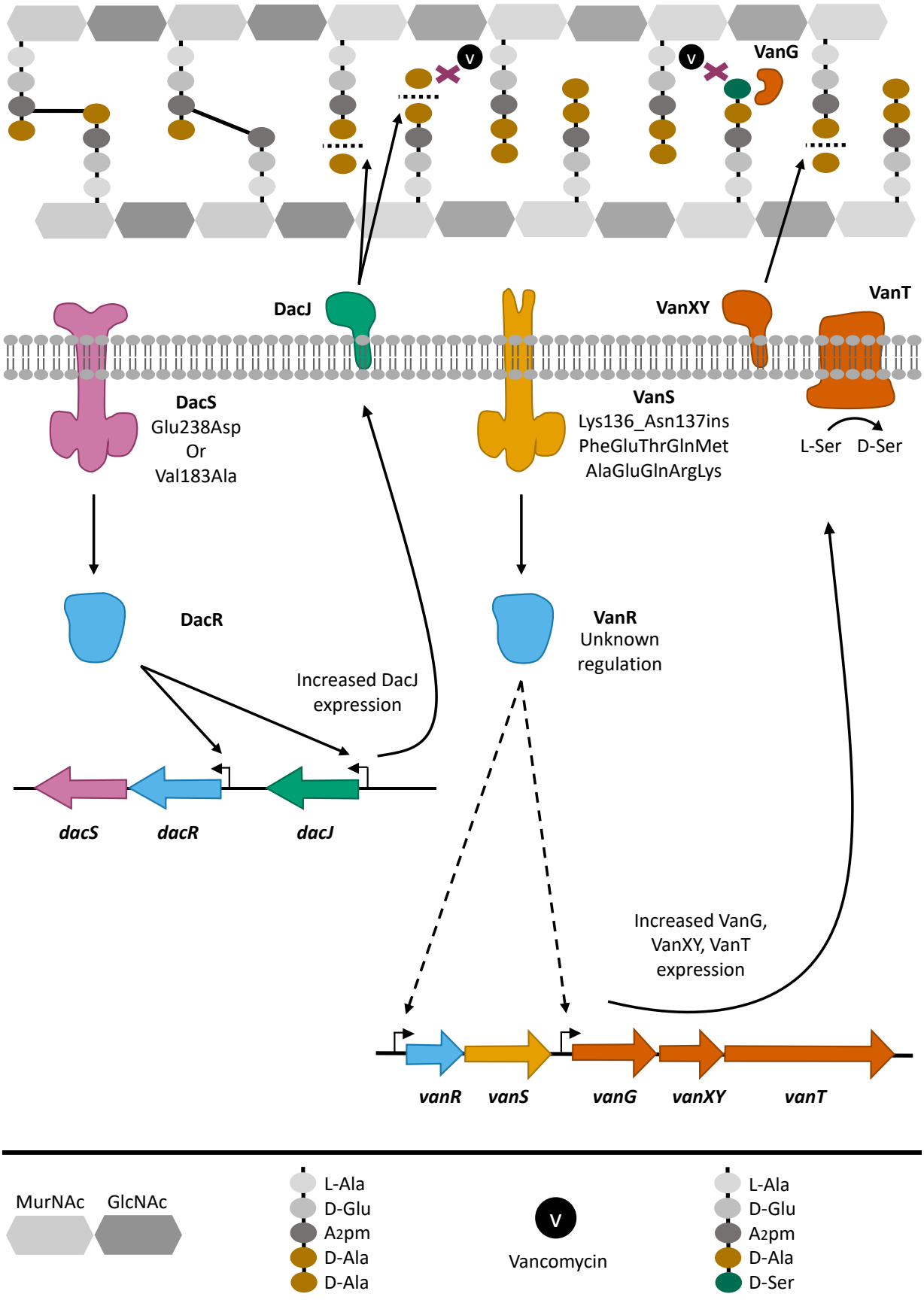


Figure 7.1 - Mechanisms of Vancomycin Resistance in C. difficile

dacS and *vanS* mediated vancomycin resistance in *C. difficile*. The mutations *dacS*c.714G>T (Bc1) and *dacS*c.548T>C (Bc8/9) increase the activity of the DacRS TCS, a positive regulator of the *dacJRS* gene cluster, leading to increased expression of *dacJRS*. Increased expression of DacJ, a putative D,D-carboxypeptidase, results in vancomycin resistance through depletion of D-Ala D-Ala vancomycin binding sites. The 30 nucleotide insertion in *vanS* (*vanS*c.367_396dup) increases expression of *vanG, vanXY* and *vanT*. The mechanism of vancomycin resistance associated with the *van* genes has been previously characterised – expression of the *van* cluster results in reduced vancomycin binding through replacement of the terminal D-Ala with D-Ser in nascent peptidoglycan (Reynolds and Courvalin, 2005; Shen et al., 2020).

Bibliography

Adams, C.M., Eckenroth, B.E., Putnam, E.E., Doublié, S., Shen, A., 2013. Structural and Functional Analysis of the CspB Protease Required for *Clostridium* Spore Germination. PLOS Pathogens 9, e1003165. https://doi.org/10.1371/journal.ppat.1003165

Adams, D.C., Collyer, M.L., 2009. A General Framework For The Analysis Of Phenotypic Trajectories In Evolutionary Studies. Evolution 63, 1143–1154. https://doi.org/10.1111/j.1558-5646.2009.00649.x

Ahmed, M.O., Baptiste, K.E., 2018. Vancomycin-Resistant *Enterococci*: A Review of Antimicrobial Resistance Mechanisms and Perspectives of Human and Animal Health. Microbial Drug Resistance 24, 590–606. https://doi.org/10.1089/mdr.2017.0147

Akhmanova, A., Steinmetz, M.O., 2015. Control of microtubule organization and dynamics: two ends in the limelight. Nat Rev Mol Cell Biol 16, 711–726. https://doi.org/10.1038/nrm4084

Aktories, K., Lang, A.E., Schwan, C., Mannherz, H.G., 2011. Actin as target for modification by bacterial protein toxins. The FEBS Journal 278, 4526–4543. https://doi.org/10.1111/j.1742-4658.2011.08113.x

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ammam, F., Marvaud, J.-C., Lambert, T., 2012. Distribution of the *vanG*-like gene cluster in *Clostridium difficile* clinical isolates. Can. J. Microbiol. 58, 547–551. https://doi.org/10.1139/w2012-002

Ammam, F., Meziane-cherif, D., Mengin-Lecreulx, D., Blanot, D., Patin, D., Boneca, I.G., Courvalin, P., Lambert, T., Candela, T., 2013. The functional *vanGCd* cluster of *Clostridium difficile* does not confer vancomycin resistance. Molecular Microbiology 89, 612–625. https://doi.org/10.1111/mmi.12299

Ammam, F., Patin, D., Coullon, H., Blanot, D., Lambert, T., Mengin-Lecreulx, D., Candela, T., 2020. AsnB is responsible for peptidoglycan precursor amidation in *Clostridium difficile* in the presence of vancomycin. Microbiology 166, 567–578. https://doi.org/10.1099/mic.0.000917

Anderson, D.M., Sheedlo, M.J., Jensen, J.L., Lacy, D.B., 2020. Structural insights into the transition of *Clostridioides difficile* binary toxin from prepore to pore. Nat Microbiol 5, 102–107. https://doi.org/10.1038/s41564-019-0601-8

Andrews, Simon, 2019. FastQC - A quality control tool for high throughput sequence data.

Anwar, F., Roxas, B.A.P., Shehab, K.W., Ampel, N.M., Viswanathan, V.K., Vedantam, G., 2022. Low-toxin *Clostridioides difficile* RT027 strains exhibit robust virulence. Emerg Microbes Infect 11, 1982–1993. https://doi.org/10.1080/22221751.2022.2105260

Artsimovitch, I., Seddon, J., Sears, P., 2012. Fidaxomicin Is an Inhibitor of the Initiation of Bacterial RNA Synthesis. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 55, S127. https://doi.org/10.1093/cid/cis358

Ashton, P.M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., Wain, J., O'Grady, J., 2015. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. Nat Biotechnol 33, 296–300. https://doi.org/10.1038/nbt.3103

Bagel, S., Hüllen, V., Wiedemann, B., Heisig, P., 1999. Impact of *gyrA* and *parC* Mutations on Quinolone Resistance, Doubling Time, and Supercoiling Degree of *Escherichia coli*. Antimicrob Agents Chemother 43, 868–875.

Bailey, S.F., Alonso Morales, L.A., Kassen, R., 2021. Effects of Synonymous Mutations beyond Codon Bias: The Evidence for Adaptive Synonymous Substitutions from Microbial Evolution Experiments. Genome Biol Evol 13, evab141. https://doi.org/10.1093/gbe/evab141

Banawas, S.S., 2018. *Clostridium difficile* Infections: A Global Overview of Drug Sensitivity and Resistance Mechanisms. BioMed Research International 2018, 8414257. https://doi.org/10.1155/2018/8414257

Band, V.I., Weiss, D.S., 2019. Heteroresistance: A cause of unexplained antibiotic treatment failure? PLoS Pathog 15, e1007726. https://doi.org/10.1371/journal.ppat.1007726

Barth, H., Pfeifer, G., Hofmann, F., Maier, E., Benz, R., Aktories, K., 2001. Low pH-induced Formation of Ion Channels by *Clostridium difficile* Toxin B in Target Cells *. Journal of Biological Chemistry 276, 10670–10676. https://doi.org/10.1074/jbc.M009445200

Basson, A.R., Zhou, Y., Seo, B., Rodriguez-Palacios, A., Cominelli, F., 2020. Autologous fecal microbiota transplantation for the treatment of inflammatory bowel disease. Translational Research 226, 1–11. https://doi.org/10.1016/j.trsl.2020.05.008

Beceiro, A., Tomás, M., Bou, G., 2013. Antimicrobial Resistance and Virulence: a Successful or Deleterious Association in the Bacterial World? Clin Microbiol Rev 26, 185–230. https://doi.org/10.1128/CMR.00059-12

Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J.I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F.A., Kreshuk, A., 2019. ilastik: interactive machine learning for (bio)image analysis. Nat Methods 16, 1226–1232. https://doi.org/10.1038/s41592-019-0582-9

Bernard, E., Rolain, T., Courtin, P., Hols, P., Chapot-Chartier, M.-P., 2011. Identification of the Amidotransferase AsnB1 as Being Responsible for meso-Diaminopimelic Acid Amidation in *Lactobacillus plantarum* Peptidoglycan ▿. J Bacteriol 193, 6323–6330. https://doi.org/10.1128/JB.05060-11

Bernstein, J.A., Lin, P.-H., Cohen, S.N., Lin-Chao, S., 2004. Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. Proc Natl Acad Sci U S A 101, 2758–2763. https://doi.org/10.1073/pnas.0308747101

Bilverstone, T.W., Garland, M., Cave, R.J., Kelly, M.L., Tholen, M., Bouley, D.M., Kaye, P., Minton, N.P., Bogyo, M., Kuehne, S.A., Melnyk, R.A., 2020. The glucosyltransferase activity of *C. difficile* Toxin B is required for disease pathogenesis. PLOS Pathogens 16, e1008852. https://doi.org/10.1371/journal.ppat.1008852

Boekhoud, I.M., Hornung, B.V.H., Sevilla, E., Harmanus, C., Bos-Sanders, I.M.J.G., Terveer, E.M., Bolea, R., Corver, J., Kuijper, E.J., Smits, W.K., 2020. Plasmid-mediated metronidazole resistance in *Clostridioides difficile*. Nat Commun 11, 598. https://doi.org/10.1038/s41467-020-14382-1

Boekhoud, I.M., Sidorov, I., Nooij, S., Harmanus, C., Bos-Sanders, I.M.J.G., Viprey, V., Spittal, W., Clark, E., Davies, K., Freeman, J., Kuijper, E.J., Smits, W.K., the COMBACTE-CDI Consortium, 2021. Haem is crucial for medium-dependent metronidazole resistance in clinical isolates of *Clostridioides difficile*. Journal of Antimicrobial Chemotherapy 76, 1731–1740. https://doi.org/10.1093/jac/dkab097

Bourrel, A.S., Poirel, L., Royer, G., Darty, M., Vuillemin, X., Kieffer, N., Clermont, O., Denamur, E., Nordmann, P., Decousser, J.-W., IAME Resistance Group, 2019. Colistin resistance in Parisian inpatient faecal *Escherichia coli* as the result of two distinct evolutionary pathways. Journal of Antimicrobial Chemotherapy 74, 1521–1530. https://doi.org/10.1093/jac/dkz090

Buddle, J.E., Fagan, R.P., 2023. Pathogenicity and virulence of *Clostridioides difficile*. Virulence 14, 2150452. https://doi.org/10.1080/21505594.2022.2150452

Buddle, J.E., Wright, R.C.T., Turner, C.E., Chaudhuri, R.R., Brockhurst, M.A., Fagan, R.P., 2023. Multiple evolutionary pathways lead to vancomycin resistance in *Clostridioides difficile*. https://doi.org/10.1101/2023.09.15.557922

Buffie, C.G., Bucci, V., Stein, R.R., McKenney, P.T., Ling, L., Gobourne, A., No, D., Liu, H., Kinnebrew, M., Viale, A., Littmann, E., van den Brink, M.R.M., Jenq, R.R., Taur, Y., Sander, C., Cross, J.R., Toussaint, N.C., Xavier, J.B., Pamer, E.G., 2015. Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. Nature 517, 205–208. https://doi.org/10.1038/nature13828

Buffie, C.G., Jarchum, I., Equinda, M., Lipuma, L., Gobourne, A., Viale, A., Ubeda, C., Xavier, J., Pamer, E.G., 2012. Profound Alterations of Intestinal Microbiota following a Single Dose of Clindamycin Results in Sustained Susceptibility to *Clostridium difficile*-Induced Colitis. Infection and Immunity 80, 62–73. https://doi.org/10.1128/iai.05496-11

Burke, K.E., Lamont, J.T., 2014. *Clostridium difficile* Infection: A Worldwide Disease 8, 1–6. https://doi.org/10.5009/gnl.2014.8.1.1

Cadzow, M., Boocock, J., Nguyen, H.T., Wilcox, P., Merriman, T.R., Black, M.A., 2014. A bioinformatics workflow for detecting signatures of selection in genomic data. Front. Genet. 5. https://doi.org/10.3389/fgene.2014.00293

Cameron, D.R., Mortin, L.I., Rubio, A., Mylonakis, E., Moellering Jr, R.C., Eliopoulos, G.M., Peleg, A.Y., 2015. Impact of daptomycin resistance on *Staphylococcus aureus* virulence. Virulence 6, 127–131. https://doi.org/10.1080/21505594.2015.1011532

Cameron, T.A., Matz, L.M., Lay, N.R.D., 2018. Polynucleotide phosphorylase: Not merely an RNase but a pivotal post-transcriptional regulator. PLOS Genetics 14, e1007654. https://doi.org/10.1371/journal.pgen.1007654

Cao, X., Boyaci, H., Chen, J., Bao, Y., Landick, R., Campbell, E.A., 2022. Basis of narrow-spectrum activity of fidaxomicin on *Clostridioides difficile*. Nature 604, 541–545. https://doi.org/10.1038/s41586-022-04545-z

Carter, G.P., Lyras, D., Allen, D.L., Mackin, K.E., Howarth, P.M., O'Connor, J.R., Rood, J.I., 2007. Binary Toxin Production in *Clostridium difficile* Is Regulated by CdtR, a LytTR Family Response Regulator. Journal of Bacteriology 189, 7290–7301. https://doi.org/10.1128/jb.00731-07

Cartman, S.T., Kelly, M.L., Heeg, D., Heap, J.T., Minton, N.P., 2012. Precise Manipulation of the *Clostridium difficile* Chromosome Reveals a Lack of Association between the *tcdC* Genotype and Toxin Production. Appl Environ Microbiol 78, 4683–4690. https://doi.org/10.1128/AEM.00249-12

Casadesús, J., Low, D., 2006. Epigenetic Gene Regulation in the Bacterial World. Microbiology and Molecular Biology Reviews 70, 830–856. https://doi.org/10.1128/mmbr.00016-06

CDC, 2024. 2019 Antibiotic Resistance Threats Report [WWW Document]. Antimicrobial Resistance. URL https://www.cdc.gov/antimicrobial-resistance/data-research/threats/index.html (accessed 6.6.24).

Chen, L., Duan, K., 2016. A PhoPQ-Regulated ABC Transporter System Exports Tetracycline in *Pseudomonas aeruginosa*. Antimicrob Agents Chemother 60, 3016–3024. https://doi.org/10.1128/AAC.02986-15

Cherny, S.S., Nevo, D., Baraz, A., Baruch, S., Lewin-Epstein, O., Stein, G.Y., Obolski, U., 2021. Revealing antibiotic cross-resistance patterns in hospitalized patients through Bayesian network modelling. Journal of Antimicrobial Chemotherapy 76, 239–248. https://doi.org/10.1093/jac/dkaa408

Chilton, C.H., Crowther, G.S., Ashwin, H., Longshaw, C.M., Wilcox, M.H., 2016. Association of Fidaxomicin with *C. difficile* Spores: Effects of Persistence on Subsequent Spore Recovery, Outgrowth and Toxin Production. PLOS ONE 11, e0161200. https://doi.org/10.1371/journal.pone.0161200

Chilton, C.H., Crowther, G.S., Todhunter, S.L., Ashwin, H., Longshaw, C.M., Karas, A., Wilcox, M.H., 2015. Efficacy of alternative fidaxomicin dosing regimens for treatment of simulated *Clostridium difficile* infection in an in vitro human gut model. J Antimicrob Chemother 70, 2598–2607. https://doi.org/10.1093/jac/dkv156

Cho, K.H., 2017. The Structure and Function of the Gram-Positive Bacterial RNA Degradosome. Front Microbiol 8, 154. https://doi.org/10.3389/fmicb.2017.00154

Chodkowski, J.L., Shade, A., 2023. A coevolution experiment between *Flavobacterium johnsoniae* and *Burkholderia thailandensis* reveals parallel mutations that reduce antibiotic susceptibility. Microbiology (Reading) 169, 001267. https://doi.org/10.1099/mic.0.001267

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin) 6, 80–92. https://doi.org/10.4161/fly.19695

CLSI, 2024. CLSI & Antimicrobial Susceptibility Testing (AST) [WWW Document]. Clinical & Laboratory Standards Institute. URL https://clsi.org/meetings/susceptibility-testing-subcommittees/clsi-and-ast/ (accessed 5.29.24).

Cole, S.A., Stahl, T.J., 2015. Persistent and Recurrent *Clostridium difficile* Colitis. Clin Colon Rectal Surg 28, 65–69. https://doi.org/10.1055/s-0035-1547333

Collins, J., Auchtung, J.M., Schaefer, L., Eaton, K.A., Britton, R.A., 2015. Humanized microbiota mice as a model of recurrent *Clostridium difficile* disease. Microbiome 3, 35. https://doi.org/10.1186/s40168-015-0097-2

Cong, Y., Yang, S., Rao, X., 2019. Vancomycin resistant *Staphylococcus aureus* infections: A review of case updating and clinical features. J Adv Res 21, 169–176. https://doi.org/10.1016/j.jare.2019.10.005

Coullon, H., Candela, T., 2022. *Clostridioides difficile* peptidoglycan modifications. Current Opinion in Microbiology 65, 156–161. https://doi.org/10.1016/j.mib.2021.11.010

Coullon, H., Rifflet, A., Wheeler, R., Janoir, C., Boneca, I.G., Candela, T., 2020. Peptidoglycan analysis reveals that synergistic deacetylase activity in vegetative *Clostridium difficile* impacts the host response. Journal of Biological Chemistry 295, 16785–16796. https://doi.org/10.1074/jbc.RA119.012442

Cruz, M.P., 2012. Fidaxomicin (Dificid), a Novel Oral Macrocyclic Antibacterial Agent For the Treatment of *Clostridium difficile*–Associated Diarrhea in Adults. P T 37, 278–281.

Cui, L., Murakami, H., Kuwahara-Arai, K., Hanaki, H., Hiramatsu, K., 2000. Contribution of a Thickened Cell Wall and Its Glutamine Nonamidated Component to the Vancomycin Resistance Expressed by *Staphylococcus aureus* Mu50. Antimicrob Agents Chemother 44, 2276–2285.

Curry, S.R., Marsh, J.W., Muto, C.A., O'Leary, M.M., Pasculle, A.W., Harrison, L.H., 2007. *tcdC* Genotypes Associated with Severe TcdC Truncation in an Epidemic Clone and Other Strains of *Clostridium difficile*. Journal of Clinical Microbiology 45, 215–221. https://doi.org/10.1128/jcm.01599-06

Czepiel, J., Dróżdż, M., Pituch, H., Kuijper, E.J., Perucki, W., Mielimonka, A., Goldman, S., Wultańska, D., Garlicki, A., Biesiada, G., 2019. *Clostridium difficile* infection: review. Eur J Clin Microbiol Infect Dis 38, 1211–1221. https://doi.org/10.1007/s10096-019-03539-6

D'Aquila, P., De Rango, F., Paparazzo, E., Passarino, G., Bellizzi, D., 2023. Epigenetic-Based Regulation of Transcriptome in *Escherichia coli* Adaptive Antibiotic Resistance. Microbiol Spectr 11, e04583-22. https://doi.org/10.1128/spectrum.04583-22

Darkoh, C., Keita, K., Odo, C., Oyaro, M., Brown, E.L., Arias, C.A., Hanson, B.M., DuPont, H.L., 2021. Emergence of Clinical *Clostridioides difficile* Isolates With Decreased Susceptibility to Vancomycin. Clin Infect Dis 74, 120–126. https://doi.org/10.1093/cid/ciaa912

De Lay, N., Gottesman, S., 2011. Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. RNA 17, 1172–1189. https://doi.org/10.1261/rna.2531211

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, 2704–2711. https://doi.org/10.1128/IAI.00147-12

Deatherage, D.E., Barrick, J.E., 2014. Identification of mutations in laboratory evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol 1151, 165–188. https://doi.org/10.1007/978-1-4939-0554-6_12

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, 10.1128/mbio.02383-14. https://doi.org/10.1128/mbio.02383-14

Deshpande, A., Wu, X., Huo, W., Palmer, K.L., Hurdle, J.G., 2020. Chromosomal Resistance to Metronidazole in *Clostridioides difficile* Can Be Mediated by Epistasis between Iron Homeostasis and Oxidoreductases. Antimicrob Agents Chemother 64, e00415-20. https://doi.org/10.1128/AAC.00415-20

Djordjevic, S.P., Jarocki, V.M., Seemann, T., Cummins, M.L., Watt, A.E., Drigo, B., Wyrsch, E.R., Reid, C.J., Donner, E., Howden, B.P., 2024. Genomic surveillance for antimicrobial resistance — a One Health perspective. Nat Rev Genet 25, 142–157. https://doi.org/10.1038/s41576-023-00649-y

Dong, D., Chen, X., Jiang, C., Zhang, L., Cai, G., Han, L., Wang, X., Mao, E., Peng, Y., 2014. Genetic analysis of Tn*916*-like elements conferring tetracycline resistance in clinical isolates of *Clostridium difficile*. Int J Antimicrob Agents 43, 73–77. https://doi.org/10.1016/j.ijantimicag.2013.09.004

Dorst, A., Berg, R., Gertzen, C.G.W., Schäfle, D., Zerbe, K., Gwerder, M., Schnell, S.D., Sander, P., Gohlke, H., Gademann, K., 2020. Semisynthetic Analogs of the Antibiotic Fidaxomicin—Design, Synthesis, and Biological Evaluation. ACS Med. Chem. Lett. 11, 2414–2420. https://doi.org/10.1021/acsmedchemlett.0c00381

Drawz, S.M., Bonomo, R.A., 2010. Three Decades of β-Lactamase Inhibitors. Clin Microbiol Rev 23, 160−201. https://doi.org/10.1128/CMR.00037-09

Dridi, L., Tankovic, J., Burghoffer, B., Barbut, F., Petit, J.-C., 2002. *gyrA* and *gyrB* Mutations Are Implicated in Cross-Resistance to Ciprofloxacin and Moxifloxacin in *Clostridium difficile*. Antimicrob Agents Chemother 46, 3418–3421. https://doi.org/10.1128/AAC.46.11.3418-3421.2002

Dsouza, M., Menon, R., Crossette, E., Bhattarai, S.K., Schneider, J., Kim, Y.-G., Reddy, S., Caballero, S., Felix, C., Cornacchione, L., Hendrickson, J., Watson, A.R., Minot, S.S., Greenfield, N., Schopf, L., Szabady, R., Patarroyo, J., Smith, W., Harrison, P., Kuijper, E.J., Kelly, C.P., Olle, B., Bobilev, D., Silber, J.L., Bucci, V., Roberts, B., Faith, J., Norman, J.M., 2022. Colonization of the live biotherapeutic product VE303 and modulation of the microbiota and metabolites in healthy volunteers. Cell Host & Microbe 30, 583-598.e8. https://doi.org/10.1016/j.chom.2022.03.016

Ducarmon, Q.R., Zwittink, R.D., Hornung, B.V.H., van Schaik, W., Young, V.B., Kuijper, E.J., 2019. Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. Microbiology and Molecular Biology Reviews 83, 10.1128/mmbr.00007-19. https://doi.org/10.1128/mmbr.00007-19

Ducret, A., Quardokus, E.M., Brun, Y.V., 2016. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. Nat Microbiol 1, 1–7. https://doi.org/10.1038/nmicrobiol.2016.77

Dyrkell, F., Giske, C.G., Fang, H., 2021. Epidemiological typing of ST80 vancomycin-resistant *Enterococcus faecium*: core genome multilocus sequence typing versus single nucleotide polymorphism-based typing. Journal of Global Antimicrobial Resistance 25, 119–123. https://doi.org/10.1016/j.jgar.2021.03.005

Dzyubak, E., Yap, M.-N.F., 2016. The Expression of Antibiotic Resistance Methyltransferase Correlates with mRNA Stability Independently of Ribosome Stalling. Antimicrobial Agents and Chemotherapy 60, 7178–7188. https://doi.org/10.1128/aac.01806-16

ECDPC, 2023. Facts on *Clostridioides difficile* infections. URL https://www.ecdc.europa.eu/en/clostridium-difficile-infections/facts

Edwards, A.N., McBride, S.M., 2023. The RgaS-RgaR two-component system promotes *Clostridioides difficile* sporulation through a small RNA and the Agr1 system. bioRxiv 2023.06.26.546640. https://doi.org/10.1101/2023.06.26.546640

Edwards, A.N., McBride, S.M., 2014. Initiation of sporulation in *Clostridium difficile*: a twist on the classic model. FEMS Microbiology Letters 358, 110–118. https://doi.org/10.1111/1574-6968.12499

Egan, A.J.F., Biboy, J., van't Veer, I., Breukink, E., Vollmer, W., 2015. Activities and regulation of peptidoglycan synthases. Philosophical Transactions of the Royal Society B: Biological Sciences 370, 20150031. https://doi.org/10.1098/rstb.2015.0031

Egan, A.J.F., Errington, J., Vollmer, W., 2020. Regulation of peptidoglycan synthesis and remodelling. Nat Rev Microbiol 18, 446–460. https://doi.org/10.1038/s41579-020-0366-3

Egerer, M., Giesemann, T., Herrmann, C., Aktories, K., 2009. Autocatalytic Processing of *Clostridium difficile* Toxin B: Binding Of Inositol Hexakisphosphate*. Journal of Biological Chemistry 284, 3389–3395. https://doi.org/10.1074/jbc.M806002200

Enany, S., 2017. *Clostridium difficile*: A Comprehensive Overview. BoD – Books on Demand.

Ernst, K., Eberhardt, N., Mittler, A.-K., Sonnabend, M., Anastasia, A., Freisinger, S., Schiene-Fischer, C., Malešević, M., Barth, H., 2018. Pharmacological Cyclophilin Inhibitors Prevent Intoxication of Mammalian Cells with *Bordetella pertussis* Toxin. Toxins (Basel) 10, 181. https://doi.org/10.3390/toxins10050181

Eubank, T.A., Dureja, C., Garey, K.W., Hurdle, J.G., Gonzales-Luna, A.J., 2024. Reduced Vancomycin Susceptibility in *Clostridioides difficile* Is Associated With Lower Rates of Initial Cure and Sustained Clinical Response. Clinical Infectious Diseases ciae087. https://doi.org/10.1093/cid/ciae087

Evers, S., Courvalin, P., 1996. Regulation of VanB-type vancomycin resistance gene expression by the VanS(B)-VanR (B) two-component regulatory system in *Enterococcus faecalis* V583. J Bacteriol 178, 1302–1309.

Fagan, R.P., Fairweather, N.F., 2011. *Clostridium difficile* Has Two Parallel and Essential Sec Secretion Systems *. Journal of Biological Chemistry 286, 27483–27493. https://doi.org/10.1074/jbc.M111.263889

Fan, Z., Chen, H., Li, M., Pan, X., Fu, W., Ren, H., Chen, R., Bai, F., Jin, Y., Cheng, Z., Jin, S., Wu, W., 2019. *Pseudomonas aeruginosa* Polynucleotide Phosphorylase Contributes to Ciprofloxacin Resistance by Regulating PrtR. Front Microbiol 10, 1762. https://doi.org/10.3389/fmicb.2019.01762

Farooq, P.D., Urrunaga, N.H., Tang, D.M., von Rosenvinge, E.C., 2015. Pseudomembranous Colitis. Dis Mon 61, 181–206. https://doi.org/10.1016/j.disamonth.2015.01.006

Farrow, K.A., Lyras, D., Rood, J.I., 2000. The Macrolide-Lincosamide-Streptogramin B Resistance Determinant from *Clostridium difficile* 630 Contains Two *erm(B)* Genes. Antimicrob Agents Chemother 44, 411–413.

Foucault, M.-L., Depardieu, F., Courvalin, P., Grillot-Courvalin, C., 2010. Inducible expression eliminates the fitness cost of vancomycin resistance in *enterococci*. Proc Natl Acad Sci U S A 107, 16964–16969. https://doi.org/10.1073/pnas.1006855107

Fuchs, M., Lamm-Schmidt, V., Sulzer, J., Ponath, F., Jenniches, L., Kirk, J.A., Fagan, R.P., Barquist, L., Vogel, J., Faber, F., 2021. An RNA-centric global view of *Clostridioides difficile* reveals broad activity of Hfq in a clinically important gram-positive bacterium. Proceedings of the National Academy of Sciences 118, e2103579118. https://doi.org/10.1073/pnas.2103579118

Galinier, A., Delan-Forino, C., Foulquier, E., Lakhal, H., Pompeo, F., 2023. Recent Advances in Peptidoglycan Synthesis and Regulation in Bacteria. Biomolecules 13, 720. https://doi.org/10.3390/biom13050720

Gallagher, J.C., Reilly, J.P., Navalkele, B., Downham, G., Haynes, K., Trivedi, M., 2015. Clinical and Economic Benefits of Fidaxomicin Compared to Vancomycin for *Clostridium difficile* Infection. Antimicrobial Agents and Chemotherapy 59, 7007–7010. https://doi.org/10.1128/aac.00939-15

Gallagher, L.A., Coughlan, S., Black, N.S., Lalor, P., Waters, E.M., Wee, B., Watson, M., Downing, T., Fitzgerald, J.R., Fleming, G.T.A., O'Gara, J.P., 2017. Tandem Amplification of the *Staphylococcal* Cassette Chromosome *mec* Element Can Drive High-Level Methicillin Resistance in Methicillin-Resistant *Staphylococcus aureus*. Antimicrobial Agents and Chemotherapy 61, 10.1128/aac.00869-17. https://doi.org/10.1128/aac.00869-17

Galley, N.F., Greetham, D., Alamán-Zárate, M.G., Williamson, M.P., Evans, C.A., Spittal, W.D., Buddle, J.E., Freeman, J., Davis, G.L., Dickman, M.J., Wilcox, M.H., Lovering, A.L., Fagan, R.P., Mesnage, S., 2024. *Clostridioides difficile* canonical L,D-transpeptidases catalyze a novel type of peptidoglycan cross-links and are not required for beta-lactam resistance. Journal of Biological Chemistry 300. https://doi.org/10.1016/j.jbc.2023.105529

Gerding, D.N., Johnson, S., Rupnik, M., Aktories, K., 2014. *Clostridium difficile* binary toxin CDT. Gut Microbes 5, 15–27. https://doi.org/10.4161/gmic.26854

Gerhard, R., Nottrott, S., Schoentaube, J., Tatge, H., Olling, A., Just, I., 2008. Glucosylation of Rho GTPases by *Clostridium difficile* toxin A triggers apoptosis in intestinal epithelial cells. Journal of Medical Microbiology 57, 765–770. https://doi.org/10.1099/jmm.0.47769-0

Ghosh, A.S., Chowdhury, C., Nelson, D.E., 2008. Physiological functions of D-alanine carboxypeptidases in *Escherichia coli*. Trends in Microbiology 16, 309–317. https://doi.org/10.1016/j.tim.2008.04.006

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6, 343–345. https://doi.org/10.1038/nmeth.1318

Goenka, S.D., Gorzynski, J.E., Shafin, K., Fisk, D.G., Pesout, T., Jensen, T.D., Monlong, J., Chang, P.-C., Baid, G., Bernstein, J.A., Christle, J.W., Dalton, K.P., Garalde, D.R., Grove, M.E., Guillory, J., Kolesnikov, A., Nattestad, M., Ruzhnikov, M.R.Z., Samadi, M., Sethia, A., Spiteri, E., Wright, C.J., Xiong, K., Zhu, T., Jain, M., Sedlazeck, F.J., Carroll, A., Paten, B., Ashley, E.A., 2022. Accelerated identification of disease-causing variants with ultra-rapid nanopore genome sequencing. Nat Biotechnol 40, 1035–1041. https://doi.org/10.1038/s41587-022-01221-5

Gómez, S., Chaves, F., Orellana, M.A., 2017. Clinical, epidemiological and microbiological characteristics of relapse and re-infection in *Clostridium difficile* infection. Anaerobe 48, 147–151. https://doi.org/10.1016/j.anaerobe.2017.08.012

Gonzales-Luna, A.J., Olaitan, A.O., Shen, W.-J., Deshpande, A., Carlson, T.J., Dotson, K.M., Lancaster, C., Begum, K., Alam, M.J., Hurdle, J.G., Garey, K.W., 2021. Reduced Susceptibility to Metronidazole Is Associated With Initial Clinical Failure in *Clostridioides difficile* Infection. Open Forum Infectious Diseases 8. https://doi.org/10.1093/ofid/ofab365

Gostev, V., Kalinogorskaya, O., Sopova, J., Sulian, O., Chulkova, P., Velizhanina, M., Tsvetkova, I., Ageevets, I., Ageevets, V., Sidorenko, S., 2023. Adaptive Laboratory Evolution of *Staphylococcus aureus* Resistance to Vancomycin and Daptomycin: Mutation Patterns and Cross-Resistance. Antibiotics 12, 928. https://doi.org/10.3390/antibiotics12050928

Gourlé, H., Karlsson-Lindsjö, O., Hayer, J., Bongcam-Rudloff, E., 2019. Simulating Illumina metagenomic data with InSilicoSeq. Bioinformatics 35, 521–522. https://doi.org/10.1093/bioinformatics/bty630

Gouy, M., Gautier, C., 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res 10, 7055–7074.

Govind, R., Fitzwater, L., Nichols, R., 2015. Observations on the Role of TcdE Isoforms in *Clostridium difficile* Toxin Secretion. Journal of Bacteriology 197, 2600–2609. https://doi.org/10.1128/jb.00224-15

Grantham, R., Gautier, C., Gouy, M., Mercier, R., Pavé, A., 1980. Codon catalog usage and the genome hypothesis. Nucleic Acids Res 8, r49–r62.

Greentree, D.H., Rice, L.B., Donskey, C.J., 2022. Houston, We Have a Problem: Reports of *Clostridioides difficile* Isolates With Reduced Vancomycin Susceptibility. Clinical Infectious Diseases 75, 1661–1664. https://doi.org/10.1093/cid/ciac444

Guérillot, R., Gonçalves da Silva, A., Monk, I., Giulieri, S., Tomita, T., Alison, E., Porter, J., Pidot, S., Gao, W., Peleg, A.Y., Seemann, T., Stinear, T.P., Howden, B.P., 2018. Convergent Evolution Driven by Rifampin Exacerbates the Global Burden of Drug-Resistant *Staphylococcus aureus*. mSphere 3, 10.1128/msphere.00550-17. https://doi.org/10.1128/msphere.00550-17

Gupta, S., Allen-Vercoe, E., Petrof, E.O., 2016. Fecal microbiota transplantation: in perspective. Therap Adv Gastroenterol 9, 229–239. https://doi.org/10.1177/1756283X15607414

Gushchin, I., Orekhov, P., Melnikov, I., Polovinkin, V., Yuzhakova, A., Gordeliy, V., 2020. Sensor Histidine Kinase NarQ Activates via Helical Rotation, Diagonal Scissoring, and Eventually Piston-Like Shifts. International Journal of Molecular Sciences 21, 3110. https://doi.org/10.3390/ijms21093110

Hachmann, A.-B., Sevim, E., Gaballa, A., Popham, D.L., Antelmann, H., Helmann, J.D., 2011. Reduction in Membrane Phosphatidylglycerol Content Leads to Daptomycin Resistance in *Bacillus subtilis*. Antimicrobial Agents and Chemotherapy 55, 4326–4337. https://doi.org/10.1128/aac.01819-10

Hallgren, J., Tsirigos, K., Pedersen, M.D., Almagro Armenteros, J.J., Marcatili, P., Nielsen, H., Krogh, A., Winther, O., 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural network. https://doi.org/10.1101/2022.04.08.487609

Harrell, F., Dupont, C., 2019. Hmisc: Harrell Miscellaneous.

Hartkoorn, R.C., Sala, C., Neres, J., Pojer, F., Magnet, S., Mukherjee, R., Uplekar, S., Boy-Röttger, S., Altmann, K.-H., Cole, S.T., 2012. Towards a new tuberculosis drug: pyridomycin – nature's isoniazid. EMBO Mol Med 4, 1032–1042. https://doi.org/10.1002/emmm.201201689

Hartley-Tassell, L.E., Awad, M.M., Seib, K.L., Scarselli, M., Savino, S., Tiralongo, J., Lyras, D., Day, C.J., Jennings, M.P., 2019. Lectin Activity of the TcdA and TcdB Toxins of *Clostridium difficile*. Infect Immun 87, e00676-18. https://doi.org/10.1128/IAI.00676-18

He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D.J., Martin, M.J., Connor, T.R., Harris, S.R., Fairley, D., Bamford, K.B., D'Arc, S., Brazier, J., Brown, D., Coia, J.E., Douce, G., Gerding, D., Kim, H.J., Koh, T.H., Kato, H., Senoh, M., Louie, T., Michell, S., Butt, E., Peacock, S.J., Brown, N.M., Riley, T., Songer, G., Wilcox, M., Pirmohamed, M., Kuijper, E., Hawkey, P., Wren, B.W., Dougan, G., Parkhill, J.,

Lawley, T.D., 2013. Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*. Nat Genet 45, 109–113. https://doi.org/10.1038/ng.2478

Heimann, S.M., Cruz Aguilar, M.R., Mellinghof, S., Vehreschild, M.J.G.T., 2018. Economic burden and cost-effective management of *Clostridium difficile* infections. Médecine et Maladies Infectieuses 48, 23–29. https://doi.org/10.1016/j.medmal.2017.10.010

Hemmasi, S., Czulkies, B.A., Schorch, B., Veit, A., Aktories, K., Papatheodorou, P., 2015. Interaction of the *Clostridium difficile* Binary Toxin CDT and Its Host Cell Receptor, Lipolysis-stimulated Lipoprotein Receptor (LSR) *. Journal of Biological Chemistry 290, 14031–14044. https://doi.org/10.1074/jbc.M115.650523

Herbert, M., O'Keeffe, T.A., Purdy, D., Elmore, M., Minton, N.P., 2003. Gene transfer into *Clostridium difficile* CD630 and characterisation of its methylase genes. FEMS Microbiology Letters 229, 103–110. https://doi.org/10.1016/S0378-1097(03)00795-X

Hirakawa, H., Kurushima, J., Hashimoto, Y., Tomita, H., 2020. Progress Overview of Bacterial Two-Component Regulatory Systems as Potential Targets for Antimicrobial Chemotherapy. Antibiotics (Basel) 9, 635. https://doi.org/10.3390/antibiotics9100635

Ho, T.D., Williams, K.B., Chen, Y., Helm, R.F., Popham, D.L., Ellermeier, C.D., 2014. *Clostridium difficile* Extracytoplasmic Function σ Factor σV Regulates Lysozyme Resistance and Is Necessary for Pathogenesis in the Hamster Model of Infection. Infect Immun 82, 2345−2355. https://doi.org/10.1128/IAI.01483-13

Holberger, L.E., Hayes, C.S., 2009. Ribosomal Protein S12 and Aminoglycoside Antibiotics Modulate A-site mRNA Cleavage and Transfer-Messenger RNA Activity in *Escherichia coli*. J Biol Chem 284, 32188–32200. https://doi.org/10.1074/jbc.M109.062745

Hu, J., Zhang, X., Liu, X., Chen, C., Sun, B., 2015. Mechanism of Reduced Vancomycin Susceptibility Conferred by *walK* Mutation in Community-Acquired Methicillin-Resistant *Staphylococcus aureus* Strain MW2. Antimicrob Agents Chemother 59, 1352–1355. https://doi.org/10.1128/AAC.04290-14

Jahn, L.J., Munck, C., Ellabaan, M.M.H., Sommer, M.O.A., 2017. Adaptive Laboratory Evolution of Antibiotic Resistance Using Different Selection Regimes Lead to Similar Phenotypes and Genotypes. Front. Microbiol. 8. https://doi.org/10.3389/fmicb.2017.00816

Jank, T., Aktories, K., 2008. Structure and mode of action of clostridial glucosylating toxins: the ABCD model. Trends in Microbiology 16, 222–229. https://doi.org/10.1016/j.tim.2008.01.011

Jank, T., Giesemann, T., Aktories, K., 2007. Rho-glucosylating *Clostridium difficile* toxins A and B: new insights into structure and function. Glycobiology 17, 15R-22R. https://doi.org/10.1093/glycob/cwm004

Janoir, C., 2016. Virulence factors of *Clostridium difficile* and their role during infection. Anaerobe 37, 13–24. https://doi.org/10.1016/j.anaerobe.2015.10.009

Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. Philos Trans A Math Phys Eng Sci 374, 20150202. https://doi.org/10.1098/rsta.2015.0202

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Juraschek, K., Borowiak, M., Tausch, S.H., Malorny, B., Käsbohrer, A., Otani, S., Schwarz, S., Meemken, D., Deneke, C., Hammerl, J.A., 2021. Outcome of Different Sequencing and Assembly Approaches on the Detection of Plasmids and Localization of Antimicrobial Resistance Genes in Commensal *Escherichia coli*. Microorganisms 9, 598. https://doi.org/10.3390/microorganisms9030598

Just, I., Selzer, J., Wilm, M., Eichel-Streiber, C. von, Mann, M., Aktories, K., 1995. Glucosylation of Rho proteins by *Clostridium difficile* toxin B. Nature 375, 500–503. https://doi.org/10.1038/375500a0

Kaewprasert, O., Nonghanphithak, D., Chetchotisakd, P., Namwat, W., Ong, R.T.-H., Faksri, K., 2022. Whole-Genome Sequencing and Drug-Susceptibility Analysis of Serial *Mycobacterium abscessus* Isolates from Thai Patients. Biology (Basel) 11, 1319. https://doi.org/10.3390/biology11091319

Kanehisa, M., Sato, Y., Kawashima, M., 2022. KEGG mapping tools for uncovering hidden features in biological data. Protein Sci 31, 47–53. https://doi.org/10.1002/pro.4172

Kassambara, A., Mundt, F., 2020. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10, 845–858. https://doi.org/10.1038/nprot.2015.053

Keshri, V., Arbuckle, K., Chabrol, O., Rolain, J., Raoult, D., Pontarotti, P., 2019. The functional convergence of antibiotic resistance in β-lactamases is not conferred by a simple convergent substitution of amino acid. Evol Appl 12, 1812−1822. https://doi.org/10.1111/eva.12835

Khozov, A.A., Bubnov, D.M., Plisov, E.D., Vybornaya, T.V., Yuzbashev, T.V., Agrimi, G., Messina, E., Stepanova, A.A., Kudina, M.D., Alekseeva, N.V., Netrusov, A.I., Sineoky, S.P., 2023. A study on L-threonine and L-serine uptake in *Escherichia coli* K-12. Front Microbiol 14, 1151716. https://doi.org/10.3389/fmicb.2023.1151716

Kimura, M., 1985. The role of compensatory neutral mutations in molecular evolution. J. Genet. 64, 7–19. https://doi.org/10.1007/BF02923549

Kirk, J.A., Gebhart, D., Buckley, A.M., Lok, S., Scholl, D., Douce, G.R., Govoni, G.R., Fagan, R.P., 2017. New Class of Precision Antimicrobials Redefines Role of *Clostridium difficile* S-layer in Virulence and Viability. Sci Transl Med 9, eaah6813. https://doi.org/10.1126/scitranslmed.aah6813

Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., Ding, L., 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25, 2283–2285. https://doi.org/10.1093/bioinformatics/btp373

Köhler, T., van Delden, C., Curty, L.K., Hamzehpour, M.M., Pechere, J.-C., 2001. Overexpression of the MexEF-OprN Multidrug Efflux System Affects Cell-to-Cell Signaling in *Pseudomonas aeruginosa*. J Bacteriol 183, 5213–5222. https://doi.org/10.1128/JB.183.18.5213-5222.2001

Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P.A., 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37, 540–546. https://doi.org/10.1038/s41587-019-0072-8

Kolte, B., Nübel, U., 2024. Genetic determinants of resistance to antimicrobial therapeutics are rare in publicly available *Clostridioides difficile* genome sequences. Journal of Antimicrobial Chemotherapy dkae101. https://doi.org/10.1093/jac/dkae101

Korendovych, I.V., 2018. Rational and Semirational Protein Design. Methods Mol Biol 1685, 15–23. https://doi.org/10.1007/978-1-4939-7366-8_2

Lawley, T.D., Clare, S., Walker, A.W., Goulding, D., Stabler, R.A., Croucher, N., Mastroeni, P., Scott, P., Raisen, C., Mottram, L., Fairweather, N.F., Wren, B.W., Parkhill, J., Dougan, G., 2009. Antibiotic

Treatment of *Clostridium difficile* Carrier Mice Triggers a Supershedder State, Spore-Mediated Transmission, and Severe Disease in Immunocompromised Hosts. Infection and Immunity 77, 3661–3669. https://doi.org/10.1128/iai.00558-09

Lawley, T.D., Clare, S., Walker, A.W., Stares, M.D., Connor, T.R., Raisen, C., Goulding, D., Rad, R., Schreiber, F., Brandt, C., Deakin, L.J., Pickard, D.J., Duncan, S.H., Flint, H.J., Clark, T.G., Parkhill, J., Dougan, G., 2012. Targeted Restoration of the Intestinal Microbiota with a Simple, Defined Bacteriotherapy Resolves Relapsing *Clostridium difficile* Disease in Mice. PLOS Pathogens 8, e1002995. https://doi.org/10.1371/journal.ppat.1002995

Lázár, V., Nagy, I., Spohn, R., Csörgő, B., Györkei, Á., Nyerges, Á., Horváth, B., Vörös, A., Busa-Fekete, R., Hrtyan, M., Bogos, B., Méhi, O., Fekete, G., Szappanos, B., Kégl, B., Papp, B., Pál, C., 2014. Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. Nat Commun 5, 4352. https://doi.org/10.1038/ncomms5352

Lee, W.J., Lattimer, L.D.N., Stephen, S., Borum, M.L., Doman, D.B., 2015. Fecal Microbiota Transplantation: A Review of Emerging Indications Beyond Relapsing *Clostridium difficile* Toxin Colitis. Gastroenterol Hepatol (N Y) 11, 24–32.

Leeds, J.A., Sachdeva, M., Mullin, S., Barnes, S.W., Ruzin, A., 2014. *In vitro* selection, via serial passage, of *Clostridium difficile* mutants with reduced susceptibility to fidaxomicin or vancomycin. Journal of Antimicrobial Chemotherapy 69, 41–44. https://doi.org/10.1093/jac/dkt302

Leitsch, D., 2019. A review on metronidazole: an old warhorse in antimicrobial chemotherapy. Parasitology 146, 1167–1178. https://doi.org/10.1017/S0031182017002025

Lenski, R.E., 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J 11, 2181–2194. https://doi.org/10.1038/ismej.2017.69

Lessa, F., Mu Yi, Bamberg Wendy M., Beldavs Zintars G., Dumyati Ghinwa K., Dunn John R., Farley Monica M., Holzbauer Stacy M., Meek James I., Phipps Erin C., Wilson Lucy E., Winston Lisa G., Cohen Jessica A., Limbago Brandi M., Fridkin Scott K., Gerding Dale N., McDonald L. Clifford, 2015. Burden of *Clostridium difficile* Infection in the United States. New England Journal of Medicine 372, 825–834. https://doi.org/10.1056/NEJMoa1408913

Leyn, S.A., Zlamal, J.E., Kurnasov, O.V., Li, X., Elane, M., Myjak, L., Godzik, M., de Crecy, A., Garcia-Alcalde, F., Ebeling, M., Osterman, A.L., 2021. Experimental evolution in morbidostat reveals converging genomic trajectories on the path to triclosan resistance. Microb Genom 7, 000553. https://doi.org/10.1099/mgen.0.000553

Li, H., Chen, T., Yu, L., Guo, H., Chen, L., Chen, Y., Chen, M., Zhao, J., Yan, H., Zhou, L., Wang, W., 2020. Genome-wide DNA methylation and transcriptome and proteome changes in *Mycobacterium tuberculosis* with para-aminosalicylic acid resistance. Chemical Biology & Drug Design 95, 104–112. https://doi.org/10.1111/cbdd.13625

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Q., Chen, S., Zhu, K., Huang, X., Huang, Y., Shen, Z., Ding, S., Gu, D., Yang, Q., Sun, H., Hu, F., Wang, H., Cai, J., Ma, B., Zhang, R., Shen, J., 2022. Collateral sensitivity to pleuromutilins in vancomycin-resistant *Enterococcus faecium*. Nat Commun 13, 1888. https://doi.org/10.1038/s41467-022-29493-0

Liao, F., Li, W., Gu, W., Zhang, W., Liu, X., Fu, X., Xu, W., Wu, Y., Lu, J., 2018. A retrospective study of community-acquired *Clostridium difficile* infection in southwest China. Sci Rep 8, 3992. https://doi.org/10.1038/s41598-018-21762-7

Lim, H.N., Lee, Y., Hussein, R., 2011. Fundamental relationship between operon organization and gene expression. Proc Natl Acad Sci U S A 108, 10626–10631. https://doi.org/10.1073/pnas.1105692108

Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebright, Y.W., Ebright, R.Y., Sineva, E., Gigliotti, M., Srivastava, A., Mandal, S., Jiang, Y., Liu, Y., Yin, R., Zhang, Z., Eng, E.T., Thomas, D., Donadio, S., Zhang, H., Zhang, C., Kapanidis, A.N., Ebright, R.H., 2018. Structural Basis Of Transcription Inhibition By Fidaxomicin (Lipiarmycin A3). Mol Cell 70, 60-71.e15. https://doi.org/10.1016/j.molcel.2018.02.026

Liu, Y., Beyer, A., Aebersold, R., 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell 165, 535–550. https://doi.org/10.1016/j.cell.2016.03.014

Loo, V.G., Poirier Louise, Miller Mark A., Oughton Matthew, Libman Michael D., Michaud Sophie, Bourgault Anne-Marie, Nguyen Tuyen, Frenette Charles, Kelly Mirabelle, Vibien Anne, Brassard Paul, Fenn Susan, Dewar Ken, Hudson Thomas J., Horn Ruth, René Pierre, Monczak Yury, Dascal André, 2005. A Predominantly Clonal Multi-Institutional Outbreak of *Clostridium difficile*–Associated

Diarrhea with High Morbidity and Mortality. New England Journal of Medicine 353, 2442–2449. https://doi.org/10.1056/NEJMoa051639

Louie Thomas J., Miller Mark A., Mullane Kathleen M., Weiss Karl, Lentnek Arnold, Golan Yoav, Gorbach Sherwood, Sears Pamela, Shue Youe-Kong, 2011. Fidaxomicin versus Vancomycin for *Clostridium difficile* Infection. New England Journal of Medicine 364, 422–431. https://doi.org/10.1056/NEJMoa0910812

Lynch, T., Chong, P., Zhang, J., Hizon, R., Du, T., Graham, M.R., Beniac, D.R., Booth, T.F., Kibsey, P., Miller, M., Gravel, D., Mulvey, M.R., Program (CNISP), C.N.I.S., 2013. Characterization of a Stable, Metronidazole-Resistant *Clostridium difficile* Clinical Isolate. PLOS ONE 8, e53757. https://doi.org/10.1371/journal.pone.0053757

Lyon, S.A., Hutton, M.L., Rood, J.I., Cheung, J.K., Lyras, D., 2016. CdtR Regulates TcdA and TcdB Production in *Clostridium difficile*. PLOS Pathogens 12, e1005758. https://doi.org/10.1371/journal.ppat.1005758

Madan, R., Petri, W.A.J., 2012. Immune responses to *Clostridium difficile* infection. Trends in Molecular Medicine 18, 658–666. https://doi.org/10.1016/j.molmed.2012.09.005

Majumdar, A., Govind, R., 2022. Regulation of *Clostridioides difficile* toxin production. Curr Opin Microbiol 65, 95–100. https://doi.org/10.1016/j.mib.2021.10.018

Mani, N., Dupuy, B., 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proceedings of the National Academy of Sciences 98, 5844–5849. https://doi.org/10.1073/pnas.101126598

Mansfield, M.J., Tremblay, B.J.-M., Zeng, J., Wei, X., Hodgins, H., Worley, J., Bry, L., Dong, M., Doxey, A.C., 2020. Phylogenomics of 8,839 *Clostridioides difficile* genomes reveals recombination-driven evolution and diversification of toxin A and B. PLOS Pathogens 16, e1009181. https://doi.org/10.1371/journal.ppat.1009181

Marchandin, H., Anjou, C., Poulen, G., Freeman, J., Wilcox, M., Jean-Pierre, H., Barbut, F., 2023. In vivo emergence of a still uncommon resistance to fidaxomicin in the urgent antimicrobial resistance threat *Clostridioides difficile*. Journal of Antimicrobial Chemotherapy 78, 1992–1999. https://doi.org/10.1093/jac/dkad194

Marinus, M.G., Casadesus, J., 2009. Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. FEMS Microbiol Rev 33, 488–503. https://doi.org/10.1111/j.1574-6976.2008.00159.x

Markkanen, M.A., Haukka, K., Pärnänen, K.M.M., Dougnon, V.T., Bonkoungou, I.J.O., Garba, Z., Tinto, H., Sarekoski, A., Karkman, A., Kantele, A., Virta, M.P.J., 2023. Metagenomic Analysis of the Abundance and Composition of Antibiotic Resistance Genes in Hospital Wastewater in Benin, Burkina Faso, and Finland. mSphere 8, e0053822. https://doi.org/10.1128/msphere.00538-22

Martin, H., Willey, B., Low, D.E., Staempfli, H.R., McGeer, A., Boerlin, P., Mulvey, M., Weese, J.S., 2008. Characterization of *Clostridium difficile* Strains Isolated from Patients in Ontario, Canada, from 2004 to 2006. J Clin Microbiol 46, 2999–3004. https://doi.org/10.1128/JCM.02437-07

Martini, M.C., Hicks, N.D., Xiao, J., Alonso, M.N., Barbier, T., Sixsmith, J., Fortune, S.M., Shell, S.S., 2022. Loss of RNase J leads to multi-drug tolerance and accumulation of highly structured mRNA fragments in *Mycobacterium tuberculosis*. PLOS Pathogens 18, e1010705. https://doi.org/10.1371/journal.ppat.1010705

Massa, S.M., Sharma, A.D., Siletti, C., Tu, Z., Godfrey, J.J., Gutheil, W.G., Huynh, T.N., 2020. c-di-AMP Accumulation Impairs Muropeptide Synthesis in *Listeria monocytogenes*. J Bacteriol 202, e00307-20. https://doi.org/10.1128/JB.00307-20

Matamouros, S., England, P., Dupuy, B., 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Molecular Microbiology 64, 1274–1288. https://doi.org/10.1111/j.1365-2958.2007.05739.x

McDonald, M.J., 2019. Microbial Experimental Evolution – a proving ground for evolutionary theory and a tool for discovery. EMBO reports 20, e46992. https://doi.org/10.15252/embr.201846992

McGovern, B.H., Ford, C.B., Henn, M.R., Pardi, D.S., Khanna, S., Hohmann, E.L., O'Brien, E.J., Desjardins, C.A., Bernardo, P., Wortman, J.R., Lombardo, M.-J., Litcofsky, K.D., Winkler, J.A., McChalicher, C.W.J., Li, S.S., Tomlinson, A.D., Nandakumar, M., Cook, D.N., Pomerantz, R.J., Auninš, J.G., Trucksis, M., 2021. SER-109, an Investigational Microbiome Drug to Reduce Recurrence After *Clostridioides difficile* Infection: Lessons Learned From a Phase 2 Trial. Clinical Infectious Diseases 72, 2132–2140. https://doi.org/10.1093/cid/ciaa387

Meacock, O.J., Durham, W.M., 2023. Tracking bacteria at high density with FAST, the Feature-Assisted Segmenter/Tracker. PLOS Computational Biology 19, e1011524. https://doi.org/10.1371/journal.pcbi.1011524

Mediati, D.G., Wu, S., Wu, W., Tree, J.J., 2021. Networks of Resistance: Small RNA Control of Antibiotic Resistance. Trends Genet 37, 35–45. https://doi.org/10.1016/j.tig.2020.08.016

Mehta, S., Cuirolo, A.X., Plata, K.B., Riosa, S., Silverman, J.A., Rubio, A., Rosato, R.R., Rosato, A.E., 2012. VraSR Two-Component Regulatory System Contributes to *mprF*-Mediated Decreased Susceptibility to Daptomycin in *In Vivo*-Selected Clinical Strains of Methicillin-Resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 56, 92–102. https://doi.org/10.1128/AAC.00432-10

Melnyk, A.H., Wong, A., Kassen, R., 2015. The fitness costs of antibiotic resistance mutations. Evol Appl 8, 273–283. https://doi.org/10.1111/eva.12196

Meziane-Cherif, D., Stogios, P.J., Evdokimova, E., Egorova, O., Savchenko, A., Courvalin, P., 2015. Structural and Functional Adaptation of Vancomycin Resistance VanT Serine Racemases. mBio 6, 10.1128/mbio.00806-15. https://doi.org/10.1128/mbio.00806-15

Miles-Jay, A., Snitkin, E.S., Lin, M.Y., Shimasaki, T., Schoeny, M., Fukuda, C., Dangana, T., Moore, N., Sansom, S.E., Yelin, R.D., Bell, P., Rao, K., Keidan, M., Standke, A., Bassis, C., Hayden, M.K., Young, V.B., 2023. Longitudinal genomic surveillance of carriage and transmission of *Clostridioides difficile* in an intensive care unit. Nat Med 29, 2526–2534. https://doi.org/10.1038/s41591-023-02549-4

Molleti, R.R., Bidell, M.R., Tatara, A.W., 2023. Fidaxomicin-Associated Hypersensitivity Reactions: Report of a Morbilliform Drug Eruption. Journal of Pharmacy Practice 36, 993–997. https://doi.org/10.1177/08971900221078780

Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S., Stewart, P.D.S., 2014. Membrane protein structure determination — The next generation. Biochim Biophys Acta 1838, 78–87. https://doi.org/10.1016/j.bbamem.2013.07.010

Moss, E.L., Maghini, D.G., Bhatt, A.S., 2020. Complete, closed bacterial genomes from microbiomes using nanopore sequencing. Nat Biotechnol 38, 701–707. https://doi.org/10.1038/s41587-020-0422-6

Müh, U., Ellermeier, C.D., Weiss, D.S., 2022. The WalRK Two-Component System Is Essential for Proper Cell Envelope Biogenesis in *Clostridioides difficile*. J Bacteriol 204, e00121-22. https://doi.org/10.1128/jb.00121-22

Muhammad, J.S., Khan, N.A., Maciver, S.K., Alharbi, A.M., Alfahemi, H., Siddiqui, R., 2022. Epigenetic-Mediated Antimicrobial Resistance: Host versus Pathogen Epigenetic Alterations. Antibiotics 11, 809. https://doi.org/10.3390/antibiotics11060809

Myers, K.S., Park, D.M., Beauchene, N.A., Kiley, P.J., 2015. Defining Bacterial Regulons Using ChIP-seq Methods. Methods 86, 80–88. https://doi.org/10.1016/j.ymeth.2015.05.022

Napolitano, L.M., Edmiston, C.E., 2017. *Clostridium difficile* disease: Diagnosis, pathogenesis, and treatment update. Surgery 162, 325–348. https://doi.org/10.1016/j.surg.2017.01.018

Neiditch, M.B., Federle, M.J., Pompeani, A.J., Kelly, R.C., Swem, D.L., Jeffrey, P.D., Bassler, B.L., Hughson, F.M., 2006. Ligand-Induced Asymmetry in Histidine Sensor Kinase Complex Regulates Quorum Sensing. Cell 126, 1095–1108. https://doi.org/10.1016/j.cell.2006.07.032

Nelson, R.L., Suda, K.J., Evans, C.T., 2017. Antibiotic treatment for *Clostridium difficile*-associated diarrhoea in adults. Cochrane Database Syst Rev 3, CD004610. https://doi.org/10.1002/14651858.CD004610.pub5

NICE, 2021. Rationales | *Clostridioides difficile* infection: antimicrobial prescribing | Guidance | NICE [WWW Document]. URL https://www.nice.org.uk/guidance/ng199/chapter/Rationales (accessed 4.15.24).

Nicholson, M.R., Thomsen, I.P., Slaughter, J.C., Creech, C.B., Edwards, K.M., 2015. Novel Risk Factors for Recurrent *Clostridium difficile* Infection in Children. Journal of Pediatric Gastroenterology and Nutrition 60, 18–22. https://doi.org/10.1097/MPG.0000000000000553

Noda, M., Matoba, Y., Kumagai, T., Sugiyama, M., 2005. A novel assay method for an amino acid racemase reaction based on circular dichroism. Biochem J 389, 491–496. https://doi.org/10.1042/BJ20041649

Nood, E. van, Vrieze Anne, Nieuwdorp Max, Fuentes Susana, Zoetendal Erwin G., de Vos Willem M., Visser Caroline E., Kuijper Ed J., Bartelsman Joep F.W.M., Tijssen Jan G.P., Speelman Peter, Dijkgraaf Marcel G.W., Keller Josbert J., 2013. Duodenal Infusion of Donor Feces for Recurrent *Clostridium difficile*. New England Journal of Medicine 368, 407–415. https://doi.org/10.1056/NEJMoa1205037

Nowak, A., Hedenstierna, M., Ursing, J., Lidman, C., Nowak, P., 2019. Efficacy of Routine Fecal Microbiota Transplantation for Treatment of Recurrent *Clostridium difficile* Infection: A Retrospective Cohort Study. International Journal of Microbiology 2019, 7395127. https://doi.org/10.1155/2019/7395127

Oezguen, N., Power, T.D., Urvil, P., Feng, H., Pothoulakis, C., Stamler, J.S., Braun, W., Savidge, T.C., 2012. *Clostridial* toxins: Sensing a target in a hostile gut environment. Gut Microbes 3, 35–41. https://doi.org/10.4161/gmic.19250

Olaitan, A.O., Dureja, C., Youngblom, M.A., Topf, M.A., Shen, W.-J., Gonzales-Luna, A.J., Deshpande, A., Hevener, K.E., Freeman, J., Wilcox, M.H., Palmer, K.L., Garey, K.W., Pepperell, C.S., Hurdle, J.G., 2023. Decoding a cryptic mechanism of metronidazole resistance among globally disseminated

fluoroquinolone-resistant *Clostridioides difficile*. Nat Commun 14, 4130. https://doi.org/10.1038/s41467-023-39429-x

Oliveira, P.H., Ribis, J.W., Garrett, E.M., Trzilova, D., Kim, A., Sekulovic, O., Mead, E.A., Pak, T., Zhu, S., Deikus, G., Touchon, M., Lewis-Sandari, M., Beckford, C., Zeitouni, N.E., Altman, D.R., Webster, E., Oussenko, I., Bunyavanich, S., Aggarwal, A.K., Bashir, A., Patel, G., Wallach, F., Hamula, C., Huprikar, S., Schadt, E.E., Sebra, R., van Bakel, H., Kasarskis, A., Tamayo, R., Shen, A., Fang, G., 2020. Epigenomic characterization of *Clostridioides difficile* finds a conserved DNA methyltransferase that mediates sporulation and pathogenesis. Nat Microbiol 5, 166–180. https://doi.org/10.1038/s41564-019-0613-4

ONT, 2024. nanoporetech/medaka.

Ormsby, M.J., Vaz, F., Kirk, J.A., Barwinska-Sendra, A., Hallam, J.C., Lanzoni-Mangutchi, P., Cole, J., Chaudhuri, R.R., Salgado, P.S., Fagan, R.P., Douce, G.R., 2023. An intact S-layer is advantageous to *Clostridioides difficile* within the host. PLoS Pathog 19, e1011015. https://doi.org/10.1371/journal.ppat.1011015

Ott, S.J., Waetzig, G.H., Rehman, A., Moltzau-Anderson, J., Bharti, R., Grasis, J.A., Cassidy, L., Tholey, A., Fickenscher, H., Seegert, D., Rosenstiel, P., Schreiber, S., 2017. Efficacy of Sterile Fecal Filtrate Transfer for Treating Patients With *Clostridium difficile* Infection. Gastroenterology 152, 799-811.e7. https://doi.org/10.1053/j.gastro.2016.11.010

Owens, R.C., Jr., Donskey, C.J., Gaynes, R.P., Loo, V.G., Muto, C.A., 2008. Antimicrobial-Associated Risk Factors for *Clostridium difficile* Infection. Clinical Infectious Diseases 46, S19–S31. https://doi.org/10.1086/521859

Palmer, A.C., Kishony, R., 2013. Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. Nat Rev Genet 14, 243–248. https://doi.org/10.1038/nrg3351

Palmer, K.L., Daniel, A., Hardy, C., Silverman, J., Gilmore, M.S., 2011. Genetic Basis for Daptomycin Resistance in *Enterococci*. Antimicrobial Agents and Chemotherapy 55, 3345–3356. https://doi.org/10.1128/aac.00207-11

Pannullo, A.G., Guan, Z., Goldfine, H., Ellermeier, C.D., 2023a. HexSDF Is Required for Synthesis of a Novel Glycolipid That Mediates Daptomycin and Bacitracin Resistance in *C. difficile*. mBio 14, e03397-22. https://doi.org/10.1128/mbio.03397-22

Pannullo, A.G., Zbylicki, B.R., Ellermeier, C.D., 2023b. Identification of DraRS in *Clostridioides difficile*, a Two-Component Regulatory System That Responds to Lipid II-Interacting Antibiotics. Journal of Bacteriology 205, e00164-23. https://doi.org/10.1128/jb.00164-23

Papatheodorou, P., Carette, J.E., Bell, G.W., Schwan, C., Guttenberg, G., Brummelkamp, T.R., Aktories, K., 2011. Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin *Clostridium difficile* transferase (CDT). Proceedings of the National Academy of Sciences 108, 16422–16427. https://doi.org/10.1073/pnas.1109772108

Papatheodorou, P., Zamboglou, C., Genisyuerek, S., Guttenberg, G., Aktories, K., 2010. *Clostridial* Glucosylating Toxins Enter Cells via Clathrin-Mediated Endocytosis. PLOS ONE 5, e10673. https://doi.org/10.1371/journal.pone.0010673

Papkou, A., Hedge, J., Kapel, N., Young, B., MacLean, R.C., 2020. Efflux pump activity potentiates the evolution of antibiotic resistance across *S. aureus* isolates. Nat Commun 11, 3970. https://doi.org/10.1038/s41467-020-17735-y

Patel, D., Senecal, J., Spellberg, B., Morris, A.M., Saxinger, L., Footer, B.W., McDonald, E.G., Lee, T.C., 2023. Fidaxomicin to prevent recurrent *Clostridioides difficile*: what will it cost in the USA and Canada? JAC Antimicrob Resist 5, dlac138. https://doi.org/10.1093/jacamr/dlac138

Payelleville, A., Brillard, J., 2021. Novel Identification of Bacterial Epigenetic Regulations Would Benefit From a Better Exploitation of Methylomic Data. Front. Microbiol. 12. https://doi.org/10.3389/fmicb.2021.685670

Peláez, T., Cercenado, E., Alcalá, L., Marín, M., Martín-López, A., Martínez-Alarcón, J., Catalán, P., Sánchez-Somolinos, M., Bouza, E., 2008. Metronidazole Resistance in *Clostridium difficile* Is Heterogeneous. J Clin Microbiol 46, 3028–3032. https://doi.org/10.1128/JCM.00524-08

Peltier, J., Courtin, P., El Meouche, I., Catel-Ferreira, M., Chapot-Chartier, M.-P., Lemée, L., Pons, J.-L., 2013. Genomic and expression analysis of the *vanG*-like gene cluster of *Clostridium difficile*. Microbiology (Reading) 159, 1510–1520. https://doi.org/10.1099/mic.0.065060-0

Peltier, J., Courtin, P., El Meouche, I., Lemée, L., Chapot-Chartier, M.-P., Pons, J.-L., 2011. *Clostridium difficile* Has an Original Peptidoglycan Structure with a High Level of N-Acetylglucosamine Deacetylation and Mainly 3-3 Cross-links. J Biol Chem 286, 29053–29062. https://doi.org/10.1074/jbc.M111.259150

Peng, Y., Li, C., Yang, Z., Shi, W., 2020. Adverse reactions of vancomycin in humans. Medicine (Baltimore) 99, e22376. https://doi.org/10.1097/MD.0000000000022376

Peng, Z., Jin, D., Kim, H.B., Stratton, C.W., Wu, B., Tang, Y.-W., Sun, X., 2017. Update on Antimicrobial Resistance in *Clostridium difficile*: Resistance Mechanisms and Antimicrobial Susceptibility Testing. J Clin Microbiol 55, 1998–2008. https://doi.org/10.1128/JCM.02250-16

Pérez-Pérez, J.M., Candela, H., Micol, J.L., 2009. Understanding synergy in genetic interactions. Trends Genet 25, 368–376. https://doi.org/10.1016/j.tig.2009.06.004

Pettersen, J.S., Nielsen, F.D., Andreassen, P.R., Møller-Jensen, J., Jørgensen, M.G., 2024. A comprehensive analysis of *pneumococcal* two-component system regulatory networks. NAR Genomics and Bioinformatics 6, lqae039. https://doi.org/10.1093/nargab/lqae039

PHE, 2024. MRSA, MSSA and Gram-negative bacteraemia and CDI: annual report [WWW Document]. GOV.UK. URL https://www.gov.uk/government/statistics/mrsa-mssa-and-e-coli-bacteraemia-and-c-difficile-infection-annual-epidemiological-commentary (accessed 6.6.24).

Piggot, P.J., Hilbert, D.W., 2004. Sporulation of *Bacillus subtilis*. Current Opinion in Microbiology, Growth and development 7, 579–586. https://doi.org/10.1016/j.mib.2004.10.001

Popovic, N., Korac, M., Nesic, Z., Milosevic, B., Urosevic, A., Jevtovic, D., Mitrovic, N., Markovic, A., Jordovic, J., Katanic, N., Barac, A., Milosevic, I., 2018. Oral teicoplanin versus oral vancomycin for the treatment of severe *Clostridium difficile* infection: a prospective observational study. Eur J Clin Microbiol Infect Dis 37, 745–754. https://doi.org/10.1007/s10096-017-3169-3

Pruitt, R.N., Chambers, M.G., Ng, K.K.-S., Ohi, M.D., Lacy, D.B., 2010. Structural organization of the functional domains of *Clostridium difficile* toxins A and B. Proceedings of the National Academy of Sciences 107, 13467–13472. https://doi.org/10.1073/pnas.1002199107

Prunier, A.-L., Leclercq, R., 2005. Role of *mutS* and *mutL* Genes in Hypermutability and Recombination in *Staphylococcus aureus*. J Bacteriol 187, 3455–3464. https://doi.org/10.1128/JB.187.10.3455-3464.2005

Pu, M., Cho, J.M., Cunningham, S.A., Behera, G.K., Becker, S., Amjad, T., Greenwood-Quaintance, K.E., Mendes-Soares, H., Jones-Hall, Y., Jeraldo, P.R., Chen, J., Dunny, G., Patel, R., Kashyap, P.C., 2021. Plasmid Acquisition Alters Vancomycin Susceptibility in *Clostridioides difficile*. Gastroenterology 160, 941-945.e8. https://doi.org/10.1053/j.gastro.2020.10.046

Purdy, D., O'Keeffe, T.A.T., Elmore, M., Herbert, M., McLeod, A., Bokori-Brown, M., Ostrowski, A., Minton, N.P., 2002. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Molecular Microbiology 46, 439–452. https://doi.org/10.1046/j.1365-2958.2002.03134.x

Qa'Dan, M., Spyres, L.M., Ballard, J.D., 2000. pH-Induced Conformational Changes in *Clostridium difficile* Toxin B. Infect Immun 68, 2470–2474.

Quandt, E.M., Deatherage, D.E., Ellington, A.D., Georgiou, G., Barrick, J.E., 2014. Recursive genomewide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in *Escherichia coli*. Proc Natl Acad Sci U S A 111, 2217–2222. https://doi.org/10.1073/pnas.1314561111

Quendera, A.P., Pinto, S.N., Pobre, V., Antunes, W., Bonifácio, V.D.B., Arraiano, C.M., Andrade, J.M., 2023. The ribonuclease PNPase is a key regulator of biofilm formation in *Listeria monocytogenes* and affects invasion of host cells. NPJ Biofilms Microbiomes 9, 34. https://doi.org/10.1038/s41522-023-00397-1

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Quraishi, M.N., Widlak, M., Bhala, N., Moore, D., Price, M., Sharma, N., Iqbal, T.H., 2017. Systematic review with meta-analysis: the efficacy of faecal microbiota transplantation for the treatment of recurrent and refractory *Clostridium difficile* infection. Alimentary Pharmacology & Therapeutics 46, 479–493. https://doi.org/10.1111/apt.14201

Rajabally, Y.A., Afzal, S., 2019. Clinical and economic comparison of an individualised immunoglobulin protocol vs. standard dosing for chronic inflammatory demyelinating polyneuropathy. J Neurol 266, 461–467. https://doi.org/10.1007/s00415-018-9157-4

Rajeev, L., Garber, M.E., Mukhopadhyay, A., 2020. Tools to map target genes of bacterial two-component system response regulators. Environmental Microbiology Reports 12, 267–276. https://doi.org/10.1111/1758-2229.12838

Ransom, E.M., Ellermeier, C.D., Weiss, D.S., 2015. Use of mCherry Red Fluorescent Protein for Studies of Protein Localization and Gene Expression in *Clostridium difficile*. Appl Environ Microbiol 81, 1652–1660. https://doi.org/10.1128/AEM.03446-14

Razavi, B., Apisarnthanarak, A., Mundy, L.M., 2007. *Clostridium difficile*: Emergence of Hypervirulence and Fluoroquinolone Resistance. Infection 35, 300–307. https://doi.org/10.1007/s15010-007-6113-0

Reigadas Ramírez, E., Bouza, E.S., 2018. Economic Burden of *Clostridium difficile* Infection in European Countries, in: Mastrantonio, P., Rupnik, M. (Eds.), Updates on *Clostridium difficile* in

Europe: Advances in Microbiology, Infectious Diseases and Public Health Volume 8. Springer International Publishing, Cham, pp. 1–12. https://doi.org/10.1007/978-3-319-72799-8_1

Reynolds, M.G., 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. Genetics 156, 1471–1481.

Reynolds, P.E., Courvalin, P., 2005. Vancomycin Resistance in *Enterococci* Due to Synthesis of Precursors Terminating in d-Alanyl-d-Serine. Antimicrob Agents Chemother 49, 21–25. https://doi.org/10.1128/AAC.49.1.21-25.2005

Rieux, V., Carbon, C., Azoulay-Dupuis, E., 2001. Complex Relationship between Acquisition of β-Lactam Resistance and Loss of Virulence in *Streptococcus pneumoniae*. The Journal of Infectious Diseases 184, 66–72. https://doi.org/10.1086/320992

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nat Biotechnol 29, 24–26. https://doi.org/10.1038/nbt.1754

Roemhild, R., Andersson, D.I., 2021. Mechanisms and therapeutic potential of collateral sensitivity to antibiotics. PLoS Pathog 17, e1009172. https://doi.org/10.1371/journal.ppat.1009172

Roemhild, R., Linkevicius, M., Andersson, D.I., 2020. Molecular mechanisms of collateral sensitivity to the antibiotic nitrofurantoin. PLoS Biol 18, e3000612. https://doi.org/10.1371/journal.pbio.3000612

Roldan, G.A., Cui, A.X., Pollock, N.R., 2018. Assessing the Burden of *Clostridium difficile* Infection in Low- and Middle-Income Countries. Journal of Clinical Microbiology 56, 10.1128/jcm.01747-17. https://doi.org/10.1128/jcm.01747-17

Ruan, Z., Wu, J., Chen, H., Draz, M.S., Xu, J., He, F., 2020. Hybrid Genome Assembly and Annotation of a Pandrug-Resistant *Klebsiella pneumoniae* Strain Using Nanopore and Illumina Sequencing. Infect Drug Resist 13, 199–206. https://doi.org/10.2147/IDR.S240404

Rubinstein, E., Keynan, Y., 2014. Vancomycin Revisited – 60 Years Later. Front. Public Health 2. https://doi.org/10.3389/fpubh.2014.00217

Rupnik, M., Wilcox, M.H., Gerding, D.N., 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol 7, 526–536. https://doi.org/10.1038/nrmicro2164

Saha, S., Kapoor, S., Tariq, R., Schuetz, A.N., Tosh, P.K., Pardi, D.S., Khanna, S., 2019. Increasing antibiotic resistance in *Clostridioides difficile*: A systematic review and meta-analysis. Anaerobe 58, 35–46. https://doi.org/10.1016/j.anaerobe.2019.102072

Santos-Lopez, A., Marshall, C.W., Scribner, M.R., Snyder, D.J., Cooper, V.S., 2019. Evolutionary pathways to antibiotic resistance are dependent upon environmental structure and bacterial lifestyle. eLife 8, e47612. https://doi.org/10.7554/eLife.47612

Sassi, M., Guérin, F., Lesec, L., Isnard, C., Fines-Guyon, M., Cattoir, V., Giard, J.-C., 2018. Genetic characterization of a VanG-type vancomycin-resistant *Enterococcus faecium* clinical isolate. Journal of Antimicrobial Chemotherapy 73, 852–855. https://doi.org/10.1093/jac/dkx510

Sayedy, L., Kothari, D., Richards, R.J., 2010. Toxic megacolon associated *Clostridium difficile* colitis. World J Gastrointest Endosc 2, 293–297. https://doi.org/10.4253/wjge.v2.i8.293

Sbahi, H., Palma, J.A.D., 2016. Faecal microbiota transplantation: applications and limitations in treating gastrointestinal disorders. BMJ Open Gastroenterology 3, e000087. https://doi.org/10.1136/bmjgast-2016-000087

Schaller, G.E., Shiu, S.-H., Armitage, J.P., 2011. Two-Component Systems and Their Co-Option for Eukaryotic Signal Transduction. Current Biology 21, R320–R330. https://doi.org/10.1016/j.cub.2011.02.045

Schulz zur Wiesch, P., Engelstädter, J., Bonhoeffer, S., 2010. Compensation of Fitness Costs and Reversibility of Antibiotic Resistance Mutations. Antimicrobial Agents and Chemotherapy 54, 2085–2095. https://doi.org/10.1128/aac.01460-09

Schwan, C., Kruppke, A.S., Nölke, T., Schumacher, L., Koch-Nolte, F., Kudryashev, M., Stahlberg, H., Aktories, K., 2014. *Clostridium difficile* toxin CDT hijacks microtubule organization and reroutes vesicle traffic to increase pathogen adherence. Proceedings of the National Academy of Sciences 111, 2313–2318. https://doi.org/10.1073/pnas.1311589111

Schwan, C., Stecher, B., Tzivelekidis, T., Ham, M. van, Rohde, M., Hardt, W.-D., Wehland, J., Aktories, K., 2009. *Clostridium difficile* Toxin CDT Induces Formation of Microtubule-Based Protrusions and Increases Adherence of Bacteria. PLOS Pathogens 5, e1000626. https://doi.org/10.1371/journal.ppat.1000626

Schwanbeck, J., Riedel, T., Laukien, F., Schober, I., Oehmig, I., Zimmermann, O., Overmann, J., Groß, U., Zautner, A.E., Bohne, W., 2019. Characterization of a clinical *Clostridioides difficile* isolate with

markedly reduced fidaxomicin susceptibility and a V1143D mutation in *rpoB*. Journal of Antimicrobial Chemotherapy 74, 6–10. https://doi.org/10.1093/jac/dky375

Schwedt, I., Wang, M., Gibhardt, J., Commichau, F.M., 2023. Cyclic di-AMP, a multifaceted regulator of central metabolism and osmolyte homeostasis in *Listeria monocytogenes*. Microlife 4, uqad005. https://doi.org/10.1093/femsml/uqad005

Schwengers, O., Jelonek, L., Dieckmann, M.A., Beyvers, S., Blom, J., Goesmann, A., 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom 7, 000685. https://doi.org/10.1099/mgen.0.000685

Sebaihia, M., Wren, B.W., Mullany, P., Fairweather, N.F., Minton, N., Stabler, R., Thomson, N.R., Roberts, A.P., Cerdeño-Tárraga, A.M., Wang, H., Holden, M.T., Wright, A., Churcher, C., Quail, M.A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P., Dowd, L., Fraser, A., Feltwell, T., Hance, Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabbinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B., Parkhill, J., 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet 38, 779–786. https://doi.org/10.1038/ng1830

Seemann, T., 2015. Snippy: fast bacterial variant calling from NGS reads.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Selim, S., 2022. Mechanisms of gram-positive vancomycin resistance (Review). Biomedical Reports 16, 1–6. https://doi.org/10.3892/br.2021.1490

Shen, W.-J., Deshpande, A., Hevener, K.E., Endres, B.T., Garey, K.W., Palmer, K.L., Hurdle, J.G., 2020. Constitutive expression of the cryptic *vanGCd* operon promotes vancomycin resistance in *Clostridioides difficile* clinical isolates. J Antimicrob Chemother 75, 859–867. https://doi.org/10.1093/jac/dkz513

Sinel, C., Augagneur, Y., Sassi, M., Bronsard, J., Cacaci, M., Guérin, F., Sanguinetti, M., Meignen, P., Cattoir, V., Felden, B., 2017. Small RNAs in vancomycin-resistant *Enterococcus faecium* involved in daptomycin response and resistance. Sci Rep 7, 11067. https://doi.org/10.1038/s41598-017-11265-2

Smith, J.M., Haigh, J., 1974. The hitch-hiking effect of a favourable gene. Genet Res 23, 23–35.

Song, J.H., Kim, Y.S., 2019. Recurrent *Clostridium difficile* Infection: Risk Factors, Treatment, and Prevention. Gut and Liver 13, 16–24. https://doi.org/10.5009/gnl18071

Spigaglia, P., 2016. Recent advances in the understanding of antibiotic resistance in *Clostridium difficile* infection. Therapeutic Advances in Infection 3, 23–42. https://doi.org/10.1177/2049936115622891

Spigaglia, P., Barbanti, F., Mastrantonio, Paola, on behalf of the European Study Group on *Clostridium difficile* (ESGCD), Ackermann, G., Balmelli, C., Barbut, F., Bouza, E., Brazier, J., Delmée, M., Drudy, D., Kuijper, E., Ladas, H., Mastrantonio, P., Nagy, E., Pituch, H., Poxton, I., Rupnik, M., Wullt, M., Yücesoy, M., 2011. Multidrug resistance in European *Clostridium difficile* clinical isolates. Journal of Antimicrobial Chemotherapy 66, 2227–2234. https://doi.org/10.1093/jac/dkr292

Spigaglia, P., Mastrantonio, P., 2002. Molecular Analysis of the Pathogenicity Locus and Polymorphism in the Putative Negative Regulator of Toxin Production (TcdC) among *Clostridium difficile* Clinical Isolates. J Clin Microbiol 40, 3470–3475. https://doi.org/10.1128/JCM.40.9.3470-3475.2002

Sprouffske, K., Wagner, A., 2016. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC Bioinformatics 17, 172. https://doi.org/10.1186/s12859-016-1016-7

Stabler, R.A., He, M., Dawson, L., Martin, M., Valiente, E., Corton, C., Lawley, T.D., Sebaihia, M., Quail, M.A., Rose, G., Gerding, D.N., Gibert, M., Popoff, M.R., Parkhill, J., Dougan, G., Wren, B.W., 2009. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biology 10, R102. https://doi.org/10.1186/gb-2009-10-9-r102

Stevens, V., Dumyati, G., Fine, L.S., Fisher, S.G., van Wijngaarden, E., 2011. Cumulative Antibiotic Exposures Over Time and the Risk of *Clostridium difficile* Infection. Clinical Infectious Diseases 53, 42–48. https://doi.org/10.1093/cid/cir301

Stogios, P.J., Savchenko, A., 2020. Molecular mechanisms of vancomycin resistance. Protein Science 29, 654–669. https://doi.org/10.1002/pro.3819

Sulaiman, J.E., Lam, H., 2021. Novel Daptomycin Tolerance and Resistance Mutations in Methicillin-Resistant *Staphylococcus aureus* from Adaptive Laboratory Evolution. mSphere 6, 10.1128/msphere.00692-21. https://doi.org/10.1128/msphere.00692-21

Tejada-Arranz, A., de Crecy-Lagard, V., de Reuse, H., 2020. Bacterial RNA degradosomes. Trends Biochem Sci 45, 42–57. https://doi.org/10.1016/j.tibs.2019.10.002

Thitiananpakorn, K., Aiba, Y., Tan, X.-E., Watanabe, S., Kiga, K., Sato'o, Y., Boonsiri, T., Li, F.-Y., Sasahara, T., Taki, Y., Azam, A.H., Zhang, Y., Cui, L., 2020. Association of *mprF* mutations with cross-resistance to daptomycin and vancomycin in methicillin-resistant *Staphylococcus aureus* (MRSA). Scientific Reports 10. https://doi.org/10.1038/s41598-020-73108-x

Toprak, E., Veres, A., Michel, J.-B., Chait, R., Hartl, D.L., Kishony, R., 2012. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat Genet 44, 101–105. https://doi.org/10.1038/ng.1034

Trampari, E., Prischi, F., Vargiu, A.V., Abi-Assaf, J., Bavro, V.N., Webber, M.A., 2023. Functionally distinct mutations within AcrB underpin antibiotic resistance in different lifestyles. npj Antimicrob Resist 1, 1–13. https://doi.org/10.1038/s44259-023-00001-8

Tran, T.T., Munita, J.M., Arias, C.A., 2015. Mechanisms of drug resistance: daptomycin resistance. Annals of the New York Academy of Sciences 1354, 32–53. https://doi.org/10.1111/nyas.12948

Turner, C.E., Kurupati, P., Jones, M.D., Edwards, R.J., Sriskandan, S., 2009. Emerging role of the IL-8 cleaving enzyme SpyCEP in Clinical *Streptococcus pyogenes* infection. J Infect Dis 200, 555–563. https://doi.org/10.1086/603541

UKHSA, 2022. *Clostridioides difficile*: guidance, data and analysis [WWW Document]. GOV.UK. URL https://www.gov.uk/government/collections/clostridium-difficile-guidance-data-and-analysis (accessed 6.6.24).

Ulrich, L.E., Zhulin, I.B., 2010. The MiST2 database: a comprehensive genomics resource on microbial signal transduction. Nucleic Acids Research 38, D401–D407. https://doi.org/10.1093/nar/gkp940

van Eijk, E., Anvar, S.Y., Browne, H.P., Leung, W.Y., Frank, J., Schmitz, A.M., Roberts, A.P., Smits, W.K., 2015. Complete genome sequence of the *Clostridium difficile* laboratory strain 630*Δerm* reveals differences from strain 630, including translocation of the mobile element CTn*5*. BMC Genomics 16, 31. https://doi.org/10.1186/s12864-015-1252-7

Vaudaux, P., Francois, P., Berger-Bächi, B., Lew, D.P., 2001. *In vivo* emergence of subpopulations expressing teicoplanin or vancomycin resistance phenotypes in a glycopeptide-susceptible, methicillin-resistant strain of *Staphylococcus aureus*. Journal of Antimicrobial Chemotherapy 47, 163–170. https://doi.org/10.1093/jac/47.2.163

Venugopal, A.A., Johnson, S., 2012. Fidaxomicin: a novel macrocyclic antibiotic approved for treatment of *Clostridium difficile* infection. Clin Infect Dis 54, 568–574. https://doi.org/10.1093/cid/cir830

Vollmer, W., Tomasz, A., 2000. The *pgdA* Gene Encodes for a Peptidoglycan*N*-Acetylglucosamine Deacetylase in *Streptococcus pneumoniae* *. Journal of Biological Chemistry 275, 20496–20501. https://doi.org/10.1074/jbc.M910189199

Voth, D.E., Ballard, J.D., 2005. *Clostridium difficile* Toxins: Mechanism of Action and Role in Disease. Clinical Microbiology Reviews 18, 247–263. https://doi.org/10.1128/cmr.18.2.247-263.2005

Warden, C.D., Adamson, A.W., Neuhausen, S.L., Wu, X., 2014. Detailed comparison of two popular variant calling packages for exome and targeted exon studies. PeerJ 2, e600. https://doi.org/10.7717/peerj.600

Warny, M., Pepin, J., Fang, A., Killgore, G., Thompson, A., Brazier, J., Frost, E., McDonald, L.C., 2005. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. The Lancet 366, 1079–1084. https://doi.org/10.1016/S0140-6736(05)67420-X

Wasels, F., Kuehne, S.A., Cartman, S.T., Spigaglia, P., Barbanti, F., Minton, N.P., Mastrantonio, P., 2015. Fluoroquinolone Resistance Does Not Impose a Cost on the Fitness of *Clostridium difficile In Vitro*. Antimicrob Agents Chemother 59, 1794–1796. https://doi.org/10.1128/AAC.04503-14

Webb, B.J., Subramanian, A., Lopansri, B., Goodman, B., Jones, P.B., Ferraro, J., Stenehjem, E., Brown, S.M., 2020. Antibiotic Exposure and Risk for Hospital-Associated *Clostridioides difficile* Infection. Antimicrobial Agents and Chemotherapy 64, 10.1128/aac.02169-19. https://doi.org/10.1128/aac.02169-19

Wei, T., Simko, V., 2021. R package “corrplot”: Visualization of a Correlation Matrix.

Wick, R.R., Holt, K.E., 2022. Polypolish: Short-read polishing of long-read bacterial genome assemblies. PLOS Computational Biology 18, e1009802. https://doi.org/10.1371/journal.pcbi.1009802

Wick, R.R., Judd, L.M., Holt, K.E., 2023. Assembling the perfect bacterial genome using Oxford Nanopore and Illumina sequencing. PLOS Computational Biology 19, e1010905. https://doi.org/10.1371/journal.pcbi.1010905

Wickramage, I., Peng, Z., Chakraborty, S., Harmanus, C., Kuijper, E.J., Alrabaa, S., Smits, W.K., Sun, X., 2023. The *vanRCd* Mutation 343A>G, Resulting in a Thr115Ala Substitution, Is Associated with an Elevated Minimum Inhibitory Concentration (MIC) of Vancomycin in *Clostridioides difficile* Clinical Isolates from Florida. Microbiology Spectrum 11, e03777-22. https://doi.org/10.1128/spectrum.03777-22

Wong, J.L.C., David, S., Sanchez-Garrido, J., Woo, J.Z., Low, W.W., Morecchiato, F., Giani, T., Rossolini, G.M., Beis, K., Brett, S.J., Clements, A., Aanensen, D.M., Rouskin, S., Frankel, G., 2022. Recurrent emergence of *Klebsiella pneumoniae* carbapenem resistance mediated by an inhibitory *ompK36* mRNA secondary structure. Proceedings of the National Academy of Sciences 119, e2203593119. https://doi.org/10.1073/pnas.2203593119

Wren, B.W., Tabaqchali, S., 1987. Restriction endonuclease DNA analysis of *Clostridium difficile*. Journal of Clinical Microbiology 25, 2402. https://doi.org/10.1128/jcm.25.12.2402-2404.1987

Wright, R.C.T., Friman, V.-P., Smith, M.C.M., Brockhurst, M.A., 2019. Resistance Evolution against Phage Combinations Depends on the Timing and Order of Exposure. mBio 10, e01652-19. https://doi.org/10.1128/mBio.01652-19

Wu, N., Zhang, Yumeng, Zhang, S., Yuan, Y., Liu, S., Xu, T., Cui, P., Zhang, W., Zhang, Ying, 2022. Polynucleotide Phosphorylase Mediates a New Mechanism of Persister Formation in *Escherichia coli*. Microbiology Spectrum 11, e01546-22. https://doi.org/10.1128/spectrum.01546-22

Xu, X., Godoy-Ruiz, R., Adipietro, K.A., Peralta, C., Ben-Hail, D., Varney, K.M., Cook, M.E., Roth, B.M., Wilder, P.T., Cleveland, T., Grishaev, A., Neu, H.M., Michel, S.L.J., Yu, W., Beckett, D., Rustandi, R.R., Lancaster, C., Loughney, J.W., Kristopeit, A., Christanti, S., Olson, J.W., MacKerell, A.D., Georges, A. des, Pozharski, E., Weber, D.J., 2020. Structure of the cell-binding component of the *Clostridium difficile* binary toxin reveals a di-heptamer macromolecular assembly. Proceedings of the National Academy of Sciences 117, 1049–1058. https://doi.org/10.1073/pnas.1919490117

Yan, H., Wang, Q., Teng, M., Li, X., 2019. The DNA-binding mechanism of the TCS response regulator ArlR from *Staphylococcus aureus*. Journal of Structural Biology 208, 107388. https://doi.org/10.1016/j.jsb.2019.09.005

Yoon, E.-J., Nait Chabane, Y., Goussard, S., Snesrud, E., Courvalin, P., Dé, E., Grillot-Courvalin, C., 2015. Contribution of Resistance-Nodulation-Cell Division Efflux Systems to Antibiotic Resistance and Biofilm Formation in *Acinetobacter baumannii*. mBio 6, e00309-15. https://doi.org/10.1128/mBio.00309-15

Zamorano, L., Moyà, B., Juan, C., Mulet, X., Blázquez, J., Oliver, A., 2014. The *Pseudomonas aeruginosa* CreBC Two-Component System Plays a Major Role in the Response to β-Lactams, Fitness, Biofilm Growth, and Global Regulation. Antimicrob Agents Chemother 58, 5084−5095. https://doi.org/10.1128/AAC.02556-14

Zeng, D., Debabov, D., Hartsell, T.L., Cano, R.J., Adams, S., Schuyler, J.A., McMillan, R., Pace, J.L., 2016. Approved Glycopeptide Antibacterial Drugs: Mechanism of Action and Resistance. Cold Spring Harb Perspect Med 6, a026989. https://doi.org/10.1101/cshperspect.a026989

Zhanel, G.G., Walkty, A.J., Karlowsky, J.A., 2015. Fidaxomicin: A novel agent for the treatment of *Clostridium difficile* infection. The Canadian Journal of Infectious Diseases & Medical Microbiology 26, 305. https://doi.org/10.1155/2015/934594

Zhao, J., Zhang, M., Hui, W., Zhang, Y., Wang, J., Wang, S., Kwok, L.-Y., Kong, J., Zhang, H., Zhang, W., 2023. Roles of adenine methylation in the physiology of *Lacticaseibacillus paracasei*. Nat Commun 14, 2635. https://doi.org/10.1038/s41467-023-38291-1

Zhu, Y., Pham, T.H., Nhiep, T.H.N., Vu, N.M.T., Marcellin, E., Chakrabortti, A., Wang, Y., Waanders, J., Lo, R., Huston, W.M., Bansal, N., Nielsen, L.K., Liang, Z.-X., Turner, M.S., 2016. Cyclic-di-AMP synthesis by the diadenylate cyclase CdaA is modulated by the peptidoglycan biosynthesis enzyme GlmM in *Lactococcus lactis*. Molecular Microbiology 99, 1015–1027. https://doi.org/10.1111/mmi.13281

Appendix I – Strain List

Table A1 - Strain List

Strain	Characteristics	Source
General Strains – C. difficile		
R20291	*C. difficile* ribotype 027 strain isolated during an outbreak at Stoke Mandeville hospital, UK in 2006.	(Stabler et al., 2009)
R20291Δ*PaLoc*	R20291 with the entire pathogenicity locus (*tcdD, tcdB, tcdE, tcdA, tcdC)*, except the first codon of *tcdC*, deleted.	This study
R20291Δ*PaLoc*Δ*mutSL*	R20291Δ*PaLoc* with the entire *mutSL* locus (*mutS, mutL*), except the first codon of *mutS* and the last 2 codons of *mutL*, deleted.	This study
General Strains – E. coli		
CA434	*E. coli* conjugative donor. HB101 carrying R702.	(Purdy et al., 2002)
NEB5α	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ (lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17.*	New England Biolabs
Barcoded Strains		
R20291Δ*PaLoc pyrE::*barcode 1	R20291Δ*PaLoc* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 1 (AAGTCCTCG)	This study
R20291Δ*PaLoc pyrE::*barcode 2	R20291Δ*PaLoc* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 2 (TCTTGACCG)	This study
R20291Δ*PaLoc pyrE::*barcode 3	R20291Δ*PaLoc* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 3 (AACAACACC)	This study
R20291Δ*PaLoc pyrE::*barcode 4	R20291Δ*PaLoc* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 4 (AACAGGTGG)	This study
R20291Δ*PaLoc pyrE::*barcode 5	R20291Δ*PaLoc* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 5 (ACCGATTAG)	This study

R20291Δ*PaLocΔmutSL pyrE::*barcode 7	R20291Δ*PaLocΔmutSL* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 7 (CCTCCAACT)	This study
R20291Δ*PaLocΔmutSL pyrE::*barcode 8	R20291Δ*PaLocΔmutSL* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 8 (CGAGGACAT)	This study
R20291Δ*PaLocΔmutSL pyrE::*barcode 9	R20291Δ*PaLocΔmutSL* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 9 (CTGGTTCTA)	This study
R20291Δ*PaLocΔmutSL pyrE::*barcode 10	R20291Δ*PaLocΔmutSL* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 10 (GGATGTTGG)	This study
R20291Δ*PaLocΔmutSL pyrE::*barcode 11	R20291Δ*PaLocΔmutSL* with a 218 bp insertion between *CD0188* (*pyrE*) and *CD0189,* including 9 bp Barcode 11 (GTCACCAGT)	This study
Evolved Strains		
Bc1	R20291Δ*PaLoc pyrE::*barcode 1 isolated after 60 days of vancomycin selection pressure.	This study
Bc2	R20291Δ*PaLoc pyrE::*barcode 2 isolated after 60 days of vancomycin selection pressure.	This study
Bc3	R20291Δ*PaLoc pyrE::*barcode 3 isolated after 60 days of vancomycin selection pressure.	This study
Bc4	R20291Δ*PaLoc pyrE::*barcode 4 isolated after 60 days of vancomycin selection pressure.	This study
Bc5	R20291Δ*PaLoc pyrE::*barcode 5 isolated after 60 days of vancomycin selection pressure.	This study
Bc7	R20291Δ*PaLocΔmutSL pyrE::*barcode 7 isolated after 60 days of vancomycin selection pressure.	This study
Bc8	R20291Δ*PaLocΔmutSL pyrE::*barcode 8 isolated after 60 days of vancomycin selection pressure.	This study
Bc9	R20291Δ*PaLocΔmutSL pyrE::*barcode 9 isolated after 60 days of vancomycin selection pressure.	This study
Bc10	R20291Δ*PaLocΔmutSL pyrE::*barcode 10 isolated after 60 days of vancomycin selection pressure.	This study
Bc11	R20291Δ*PaLocΔmutSL pyrE::*barcode 11 isolated after 60 days of vancomycin selection pressure.	This study
Engineered Mutant Strains		
Bc1Δ*dacJ*	Evolved endpoint isolate Bc1 with the entire *dacJ* ORF, except for the first and last codon, deleted.	This study

R20291Δ*PaLoc comR*c.836_856del	R20291Δ*PaLoc* with the *comR* 836_856del identified in Evolved R20291Δ*PaLoc pyrE*::barcode 3.	This study
R20291Δ*PaLoc comR*c.836_856del *vanT*c.688G>A	R20291Δ*PaLoc* with the *comR* 836_856del and the *vanT* 688G>A point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 3.	This study
R20291Δ*PaLoc comR*c.1337G>T	R20291Δ*PaLoc* with the *comR* 1337G>T point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 4.	This study
R20291Δ*PaLoc comR*c.1337G>T *vanT*c.271C>A	R20291Δ*PaLoc* with the *comR* 1337G>T and the *vanT* 271C>A point mutations identified in Evolved R20291Δ*PaLoc pyrE*::barcode 4.	This study
R20291Δ*PaLocΔdacRS*	R20291Δ*PaLoc* with the complete ORFs for *dacR* and *dacS* deleted.	This study
R20291Δ*PaLoc dacS*c.548T>C	R20291Δ*PaLoc* with the *dacS* 548T>C point mutation identified in Evolved R20291Δ*PaLocΔmutSL pyrE::*barcodes 8 and 9.	This study
R20291Δ*PaLoc dacS*c.548T>C *vanT*c.347G>A	R20291Δ*PaLoc* with the *dacS* 548T>C point mutation identified in Evolved R20291Δ*PaLocΔmutSL pyrE::*barcodes 8 and 9, and the *vanT* 347G>A point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 5.	This study
R20291Δ*PaLoc dacS*c.714G>T	R20291Δ*PaLoc* with the *dacS* 714G>T point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 1.	This study
R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup	R20291Δ*PaLoc* with the *dacS* 714G>T point mutation and the *vanS* 30 bp duplication identified in Evolved R20291Δ*PaLoc pyrE*::barcode 1.	This study
R20291Δ*PaLoc dacS*c.714G>T *vanS*c.367_396dup 1,197,357_1,197,400del	R20291Δ*PaLoc* with the *dacS* 714G>T point mutation, the *vanS* 30 bp duplication and the intergenic 44 bp deletion identified in Evolved R20291Δ*PaLoc pyrE*::barcode 1.	This study
R20291Δ*PaLoc dacS*c.714G>T 1,197,357_1,197,400del	R20291Δ*PaLoc* with the *dacS* 714G>T point mutation and the intergenic 44 bp deletion identified in Evolved R20291Δ*PaLoc pyrE*::barcode 1.	This study
R20291Δ*PaLoc vanS*c.367_396dup	R20291Δ*PaLoc* with the internal *vanS* 30 bp duplication identified in Evolved R20291Δ*PaLoc pyrE*::barcode 1.	This study
R20291Δ*PaLoc vanS*c.367_396dup 1,197,357_1,197,400del	R20291Δ*PaLoc* with the *vanS* 30 bp duplication and the intergenic 44 bp deletion identified in Evolved R20291Δ*PaLoc pyrE*::barcode 1.	This study
R20291Δ*PaLoc vanT*c.271C>A	R20291Δ*PaLoc* with the *vanT* 271C>A point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 4.	This study
R20291Δ*PaLoc vanT*c.347G>A	R20291Δ*PaLoc* with the *vanT* 347G>A point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 5.	This study

R20291Δ*PaLoc* *vanT*c.359A>G	R20291Δ*PaLoc* with the *vanT* 359A>G point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 11.	This study
R20291Δ*PaLoc* *vanT*c.688G>A	R20291Δ*PaLoc* with the *vanT* 688G>A point mutation identified in Evolved R20291Δ*PaLoc pyrE*::barcode 3.	This study
R20291Δ*PaLoc* 1,197,357_1,197,400del	R20291Δ*PaLoc* with the intergenic 44 bp deletion identified in Evolved R20291Δ*PaLoc pyrE*::barcode 1.	This study

Appendix II – Plasmid List

Table A2 - Plasmid List

Plasmid	Characteristics	Source
General Plasmids		
pJAK112	pMTL-SC7215 based vector with added BamHI and SacI restriction sites for cloning.	(Fuchs et al., 2021)
pJAK143	*PaLoc* deletion – 1200bp homology arms upstream and downstream of the *PaLoc* for deletion of the entire pathogenicity locus (*tcdD, tcdB, tcdE, tcdA, dxtA)*, except the first codon of *dxtA*.	(Ormsby et al., 2023)
pJEB002	*mutSL* deletion – 1200 bp homology arms upstream and downstream of *mutSL* for deletion of the entire *mutSL* locus (*mutS, mutL*), except the first codon of *mutS* and the last 2 codons of *mutL.*	This study
pMTL-SC7215	Allele exchange vector for *codA*-based selection.	(Cartman et al., 2012)
Plasmids for Barcoding		
pJAK081	pMTL-SC7215 based vector with 1,200 bp homology arms for insertion of DNA sequences between *CD0188* (*pyrE*) and *CD0189* in the *C. difficile* R20291 genome.	This study
pJAK201	*pyrE::*barcode 1 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode 1 (AAGTCCTCG).	This study
pJAK202	*pyrE::*barcode 2 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode 2 (TCTTGACCG).	This study
pJAK203	*pyrE::*barcode 3 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode (AACAACACC).	This study
pJAK204	*pyrE::*barcode 4 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode 4 (AACAGGTGG).	This study
pJAK205	*pyrE::*barcode 5 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode 5 (ACCGATTAG).	This study
pJAK207	*pyrE::*barcode 7 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode 7 (CCTCCAACT).	This study
pJAK208	*pyrE::*barcode 8 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode 8 (CGAGGACAT).	This study
pJAK209	*pyrE::*barcode 9 – pJAK081 based vector, with a 218 bp insertion containing 9 nt Barcode 9 (CTGGTTCTA).	This study

pJAK210	*pyrE::*barcode 10 – pJAK081 based vector*,* with a 218 bp insertion containing 9 nt Barcode 10 (GGATGTTGG).	This study
pJAK211	*pyrE::*barcode 11 – pJAK081 based vector*,* with a 218 bp insertion containing 9 nt Barcode e 11 (GTCACCAGT).	This study
Plasmids for Engineering Mutations		
pJEB008	*vanT*c.347G>A – pJAK112 based vector containing 1900 bp homology arms centred on the *vanT*c.347G>A point mutation.	This study
pJEB012	*vanT*c.359A>G – pJAK112 based vector containing 1900 bp homology arms centred on the *vanT*c.359A>G point mutation.	This study
pJEB019	*dacS*c.548T>C – pJAK112 based vector containing 1926 bp homology arms centred on the *dacS*c.548T>C point mutation.	This study
pJEB021	*comR*c.836_856del – pJAK112 based vector containing 1909 bp homology arms centred around *comR*c.836_856del.	This study
pJEB022	*comR*c.1337G>T – pJAK112 based vector containing 1900 bp homology arms centred on the *comR*c.1337G>T point mutation.	This study
pJEB026	*dacS*c.714G>T – pJAK112 based vector containing 1926 bp homology arms centred on the *dacS*c.714G>T point mutation.	This study
pJEB027	*vanT*c.688G>A – pJAK112 based vector containing 1889 bp homology arms centred on the *vanT*c.688G>A point mutation.	This study
pJEB031	*vanT*c.271C>A – pJAK112 based vector containing 1900 bp homology arms centred on the *vanT*c.271C>A point mutation.	This study
pJEB033	1,197,357_1,197,400del – pJAK112 based vector containing 1884 bp homology arms centred on the 1,197,357_1,197,400 deletion.	This study
pJEB034	*vanS*c.367_396dup – pJAK112 based vector containing 1942 bp homology arms centred on an intergenic 30 bp insertion in *vanS*.	This study
pLMT002	*dacRS* deletion – pJAK112 based vector containing 975 and 852 bp homology arms up and downstream respectively of *dacRS* for deletion of all except the first codon of *dacR* and last codon of *dacS.*	This study
pLMT003	*dacJ* deletion – pJAK217 based vector containing 975 bp homology arms upstream and downstream of *dacJ* for deletion of all except the first and last codon of *dacJ.*	This study
Plasmids for Overexpression		
pJEB013	*vanT* – pRPF185 based vector containing WT *vanT* for tetracycline-inducible overexpression.	This study
pJEB014	*vanT*c.271C>A – pRPF185 based vector containing *vanT*c.271C>A for tetracycline-inducible overexpression.	This study

pJEB015	*vanT*c.347G>A – pRPF185 based vector containing *vanT*c.347G>A for tetracycline-inducible overexpression.	This study
pJEB016	*vanT*c.359A>G – pRPF185 based vector containing *vanT*c.359A>G for tetracycline-inducible overexpression.	This study
pJEB017	*vanT*c.688G>A – pRPF185 based vector containing *vanT*c.688G>A for tetracycline-inducible overexpression.	This study
pRPF185	*C. difficile* inducible expression system. Shuttle plasmid containing tetracycline-inducible *gusA*	(Fagan and Fairweather, 2011)
qRT-PCR Plasmids		
pJEB029	qPCR 1 – pUC-GW-Kan vector including ~200 bp fragments of *rpoA, dacS, dacR, dacJ* and *rnpA* for qRT-PCR.	This study
pJEB032	qPCR 2 – pUC-GW-Kan vector including ~200 bp fragments of *rpoA, vanR, vanS, vanG, vanXY* and *vanT* for qRT-PCR.	This study

Appendix III – Primer List

Table A3 - Primer List

Oligonucleotide	Sequence	Use
Primers for Cloning		
RF920	CGTAGAAATACGGTGTTTTTTGTTACCCTATGGAATTTAGATATAAAAACCAATTC	Amplification of homology arm upstream of *PaLoc* with RF921
RF921	ATTTATTTTGGTGTGGACAACATTGGAATTAAATCAG	Amplification of homology arm upstream of *PaLoc* with RF920
RF922	AATTCCAATGTTGTCCACACCAAAATAAATGCC	Amplification of homology arm downstream of *PaLoc* with RF923
RF923	GGGATTTTGGTCATGAGATTATCAAAAAGGCCCAACTATGGAAAAACC	Amplification of homology arm downstream of *PaLoc* with RF922
RF2066	AATACGGTGTTTTTTGTTACCCTAGAGCTCCCACTTATAATTTCTAATGAAACTGTG	Amplification of homology arm upstream of *mutSL* with RF2067
RF2067	CCAAATATTTTACATCATTATCAAACCTCCTTCTTTTC	Amplification of homology arm upstream of *mutSL* with RF2066
RF2068	GGAGGTTTGATAATGATGTAAAATATTTGGATATTTAAAATATATGGAAAG	Amplification of homology arm downstream of *mutSL* with RF2069
RF2069	TTGGTCATGAGATTATCAAAAAGGGGATCCGCCCTTTAACTTGCACTC	Amplification of homology arm downstream of *mutSL* with RF2068
Primers for Barcoding		
RF1810	GAAAAAGGCTTCTCTCATGAGAAG	To linearise pJAK081 to add barcode fragments
RF1811	GGTACCATAAAAATAAGAAGCCTGC	To linearise pJAK081 to add barcode fragments
RF1902	ACC GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 3
RF1903	GTTGTT AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 3
RF1904	GTGG GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 4

RF1905	CTGTT AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 4
RF1906	GATTAG GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 5
RF1907	GGT AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 5
RF1912	CAACT GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 7
RF1913	GAGG AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 7
RF1914	GACAT GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 8
RF1915	CTCG AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 8
RF1916	GTTCTA GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 9
RF1917	CAG AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 9
RF1918	TTGG GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 10
RF1919	CATCC AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 10
RF1920	CAGT GAAAAAGGCTTCTCTCATGAGAAG	Inverse PCR of pJAK201 to introduce barcode 11
RF1921	GTGAC AAATGGAAGATGGAATAGAAGTAAGC	Inverse PCR of pJAK201 to introduce barcode 11
Primers for Overexpression		
RF2364	GATCGAGCTCGGACTAAGAAATGGAGGAA CAAG	Amplification of *vanT* with RF2365
RF2365	GATCGGATCCGAGTTACTTCTTTATAACATT TAACCTTC	Amplification of *vanT* with RF2364
Primers for qPCR		
RF2504	CATCATTACCAGGTGTAGCAGTG	Amplification of ~200bp *rpoA* fragment for qPCR
RF2505	GGAGGACAGATTATATCTGCACC	Amplification of ~200bp *rpoA* fragment for qPCR
RF2506	CAATCACATCATTAGCAATTTATTCCATG	Amplification of ~200bp *dacS* fragment for qPCR

RF2507	GTTCATCAATATCATCCTTTTCTTTATCC	Amplification of ~200bp *dacS* fragment for qPCR
RF2508	GGATGGGATAGAAGTTTGTAGAAAAG	Amplification of ~200bp *dacR* fragment for qPCR
RF2509	CTCTTCTAATCAGTGATTTCACTCTC	Amplification of ~200bp *dacR* fragment for qPCR
RF2510	CAACATGATTCAGAACAAGATGTTGAG	Amplification of ~200bp *dacJ* fragment for qPCR
RF2511	GCTTGCTTAACTAAATCTTCAACTGC	Amplification of ~200bp *dacJ* fragment for qPCR
RF2545	GGTAAGGAAGCTCTAGAATGTATTG	Amplification of ~200bp *vanR* fragment for qPCR
RF2546	GCAACAACTTCCAAAGGGTTAAATG	Amplification of ~200bp *vanR* fragment for qPCR
RF2547	GCTTTTCGTATGGAATATAAAGCTGC	Amplification of ~200bp *vanS* fragment for qPCR
RF2548	CTTTTCTATTGCCAATAACTCTGGAG	Amplification of ~200bp *vanS* fragment for qPCR
RF2549	GCAGTACTTGAGAATTTGAATACGG	Amplification of ~200bp *vanG* fragment for qPCR
RF2550	GGAATATGATTCTGAGAAACAGCATC	Amplification of ~200bp *vanG* fragment for qPCR
RF2551	CTGTAGATACAAGGTTTCCAAGTATTC	Amplification of ~200bp *vanXY* fragment for qPCR
RF2552	CTGTGATTTGGAAGTGCTACAAAC	Amplification of ~200bp *vanXY* fragment for qPCR
RF2553	CGTATTGCTGTGCCATTCTTTTTTATG	Amplification of ~200bp *vanT* fragment for qPCR
RF2554	CCAATAATTGATGCAGGTAGATACC	Amplification of ~200bp *vanT* fragment for qPCR

Appendix IV – Accession Numbers

Table A4 - Accession Numbers

Project Number	Description	Sequences
PRJEB66266	All Illumina data associated with this thesis.	10 parental barcoded strains, 10 evolved isolates and 10 controls, pooled population sequencing P10, 20, 30 and respective pooled controls.
PRJEB66266	All Nanopore data associated with this thesis.	WT-derived evolved isolates Bc1, 2, 3, 4 and 5.

Appendix V – Isolate Variants

Table A5 - Isolate Variants

BARCODE	POS	REF	ALT	TYPE	CONSEQUENCE	LOCUS TAG	GENE NAME	GENE FUNCTION	NT	AA
Bc1	4085308	C	A	snp	missense_variant	CDR20291_3437	dacS	two-component sensor histidine kinase	714G>T	Glu238Asp
Bc2	3730727	C	A	snp	missense_variant	CDR20291_3124	CDR20291_3124	sensor protein	692G>T	Arg231Leu
Bc2	3808825	TAT	ATCTGC	complex	missense_variant&conservative_inframe_insertion	CDR20291_3193	bclA3	putative exosporium glycoprotein	640_642delATAinsGCAGAT	Ile214delinsAlaAsp
Bc2	4085224	T	A	snp	missense_variant	CDR20291_3437	dacS	two-component sensor histidine kinase	798A>T	Lys266Asn
Bc3	1384166	AGAAGAAACTTTAGCTCACTTT	A	del	disruptive_inframe_deletion	CDR20291_1159	comR	polyribonucleotide nucleotidyltransferase	836_856delCTCACTTTGAAGAAACTTTAG	Ala279_Leu285del
Bc3	1799093	G	A	snp	missense_variant	CDR20291_1526	vanT	serine/alanine racemase	688G>A	Glu230Lys
Bc4	149615	C	G	snp	missense_variant	CDR20291_0108	CDR20291_0108	putative membrane protein	173C>G	Ala58Gly
Bc4	1384681	G	T	snp	missense_variant	CDR20291_1159	comR	polyribonucleotide nucleotidyltransferase	1337G>T	Cys446Phe
Bc4	1798676	C	A	snp	missense_variant	CDR20291_1526	vanT	serine/alanine racemase	271C>A	Pro91Thr
Bc5	973962	G	A	snp	missense_variant	CDR20291_0801	maa	maltose O-acetyltransferase	71G>A	Arg24Lys
Bc5	1798752	G	A	snp	missense_variant	CDR20291_1526	vanT	serine/alanine racemase	347G>A	Gly116Glu
Bc5	4190468	ATTACAAAAGGGCCGCTGCA	A	del	intergenic_variant					

		AAGGTGGCCTTTTGTGTAAATAATTACAATTCATAGCTGTGTTTATGGAGGTTAAC								
Bc7	8347	T	C	snp	intergenic_variant					
Bc7	251284	GA	G	del	frameshift_variant	CDR202 91_019 7	CDR2029 1_0197	putative uncharacterized protein (pseudogene)	725delA	Lys242fs
Bc7	933007	C	CT	ins	intergenic_variant					
Bc7	1202386	G	A	snp	missense_variant	CDR202 91_098 5	mrdA	putative penicillin-binding protein	1009G>A	Ala337Thr
Bc7	1479264	GA	G	del	frameshift_variant	CDR202 91_124 9	CDR2029 1_1249	putative nonribosomal peptide synthetase	489delA	Gly164fs
Bc7	1595613	T	TA	ins	frameshift_variant	CDR202 91_134 9	murJ	putative membrane protein (virulence factor (MviN) homologue)	616dupA	Thr206fs
Bc7	1619675	TA	T	del	intergenic_variant					
Bc7	2436981	AT	A	del	frameshift_variant	CDR202 91_208 1	yecS	probable amino-acid ABC transporter, permease protein	636delA	Lys212fs
Bc7	2771469	TA	T	del	frameshift_variant	CDR202 91_236 8	CDR2029 1_2368	putative competence membrane protein (pseudogene)	1213delT	Tyr405fs
Bc7	3024436	CT	C	del	frameshift_variant	CDR202 91_257 9	hpt	putative phosphoribosyltransferase	51delA	Val18fs

Bc7	3336373	AC	A	del	intergenic_variant					
Bc7	3728936	GT	G	del	intergenic_variant					
Bc7	3731131	CA	C	del	frameshift_variant	CDR202 91_312 4	CDR2029 1_3124	sensor protein	287del T	Leu96fs
Bc7	3750208	CT	C	del	frameshift_variant	CDR202 91_314 1	norR	putative transcription antitermin ator	244del A	Ser82fs
Bc7	4086199	T	C	snp	missense_variant	CDR202 91_343 8	dacR	two-component response regulator	532A> G	Thr178Al a
Bc7	2830047	T	C	snp	synonymous_variant	CDR202 91_241 5	gbeA	glycogen branching enzyme	1020T >C	His340Hi s
Bc8	23157	TA	T	del	intragenic_variant	16S_rRN A	16S_rRN A			
Bc8	544666	A	G	snp	missense_variant	CDR202 91_044 9	CDR2029 1_0449	putative transcriptional regulator	809A> G	Asp270G ly
Bc8	749566	T	C	snp	missense_variant	CDR202 91_060 7	CDR2029 1_0607	probable transporter	461T> C	Val154Al a
Bc8	1199044	C	T	snp	missense_variant	CDR202 91_098 2	mreB2	rod shape-determining protein	146C> T	Ala49Val
Bc8	1201871	T	A	snp	missense_variant	CDR202 91_098 5	CDR2029 1_0985	putative penicillin-binding protein	494T> A	Ile165Lys
Bc8	1285414	A	G	snp	intragenic_variant	16S_rRN A	16S_rRN A			
Bc8	1479264	GA	G	del	frameshift_variant	CDR202 91_124 9	CDR2029 1_1249	putative nonribosomal peptide synthetase	489del A	Gly164fs
Bc8	1661467	A	G	snp	synonymous_variant	CDR202 91_140 5	CDR2029 1_1405	putative polysaccharide deacetylase	255A> G	Pro85Pro
Bc8	1795590	C	T	snp	missense_variant	CDR202 91_152 3	vanS	two-component sensor histidine kinase	449C> T	Pro150L eu
Bc8	2167744	C	CA	ins	intragenic_variant	16S_rRN A	16S_rRN A			

Bc8	2204957	A	G	snp	synonymous_variant	CDR202 91_188 8	CDR2029 1_1888	putative membrane protein	615T> C	Gly205Gly
Bc8	2639094	A	G	snp	intragenic_variant	16S_rRN A	16S_rRN A			
Bc8	2873069	TTC	T	del	intragenic_variant	16S_rRN A	16S_rRN A			
Bc8	3166456	CA	C	del	intragenic_variant	16S_rRN A	16S_rRN A			
Bc8	3467431	AT	A	del	frameshift_variant	CDR202 91_292 4	CDR2029 1_2924	putative glycosyl hydrolase	11delA	Asn4fs
Bc8	3800064	AT	A	del	frameshift_variant	CDR202 91_318 6	CDR2029 1_3186	membrane protein, putative	1654delA	Ile552fs
Bc8	3815317	T	C	snp	synonymous_variant	CDR202 91_320 0	CDR2029 1_3200	putative transcriptional regulator	69T>C	Tyr23Tyr
Bc8	3872236	T	C	snp	intragenic_variant	23S_rRN A	23S_rRN A			
Bc8	4032859	T	C	snp	synonymous_variant	CDR202 91_339 0	greA	transcription elongation factor grea	177A> G	Ala59Ala
Bc8	4085474	A	G	snp	missense_variant	CDR202 91_343 7	dacS	two-component sensor histidine kinase	548T> C	Val183Ala
Bc8	4107320	T	C	snp	missense_variant	CDR202 91_345 9	CDR2029 1_3459	putative conjugative transposon replication initiation factor	241T> C	Tyr81His
Bc9	571745	AT	A	del	intragenic_variant	16S_rRN A	16S_rRN A			
Bc9	785084	A	G	snp	missense_variant	CDR202 91_063 3	CDR2029 1_0633	putative signaling protein	941A> G	Asp314Gly
Bc9	1095284	C	A	snp	synonymous_variant	CDR202 91_089 1	CDR2029 1_0891	cell surface protein (putativeN-acetylmuramoyl-L-alanine amidase	2019C >A	Gly673Gly
Bc9	1199434	T	C	snp	missense_variant	CDR202 91_098 2	mreB2	rod shape-determining protein	536T> C	Val179Ala
Bc9	1591582	AT	A	del	frameshift_variant	CDR202 91_134 6	CDR2029 1_1346	putative exported protein	1058delA	Asn353fs

Bc9	1920527	TA	T	del	frameshift_variant	CDR20291_1635	CDR20291_1635	two-component sensor histidine kinase	14delA	Lys5fs
Bc9	2571555	AT	A	del	intragenic_variant	16S_rRNA	16S_rRNA			
Bc9	2910395	A	AT	ins	frameshift_variant	CDR20291_2482	pyrR	PyrR bifunctional protein	452dupA	Asn151fs
Bc9	2946680	CA	C	del	frameshift_variant	CDR20291_2512	CDR20291_2512	putative exported protein	1209delT	Phe403fs
Bc9	3680965	G	A	snp	missense_variant	CDR20291_3082	sdaB	L-serine dehydratase	52C>T	His18Tyr
Bc9	3731131	CA	C	del	frameshift_variant	CDR20291_3124	CDR20291_3124	sensor protein	287delT	Leu96fs
Bc9	4079340	GT	G	del	intergenic_variant				4079341delT	
Bc9	4085474	A	G	snp	missense_variant	CDR20291_3437	dacS	two-component sensor histidine kinase	548T>C	Val183Ala
Bc9	4087103	T	C	snp	missense_variant	CDR20291_3439	dacJ	D-alanyl-D-alanine carboxypeptidase	478A>G	Thr160Ala
Bc9	4146836	A	G	snp	synonymous_variant	CDR20291_3496	CDR20291_3496	conserved hypothetical protein	864T>C	Arg288Arg
Bc10	31680	C	T	snp	intergenic_variant					
Bc10	166938	A	AG	ins	intergenic_variant					
Bc10	251284	GA	G	del	frameshift_variant	CDR20291_0197	CDR20291_0197	putative uncharacterized protein (pseudogene)	725delA	Lys242fs
Bc10	1199231	T	A	snp	missense_variant	CDR20291_0982	mreB2	rod shape-determining protein	333T>A	Ser111Arg
Bc10	1259778	A	AT	ins	frameshift_variant	CDR20291_1043	ispA	geranyltranstransferase	848dupT	Leu283fs
Bc10	1619675	TA	T	del	intergenic_variant					
Bc10	1786889	A	G	snp	missense_variant	CDR20291_1516	CDR20291_1516	ABC transporter, ATP-binding protein	581A>G	His194Arg

Bc10	2480295	AT	A	del	frameshift_variant	CDR20291_2114	CDR20291_2114	putative flavodoxin	251delA	Asn84fs
Bc10	2649705	AT	A	del	frameshift_variant	CDR20291_2259	nadA	quinolinate synthetase A	50delA	Asn17fs
Bc10	2748910	T	C	snp	missense_variant	CDR20291_2349	msmX	ABC transporter, ATP-binding protein	473T>C	Met158Thr
Bc10	2767834	C	CA	ins	intergenic_variant					
Bc10	3113632	AT	A	del	intergenic_variant					
Bc10	3566384	T	C	snp	intergenic_variant					
Bc10	3662106	GT	G	del	intergenic_variant					
Bc10	3730065	A	AT	ins	frameshift_variant	CDR20291_3123	btuD	ABC transporter, ATP-binding protein	33dupA	Phe12fs
Bc10	3731131	CA	C	del	frameshift_variant	CDR20291_3124	CDR20291_3124	sensor protein	287delT	Leu96fs
Bc10	3822970	AT	A	del	frameshift_variant	CDR20291_3206	CDR20291_3206	ABC transporter, permease protein	11delA	Asn4fs
Bc10	382372	G	A	snp	intragenic_variant	23S_rRNA	23S_rRNA			
Bc10	4086199	T	C	snp	missense_variant	CDR20291_3438	dacR	two-component response regulator	532A>G	Thr178Ala
Bc10	535956	A	G	snp	synonymous_variant	CDR20291_0445	CDR20291_0445	putative membrane protein	222A>G	Ser74Ser
Bc10	1188734	A	G	snp	synonymous_variant	CDR20291_0969	CDR20291_0969	hypothetical protein	378A>G	Ser126Ser
Bc10	2095116	G	A	snp	synonymous_variant	CDR20291_1792	CDR20291_1792	putative cell surface protein	1047G>A	Thr349Thr
Bc10	2564518	T	C	snp	synonymous_variant	CDR20291_2181	CDR20291_2181	putative sigma-54 interacting transcription antiterminator	2466A>G	Pro822Pro

Bc10	2694070	T	C	snp	synonymous_variant	CDR20291_2298	CDR20291_2298	putative multidrug efflux pump, outer membrane protein	192T>C	Ser64Ser
Bc10	3226533	T	C	snp	synonymous_variant	CDR20291_2725	glyA	serine hydroxymethyltransferase	120A>G	Val40Val
Bc10	4115049	C	T	snp	synonymous_variant	CDR20291_3465	CDR20291_3465	conjugative transposon protein	1305C>T	Asn435Asn
Bc11	434673	T	A	snp	missense_variant	CDR20291_0362	asnB1	asparagine synthetase	341T>A	Ile114Asn
Bc11	734846	C	CA	ins	frameshift_variant	CDR20291_0594	CDR20291_0594	putative glyoxalase	54dupA	Phe19fs
Bc11	1511820	AT	A	del	frameshift_variant	CDR20291_1273	rutB	putative isochorismatase	311delA	Asn104fs
Bc11	1798764	A	G	snp	missense_variant	CDR20291_1526	vanT	serine/alanine racemase	359A>G	His120Arg
Bc11	1806760	A	G	snp	intergenic_variant					
Bc11	1932986	CA	C	del	frameshift_variant	CDR20291_1645	CDR20291_1645	cell surface protein (putative cell surface-associated cysteine protease)	24delA	Val9fs
Bc11	2025318	TA	T	del	frameshift_variant	CDR20291_1728	aroB	3-dehydroquinate synthase	97delA	Ile33fs
Bc11	2084450	G	A	snp	missense_variant	CDR20291_1786	CDR20291_1786	putative uncharacterized protein	250G>A	Gly84Arg
Bc11	2678026	AT	A	del	frameshift_variant	CDR20291_2283	CDR20291_2283	putative transglycosylase	113delA	Asn38fs
Bc11	2728777	AT	A	del	frameshift_variant	CDR20291_2328	CDR20291_2328	putative Mg2+ transporter	1267delA	Met423fs
Bc11	3039561	AT	A	del	intergenic_variant					
Bc11	3070068	T	C	snp	missense_variant	CDR20291_2609	CDR20291_2609	putative membrane protein	878A>G	Asn293Ser

Bc11	3073126	A	AT	ins	frameshift_variant	CDR20291_2611	CDR20291_2611	two-component response regulator	25dupA	Ile9fs
Bc11	3075606	AT	A	del	frameshift_variant	CDR20291_2614	CDR20291_2614	putative UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	581delA	Asn194fs
Bc11	3120023	GT	G	del	frameshift_variant	CDR20291_2649	CDR20291_2649	putative N-acetylmuramoyl-L-alanine amidase	12delA	Lys4fs
Bc11	3375243	AT	A	del	frameshift_variant	CDR20291_2852	CDR20291_2852	putative transcriptional antiterminator	1895delA	Asn632fs
Bc11	3399917	T	C	snp	intergenic_variant					
Bc11	3680127	ACTG	A	del	disruptive_inframe_deletion	CDR20291_3082	sdaB	L-serine dehydratase	887_889delCAG	Ala296del
Bc11	3953261	AT	A	del	frameshift_variant	CDR20291_3321	hemK	protein methyltransferase	461delA	Asn154fs
Bc11	4047709	C	CA	ins	intergenic_variant					
Bc11	4129099	G	A	snp	missense_variant	CDR20291_3478	CDR20291_3478	transposase-like protein b	5C>T	Ala2Val
Bc11	4190617	CT	C	del	frameshift_variant	CDR20291_3541	rpmH	50S ribosomal protein L34	64delA	Arg22fs
Bc11	1711702	G	A	snp	synonymous_variant	CDR20291_1449	CDR20291_1449	putative phage tail tape measure protein	1833G>A	Lys611Lys
Bc11	2491603	T	C	snp	synonymous_variant	CDR20291_2124	aroD	3-dehydroquinate dehydratase	462A>G	Pro154Pro

Appendix VI – Population Variants

Table A6 - Population Variants

PASSAGE	BAR CODE	POS	REF	ALT	TYPE	CONSEQUENCE	LOCUS TAG	GENE NAME	NT	AA	FREQ
10	Bc1	29915	A	G	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29915A>G		30.30%
10	Bc1	29932	C	T	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29932C>T		5.17%
10	Bc1	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960T>C		28.40%
10	Bc1	31333	T	C	intergenic_variant						6.36%
10	Bc1	826072	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	163A>G	Lys55Glu	5.88%
10	Bc1	1550844	T	TA	frameshift_variant	HIGH	CDR20291_1312	CDR20291_1312	267dupA	Val90fs	6.45%
10	Bc1	1689708	C	CT	intergenic_variant						5%
10	Bc1	1795507	T	TGAACAACGTAAATTTGAGACACAAATGGCA	disruptive_inframe_insertion	MODERATE	CDR20291_1523	vanS	378_407dupATTTGAGACACAATGGCAGAACAACGTAA	Lys136_Asn137insPheGluThrGlnMetAlaGluGlnArgLys	16.87%
10	Bc1	2316228	G	T	synonymous_variant	LOW	CDR20291_1983	CDR20291_1983	462C>A	Ser154Ser	5.74%
10	Bc1	2535203	C	T	intergenic_variant						5.26%
10	Bc1	3849112	A	AT	intergenic_variant						7.60%
10	Bc1	4085308	C	A	missense_variant	MODERATE	CDR20291_3437	dacS	714G>T	Glu238Asp	100%
10	Bc10	23157	TA	T	intergenic_variant						10.84%
10	Bc10	78649	CA	C	intergenic_variant						5.41%
10	Bc10	772099	GA	G	frameshift_variant	HIGH	CDR20291_0624	CDR20291_0624	324delA	Lys108fs	6.10%

10	Bc10	826072	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	163A>G	Lys55Glu	6.94%
10	Bc10	2647261	G	A	missense_variant	MODERATE	CDR20291_2257	nadC	286C>T	Arg96Cys	7.14%
10	Bc10	3669320	G	A	intergenic_variant						8.70%
10	Bc10	3728936	G	GT	intergenic_variant						8.99%
10	Bc10	3728967	A	AT	intergenic_variant						6.67%
10	Bc10	3730065	AT	A	frameshift_variant	HIGH	CDR20291_3123	CDR20291_3123	33delA	Lys11fs	27.66%
10	Bc10	4086199	T	C	missense_variant	MODERATE	CDR20291_3438	dacR	532A>G	Thr178Ala	100%
10	Bc10	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	95.79%
10	Bc11	31600	T	G	intergenic_variant						16.33%
10	Bc11	31602	T	A	intergenic_variant						14.57%
10	Bc11	132638	T	C	intergenic_variant						5.23%
10	Bc11	147606	G	A	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	14760 6G>A		7.60%
10	Bc11	251284	GA	G	frameshift_variant	HIGH	CDR20291_0197	CDR20291_0197	725delA	Lys242fs	5.37%
10	Bc11	489013	G	T	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	433C>A	Gln145Lys	9.57%
10	Bc11	489163	C	T	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	283G>A	Glu95Lys	9.43%
10	Bc11	528739	C	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3284C>A	Ala1095Asp	5.08%
10	Bc11	660827	AT	A	intergenic_variant						6.23%
10	Bc11	825947	C	T	synonymous_variant	LOW	CDR20291_0667	CDR20291_0667	288G>A	Leu96Leu	6.81%
10	Bc11	1470820	GA	G	frameshift_variant	HIGH	CDR20291_1240	CDR20291_1240	711delA	Lys237fs	31.49%
10	Bc11	1511820	AT	A	frameshift_variant	HIGH	CDR20291_1273	CDR20291_1273	311delA	Asn104fs	7.87%
10	Bc11	1795694	C	T	missense_variant	MODERATE	CDR20291_1523	vanS	553C>T	Arg185Cys	28.64%
10	Bc11	1798542	G	A	missense_variant	MODERATE	CDR20291_1526	vanT	137G>A	Arg46His	39.27%
10	Bc11	1806760	A	G	intergenic_variant						9.97%
10	Bc11	2025318	TA	T	frameshift_variant	HIGH	CDR20291_1728	aroB	97delA	Ile33fs	12.76%

10	Bc11	2391896	T	TA	frameshift_variant	HIGH	CDR20291_2040	CDR20291_2040	63dupT	Ile22fs	52.27%
10	Bc11	2418630	AT	A	frameshift_variant	HIGH	CDR20291_2064	gabT	45delA	Lys15fs	31.20%
10	Bc11	2491603	T	C	synonymous_variant	LOW	CDR20291_2124	aroD	462A>G	Pro154Pro	8.26%
10	Bc11	2889250	C	T	missense_variant	MODERATE	CDR20291_2464	stk	1264G>A	Ala422Thr	33.05%
10	Bc11	3070068	T	C	missense_variant	MODERATE	CDR20291_2609	CDR20291_2609	878A>G	Asn293Ser	11.74%
10	Bc11	3073126	A	AT	frameshift_variant	HIGH	CDR20291_2611	CDR20291_2611	25dupA	Ile9fs	34.18%
10	Bc11	3075606	AT	A	frameshift_variant	HIGH	CDR20291_2614	CDR20291_2614	581delA	Asn194fs	18.96%
10	Bc11	3301535	T	C	missense_variant	MODERATE	CDR20291_2788	ntpB	475A>G	Asn159Asp	29.95%
10	Bc11	3399917	T	C	intergenic_variant						99.19%
10	Bc11	3680127	ACTG	A	disruptive_inframe_deletion	MODERATE	CDR20291_3082	sdaB	887_889delCAG	Ala296del	55.05%
10	Bc11	4084629	TA	T	intergenic_variant						34.01%
10	Bc11	4190617	CT	C	frameshift_variant	HIGH	CDR20291_3541	rpmH	64delA	Arg22fs	96.85%
10	Bc2	29765	T	C	intragenic_variant	MODIFIER	Gene_26819_29769	23S_rRNA	29765T>C		5.97%
10	Bc2	29869	A	T	intergenic_variant						16.20%
10	Bc2	30004	A	G	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	30004A>G		9.47%
10	Bc2	30006	G	A	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	30006G>A		16.73%
10	Bc2	30020	A	G	intergenic_variant						11.11%
10	Bc2	30037	T	C	intergenic_variant						9.36%
10	Bc2	32998	A	AAAT	intergenic_variant						12.86%
10	Bc2	128090	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128090G>A		15.79%
10	Bc2	129323	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	129323G>A		8.61%
10	Bc2	129810	C	T	intragenic_variant	MODIFIER	Gene_129785_132415	23S_rRNA	129810C>T		47.06%
10	Bc2	132468	C	T	intergenic_variant						20.87%

10	Bc2	132697	T	C	intergenic_variant						34.91%
10	Bc2	133061	TA	T	intragenic_variant	MODIFIER	Gene_133028_134659	16S_rRNA	133062delA		8.22%
10	Bc2	147594	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147594C>T		6.33%
10	Bc2	147643	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147643C>T		5.46%
10	Bc2	147691	A	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147691A>G		8.30%
10	Bc2	147693	G	A	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147693G>A		9.49%
10	Bc2	147701	GT	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147702delT		8.30%
10	Bc2	525201	C	A	intergenic_variant						29.85%
10	Bc2	525407	T	G	intergenic_variant						27.64%
10	Bc2	706094	CA	C	intergenic_variant						6.21%
10	Bc2	825947	C	T	synonymous_variant	LOW	CDR20291_0667	CDR20291_0667	288G>A	Leu96Leu	7.60%
10	Bc2	826072	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	163A>G	Lys55Glu	5.76%
10	Bc2	826104	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	131A>G	Tyr44Cys	7.35%
10	Bc2	826150	TA	T	frameshift_variant	HIGH	CDR20291_0667	CDR20291_0667	84delT	Phe28fs	15.70%
10	Bc2	1974218	A	C	missense_variant	MODERATE	CDR20291_1681	CDR20291_1681	898A>C	Asn300His	5.69%
10	Bc2	3730727	C	A	missense_variant	MODERATE	CDR20291_3124	CDR20291_3124	692G>T	Arg231Leu	100%
10	Bc2	3808824	CT	C	frameshift_variant	HIGH	CDR20291_3193	bclA3	642delA	Ile214fs	90.20%
10	Bc2	3846374	TA	T	intergenic_variant						11.96%
10	Bc2	3849112	A	AT	intergenic_variant						5.78%
10	Bc2	3872941	G	A	intergenic_variant						10.66%
10	Bc2	3873178	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873178C>T		10.38%
10	Bc2	3873262	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873262C>T		24.58%

10	Bc2	4085224	T	A	missense_variant	MODERATE	CDR20291_3437	dacS	798A>T	Lys266Asn	100%
10	Bc2	3808827	T	TCTGC	frameshift_variant	HIGH	CDR20291_3193	bclA3	639_640insGCAG	Ile214fs	100%
10	Bc3	29956	C	T	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29956C>T		5.17%
10	Bc3	147413	C	T	intragenic_variant	MODIFIER	Gene_144560_147512	23S_rRNA	147413C>T		5.53%
10	Bc3	150726	G	T	stop_gained	HIGH	CDR20291_0109	CDR20291_0109	379G>T	Glu127*	6.56%
10	Bc3	581487	T	TA	intergenic_variant						83.45%
10	Bc3	581494	T	TA	intergenic_variant						95.35%
10	Bc3	706094	C	CA	intergenic_variant						5.74%
10	Bc3	825947	C	T	synonymous_variant	LOW	CDR20291_0667	CDR20291_0667	288G>A	Leu96Leu	5.62%
10	Bc3	3680122	ATGCTAC	A	conservative_inframe_deletion	MODERATE	CDR20291_3082	sdaB	889_894delGTAGCA	Val297_Ala298del	20.70%
10	Bc4	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960T>C		25.25%
10	Bc4	33003	G	A	intergenic_variant						5.59%
10	Bc4	129323	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	12932 3G>A		10.59%
10	Bc4	149615	C	G	missense_variant	MODERATE	CDR20291_0108	CDR20291_0108	173C>G	Ala58Gly	7.88%
10	Bc4	1384681	G	T	missense_variant	MODERATE	CDR20291_1159	comR	1337G>T	Cys446Phe	95.91%
10	Bc4	1795695	G	T	missense_variant	MODERATE	CDR20291_1523	vanS	554G>T	Arg185Leu	45.93%
10	Bc4	4047709	CA	C	intergenic_variant						5.70%
10	Bc5	29915	A	G	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29915A>G		25%
10	Bc5	29919	G	A	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29919G>A		6.17%
10	Bc5	29956	C	T	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29956C>T		8.86%
10	Bc5	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960T>C		21.33%
10	Bc5	30037	T	C	intergenic_variant						12.12%

10	Bc5	128090	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128090G>A			28.57%
10	Bc5	147691	A	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147691A>G			15.62%
10	Bc5	147693	G	A	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147693G>A			15.62%
10	Bc5	147701	GT	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147702delT			17.86%
10	Bc5	1798752	G	A	missense_variant	MODERATE	CDR20291_1526	vanT	347G>A	Gly116Glu		10.20%
10	Bc5	3621061	C	T	missense_variant	MODERATE	CDR20291_3034	CDR20291_3034	764G>A	Gly255Asp		100%
10	Bc7	29502	C	T	intragenic_variant	MODIFIER	Gene_26819_29769	23S_rRNA	29502C>T			6.72%
10	Bc7	32998	A	AAAT	intergenic_variant							9.80%
10	Bc7	34459	A	G	intergenic_variant							9.48%
10	Bc7	132638	T	C	intergenic_variant							5.66%
10	Bc7	359692	A	G	missense_variant	MODERATE	CDR20291_0303	rbsB	659A>G	Gln220Arg		100%
10	Bc7	451877	T	C	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	564T>C	Ile188Ile		20.89%
10	Bc7	528495	T	G	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3040T>G	Ser1014Ala		87.64%
10	Bc7	528738	G	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3283G>A	Ala1095Thr		5.21%
10	Bc7	528739	C	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3284C>A	Ala1095Asp		5.82%
10	Bc7	613231	TTATATTGG	T	intergenic_variant							72.07%
10	Bc7	830581	A	G	intergenic_variant							99.54%
10	Bc7	1096908	G	A	missense_variant	MODERATE	CDR20291_0892	CDR20291_0892	1112G>A	Gly371Glu		28.89%
10	Bc7	1096909	A	T	synonymous_variant	LOW	CDR20291_0892	CDR20291_0892	1113A>T	Gly371Gly		28.79%
10	Bc7	1096918	G	A	synonymous_variant	LOW	CDR20291_0892	CDR20291_0892	1122G>A	Lys374Lys		30.81%
10	Bc7	1096923	G	A	missense_variant	MODERATE	CDR20291_0892	CDR20291_0892	1127G>A	Ser376Asn		28.35%
10	Bc7	1096945	C	T	synonymous_variant	LOW	CDR20291_0892	CDR20291_0892	1149C>T	Tyr383Tyr		33.66%
10	Bc7	1292784	A	AAT	intergenic_variant							5.45%

10	Bc7	1619675	TA	T	intergenic_variant						6.45%
10	Bc7	1974218	A	C	missense_variant	MODERATE	CDR20291_1681	CDR20291_1681	898A>C	Asn300His	6.38%
10	Bc7	2455225	T	C	missense_variant	MODERATE	CDR20291_2094	mapA	1961A>G	Tyr654Cys	100%
10	Bc7	3183352	T	C	synonymous_variant	LOW	CDR20291_2691	tgt	114A>G	Pro38Pro	100%
10	Bc7	3280790	GA	G	frameshift_variant	HIGH	CDR20291_2769	CDR20291_2769	1661delA	Asn554fs	94.29%
10	Bc7	3336373	AC	A	intergenic_variant						90.77%
10	Bc7	3728936	GT	G	intergenic_variant						93.50%
10	Bc7	3873262	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873262C>T		23.58%
10	Bc7	4086199	T	C	missense_variant	MODERATE	CDR20291_3438	dacR	532A>G	Thr178Ala	97.84%
10	Bc7	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	97.24%
10	Bc8	31333	T	C	intergenic_variant						6.22%
10	Bc8	31600	T	G	intergenic_variant						18.97%
10	Bc8	132638	T	C	intergenic_variant						5.13%
10	Bc8	147691	A	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147691A>G		12.28%
10	Bc8	147701	GT	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147702delT		12.16%
10	Bc8	528739	C	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3284C>A	Ala1095Asp	6.36%
10	Bc8	538059	CAG	C	frameshift_variant	HIGH	CDR20291_0446	CDR20291_0446	905_906delGA	Arg302fs	8.68%
10	Bc8	749566	T	C	missense_variant	MODERATE	CDR20291_0607	CDR20291_0607	461T>C	Val154Ala	100%
10	Bc8	864745	C	T	intergenic_variant						9.85%
10	Bc8	1199044	C	T	missense_variant	MODERATE	CDR20291_0982	mreB2	146C>T	Ala49Val	9.71%
10	Bc8	1199091	A	G	missense_variant	MODERATE	CDR20291_0982	mreB2	193A>G	Lys65Glu	5.71%
10	Bc8	1199494	A	G	missense_variant	MODERATE	CDR20291_0982	mreB2	596A>G	Asp199Gly	9.01%
10	Bc8	1292784	AAT	A	intergenic_variant						14.69%

10	Bc8	1550844	TA	T	frameshift_variant	HIGH	CDR20291_1312	CDR20291_1312	267delA	Val90fs	5.04%
10	Bc8	2408126	GT	G	frameshift_variant	HIGH	CDR20291_2055	CDR20291_2055	48delA	Lys16fs	23.91%
10	Bc8	3213230	CT	C	frameshift_variant	HIGH	CDR20291_2718	CDR20291_2718	363delA	Glu122fs	9.67%
10	Bc8	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	97.55%
10	Bc8	4085474	A	G	missense_variant	MODERATE	CDR20291_3437	dacS	548T>C	Val183Ala	100%
10	Bc9	29765	T	C	intragenic_variant	MODIFIER	Gene_26819_29769	23S_rRNA	29765T>C		5.36%
10	Bc9	31333	T	C	intergenic_variant						8.16%
10	Bc9	128090	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128090G>A		23.60%
10	Bc9	129810	C	T	intragenic_variant	MODIFIER	Gene_129785_132415	23S_rRNA	129810C>T		48.59%
10	Bc9	132468	C	T	intergenic_variant						13.73%
10	Bc9	132697	T	C	intergenic_variant						26.25%
10	Bc9	143974	G	A	intergenic_variant						31.36%
10	Bc9	144756	A	G	intragenic_variant	MODIFIER	Gene_144560_147512	23S_rRNA	144756A>G		12.35%
10	Bc9	147243	C	T	intragenic_variant	MODIFIER	Gene_144560_147512	23S_rRNA	147243C>T		86.19%
10	Bc9	528261	T	G	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	2806T>G	Ser936Ala	5.12%
10	Bc9	528495	T	G	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3040T>G	Ser1014Ala	85.29%
10	Bc9	528504	G	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3049G>A	Ala1017Thr	86.34%
10	Bc9	528505	C	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3050C>A	Ala1017Asp	86.39%
10	Bc9	581480	T	TA	intergenic_variant						73.93%
10	Bc9	581487	T	TA	intergenic_variant						80.27%
10	Bc9	581494	T	TA	intergenic_variant						94.74%
10	Bc9	613237	TGG	T	intergenic_variant						12%
10	Bc9	657549	G	GT	frameshift_variant	HIGH	CDR20291_0538	CDR20291_0538	478dupA	Thr160fs	10.19%

10	Bc9	826072	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	163A>G	Lys55Glu	6.45%
10	Bc9	1093913	A	G	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	648A>G	Pro216Pro	99.34%
10	Bc9	1093916	A	G	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	651A>G	Leu217Leu	100%
10	Bc9	1093922	T	C	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	657T>C	Thr219Thr	100%
10	Bc9	1094265	A	G	missense_variant	MODERATE	CDR20291_0891	CDR20291_0891	1000A>G	Asn334Asp	58.43%
10	Bc9	1094282	C	T	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	1017C>T	Asn339Asn	65.67%
10	Bc9	1094444	C	T	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	1179C>T	Asn393Asn	100%
10	Bc9	1094468	T	G	missense_variant	MODERATE	CDR20291_0891	CDR20291_0891	1203T>G	Asp401Glu	100%
10	Bc9	1094475	T	A	missense_variant	MODERATE	CDR20291_0891	CDR20291_0891	1210T>A	Leu404Ile	100%
10	Bc9	1170707	GT	G	intergenic_variant						5.12%
10	Bc9	1199434	T	C	missense_variant	MODERATE	CDR20291_0982	mreB2	536T>C	Val179Ala	96.39%
10	Bc9	1231049	TA	T	intergenic_variant						18.70%
10	Bc9	1302247	AAT	A	intergenic_variant						9.98%
10	Bc9	2946680	CA	C	frameshift_variant	HIGH	CDR20291_2512	CDR20291_2512	1209delT	Phe403fs	92.54%
10	Bc9	2970867	G	A	intergenic_variant						10.43%
10	Bc9	3431722	T	C	missense_variant	MODERATE	CDR20291_2898	xylA	217A>G	Thr73Ala	8.12%
10	Bc9	3873178	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873178C>T		11.17%
10	Bc9	3891493	T	C	synonymous_variant	LOW	CDR20291_3257	cbiP	1158A>G	Gly386Gly	11.13%
10	Bc9	4079340	GT	G	intergenic_variant						49.72%
10	Bc9	4085474	A	G	missense_variant	MODERATE	CDR20291_3437	dacS	548T>C	Val183Ala	98.54%
10	Bc9	4146836	A	G	synonymous_variant	LOW	CDR20291_3496	CDR20291_3496	864T>C	Arg288Arg	98.76%
10	Bc9	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	97.74%
20	Bc1	29502	C	T	intragenic_variant	MODIFIER	Gene_26819_29769	23S_rRNA	29502C>T		5%
20	Bc1	38142	G	GCT	intergenic_variant						5.07%

20	Bc1	42968	A	T	missense_variant	MODERATE	CDR20291_0015	clpC	72A>T	Leu24Phe	5.43%
20	Bc1	51407	T	TC	intergenic_variant						5.44%
20	Bc1	63279	T	A	intergenic_variant						20.59%
20	Bc1	63287	T	C	intergenic_variant						5.50%
20	Bc1	75568	A	T	synonymous_variant	LOW	CDR20291_0039	CDR20291_0039	720A>T	Ala240Ala	5.69%
20	Bc1	128090	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128090G>A		13.74%
20	Bc1	128972	C	T	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128972C>T		28.41%
20	Bc1	129810	C	T	intragenic_variant	MODIFIER	Gene_129785_132415	23S_rRNA	129810C>T		41.18%
20	Bc1	132468	C	T	intergenic_variant						17.99%
20	Bc1	132697	T	C	intergenic_variant						26.73%
20	Bc1	133061	TA	T	intragenic_variant	MODIFIER	Gene_133028_134659	16S_rRNA	133062delA		7.41%
20	Bc1	143974	G	A	intergenic_variant						34.94%
20	Bc1	147594	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147594C>T		6.57%
20	Bc1	186255	A	T	missense_variant	MODERATE	CDR20291_0142	secA1	1447A>T	Thr483Ser	6.76%
20	Bc1	213280	ATAGT	A	intergenic_variant						7.48%
20	Bc1	244756	G	T	intergenic_variant						10.14%
20	Bc1	249537	T	G	intergenic_variant						8.76%
20	Bc1	270953	T	A	intergenic_variant						6.52%
20	Bc1	291544	A	T	missense_variant	MODERATE	CDR20291_0233	flgL	232A>T	Thr78Ser	6.11%
20	Bc1	305314	A	T	synonymous_variant	LOW	CDR20291_0246	CDR20291_0246	246A>T	Ser82Ser	9.09%
20	Bc1	305427	C	T	missense_variant	MODERATE	CDR20291_0246	CDR20291_0246	359C>T	Ala120Val	7.23%
20	Bc1	306048	T	A	intergenic_variant						11.94%
20	Bc1	308263	A	G	intergenic_variant						18.49%

20	Bc1	308268	T	A	intergenic_variant						20.17%
20	Bc1	345893	A	G	synonymous_variant	LOW	CDR20291_0290	CDR20291_0290	18A>G	Leu6Leu	5.56%
20	Bc1	357069	A	T	missense_variant	MODERATE	CDR20291_0301	rbsR	38A>T	Glu13Val	5.26%
20	Bc1	368787	G	A	intergenic_variant						10.94%
20	Bc1	368946	T	C	intergenic_variant						8.25%
20	Bc1	450681	T	C	intron_variant	MODIFIER	CDR20291_0376	CDR20291_0376	204+15T>C		9.14%
20	Bc1	450687	C	T	intron_variant	MODIFIER	CDR20291_0376	CDR20291_0376	204+21C>T		7.87%
20	Bc1	450738	T	C	intron_variant	MODIFIER	CDR20291_0376	CDR20291_0376	204+72T>C		6.15%
20	Bc1	450760	G	A	intron_variant	MODIFIER	CDR20291_0376	CDR20291_0376	204+94G>A		5.13%
20	Bc1	450983	C	T	intron_variant	MODIFIER	CDR20291_0376	CDR20291_0376	204+317C>T		7.53%
20	Bc1	451147	A	T	intron_variant	MODIFIER	CDR20291_0376	CDR20291_0376	204+481A>T		5.91%
20	Bc1	451496	T	C	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	183T>C	Thr61Thr	9.57%
20	Bc1	451526	G	A	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	213G>A	Lys71Lys	6.50%
20	Bc1	451561	A	G	missense_variant	MODERATE	CDR20291_0377	CDR20291_0377	248A>G	Lys83Arg	6.06%
20	Bc1	451751	G	T	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	438G>T	Gly146Gly	7.82%
20	Bc1	451766	C	T	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	453C>T	Ala151Ala	11.86%
20	Bc1	451949	G	A	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	636G>A	Lys212Lys	6.06%
20	Bc1	452192	G	A	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	879G>A	Glu293Glu	5.24%
20	Bc1	452201	C	T	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	888C>T	Arg296Arg	6.70%
20	Bc1	452235	C	A	missense_variant	MODERATE	CDR20291_0377	CDR20291_0377	922C>A	Gln308Lys	5.16%
20	Bc1	452312	A	G	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	999A>G	Lys333Lys	7.39%
20	Bc1	452320	G	A	missense_variant	MODERATE	CDR20291_0377	CDR20291_0377	1007G>A	Ser336Asn	5.14%
20	Bc1	452509	G	A	missense_variant	MODERATE	CDR20291_0376	CDR20291_0376	217G>A	Val73Met	5.56%
20	Bc1	456321	G	A	intergenic_variant						6.87%

20	Bc1	478923	G	T	missense_variant	MODERATE	CDR20291_0397	CDR20291_0397	877C>A	Gln293Lys	8.49%
20	Bc1	484826	C	T	missense_variant	MODERATE	CDR20291_0401	CDR20291_0401	1990C>T	Pro664Ser	5.43%
20	Bc1	488294	C	T	intergenic_variant						5.76%
20	Bc1	488341	C	T	intergenic_variant						5.46%
20	Bc1	488445	T	C	intergenic_variant						7.93%
20	Bc1	488449	A	C	intergenic_variant						8.19%
20	Bc1	488607	T	C	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	839A>G	His280Arg	7.41%
20	Bc1	488613	T	C	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	833A>G	Lys278Arg	5%
20	Bc1	488777	C	T	synonymous_variant	LOW	CDR20291_0405	CDR20291_0405	669G>A	Lys223Lys	5.08%
20	Bc1	488901	G	A	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	545C>T	Thr182Ile	7.66%
20	Bc1	489008	A	C	synonymous_variant	LOW	CDR20291_0405	CDR20291_0405	438T>G	Gly146Gly	6.61%
20	Bc1	489089	C	T	synonymous_variant	LOW	CDR20291_0405	CDR20291_0405	357G>A	Gly119Gly	5.47%
20	Bc1	489185	G	A	synonymous_variant	LOW	CDR20291_0405	CDR20291_0405	261C>T	Asn87Asn	7.47%
20	Bc1	489188	T	A	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	258A>T	Glu86Asp	8.14%
20	Bc1	489392	T	C	synonymous_variant	LOW	CDR20291_0405	CDR20291_0405	54A>G	Gln18Gln	5.08%
20	Bc1	489867	A	G	intergenic_variant						5.33%
20	Bc1	489927	C	T	intergenic_variant						5.33%
20	Bc1	489962	A	G	intergenic_variant						6.22%
20	Bc1	490091	G	T	intergenic_variant						5.91%
20	Bc1	490092	T	C	intergenic_variant						6.67%
20	Bc1	507728	G	T	missense_variant	MODERATE	CDR20291_0424	spaK	58G>T	Ala20Ser	6.71%
20	Bc1	509411	C	T	missense_variant	MODERATE	CDR20291_0425	CDR20291_0425	2393G>A	Cys798Tyr	6.92%
20	Bc1	509413	T	A	synonymous_variant	LOW	CDR20291_0425	CDR20291_0425	2391A>T	Ala797Ala	6.67%
20	Bc1	527220	G	T	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	1765G>T	Val589Leu	5.49%

20	Bc1	567209	G	T	stop_gained	HIGH	CDR20291_0470	CDR20291_0470	262G>T	Glu88*	9.45%
20	Bc1	579887	TG	T	frameshift_variant	HIGH	CDR20291_0482	CDR20291_0482	390delG	Met130fs	7.14%
20	Bc1	603055	T	C	intron_variant	MODIFIER	CDR20291_0499	CDR20291_0499	432+70T>C		10%
20	Bc1	603725	G	A	missense_variant	MODERATE	CDR20291_0500	CDR20291_0500	4G>A	Val2Ile	6.32%
20	Bc1	603901	A	G	synonymous_variant	LOW	CDR20291_0500	CDR20291_0500	180A>G	Leu60Leu	5.76%
20	Bc1	603969	G	A	missense_variant	MODERATE	CDR20291_0500	CDR20291_0500	248G>A	Arg83Lys	6.36%
20	Bc1	604154	A	C	missense_variant	MODERATE	CDR20291_0500	CDR20291_0500	433A>C	Lys145Gln	6.47%
20	Bc1	604264	T	G	missense_variant	MODERATE	CDR20291_0500	CDR20291_0500	543T>G	His181Gln	9.72%
20	Bc1	604408	C	T	synonymous_variant	LOW	CDR20291_0500	CDR20291_0500	687C>T	Asn229Asn	7.55%
20	Bc1	604456	C	T	synonymous_variant	LOW	CDR20291_0500	CDR20291_0500	735C>T	Ala245Ala	7.29%
20	Bc1	604582	A	T	synonymous_variant	LOW	CDR20291_0500	CDR20291_0500	861A>T	Ser287Ser	5.22%
20	Bc1	604721	A	G	missense_variant	MODERATE	CDR20291_0500	CDR20291_0500	1000A>G	Asn334Asp	9.39%
20	Bc1	604817	T	C	synonymous_variant	LOW	CDR20291_0500	CDR20291_0500	1096T>C	Leu366Leu	9.30%
20	Bc1	604899	G	T	intron_variant	MODIFIER	CDR20291_0499	CDR20291_0499	433-101G>T		13.02%
20	Bc1	621303	G	A	intron_variant	MODIFIER	CDR20291_0513	CDR20291_0513	432+88G>A		6.04%
20	Bc1	622175	C	T	synonymous_variant	LOW	CDR20291_0514	CDR20291_0514	286C>T	Leu96Leu	6.15%
20	Bc1	622288	C	A	synonymous_variant	LOW	CDR20291_0514	CDR20291_0514	399C>A	Leu133Leu	6.25%
20	Bc1	622449	C	A	missense_variant	MODERATE	CDR20291_0514	CDR20291_0514	560C>A	Thr187Asn	6.10%
20	Bc1	622590	C	T	missense_variant	MODERATE	CDR20291_0514	CDR20291_0514	701C>T	Pro234Leu	7.25%
20	Bc1	622750	A	G	stop_retained_variant	LOW	CDR20291_0514	CDR20291_0514	861A>G	Ter287Ter	12.20%
20	Bc1	622862	G	A	missense_variant	MODERATE	CDR20291_0514	CDR20291_0514	973G>A	Val325Ile	7.41%
20	Bc1	623006	G	A	intron_variant	MODIFIER	CDR20291_0513	CDR20291_0513	433-33G>A		5.80%
20	Bc1	623051	G	A	missense_variant	MODERATE	CDR20291_0513	CDR20291_0513	445G>A	Val149Met	6.50%
20	Bc1	644649	A	T	intron_variant	MODIFIER	CDR20291_0529	CDR20291_0529	666+479A>T		5.47%

20	Bc1	644818	A	G	intron_variant	MODIFIER	CDR20291_0529	CDR20291_0529	666+648A>G		5.86%
20	Bc1	645669	C	T	missense_variant	MODERATE	CDR20291_0530	CDR20291_0530	845C>T	Ser282Leu	7.69%
20	Bc1	645676	A	T	synonymous_variant	LOW	CDR20291_0530	CDR20291_0530	852A>T	Ser284Ser	8.23%
20	Bc1	645815	A	G	missense_variant	MODERATE	CDR20291_0530	CDR20291_0530	991A>G	Asn331Asp	7.98%
20	Bc1	706094	CA	C	intergenic_variant						5.98%
20	Bc1	826104	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	131A>G	Tyr44Cys	9.73%
20	Bc1	864133	T	A	intergenic_variant						8.28%
20	Bc1	864154	T	C	intergenic_variant						8.78%
20	Bc1	864238	G	A	intergenic_variant						5.49%
20	Bc1	864304	T	C	intergenic_variant						5.43%
20	Bc1	864563	T	A	intergenic_variant						5.16%
20	Bc1	865038	T	G	missense_variant	MODERATE	CDR20291_0697	CDR20291_0697	204T>G	Asn68Lys	6.36%
20	Bc1	865047	A	G	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	213A>G	Lys71Lys	5.24%
20	Bc1	865182	C	T	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	348C>T	Phe116Phe	5.43%
20	Bc1	865267	A	C	missense_variant	MODERATE	CDR20291_0697	CDR20291_0697	433A>C	Lys145Gln	5.10%
20	Bc1	865332	T	C	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	498T>C	Cys166Cys	7.49%
20	Bc1	865371	T	C	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	537T>C	Asn179Asn	5.65%
20	Bc1	865609	C	T	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	775C>T	Leu259Leu	10.23%
20	Bc1	865620	T	G	missense_variant	MODERATE	CDR20291_0697	CDR20291_0697	786T>G	Asp262Glu	11.74%
20	Bc1	865623	C	T	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	789C>T	Asn263Asn	7.86%
20	Bc1	865830	T	G	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	996T>G	Val332Val	7.10%
20	Bc1	865836	T	C	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	1002T>C	Asp334Asp	6.79%
20	Bc1	865930	T	C	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	1096T>C	Leu366Leu	7.14%
20	Bc1	866043	A	C	intergenic_variant						5.07%

20	Bc1	890661	T	A	missense_variant	MODERATE	CDR20291_0722	CDR20291_0722	11T>A	Ile4Lys	6.36%
20	Bc1	895286	G	T	intergenic_variant						5.23%
20	Bc1	913267	C	T	missense_variant	MODERATE	CDR20291_0744	CDR20291_0744	131C>T	Ser44Phe	6.61%
20	Bc1	913270	T	G	missense_variant	MODERATE	CDR20291_0744	CDR20291_0744	134T>G	Ile45Ser	6.25%
20	Bc1	972090	G	A	missense_variant	MODERATE	CDR20291_0799	modB	31G>A	Val11Ile	5.77%
20	Bc1	983711	A	G	intergenic_variant						7.18%
20	Bc1	983758	A	G	intergenic_variant						5.49%
20	Bc1	983860	A	T	intergenic_variant						12.31%
20	Bc1	984122	G	A	missense_variant	MODERATE	CDR20291_0809	tlpB	4G>A	Ala2Thr	5%
20	Bc1	984123	C	T	missense_variant	MODERATE	CDR20291_0809	tlpB	5C>T	Ala2Val	6.53%
20	Bc1	984172	G	A	synonymous_variant	LOW	CDR20291_0809	tlpB	54G>A	Gln18Gln	6.49%
20	Bc1	984530	G	A	missense_variant	MODERATE	CDR20291_0809	tlpB	412G>A	Val138Ile	7.23%
20	Bc1	984545	A	G	missense_variant	MODERATE	CDR20291_0809	tlpB	427A>G	Ile143Val	9.44%
20	Bc1	984551	C	A	missense_variant	MODERATE	CDR20291_0809	tlpB	433C>A	Gln145Lys	5.51%
20	Bc1	984655	C	T	synonymous_variant	LOW	CDR20291_0809	tlpB	537C>T	Asn179Asn	6.38%
20	Bc1	984787	A	G	synonymous_variant	LOW	CDR20291_0809	tlpB	669A>G	Lys223Lys	8.42%
20	Bc1	984832	A	G	synonymous_variant	LOW	CDR20291_0809	tlpB	714A>G	Ala238Ala	6.25%
20	Bc1	984907	C	T	synonymous_variant	LOW	CDR20291_0809	tlpB	789C>T	Asn263Asn	7.57%
20	Bc1	984951	G	A	missense_variant	MODERATE	CDR20291_0809	tlpB	833G>A	Arg278Lys	5.34%
20	Bc1	985031	C	T	missense_variant	MODERATE	CDR20291_0809	tlpB	913C>T	His305Tyr	6.34%
20	Bc1	985161	T	C	missense_variant	MODERATE	CDR20291_0809	tlpB	1043T>C	Ile348Thr	6.09%
20	Bc1	985222	G	A	synonymous_variant	LOW	CDR20291_0809	tlpB	1104G>A	Thr368Thr	5.80%
20	Bc1	985296	G	T	intergenic_variant						14.29%
20	Bc1	1005720	C	A	intron_variant	MODIFIER	CDR20291_0825	CDR20291_0825	348+19C>A		5.88%

20	Bc1	1005917	C	T	intron_variant	MODIFIER	CDR20291_0825	CDR20291_0825	348+216C>T		5.65%
20	Bc1	1005952	A	G	intron_variant	MODIFIER	CDR20291_0825	CDR20291_0825	348+251A>G		10.14%
20	Bc1	1006008	T	C	intron_variant	MODIFIER	CDR20291_0825	CDR20291_0825	348+307T>C		6.77%
20	Bc1	1006012	C	T	intron_variant	MODIFIER	CDR20291_0825	CDR20291_0825	348+311C>T		7.91%
20	Bc1	1006176	A	T	intron_variant	MODIFIER	CDR20291_0825	CDR20291_0825	348+475A>T		7.53%
20	Bc1	1006447	A	G	synonymous_variant	LOW	CDR20291_0826	tlpB	105A>G	Lys35Lys	9.81%
20	Bc1	1006555	G	A	synonymous_variant	LOW	CDR20291_0826	tlpB	213G>A	Lys71Lys	5.83%
20	Bc1	1006754	G	A	missense_variant	MODERATE	CDR20291_0826	tlpB	412G>A	Val138Ile	8.41%
20	Bc1	1006769	A	G	missense_variant	MODERATE	CDR20291_0826	tlpB	427A>G	Ile143Val	10.65%
20	Bc1	1006775	C	A	missense_variant	MODERATE	CDR20291_0826	tlpB	433C>A	Gln145Lys	6.13%
20	Bc1	1006810	G	A	synonymous_variant	LOW	CDR20291_0826	tlpB	468G>A	Glu156Glu	5.29%
20	Bc1	1006879	C	T	synonymous_variant	LOW	CDR20291_0826	tlpB	537C>T	Asn179Asn	6.82%
20	Bc1	1006965	G	A	missense_variant	MODERATE	CDR20291_0826	tlpB	623G>A	Arg208Lys	5.37%
20	Bc1	1007032	G	A	synonymous_variant	LOW	CDR20291_0826	tlpB	690G>A	Arg230Arg	8.37%
20	Bc1	1007057	G	A	missense_variant	MODERATE	CDR20291_0826	tlpB	715G>A	Gly239Arg	5.43%
20	Bc1	1007077	C	T	synonymous_variant	LOW	CDR20291_0826	tlpB	735C>T	Ala245Ala	5.53%
20	Bc1	1007175	A	G	missense_variant	MODERATE	CDR20291_0826	tlpB	833A>G	Lys278Arg	6%
20	Bc1	1007209	G	A	synonymous_variant	LOW	CDR20291_0826	tlpB	867G>A	Val289Val	5.60%
20	Bc1	1007338	T	G	synonymous_variant	LOW	CDR20291_0826	tlpB	996T>G	Val332Val	7.66%
20	Bc1	1007463	G	T	intron_variant	MODIFIER	CDR20291_0825	CDR20291_0825	349-63G>T		8.06%
20	Bc1	1007520	G	T	splice_region_variant&intron_variant	LOW	CDR20291_0825	CDR20291_0825	349-6G>T		8.67%
20	Bc1	1007568	A	G	missense_variant	MODERATE	CDR20291_0825	CDR20291_0825	391A>G	Arg131Gly	12.94%
20	Bc1	1010915	T	G	intergenic_variant						5.84%

20	Bc1	1011424	G	A	missense_variant	MODERATE	CDR20291_0829	dinB	233G>A	Arg78Lys	5.22%
20	Bc1	1016490	A	T	intergenic_variant						6.32%
20	Bc1	1016493	A	T	intergenic_variant						7.53%
20	Bc1	1021276	A	C	missense_variant	MODERATE	CDR20291_0837	CDR20291_0837	367A>C	Ile123Leu	5.41%
20	Bc1	1021278	A	C	synonymous_variant	LOW	CDR20291_0837	CDR20291_0837	369A>C	Ile123Ile	5.56%
20	Bc1	1025083	C	T	synonymous_variant	LOW	CDR20291_0840	CDR20291_0840	501C>T	Asn167Asn	5.88%
20	Bc1	1035954	T	A	intergenic_variant						9.72%
20	Bc1	1080772	A	C	missense_variant	MODERATE	CDR20291_0885	CDR20291_0885	1439A>C	Asn480Thr	6.80%
20	Bc1	1129543	G	C	intergenic_variant						14.89%
20	Bc1	1129545	AT	A	intergenic_variant						5.15%
20	Bc1	1131656	C	T	missense_variant	MODERATE	CDR20291_0915	thlA1	1142C>T	Thr381Ile	7.26%
20	Bc1	1131658	C	T	missense_variant	MODERATE	CDR20291_0915	thlA1	1144C>T	Leu382Phe	7.03%
20	Bc1	1161563	T	TAA	frameshift_variant&stop_gained	HIGH	CDR20291_0948	CDR20291_0948	509_510insAA	Tyr170fs	5.26%
20	Bc1	1173874	T	C	intergenic_variant						6.21%
20	Bc1	1173925	T	C	intergenic_variant						5.10%
20	Bc1	1174195	A	G	intergenic_variant						6.21%
20	Bc1	1174597	C	T	missense_variant	MODERATE	CDR20291_0959	CDR20291_0959	5C>T	Ala2Val	8.49%
20	Bc1	1174646	A	G	synonymous_variant	LOW	CDR20291_0959	CDR20291_0959	54A>G	Gln18Gln	7.69%
20	Bc1	1174805	G	A	synonymous_variant	LOW	CDR20291_0959	CDR20291_0959	213G>A	Lys71Lys	6.63%
20	Bc1	1174971	T	C	missense_variant	MODERATE	CDR20291_0959	CDR20291_0959	379T>C	Tyr127His	9.25%
20	Bc1	1175208	T	C	missense_variant	MODERATE	CDR20291_0959	CDR20291_0959	616T>C	Phe206Leu	7%
20	Bc1	1175307	G	A	missense_variant	MODERATE	CDR20291_0959	CDR20291_0959	715G>A	Gly239Arg	8.53%
20	Bc1	1175381	T	C	synonymous_variant	LOW	CDR20291_0959	CDR20291_0959	789T>C	Asn263Asn	5.04%
20	Bc1	1175479	A	G	missense_variant	MODERATE	CDR20291_0959	CDR20291_0959	887A>G	His296Arg	16.11%

20	Bc1	1175487	A	G	missense_variant	MODERATE	CDR20291_0959	CDR20291_0959	895A>G	Lys299Glu	14.17%
20	Bc1	1175788	G	A	intergenic_variant						7.34%
20	Bc1	1227805	T	A	missense_variant	MODERATE	CDR20291_1009	etfA3	1207T>A	Ser403Thr	17.89%
20	Bc1	1232412	T	A	missense_variant	MODERATE	CDR20291_1012	ackA	1186T>A	Leu396Met	5.49%
20	Bc1	1287073	C	T	intergenic_variant						6.19%
20	Bc1	1287074	A	G	intergenic_variant						6.14%
20	Bc1	1287083	T	A	intergenic_variant						11.76%
20	Bc1	1287092	C	T	intergenic_variant						11.97%
20	Bc1	1287093	A	G	intergenic_variant						11.86%
20	Bc1	1287098	A	G	intergenic_variant						5.22%
20	Bc1	1295582	A	G	intergenic_variant						5.80%
20	Bc1	1310504	C	T	missense_variant	MODERATE	CDR20291_1089	smc	1820C>T	Thr607Ile	8.18%
20	Bc1	1351868	A	T	intergenic_variant						5.11%
20	Bc1	1351881	A	T	intergenic_variant						5.30%
20	Bc1	1351884	A	T	intergenic_variant						5.34%
20	Bc1	1353599	T	C	intergenic_variant						9.84%
20	Bc1	1356356	T	C	intergenic_variant						6.19%
20	Bc1	1356361	A	G	intergenic_variant						6.80%
20	Bc1	1356362	G	A	intergenic_variant						7.14%
20	Bc1	1437198	A	T	missense_variant	MODERATE	CDR20291_1210	CDR20291_1210	1798A>T	Ile600Phe	6.45%
20	Bc1	1442176	C	T	missense_variant	MODERATE	CDR20291_1216	CDR20291_1216	251C>T	Thr84Ile	6.42%
20	Bc1	1467913	A	T	missense_variant	MODERATE	CDR20291_1237	CDR20291_1237	1795A>T	Asn599Tyr	5.38%
20	Bc1	1473217	T	C	intergenic_variant						13.51%
20	Bc1	1473218	G	A	intergenic_variant						13.25%

20	Bc1	1501285	G	T	intergenic_variant						5.79%
20	Bc1	1501287	T	G	intergenic_variant						5.79%
20	Bc1	1513978	T	C	intergenic_variant						60.14%
20	Bc1	1515416	A	C	missense_variant	MODERATE	CDR20291_1277	CDR20291_1277	18T>G	Asp6Glu	8.89%
20	Bc1	1521347	G	T	intergenic_variant						6.20%
20	Bc1	1534958	G	C	missense_variant	MODERATE	CDR20291_1293	CDR20291_1293	615G>C	Leu205Phe	5.26%
20	Bc1	1556625	G	A	intergenic_variant						5.04%
20	Bc1	1563467	A	T	stop_gained	HIGH	CDR20291_1320	CDR20291_1320	776T>A	Leu259*	6.96%
20	Bc1	1570564	C	T	intergenic_variant						5.31%
20	Bc1	1574763	A	T	missense_variant	MODERATE	CDR20291_1330	CDR20291_1330	661A>T	Thr221Ser	5.56%
20	Bc1	1577010	A	T	synonymous_variant	LOW	CDR20291_1333	ssuA	585A>T	Pro195Pro	5.56%
20	Bc1	1580623	T	A	intergenic_variant						9.84%
20	Bc1	1580664	A	G	intergenic_variant						12.26%
20	Bc1	1580665	C	A	intergenic_variant						12.90%
20	Bc1	1580885	T	C	intergenic_variant						7.55%
20	Bc1	1580889	C	T	intergenic_variant						8.59%
20	Bc1	1580992	C	G	intergenic_variant						11.93%
20	Bc1	1581111	A	C	intergenic_variant						7.79%
20	Bc1	1581467	G	A	missense_variant	MODERATE	CDR20291_1338	CDR20291_1338	248G>A	Arg83Lys	5.38%
20	Bc1	1581509	T	C	missense_variant	MODERATE	CDR20291_1338	CDR20291_1338	290T>C	Val97Ala	5.26%
20	Bc1	1581652	A	C	missense_variant	MODERATE	CDR20291_1338	CDR20291_1338	433A>C	Lys145Gln	6.15%
20	Bc1	1581657	G	T	synonymous_variant	LOW	CDR20291_1338	CDR20291_1338	438G>T	Gly146Gly	6.57%
20	Bc1	1581835	A	C	missense_variant	MODERATE	CDR20291_1338	CDR20291_1338	616A>C	Ile206Leu	7.55%
20	Bc1	1582061	G	A	missense_variant	MODERATE	CDR20291_1338	CDR20291_1338	842G>A	Arg281Lys	5.95%

20	Bc1	1582141	A	C	missense_variant	MODERATE	CDR20291_1338	CDR20291_1338	922A>C	Lys308Gln	8.82%
20	Bc1	1582267	A	G	missense_variant	MODERATE	CDR20291_1338	CDR20291_1338	1048A>G	Asn350Asp	12.99%
20	Bc1	1582397	T	G	intergenic_variant						5.10%
20	Bc1	1595288	G	T	missense_variant	MODERATE	CDR20291_1349	CDR20291_1349	284G>T	Cys95Phe	8.33%
20	Bc1	1606429	A	C	intergenic_variant						8.73%
20	Bc1	1606430	T	C	intergenic_variant						8.46%
20	Bc1	1606437	G	A	intergenic_variant						8.94%
20	Bc1	1606438	G	T	intergenic_variant						8.33%
20	Bc1	1643110	C	T	intergenic_variant						5.07%
20	Bc1	1643112	T	A	intergenic_variant						5.34%
20	Bc1	1655103	A	T	missense_variant	MODERATE	CDR20291_1399	hisB	115A>T	Met39Leu	5.26%
20	Bc1	1669976	G	A	missense_variant	MODERATE	CDR20291_1414	ilvB	976G>A	Val326Ile	5.61%
20	Bc1	1694173	A	T	missense_variant	MODERATE	CDR20291_1433	CDR20291_1433	790A>T	Ile264Leu	8.33%
20	Bc1	1722891	A	G	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	246A>G	Glu82Glu	6.14%
20	Bc1	1722892	C	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	247C>T	Leu83Leu	6.03%
20	Bc1	1722898	C	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	253C>T	Leu85Leu	5.22%
20	Bc1	1722910	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	265G>A	Asp89Asn	6.03%
20	Bc1	1722912	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	267T>C	Asp89Asp	6.14%
20	Bc1	1722921	A	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	276A>T	Gly92Gly	5.83%
20	Bc1	1722936	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	291T>C	Val97Val	5.88%
20	Bc1	1722957	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	312T>C	Gly104Gly	6.25%
20	Bc1	1722982	A	G	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	337A>G	Asn113Asp	5.56%
20	Bc1	1722987	G	A	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	342G>A	Lys114Lys	5.38%
20	Bc1	1723004	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	359G>A	Arg120Lys	5.38%

20	Bc1	1723012	A	G	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	367A>G	Arg123Gly	5.17%
20	Bc1	1723014	A	T	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	369A>T	Arg123Ser	5.08%
20	Bc1	1723030	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	385G>A	Gly129Arg	5.26%
20	Bc1	1723056	T	G	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	411T>G	Asn137Lys	8.33%
20	Bc1	1723110	G	T	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	465G>T	Glu155Asp	8.79%
20	Bc1	1731454	T	C	intron_variant	MODIFIER	CDR20291_1468	CDR20291_1468	537+118T>C		9.55%
20	Bc1	1731640	T	C	intron_variant	MODIFIER	CDR20291_1468	CDR20291_1468	537+304T>C		8%
20	Bc1	1731644	T	C	intron_variant	MODIFIER	CDR20291_1468	CDR20291_1468	537+308T>C		6.49%
20	Bc1	1731981	A	C	intron_variant	MODIFIER	CDR20291_1468	CDR20291_1468	537+645A>C		5.13%
20	Bc1	1732072	C	T	intron_variant	MODIFIER	CDR20291_1468	CDR20291_1468	537+736C>T		5.59%
20	Bc1	1732121	A	G	intron_variant	MODIFIER	CDR20291_1468	CDR20291_1468	537+785A>G		6.25%
20	Bc1	1732499	C	A	missense_variant	MODERATE	CDR20291_1469	CDR20291_1469	70C>A	Gln24Lys	10.27%
20	Bc1	1732609	T	G	missense_variant	MODERATE	CDR20291_1469	CDR20291_1469	180T>G	His60Gln	8.81%
20	Bc1	1732696	G	A	synonymous_variant	LOW	CDR20291_1469	CDR20291_1469	267G>A	Leu89Leu	7.21%
20	Bc1	1732899	G	A	missense_variant	MODERATE	CDR20291_1469	CDR20291_1469	470G>A	Arg157Lys	5.67%
20	Bc1	1732945	G	A	synonymous_variant	LOW	CDR20291_1469	CDR20291_1469	516G>A	Glu172Glu	5.03%
20	Bc1	1732988	C	A	missense_variant	MODERATE	CDR20291_1469	CDR20291_1469	559C>A	Gln187Lys	5.26%
20	Bc1	1733062	G	T	synonymous_variant	LOW	CDR20291_1469	CDR20291_1469	633G>T	Val211Val	5.04%
20	Bc1	1769977	T	C	intergenic_variant						6.73%
20	Bc1	1771808	C	T	missense_variant	MODERATE	CDR20291_1503	CDR20291_1503	337G>A	Gly113Ser	5.05%
20	Bc1	1782978	G	T	missense_variant	MODERATE	CDR20291_1513	CDR20291_1513	1502G>T	Trp501Leu	5.34%
20	Bc1	1791090	A	C	missense_variant	MODERATE	CDR20291_1520	CDR20291_1520	390A>C	Leu130Phe	7.32%
20	Bc1	1795475	G	T	stop_gained	HIGH	CDR20291_1523	vanS	334G>T	Glu112*	6.67%
20	Bc1	1795483	A	T	missense_variant	MODERATE	CDR20291_1523	vanS	342A>T	Lys114Asn	8.54%

20	Bc1	1801508	G	A	intergenic_variant						7.41%
20	Bc1	1804628	C	A	missense_variant	MODERATE	CDR20291_1530	asnB	1459C>A	Gln487Lys	7.22%
20	Bc1	1813462	A	C	missense_variant	MODERATE	CDR20291_1537	CDR20291_1537	253A>C	Thr85Pro	14.15%
20	Bc1	1819315	T	C	synonymous_variant	LOW	CDR20291_1542	CDR20291_1542	18A>G	Gly6Gly	5.17%
20	Bc1	1827059	A	C	missense_variant	MODERATE	CDR20291_1549	CDR20291_1549	519A>C	Arg173Ser	7.69%
20	Bc1	1859942	A	T	missense_variant	MODERATE	CDR20291_1570	CDR20291_1570	669A>T	Leu223Phe	5%
20	Bc1	1864416	C	CT	frameshift_variant	HIGH	CDR20291_1576	CDR20291_1576	18dupA	Val7fs	74.49%
20	Bc1	1877684	T	C	intergenic_variant						5.26%
20	Bc1	1909449	G	A	intergenic_variant						6.78%
20	Bc1	1921150	G	A	missense_variant	MODERATE	CDR20291_1635	CDR20291_1635	629G>A	Gly210Asp	7.08%
20	Bc1	1929322	T	G	missense_variant	MODERATE	CDR20291_1641	add	477T>G	Asn159Lys	10.53%
20	Bc1	1934580	T	A	synonymous_variant	LOW	CDR20291_1645	CDR20291_1645	1611T>A	Pro537Pro	5.49%
20	Bc1	1934840	T	A	missense_variant	MODERATE	CDR20291_1645	CDR20291_1645	1871T>A	Val624Asp	6.74%
20	Bc1	1934847	T	A	missense_variant	MODERATE	CDR20291_1645	CDR20291_1645	1878T>A	Asn626Lys	5.49%
20	Bc1	1937136	C	G	missense_variant	MODERATE	CDR20291_1647	CDR20291_1647	323C>G	Ser108Cys	15.93%
20	Bc1	1969452	A	T	missense_variant	MODERATE	CDR20291_1678	CDR20291_1678	311A>T	Lys104Ile	5.31%
20	Bc1	1983788	T	A	synonymous_variant	LOW	CDR20291_1692	CDR20291_1692	360A>T	Pro120Pro	6.08%
20	Bc1	1983794	T	C	missense_variant	MODERATE	CDR20291_1692	CDR20291_1692	354A>G	Ile118Met	7.02%
20	Bc1	1985222	C	T	intergenic_variant						43.39%
20	Bc1	1985223	A	G	intergenic_variant						42.19%
20	Bc1	1990279	G	T	missense_variant	MODERATE	CDR20291_1698	CDR20291_1698	1261G>T	Val421Leu	6.45%
20	Bc1	2027649	T	A	missense_variant	MODERATE	CDR20291_1730	aroC	68T>A	Ile23Lys	5.05%
20	Bc1	2139878	G	A	missense_variant	MODERATE	CDR20291_1826	CDR20291_1826	1280C>T	Ala427Val	6.12%
20	Bc1	2166177	A	C	missense_variant	MODERATE	CDR20291_1852	CDR20291_1852	553A>C	Asn185His	5.15%

20	Bc1	2180940	GA	G	frameshift_variant	HIGH	CDR20291_1869	CDR20291_1869	388delT	Ser130fs	8.89%
20	Bc1	2187207	A	C	synonymous_variant	LOW	CDR20291_1873	CDR20291_1873	307A>C	Arg103Arg	5.21%
20	Bc1	2197382	A	T	intergenic_variant						5.32%
20	Bc1	2209825	A	C	missense_variant	MODERATE	CDR20291_1893	CDR20291_1893	608T>G	Val203Gly	5.75%
20	Bc1	2213919	C	A	missense_variant	MODERATE	CDR20291_1896	CDR20291_1896	511G>T	Asp171Tyr	6.98%
20	Bc1	2224519	A	C	intergenic_variant						5.38%
20	Bc1	2225442	A	G	intergenic_variant						5.88%
20	Bc1	2225456	C	T	intergenic_variant						5.04%
20	Bc1	2225477	A	C	intergenic_variant						5.10%
20	Bc1	2225627	T	C	intergenic_variant						7.41%
20	Bc1	2225701	C	T	intergenic_variant						6.15%
20	Bc1	2226134	A	G	synonymous_variant	LOW	CDR20291_1906	CDR20291_1906	105A>G	Lys35Lys	6.58%
20	Bc1	2226462	A	C	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	433A>C	Lys145Gln	8.29%
20	Bc1	2226550	T	C	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	521T>C	Val174Ala	5.75%
20	Bc1	2226623	C	A	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	594C>A	Asp198Glu	8.89%
20	Bc1	2226683	C	A	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	654C>A	Ser218Arg	6.56%
20	Bc1	2226722	C	T	synonymous_variant	LOW	CDR20291_1906	CDR20291_1906	693C>T	Asn231Asn	8.26%
20	Bc1	2226868	A	G	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	839A>G	His280Arg	6.17%
20	Bc1	2226883	C	T	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	854C>T	Ser285Leu	6.36%
20	Bc1	2226926	G	A	synonymous_variant	LOW	CDR20291_1906	CDR20291_1906	897G>A	Glu299Glu	6.87%
20	Bc1	2226951	A	C	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	922A>C	Lys308Gln	8.80%
20	Bc1	2227036	G	A	missense_variant	MODERATE	CDR20291_1906	CDR20291_1906	1007G>A	Ser336Asn	8.96%
20	Bc1	2227150	G	T	intergenic_variant						7.56%
20	Bc1	2227207	T	G	intergenic_variant						10.19%

20	Bc1	2236906	T	C	intergenic_variant						5.65%
20	Bc1	2238006	G	A	intergenic_variant						5.76%
20	Bc1	2238064	G	A	intergenic_variant						10.43%
20	Bc1	2246454	A	G	intergenic_variant						8.04%
20	Bc1	2279772	T	A	intergenic_variant						5.95%
20	Bc1	2281106	A	C	intergenic_variant						5.22%
20	Bc1	2290771	A	T	intergenic_variant						19.83%
20	Bc1	2313864	C	T	missense_variant	MODERATE	CDR20291_1982	pbuX	388G>A	Gly130Ser	7.45%
20	Bc1	2345621	G	C	stop_gained	HIGH	CDR20291_2006	CDR20291_2006	1400C>G	Ser467*	7.53%
20	Bc1	2359662	A	C	missense_variant	MODERATE	CDR20291_2016	CDR20291_2016	1015A>C	Lys339Gln	7.81%
20	Bc1	2390681	T	G	missense_variant	MODERATE	CDR20291_2040	CDR20291_2040	1279A>C	Lys427Gln	6.90%
20	Bc1	2426589	A	G	synonymous_variant	LOW	CDR20291_2072	msrAB	684T>C	Ser228Ser	5.49%
20	Bc1	2493617	T	A	intergenic_variant						6.48%
20	Bc1	2530547	A	T	missense_variant	MODERATE	CDR20291_2155	CDR20291_2155	202T>A	Ser68Thr	5.15%
20	Bc1	2535190	TA	T	intergenic_variant						5.95%
20	Bc1	2535203	C	T	intergenic_variant						27.55%
20	Bc1	2562720	T	G	synonymous_variant	LOW	CDR20291_2178	CDR20291_2178	276A>C	Ser92Ser	6.54%
20	Bc1	2569168	C	A	missense_variant	MODERATE	CDR20291_2183	CDR20291_2183	377C>A	Thr126Asn	7.69%
20	Bc1	2570215	G	T	missense_variant	MODERATE	CDR20291_2184	CDR20291_2184	78G>T	Lys26Asn	5.71%
20	Bc1	2573580	T	C	missense_variant	MODERATE	CDR20291_2187	CDR20291_2187	1514A>G	Glu505Gly	5.43%
20	Bc1	2580795	C	A	intergenic_variant						8.14%
20	Bc1	2580797	A	C	intergenic_variant						8.43%
20	Bc1	2581742	C	T	initiator_codon_variant	LOW	CDR20291_2196	CDR20291_2196	3G>A	Met1?	6.14%
20	Bc1	2589250	T	C	synonymous_variant	LOW	CDR20291_2205	CDR20291_2205	516T>C	Leu172Leu	5.88%

20	Bc1	2604996	C	A	missense_variant	MODERATE	CDR20291_2219	xpt	384G>T	Leu128Phe	5.26%
20	Bc1	2606168	C	A	missense_variant	MODERATE	CDR20291_2220	mtlD	753G>T	Met251Ile	8.47%
20	Bc1	2635343	C	A	stop_gained	HIGH	CDR20291_2243	trxB3	613G>T	Glu205*	8.59%
20	Bc1	2636970	A	C	stop_lost&splice_region_variant	HIGH	CDR20291_2245	CDR20291_2245	1693T>G	Ter565Gluext*?	6.10%
20	Bc1	2642049	C	T	missense_variant	MODERATE	CDR20291_2249	CDR20291_2249	212G>A	Gly71Glu	9.68%
20	Bc1	2642056	A	T	missense_variant	MODERATE	CDR20291_2249	CDR20291_2249	205T>A	Leu69Met	11.59%
20	Bc1	2694883	A	T	synonymous_variant	LOW	CDR20291_2298	CDR20291_2298	1005A>T	Leu335Leu	6.38%
20	Bc1	2694888	C	G	missense_variant	MODERATE	CDR20291_2298	CDR20291_2298	1010C>G	Thr337Arg	5.15%
20	Bc1	2735565	A	G	intergenic_variant						6.32%
20	Bc1	2752476	G	A	missense_variant	MODERATE	CDR20291_2352	glcK	401C>T	Thr134Ile	6.11%
20	Bc1	2783023	C	A	missense_variant	MODERATE	CDR20291_2378	CDR20291_2378	19G>T	Val7Leu	5.88%
20	Bc1	2799408	A	G	intergenic_variant						15.52%
20	Bc1	2799435	A	G	intergenic_variant						7.52%
20	Bc1	2799436	T	A	intergenic_variant						6.87%
20	Bc1	2816244	C	A	missense_variant	MODERATE	CDR20291_2404	CDR20291_2404	172G>T	Val58Leu	5.04%
20	Bc1	2818449	C	CT	intergenic_variant						7.14%
20	Bc1	2929770	T	A	intergenic_variant						8.33%
20	Bc1	2932908	A	G	intergenic_variant						7.32%
20	Bc1	2932994	G	A	synonymous_variant	LOW	CDR20291_2501	CDR20291_2501	1068C>T	Ser356Ser	7.78%
20	Bc1	2933149	A	G	missense_variant	MODERATE	CDR20291_2501	CDR20291_2501	913T>C	Tyr305His	5.15%
20	Bc1	2933229	C	T	missense_variant	MODERATE	CDR20291_2501	CDR20291_2501	833G>A	Arg278Lys	5.02%
20	Bc1	2933273	A	G	synonymous_variant	LOW	CDR20291_2501	CDR20291_2501	789T>C	Asn263Asn	6.98%
20	Bc1	2933468	G	T	missense_variant	MODERATE	CDR20291_2501	CDR20291_2501	594C>A	Asp198Glu	5.33%
20	Bc1	2933624	A	C	synonymous_variant	LOW	CDR20291_2501	CDR20291_2501	438T>G	Gly146Gly	7.23%

20	Bc1	2933629	G	T	missense_variant	MODERATE	CDR20291_2501	CDR20291_2501	433C>A	Gln145Lys	6.87%
20	Bc1	2933683	A	G	missense_variant	MODERATE	CDR20291_2501	CDR20291_2501	379T>C	Tyr127His	9.68%
20	Bc1	2933849	T	C	synonymous_variant	LOW	CDR20291_2501	CDR20291_2501	213A>G	Lys71Lys	7.22%
20	Bc1	2933858	T	C	synonymous_variant	LOW	CDR20291_2501	CDR20291_2501	204A>G	Lys68Lys	6.51%
20	Bc1	2934228	T	A	intergenic_variant						5.08%
20	Bc1	2934452	T	C	intergenic_variant						7.79%
20	Bc1	2934487	G	A	intergenic_variant						5.46%
20	Bc1	2934688	G	A	intergenic_variant						5.48%
20	Bc1	2948017	T	A	intergenic_variant						5.59%
20	Bc1	2948019	A	T	intergenic_variant						5.48%
20	Bc1	2948021	A	T	intergenic_variant						5.56%
20	Bc1	3010906	C	G	intergenic_variant						11.64%
20	Bc1	3010909	A	C	intergenic_variant						11.03%
20	Bc1	3010915	A	G	intergenic_variant						13.07%
20	Bc1	3010962	A	G	intergenic_variant						6.04%
20	Bc1	3012532	A	G	intergenic_variant						5.71%
20	Bc1	3012534	G	T	intergenic_variant						5.23%
20	Bc1	3012535	C	T	intergenic_variant						5.19%
20	Bc1	3012542	C	T	intergenic_variant						7.98%
20	Bc1	3012589	C	T	intergenic_variant						21.26%
20	Bc1	3012595	G	T	intergenic_variant						19.16%
20	Bc1	3012598	C	G	intergenic_variant						18.82%
20	Bc1	3012606	C	T	intergenic_variant						17.90%
20	Bc1	3041001	G	A	intergenic_variant						5.61%

20	Bc1	3041002	T	C	intergenic_variant						5.50%
20	Bc1	3042477	A	T	intergenic_variant						8.86%
20	Bc1	3045334	A	G	missense_variant	MODERATE	CDR20291_2595	CDR20291_2595	617T>C	Val206Ala	6.14%
20	Bc1	3077294	T	C	missense_variant	MODERATE	CDR20291_2615	glyA	461A>G	Asp154Gly	5%
20	Bc1	3165995	G	A	missense_variant	MODERATE	CDR20291_2683	CDR20291_2683	269C>T	Ser90Leu	5.26%
20	Bc1	3200231	A	G	synonymous_variant	LOW	CDR20291_2707	CDR20291_2707	76T>C	Leu26Leu	6.96%
20	Bc1	3210197	A	G	synonymous_variant	LOW	CDR20291_2716	tlpB	1002T>C	Asp334Asp	6.04%
20	Bc1	3210203	A	C	synonymous_variant	LOW	CDR20291_2716	tlpB	996T>G	Val332Val	5.11%
20	Bc1	3210366	C	T	missense_variant	MODERATE	CDR20291_2716	tlpB	833G>A	Arg278Lys	5.63%
20	Bc1	3210436	A	G	synonymous_variant	LOW	CDR20291_2716	tlpB	763T>C	Leu255Leu	6.57%
20	Bc1	3210509	C	T	synonymous_variant	LOW	CDR20291_2716	tlpB	690G>A	Arg230Arg	6.76%
20	Bc1	3210662	G	A	synonymous_variant	LOW	CDR20291_2716	tlpB	537C>T	Asn179Asn	7.02%
20	Bc1	3210787	C	T	missense_variant	MODERATE	CDR20291_2716	tlpB	412G>A	Val138Ile	11.01%
20	Bc1	3210788	T	C	synonymous_variant	LOW	CDR20291_2716	tlpB	411A>G	Lys137Lys	12.46%
20	Bc1	3210986	C	T	synonymous_variant	LOW	CDR20291_2716	tlpB	213G>A	Lys71Lys	7.03%
20	Bc1	3211365	T	A	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+482A>T		5.88%
20	Bc1	3211467	C	T	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+380G>A		5.36%
20	Bc1	3211504	T	C	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+343A>G		5.56%
20	Bc1	3211529	G	A	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+318C>T		7.65%
20	Bc1	3211533	A	G	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+314T>C		6.17%
20	Bc1	3211589	T	C	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+258A>G		8.72%
20	Bc1	3211708	A	G	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+139T>C		5.56%
20	Bc1	3211825	A	G	intron_variant	MODIFIER	CDR20291_2715	CDR20291_2715	117+22T>C		8.15%

20	Bc1	3215114	T	A	intergenic_variant						5.33%
20	Bc1	3215125	G	A	intergenic_variant						5.92%
20	Bc1	3219672	G	A	missense_variant	MODERATE	CDR20291_2722	CDR20291_2722	655C>T	Pro219Ser	9.45%
20	Bc1	3338134	C	T	missense_variant	MODERATE	CDR20291_2819	CDR20291_2819	402G>A	Met134Ile	9.15%
20	Bc1	3349319	G	A	missense_variant	MODERATE	CDR20291_2830	nrdE	641C>T	Ala214Val	5.38%
20	Bc1	3352729	A	C	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	847-60T>G		5.08%
20	Bc1	3352756	C	T	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	847-87G>A		5.58%
20	Bc1	3352763	A	G	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	847-94T>C		6.48%
20	Bc1	3352859	T	C	missense_variant	MODERATE	CDR20291_2834	CDR20291_2834	1048A>G	Asn350Asp	12.56%
20	Bc1	3352985	T	G	missense_variant	MODERATE	CDR20291_2834	CDR20291_2834	922A>C	Lys308Gln	7.37%
20	Bc1	3352994	A	G	missense_variant	MODERATE	CDR20291_2834	CDR20291_2834	913T>C	Tyr305His	5.32%
20	Bc1	3353238	C	T	synonymous_variant	LOW	CDR20291_2834	CDR20291_2834	669G>A	Lys223Lys	8.21%
20	Bc1	3353364	A	C	missense_variant	MODERATE	CDR20291_2834	CDR20291_2834	543T>G	His181Gln	11.11%
20	Bc1	3353469	C	A	synonymous_variant	LOW	CDR20291_2834	CDR20291_2834	438G>T	Gly146Gly	6.61%
20	Bc1	3353474	T	G	missense_variant	MODERATE	CDR20291_2834	CDR20291_2834	433A>C	Lys145Gln	9.95%
20	Bc1	3353617	A	G	missense_variant	MODERATE	CDR20291_2834	CDR20291_2834	290T>C	Val97Ala	5%
20	Bc1	3353856	C	T	synonymous_variant	LOW	CDR20291_2834	CDR20291_2834	51G>A	Lys17Lys	5.71%
20	Bc1	3354130	T	A	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	846+507A>T		5.53%
20	Bc1	3354149	A	T	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	846+488T>A		5.26%
20	Bc1	3354165	A	T	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	846+472T>A		5.77%
20	Bc1	3354174	A	G	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	846+463T>C		7.69%
20	Bc1	3354333	G	A	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	846+304C>T		5.13%
20	Bc1	3354552	T	C	intron_variant	MODIFIER	CDR20291_2833	CDR20291_2833	846+85A>G		5.10%
20	Bc1	3403569	G	T	missense_variant	MODERATE	CDR20291_2874	CDR20291_2874	798C>A	Asn266Lys	6.96%

20	Bc1	3404820	C	A	missense_variant	MODERATE	CDR20291_2875	CDR20291_2875	362G>T	Arg121Ile	6.62%
20	Bc1	3415557	G	C	missense_variant	MODERATE	CDR20291_2884	CDR20291_2884	61C>G	Gln21Glu	7.63%
20	Bc1	3475836	A	C	intergenic_variant						6.55%
20	Bc1	3475837	G	T	intergenic_variant						6.55%
20	Bc1	3485026	T	A	missense_variant	MODERATE	CDR20291_2936	bglG1	43A>T	Thr15Ser	5.22%
20	Bc1	3500766	A	AACC	intergenic_variant						10.69%
20	Bc1	3501317	T	C	intron_variant	MODIFIER	CDR20291_2949	CDR20291_2949	186+274T>C		11.11%
20	Bc1	3501412	T	C	intron_variant	MODIFIER	CDR20291_2949	CDR20291_2949	186+369T>C		7.03%
20	Bc1	3501839	A	G	missense_variant	MODERATE	CDR20291_2950	CDR20291_2950	4A>G	Ile2Val	5.12%
20	Bc1	3502247	G	A	missense_variant	MODERATE	CDR20291_2950	CDR20291_2950	412G>A	Val138Ile	9.03%
20	Bc1	3502273	T	G	synonymous_variant	LOW	CDR20291_2950	CDR20291_2950	438T>G	Gly146Gly	8.80%
20	Bc1	3502549	G	A	synonymous_variant	LOW	CDR20291_2950	CDR20291_2950	714G>A	Ala238Ala	8.68%
20	Bc1	3502757	A	C	missense_variant	MODERATE	CDR20291_2950	CDR20291_2950	922A>C	Lys308Gln	8.26%
20	Bc1	3502768	A	G	synonymous_variant	LOW	CDR20291_2950	CDR20291_2950	933A>G	Lys311Lys	5.12%
20	Bc1	3502986	G	A	intron_variant	MODIFIER	CDR20291_2949	CDR20291_2949	187-116G>A		5.66%
20	Bc1	3502994	A	G	intron_variant	MODIFIER	CDR20291_2949	CDR20291_2949	187-108A>G		11.52%
20	Bc1	3525681	A	G	intergenic_variant						18%
20	Bc1	3525684	C	T	intergenic_variant						17.11%
20	Bc1	3571838	G	A	intergenic_variant						6.25%
20	Bc1	3594799	G	A	missense_variant	MODERATE	CDR20291_3013	CDR20291_3013	547G>A	Glu183Lys	5.73%
20	Bc1	3604313	G	A	intergenic_variant						5%
20	Bc1	3659736	T	A	intergenic_variant						6.85%
20	Bc1	3672567	C	A	missense_variant	MODERATE	CDR20291_3076	snorO	1946G>T	Arg649Ile	5.08%
20	Bc1	3752605	TG	T	frameshift_variant	HIGH	CDR20291_3143	pflD	1942delC	His648fs	5.19%

20	Bc1	3752893	C	A	missense_variant	MODERATE	CDR20291_3143	pflD	1655G>T	Gly552Val	7.55%
20	Bc1	3795225	T	G	missense_variant	MODERATE	CDR20291_3180	CDR20291_3180	233A>C	Asn78Thr	5.34%
20	Bc1	3814843	A	T	synonymous_variant	LOW	CDR20291_3199	CDR20291_3199	255T>A	Pro85Pro	5.84%
20	Bc1	3822248	A	T	stop_gained	HIGH	CDR20291_3206	CDR20291_3206	734T>A	Leu245*	5.04%
20	Bc1	3822253	A	G	synonymous_variant	LOW	CDR20291_3206	CDR20291_3206	729T>C	Phe243Phe	5.13%
20	Bc1	3823626	G	A	synonymous_variant	LOW	CDR20291_3207	CDR20291_3207	219C>T	Val73Val	9.70%
20	Bc1	3834355	CA	C	intergenic_variant						5.31%
20	Bc1	3852496	G	A	missense_variant	MODERATE	CDR20291_3229	hymC	706G>A	Ala236Thr	5.22%
20	Bc1	3873178	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	38731 78C>T		16.17%
20	Bc1	3873262	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	38732 62C>T		20%
20	Bc1	3923586	A	T	intergenic_variant						6.77%
20	Bc1	3944610	A	G	intergenic_variant						5.29%
20	Bc1	3949251	T	A	missense_variant	MODERATE	CDR20291_3316	rpiB2	155A>T	Asn52Ile	8.62%
20	Bc1	3984828	T	G	missense_variant	MODERATE	CDR20291_3349	CDR20291_3349	946A>C	Ile316Leu	5.19%
20	Bc1	3987198	A	G	intergenic_variant						9.93%
20	Bc1	4035246	C	A	missense_variant	MODERATE	CDR20291_3393	CDR20291_3393	410G>T	Gly137Val	8.79%
20	Bc1	4044396	A	G	intergenic_variant						8.11%
20	Bc1	4044397	C	T	intergenic_variant						8.41%
20	Bc1	4053137	G	A	synonymous_variant	LOW	CDR20291_3413	CDR20291_3413	888C>T	Ala296Ala	5.13%
20	Bc1	4085308	C	A	missense_variant	MODERATE	CDR20291_3437	dacS	714G>T	Glu238Asp	100%
20	Bc1	4088219	A	T	missense_variant	MODERATE	CDR20291_3440	CDR20291_3440	514A>T	Ile172Leu	7.45%
20	Bc1	4127895	A	G	intron_variant	MODIFIER	CDR20291_3477	CDR20291_3477	154-76T>C		6.63%
20	Bc1	4127926	A	C	intron_variant	MODIFIER	CDR20291_3477	CDR20291_3477	154-107T>G		7.50%

20	Bc1	4127960	A	G	intron_variant	MODIFIER	CDR20291_3477	CDR20291_3477	154-141T>C		8.47%
20	Bc1	4128104	T	C	missense_variant	MODERATE	CDR20291_3478	CDR20291_3478	1000A>G	Asn334Asp	6.73%
20	Bc1	4128435	C	T	synonymous_variant	LOW	CDR20291_3478	CDR20291_3478	669G>A	Lys223Lys	6.38%
20	Bc1	4128658	G	A	missense_variant	MODERATE	CDR20291_3478	CDR20291_3478	446C>T	Pro149Leu	11.19%
20	Bc1	4128666	A	C	synonymous_variant	LOW	CDR20291_3478	CDR20291_3478	438T>G	Gly146Gly	6.80%
20	Bc1	4128921	A	G	synonymous_variant	LOW	CDR20291_3478	CDR20291_3478	183T>C	Thr61Thr	5.82%
20	Bc1	4129050	C	T	synonymous_variant	LOW	CDR20291_3478	CDR20291_3478	54G>A	Gln18Gln	6.37%
20	Bc1	4129099	G	A	missense_variant	MODERATE	CDR20291_3478	CDR20291_3478	5C>T	Ala2Val	10.48%
20	Bc1	4129100	C	T	missense_variant	MODERATE	CDR20291_3478	CDR20291_3478	4G>A	Ala2Thr	5.29%
20	Bc1	4129288	T	C	intron_variant	MODIFIER	CDR20291_3477	CDR20291_3477	153+556A>G		8.65%
20	Bc1	4129586	C	T	intron_variant	MODIFIER	CDR20291_3477	CDR20291_3477	153+258G>A		5.10%
20	Bc1	4129621	A	G	intron_variant	MODIFIER	CDR20291_3477	CDR20291_3477	153+223T>C		7.88%
20	Bc1	4137572	A	T	missense_variant	MODERATE	CDR20291_3486	CDR20291_3486	526T>A	Phe176Ile	5.13%
20	Bc1	4146249	A	T	missense_variant	MODERATE	CDR20291_3494	CDR20291_3494	221T>A	Ile74Asn	8.04%
20	Bc1	4179179	T	A	missense_variant	MODERATE	CDR20291_3528	CDR20291_3528	278A>T	Tyr93Phe	5.26%
20	Bc1	4182390	G	A	missense_variant	MODERATE	CDR20291_3531	spo0J	50C>T	Ala17Val	7.14%
20	Bc1	4184259	C	T	intergenic_variant						11.11%
20	Bc1	4184262	A	G	intergenic_variant						10.53%
20	Bc10	1619675	TA	T	intergenic_variant						96.24%
20	Bc10	29869	A	T	intergenic_variant						17.14%
20	Bc10	30004	A	G	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	30004A>G		7.87%
20	Bc10	30006	G	A	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	30006G>A		15.29%
20	Bc10	30020	A	G	intergenic_variant						8.76%

20	Bc10	30037	T	C	intergenic_variant						8.53%
20	Bc10	31333	T	C	intergenic_variant						9.69%
20	Bc10	31600	T	G	intergenic_variant						23.21%
20	Bc10	31602	T	A	intergenic_variant						21.65%
20	Bc10	31680	C	T	intergenic_variant						67.82%
20	Bc10	32530	T	C	intergenic_variant						27.08%
20	Bc10	32998	A	AAAT	intergenic_variant						12.56%
20	Bc10	128090	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128090G>A		18.87%
20	Bc10	132468	C	T	intergenic_variant						16.67%
20	Bc10	133061	TA	T	intragenic_variant	MODIFIER	Gene_133028_134659	16S_rRNA	133062delA		6.40%
20	Bc10	147413	C	T	intragenic_variant	MODIFIER	Gene_144560_147512	23S_rRNA	147413C>T		7.81%
20	Bc10	147643	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147643C>T		5.32%
20	Bc10	147691	A	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147691A>G		8.10%
20	Bc10	147701	GT	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147702delT		7.77%
20	Bc10	166938	A	AG	intergenic_variant						42.99%
20	Bc10	251284	GA	G	frameshift_variant	HIGH	CDR20291_0197	CDR20291_0197	725delA	Lys242fs	8.84%
20	Bc10	528729	T	G	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3274T>G	Ser1092Ala	5.70%
20	Bc10	528738	G	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3283G>A	Ala1095Thr	6.91%
20	Bc10	528739	C	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	3284C>A	Ala1095Asp	8.94%
20	Bc10	535956	A	G	synonymous_variant	LOW	CDR20291_0445	CDR20291_0445	222A>G	Ser74Ser	27.60%
20	Bc10	572237	A	G	intergenic_variant						21.98%
20	Bc10	572241	A	G	intergenic_variant						18.37%

20	Bc10	581480	T	TA	intergenic_variant						75.44%
20	Bc10	706094	CA	C	intergenic_variant						6.84%
20	Bc10	825954	C	T	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	281G>A	Cys94Tyr	25.63%
20	Bc10	825995	T	C	synonymous_variant	LOW	CDR20291_0667	CDR20291_0667	240A>G	Lys80Lys	34.01%
20	Bc10	825997	T	G	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	238A>C	Lys80Gln	26.62%
20	Bc10	826072	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	163A>G	Lys55Glu	7.28%
20	Bc10	826104	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	131A>G	Tyr44Cys	9.88%
20	Bc10	826150	TA	T	frameshift_variant	HIGH	CDR20291_0667	CDR20291_0667	84delT	Phe28fs	15.85%
20	Bc10	1096069	C	T	synonymous_variant	LOW	CDR20291_0892	CDR20291_0892	273C>T	Tyr91Tyr	100%
20	Bc10	1096083	G	A	missense_variant	MODERATE	CDR20291_0892	CDR20291_0892	287G>A	Gly96Glu	100%
20	Bc10	1096085	G	T	missense_variant	MODERATE	CDR20291_0892	CDR20291_0892	289G>T	Ala97Ser	100%
20	Bc10	1199231	T	A	missense_variant	MODERATE	CDR20291_0982	mreB2	333T>A	Ser111Arg	100%
20	Bc10	1292784	A	AAT	intergenic_variant						5.74%
20	Bc10	1422217	GT	G	frameshift_variant	HIGH	CDR20291_1191	CDR20291_1191	426delT	Ile143fs	13.62%
20	Bc10	1437622	C	T	missense_variant	MODERATE	CDR20291_1210	CDR20291_1210	2222C>T	Ala741Val	13.22%
20	Bc10	1634267	GA	G	frameshift_variant	HIGH	CDR20291_1382	CDR20291_1382	1461delA	Glu488fs	8.19%
20	Bc10	1660946	TA	T	intergenic_variant						5.19%
20	Bc10	1722735	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	90T>C	Asn30Asn	5.77%
20	Bc10	1722777	G	A	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	132G>A	Glu44Glu	6.06%
20	Bc10	1722780	T	G	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	135T>G	Gly45Gly	5.86%
20	Bc10	1722788	C	T	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	143C>T	Ala48Val	5.76%
20	Bc10	2112906	C	T	missense_variant	MODERATE	CDR20291_1800	CDR20291_1800	820G>A	Val274Met	10.48%
20	Bc10	2233575	AT	A	intergenic_variant						7.73%
20	Bc10	2394636	C	CT	frameshift_variant	HIGH	CDR20291_2044	rrf	465dupA	Ala156fs	10.53%

20	Bc10	2480295	AT	A	frameshift_variant	HIGH	CDR20291_2114	CDR20291_2114	251delA	Asn84fs	49.37%
20	Bc10	2564518	T	C	synonymous_variant	LOW	CDR20291_2181	CDR20291_2181	2466A>G	Pro822Pro	16.44%
20	Bc10	2580872	G	A	intergenic_variant						7.60%
20	Bc10	2694070	T	C	synonymous_variant	LOW	CDR20291_2298	CDR20291_2298	192T>C	Ser64Ser	14.22%
20	Bc10	2748910	T	C	missense_variant	MODERATE	CDR20291_2349	CDR20291_2349	473T>C	Met158Thr	16.35%
20	Bc10	2767834	C	CA	intergenic_variant						14.76%
20	Bc10	2844154	T	C	synonymous_variant	LOW	CDR20291_2426	elaC	909T>C	Phe303Phe	10.43%
20	Bc10	3113632	AT	A	intergenic_variant						63.64%
20	Bc10	3233662	A	AT	intergenic_variant						12.69%
20	Bc10	3372748	T	C	missense_variant	MODERATE	CDR20291_2848	CDR20291_2848	407A>G	His136Arg	10.78%
20	Bc10	3459079	G	A	missense_variant	MODERATE	CDR20291_2918	ascB	53C>T	Ala18Val	9.13%
20	Bc10	3668919	AT	A	frameshift_variant	HIGH	CDR20291_3073	CDR20291_3073	289delA	Met97fs	6.87%
20	Bc10	3730065	A	AT	frameshift_variant	HIGH	CDR20291_3123	CDR20291_3123	33dupA	Phe12fs	16.83%
20	Bc10	3822970	AT	A	frameshift_variant	HIGH	CDR20291_3206	CDR20291_3206	11delA	Asn4fs	13.41%
20	Bc10	3846374	TA	T	intergenic_variant						9.78%
20	Bc10	3849112	A	AT	intergenic_variant						5.86%
20	Bc10	3872941	G	A	intergenic_variant						8.94%
20	Bc10	3873178	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873178C>T		11.40%
20	Bc10	3873262	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873262C>T		22.52%
20	Bc10	3873529	G	A	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873529G>A		47.06%
20	Bc10	4070161	CT	C	frameshift_variant	HIGH	CDR20291_3426	pyrAB1	374delA	Lys125fs	5.19%
20	Bc10	4086199	T	C	missense_variant	MODERATE	CDR20291_3438	dacR	532A>G	Thr178Ala	100%
20	Bc10	4115049	C	T	synonymous_variant	LOW	CDR20291_3465	CDR20291_3465	1305C>T	Asn435Asn	11.76%

20	Bc10	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	99.11%
20	Bc11	2235737	CT	C	frameshift_variant	HIGH	CDR20291_1913	CDR20291_1913	246delA	Val83fs	96.57%
20	Bc11	2999	A	G	intergenic_variant						6.95%
20	Bc11	29765	T	C	intragenic_variant	MODIFIER	Gene_26819_29769	23S_rRNA	29765T>C		8.39%
20	Bc11	29956	C	T	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29956C>T		6.21%
20	Bc11	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960T>C		24.81%
20	Bc11	30037	T	C	intergenic_variant						11.58%
20	Bc11	30592	A	C	intergenic_variant						5.26%
20	Bc11	32530	T	C	intergenic_variant						26.47%
20	Bc11	129810	C	T	intragenic_variant	MODIFIER	Gene_129785_132415	23S_rRNA	129810C>T		40.24%
20	Bc11	132697	T	C	intergenic_variant						29.89%
20	Bc11	142172	G	A	missense_variant	MODERATE	CDR20291_0106	nrdD	2348G>A	Arg783Lys	11.38%
20	Bc11	147606	G	A	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147606G>A		7.14%
20	Bc11	434598	A	G	missense_variant	MODERATE	CDR20291_0362	CDR20291_0362	266A>G	Gln89Arg	7.19%
20	Bc11	488533	A	G	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	913T>C	Tyr305His	25.97%
20	Bc11	489013	G	T	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	433C>A	Gln145Lys	33.33%
20	Bc11	489163	C	T	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	283G>A	Glu95Lys	41.70%
20	Bc11	489871	G	A	intergenic_variant						5.34%
20	Bc11	505422	GT	G	frameshift_variant	HIGH	CDR20291_0421	spaE	68delT	Phe23fs	91.04%
20	Bc11	527246	C	G	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	1791C>G	Asp597Glu	7.83%
20	Bc11	527253	A	G	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	1798A>G	Lys600Glu	7.93%
20	Bc11	570075	GT	G	frameshift_variant	HIGH	CDR20291_0473	CDR20291_0473	710delT	Leu237fs	7.69%
20	Bc11	603139	A	G	intron_variant	MODIFIER	CDR20291_0499	CDR20291_0499	432+154A>G		5.04%

20	Bc11	734846	C	CA	frameshift_variant	HIGH	CDR20291_0594	CDR20291_0594	54dupA	Phe19fs	37.16%
20	Bc11	816714	T	TA	frameshift_variant	HIGH	CDR20291_0657	CDR20291_0657	418dupA	Ile140fs	12.64%
20	Bc11	1096945	C	T	synonymous_variant	LOW	CDR20291_0892	CDR20291_0892	1149C>T	Tyr383Tyr	25.33%
20	Bc11	1268360	TG	T	frameshift_variant	HIGH	CDR20291_1052	spo0A	8delG	Gly3fs	8.76%
20	Bc11	1390370	GA	G	frameshift_variant	HIGH	CDR20291_1165	ftsK	26delA	Lys9fs	15.33%
20	Bc11	1470820	GA	G	frameshift_variant	HIGH	CDR20291_1240	CDR20291_1240	711delA	Lys237fs	8.64%
20	Bc11	1511820	AT	A	frameshift_variant	HIGH	CDR20291_1273	CDR20291_1273	311delA	Asn104fs	80%
20	Bc11	1523756	A	AT	intergenic_variant						7.59%
20	Bc11	1525095	GA	G	frameshift_variant	HIGH	CDR20291_1284	CDR20291_1284	828delA	Lys276fs	5.96%
20	Bc11	1571700	GA	G	frameshift_variant	HIGH	CDR20291_1328	feoB1	135delA	Gly46fs	37.86%
20	Bc11	1592813	A	T	intergenic_variant						100%
20	Bc11	1647600	GA	G	frameshift_variant	HIGH	CDR20291_1391	CDR20291_1391	147delA	Lys49fs	35.62%
20	Bc11	1711702	G	A	synonymous_variant	LOW	CDR20291_1449	CDR20291_1449	1833G>A	Lys611Lys	36.24%
20	Bc11	1722891	A	G	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	246A>G	Glu82Glu	5.56%
20	Bc11	1722892	C	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	247C>T	Leu83Leu	5.10%
20	Bc11	1722898	C	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	253C>T	Leu85Leu	5.03%
20	Bc11	1722910	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	265G>A	Asp89Asn	5.95%
20	Bc11	1722912	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	267T>C	Asp89Asp	5.43%
20	Bc11	1722921	A	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	276A>T	Gly92Gly	6.38%
20	Bc11	1722936	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	291T>C	Val97Val	6.81%
20	Bc11	1722957	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	312T>C	Gly104Gly	6.84%
20	Bc11	1722982	A	G	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	337A>G	Asn113Asp	5.59%
20	Bc11	1722987	G	A	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	342G>A	Lys114Lys	5.03%
20	Bc11	1723004	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	359G>A	Arg120Lys	5.92%

20	Bc11	1723012	A	G	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	367A>G	Arg123Gly	5.95%
20	Bc11	1723014	A	T	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	369A>T	Arg123Ser	5.95%
20	Bc11	1723030	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	385G>A	Gly129Arg	5.59%
20	Bc11	1742681	GA	G	frameshift_variant	HIGH	CDR20291_1476	CDR20291_1476	1173delA	Glu392fs	5.92%
20	Bc11	1798542	G	A	missense_variant	MODERATE	CDR20291_1526	vanT	137G>A	Arg46His	5.63%
20	Bc11	1798551	T	C	missense_variant	MODERATE	CDR20291_1526	vanT	146T>C	Val49Ala	13.10%
20	Bc11	1798764	A	G	missense_variant	MODERATE	CDR20291_1526	vanT	359A>G	His120Arg	37.91%
20	Bc11	1798989	C	T	missense_variant	MODERATE	CDR20291_1526	vanT	584C>T	Pro195Leu	46.78%
20	Bc11	1806760	A	G	intergenic_variant						77.84%
20	Bc11	1834373	TA	T	intergenic_variant						9.38%
20	Bc11	1864416	C	CT	frameshift_variant	HIGH	CDR20291_1576	CDR20291_1576	18dupA	Val7fs	86.33%
20	Bc11	1932986	CA	C	frameshift_variant	HIGH	CDR20291_1645	CDR20291_1645	24delA	Val9fs	34.30%
20	Bc11	2025318	TA	T	frameshift_variant	HIGH	CDR20291_1728	aroB	97delA	Ile33fs	69.92%
20	Bc11	2052210	A	G	missense_variant	MODERATE	CDR20291_1754	CDR20291_1754	86A>G	His29Arg	14.89%
20	Bc11	2578156	AT	A	intergenic_variant						95.51%
20	Bc11	2298110	A	AT	intergenic_variant						94.23%
20	Bc11	2361948	C	A	intergenic_variant						100%
20	Bc11	2361956	T	TA	intergenic_variant						92.98%
20	Bc11	2391896	T	TA	frameshift_variant	HIGH	CDR20291_2040	CDR20291_2040	63dupT	Ile22fs	8.28%
20	Bc11	2418630	AT	A	frameshift_variant	HIGH	CDR20291_2064	gabT	45delA	Lys15fs	8.59%
20	Bc11	2491603	T	C	synonymous_variant	LOW	CDR20291_2124	aroD	462A>G	Pro154Pro	77.78%
20	Bc11	2600073	A	G	missense_variant	MODERATE	CDR20291_2216	gatA	356T>C	Val119Ala	39.55%
20	Bc11	2664937	AT	A	frameshift_variant	HIGH	CDR20291_2272	CDR20291_2272	1338delA	Lys446fs	10.44%
20	Bc11	2680786	A	AT	intergenic_variant						93.33%

20	Bc11	2889250	C	T	missense_variant	MODERATE	CDR20291_2464	stk	1264G>A	Ala422Thr	6.34%
20	Bc11	3070068	T	C	missense_variant	MODERATE	CDR20291_2609	CDR20291_2609	878A>G	Asn293Ser	84.62%
20	Bc11	3073126	A	AT	frameshift_variant	HIGH	CDR20291_2611	CDR20291_2611	25dupA	Ile9fs	6.45%
20	Bc11	3075606	AT	A	frameshift_variant	HIGH	CDR20291_2614	CDR20291_2614	581delA	Asn194fs	80%
20	Bc11	3113632	AT	A	intergenic_variant						12.99%
20	Bc11	3120023	GT	G	frameshift_variant	HIGH	CDR20291_2649	CDR20291_2649	12delA	Lys4fs	70.69%
20	Bc11	3209465	TA	T	intergenic_variant						6.43%
20	Bc11	3301535	T	C	missense_variant	MODERATE	CDR20291_2788	ntpB	475A>G	Asn159Asp	8.33%
20	Bc11	3375243	AT	A	frameshift_variant	HIGH	CDR20291_2852	CDR20291_2852	1895delA	Asn632fs	44.72%
20	Bc11	3399917	T	C	intergenic_variant						100%
20	Bc11	3680127	ACTG	A	disruptive_inframe_deletion	MODERATE	CDR20291_3082	sdaB	887_889delCAG	Ala296del	84.93%
20	Bc11	3873529	G	A	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873529G>A		51.20%
20	Bc11	4084629	TA	T	intergenic_variant						6.88%
20	Bc11	4128000	T	C	synonymous_variant	LOW	CDR20291_3478	CDR20291_3478	1104A>G	Thr368Thr	63.30%
20	Bc11	3445638	G	GA	frameshift_variant	HIGH	CDR20291_2909	CDR20291_2909	2069dupT	Lys691fs	97.06%
20	Bc11	4190617	CT	C	frameshift_variant	HIGH	CDR20291_3541	rpmH	64delA	Arg22fs	96.69%
20	Bc2	581494	T	TA	intergenic_variant						96.84%
20	Bc2	31602	T	A	intergenic_variant						18.24%
20	Bc2	128972	C	T	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128972C>T		24.09%
20	Bc2	133061	TA	T	intragenic_variant	MODIFIER	Gene_133028_134659	16S_rRNA	133062delA		7.31%
20	Bc2	147594	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147594C>T		7.39%
20	Bc2	168011	G	A	missense_variant	MODERATE	CDR20291_0126	mreB1	604G>A	Gly202Arg	39.40%
20	Bc2	581487	T	TA	intergenic_variant						83.67%

20	Bc2	706094	CA	C	intergenic_variant						5.74%
20	Bc2	825947	C	T	synonymous_variant	LOW	CDR20291_0667	CDR20291_0667	288G>A	Leu96Leu	7.18%
20	Bc2	1974218	A	C	missense_variant	MODERATE	CDR20291_1681	CDR20291_1681	898A>C	Asn300His	6.33%
20	Bc2	3399845	A	T	intergenic_variant						37.47%
20	Bc2	3730727	C	A	missense_variant	MODERATE	CDR20291_3124	CDR20291_3124	692G>T	Arg231Leu	100%
20	Bc2	3808824	CT	C	frameshift_variant	HIGH	CDR20291_3193	bclA3	642delA	Ile214fs	87.02%
20	Bc2	4085224	T	A	missense_variant	MODERATE	CDR20291_3437	dacS	798A>T	Lys266Asn	100%
20	Bc2	3808827	T	TCTGC	frameshift_variant	HIGH	CDR20291_3193	bclA3	639_640insGCAG	Ile214fs	100%
20	Bc3	29956	C	T	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29956C>T		5.56%
20	Bc3	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960T>C		25.96%
20	Bc3	1384166	AGAAGAAACTTTAGCTCACTTT	A	disruptive_inframe_deletion	MODERATE	CDR20291_1159	comR	836_856delCTCACTTTGAAGAAACTTAG	Ala279_Leu285del	16.28%
20	Bc4	581494	T	TA	intergenic_variant						98.04%
20	Bc4	29502	C	T	intragenic_variant	MODIFIER	Gene_26819_29769	23S_rRNA	29502C>T		5.05%
20	Bc4	32442	T	TA	intergenic_variant						27.78%
20	Bc4	33003	G	A	intergenic_variant						5.63%
20	Bc4	128090	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128090G>A		18.75%
20	Bc4	128972	C	T	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128972C>T		18.60%
20	Bc4	129323	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	129323G>A		8.48%
20	Bc4	129810	C	T	intragenic_variant	MODIFIER	Gene_129785_132415	23S_rRNA	129810C>T		43.13%
20	Bc4	132468	C	T	intergenic_variant						15.19%
20	Bc4	132697	T	C	intergenic_variant						27.80%

20	Bc4	143974	G	A	intergenic_variant						25.26%
20	Bc4	147594	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	14759 4C>T		7.25%
20	Bc4	149615	C	G	missense_variant	MODERATE	CDR20291_0108	CDR20291_0108	173C>G	Ala58Gly	100%
20	Bc4	581480	T	TA	intergenic_variant						69.88%
20	Bc4	581487	T	TA	intergenic_variant						83.50%
20	Bc4	1384681	G	T	missense_variant	MODERATE	CDR20291_1159	comR	1337 G>T	Cys446 Phe	100%
20	Bc4	1798676	C	A	missense_variant	MODERATE	CDR20291_1526	vanT	271C>A	Pro91Thr	99.48%
20	Bc4	1974218	A	C	missense_variant	MODERATE	CDR20291_1681	CDR20291_1681	898A>C	Asn300 His	6.49%
20	Bc4	2973938	T	G	missense_variant	MODERATE	CDR20291_2539	murG	1173A >C	Glu391 Asp	5.08%
20	Bc4	3873178	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	38731 78C>T		6.67%
20	Bc5	147413	C	T	intragenic_variant	MODIFIER	Gene_144560_147512	23S_rRNA	14741 3C>T		5.44%
20	Bc5	706094	C	CA	intergenic_variant						5.62%
20	Bc5	826072	T	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	163A>G	Lys55Glu	5.59%
20	Bc5	973962	G	A	missense_variant	MODERATE	CDR20291_0801	maa	71G>A	Arg24Lys	53.75%
20	Bc5	1798752	G	A	missense_variant	MODERATE	CDR20291_1526	vanT	347G >A	Gly116 Glu	100%
20	Bc7	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960 T>C		28%
20	Bc7	31333	T	C	intergenic_variant						7.19%
20	Bc7	32285	C	T	intergenic_variant						7.23%
20	Bc7	225326	GA	G	frameshift_variant	HIGH	CDR20291_0174	CDR20291_0174	363delA	Lys121 fs	6.47%
20	Bc7	451877	T	C	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	564T>C	Ile188Ile	30.85%
20	Bc7	528261	T	G	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	2806T >G	Ser936 Ala	5%
20	Bc7	528270	G	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	2815 G>A	Ala939 Thr	6.48%
20	Bc7	528271	C	A	missense_variant	MODERATE	CDR20291_0440	CDR20291_0440	2816C >A	Ala939 Asp	5.98%

20	Bc7	826176	A	C	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	59T>G	Phe20Cys	5.26%
20	Bc7	826179	C	T	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	56G>A	Cys19Tyr	6.06%
20	Bc7	826183	T	A	missense_variant	MODERATE	CDR20291_0667	CDR20291_0667	52A>T	Asn18Tyr	5.36%
20	Bc7	1164180	GA	G	intergenic_variant						9.18%
20	Bc7	1201945	C	T	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	568C>T	Pro190Ser	8%
20	Bc7	1201973	C	T	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	596C>T	Thr199Ile	10.20%
20	Bc7	1202386	G	A	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	1009G>A	Ala337Thr	33.77%
20	Bc7	1203199	A	G	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	1822A>G	Thr608Ala	10.66%
20	Bc7	1203458	C	T	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	2081C>T	Ala694Val	20.66%
20	Bc7	1426206	C	T	missense_variant	MODERATE	CDR20291_1195	CDR20291_1195	211G>A	Val71Ile	16.67%
20	Bc7	1479264	GA	G	frameshift_variant	HIGH	CDR20291_1249	CDR20291_1249	489delA	Gly164fs	28.82%
20	Bc7	1625338	A	G	missense_variant	MODERATE	CDR20291_1375	pyrC	77A>G	Glu26Gly	10.34%
20	Bc7	1913676	A	G	intergenic_variant						10.67%
20	Bc7	2436981	AT	A	frameshift_variant	HIGH	CDR20291_2081	CDR20291_2081	636delA	Lys212fs	50.83%
20	Bc7	2771469	TA	T	frameshift_variant	HIGH	CDR20291_2368	CDR20291_2368	1213delT	Tyr405fs	23.32%
20	Bc7	3024436	CT	C	frameshift_variant	HIGH	CDR20291_2579	hpt	51delA	Val18fs	6.33%
20	Bc7	3336373	AC	A	intergenic_variant						93.24%
20	Bc7	3728936	GT	G	intergenic_variant						93.23%
20	Bc7	3750208	CT	C	frameshift_variant	HIGH	CDR20291_3141	CDR20291_3141	244delA	Ser82fs	32.11%
20	Bc7	3788478	T	C	missense_variant	MODERATE	CDR20291_3174	hydN1	451A>G	Thr151Ala	7.37%
20	Bc7	4047709	CA	C	intergenic_variant						6.04%
20	Bc7	4086199	T	C	missense_variant	MODERATE	CDR20291_3438	dacR	532A>G	Thr178Ala	97.87%
20	Bc7	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	97.35%
20	Bc8	29869	A	T	intergenic_variant						22.99%

20	Bc8	29956	C	T	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29956C>T		5.03%
20	Bc8	29988	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29988T>C		5.20%
20	Bc8	30037	T	C	intergenic_variant						7.48%
20	Bc8	31327	A	AG	intergenic_variant						5.97%
20	Bc8	31333	T	C	intergenic_variant						8.03%
20	Bc8	31600	T	G	intergenic_variant						17.83%
20	Bc8	31602	T	A	intergenic_variant						15.38%
20	Bc8	31621	T	G	intergenic_variant						19.86%
20	Bc8	32285	C	T	intergenic_variant						8.42%
20	Bc8	32442	T	TA	intergenic_variant						24.68%
20	Bc8	32530	T	C	intergenic_variant						16.81%
20	Bc8	129810	C	T	intragenic_variant	MODIFIER	Gene_129785_132415	23S_rRNA	129810C>T		50.46%
20	Bc8	132697	T	C	intergenic_variant						24.55%
20	Bc8	133061	TA	T	intragenic_variant	MODIFIER	Gene_133028_134659	16S_rRNA	133062delA		7.77%
20	Bc8	147594	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147594C>T		5.69%
20	Bc8	147619	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147619C>T		5.30%
20	Bc8	147643	C	T	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147643C>T		9.70%
20	Bc8	147691	A	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147691A>G		11.28%
20	Bc8	147701	GT	G	intragenic_variant	MODIFIER	Gene_147585_147701	5S_rRNA	147702delT		11.63%
20	Bc8	251284	GA	G	frameshift_variant	HIGH	CDR20291_0197	CDR20291_0197	725delA	Lys242fs	22.65%
20	Bc8	525201	C	A	intergenic_variant						6.43%
20	Bc8	525407	T	G	intergenic_variant						6.75%

20	Bc8	581480	T	TA	intergenic_variant						69.51%
20	Bc8	749566	T	C	missense_variant	MODERATE	CDR20291_0607	CDR20291_0607	461T>C	Val154Ala	100%
20	Bc8	864838	A	G	missense_variant	MODERATE	CDR20291_0697	CDR20291_0697	4A>G	Ile2Val	5.84%
20	Bc8	864839	T	C	missense_variant	MODERATE	CDR20291_0697	CDR20291_0697	5T>C	Ile2Thr	5.92%
20	Bc8	1141388	G	GAT	frameshift_variant	HIGH	CDR20291_0927	CDR20291_0927	282_283dupAT	Ser95fs	84.27%
20	Bc8	1156771	A	AG	intergenic_variant						40%
20	Bc8	1199044	C	T	missense_variant	MODERATE	CDR20291_0982	mreB2	146C>T	Ala49Val	100%
20	Bc8	1201871	T	A	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	494T>A	Ile165Lys	100%
20	Bc8	1205892	A	G	synonymous_variant	LOW	CDR20291_0987	divIVB	732A>G	Leu244Leu	100%
20	Bc8	1550844	TA	T	frameshift_variant	HIGH	CDR20291_1312	CDR20291_1312	267delA	Val90fs	5.71%
20	Bc8	1689708	C	CT	intergenic_variant						6.36%
20	Bc8	1864416	C	CT	frameshift_variant	HIGH	CDR20291_1576	CDR20291_1576	18dupA	Val7fs	90.40%
20	Bc8	1974218	A	C	missense_variant	MODERATE	CDR20291_1681	CDR20291_1681	898A>C	Asn300His	6%
20	Bc8	2025283	TA	T	frameshift_variant	HIGH	CDR20291_1728	aroB	62delA	Asn21fs	19.33%
20	Bc8	2167744	CA	C	intergenic_variant						7.59%
20	Bc8	2204957	A	G	synonymous_variant	LOW	CDR20291_1888	CDR20291_1888	615T>C	Gly205Gly	100%
20	Bc8	3362520	T	TA	frameshift_variant	HIGH	CDR20291_2841	CDR20291_2841	179dupA	Asn60fs	36.67%
20	Bc8	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	94.74%
20	Bc8	3873529	G	A	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873529G>A		51.43%
20	Bc8	4085474	A	G	missense_variant	MODERATE	CDR20291_3437	dacS	548T>C	Val183Ala	100%
20	Bc9	2578156	AT	A	intergenic_variant						97.60%
20	Bc9	2674743	CT	C	intergenic_variant						98.88%
20	Bc9	29988	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29988T>C		5.56%
20	Bc9	31333	T	C	intergenic_variant						5.95%

20	Bc9	31600	T	G	intergenic_variant						20.75%
20	Bc9	31621	T	G	intergenic_variant						23.16%
20	Bc9	38555	C	T	synonymous_variant	LOW	CDR20291_0011	CDR20291_0011	231C>T	Asp77Asp	12.94%
20	Bc9	308263	A	G	intergenic_variant						6.45%
20	Bc9	308268	T	A	intergenic_variant						6.59%
20	Bc9	308443	T	A	intergenic_variant						6.10%
20	Bc9	308448	C	T	intergenic_variant						6.17%
20	Bc9	525201	C	A	intergenic_variant						5.77%
20	Bc9	1199434	T	C	missense_variant	MODERATE	CDR20291_0982	mreB2	536T>C	Val179Ala	100%
20	Bc9	1307656	C	CA	frameshift_variant	HIGH	CDR20291_1088	CDR20291_1088	79dupA	Ile27fs	5.56%
20	Bc9	1356677	GA	G	frameshift_variant	HIGH	CDR20291_1130	CDR20291_1130	167delA	Asn56fs	10.71%
20	Bc9	1559763	T	C	missense_variant	MODERATE	CDR20291_1318	CDR20291_1318	1508A>G	Asp503Gly	14.61%
20	Bc9	1595965	AT	A	frameshift_variant	HIGH	CDR20291_1349	CDR20291_1349	969delT	Phe323fs	10.10%
20	Bc9	2361948	C	A	intergenic_variant						100%
20	Bc9	2361956	T	TA	intergenic_variant						91.30%
20	Bc9	2571555	AT	A	intergenic_variant						91.51%
20	Bc9	3039134	AT	A	intergenic_variant						99.06%
20	Bc9	2680786	A	AT	intergenic_variant						87.37%
20	Bc9	2693758	G	A	missense_variant	MODERATE	CDR20291_2297	CDR20291_2297	3048G>A	Met1016Ile	5.15%
20	Bc9	2772178	CT	C	splice_acceptor_variant&splice_donor_variant&intron_variant	HIGH	CDR20291_2368	CDR20291_2368	504+1delA		90.91%
20	Bc9	2791063	G	T	missense_variant	MODERATE	CDR20291_2385	CDR20291_2385	1478C>A	Ser493Tyr	100%
20	Bc9	2887646	AT	A	intergenic_variant						91.30%
20	Bc9	2946680	CA	C	frameshift_variant	HIGH	CDR20291_2512	CDR20291_2512	1209delT	Phe403fs	92.66%

20	Bc9	3077985	CA	C	intergenic_variant						88.14%
20	Bc9	3162097	CT	C	intergenic_variant						94.19%
20	Bc9	3292465	T	C	missense_variant	MODERATE	CDR20291_2781	CDR20291_2781	174A>G	Ile58Met	100%
20	Bc9	3361914	TA	T	intergenic_variant						90.09%
20	Bc9	3472928	G	A	intergenic_variant						100%
20	Bc9	3561410	G	A	synonymous_variant	LOW	CDR20291_2987	CDR20291_2987	688C>T	Leu230Leu	38.78%
20	Bc9	3280790	GA	G	frameshift_variant	HIGH	CDR20291_2769	CDR20291_2769	1661delA	Asn554fs	97.27%
20	Bc9	3846374	TA	T	intergenic_variant						11.32%
20	Bc9	3872941	G	A	intergenic_variant						15.84%
20	Bc9	4079340	GT	G	intergenic_variant						92.74%
20	Bc9	4085474	A	G	missense_variant	MODERATE	CDR20291_3437	dacS	548T>C	Val183Ala	100%
20	Bc9	4087103	T	C	missense_variant	MODERATE	CDR20291_3439	dacJ	478A>G	Thr160Ala	94.23%
20	Bc9	4146836	A	G	synonymous_variant	LOW	CDR20291_3496	CDR20291_3496	864T>C	Arg288Arg	100%
20	Bc9	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	95.88%
30	Bc1	1197338	ATATGGAATAATTTCCATATCTCTAATGAGTACATAAAACAGAGG	A	frameshift_variant&stop_lost&splice_region_variant	HIGH	CDR20291_0979	CDR20291_0979	146_*21delCCTCTGTTTTATGTACTCATTAGAGATATGGAAATTATTCCATA	Thr49fs	29.49%
30	Bc1	1513978	T	C	intergenic_variant						10.09%
30	Bc1	1795507	T	TGAACAACGTAAATTTGAGACACAAATGGCA	disruptive_inframe_insertion	MODERATE	CDR20291_1523	vanS	378_407dupATTTGAGACACAATGGCAGAACAACGTAA	Lys136_Asn137insPheGluThrGlnMetAlaGluGlnArgLys	14.74%
30	Bc1	4085308	C	A	missense_variant	MODERATE	CDR20291_3437	dacS	714G>T	Glu238Asp	100%

30	Bc10	581494	T	TA	intergenic_variant						98.99%
30	Bc10	31680	C	T	intergenic_variant						95.34%
30	Bc10	41468	C	T	stop_gained	HIGH	CDR20291_0013	CDR20291_0013	292C>T	Gln98*	47%
30	Bc10	143974	G	A	intergenic_variant						18.82%
30	Bc10	166938	A	AG	intergenic_variant						83.12%
30	Bc10	251284	GA	G	frameshift_variant	HIGH	CDR20291_0197	CDR20291_0197	725delA	Lys242fs	16.96%
30	Bc10	535956	A	G	synonymous_variant	LOW	CDR20291_0445	CDR20291_0445	222A>G	Ser74Ser	94.77%
30	Bc10	581487	T	TA	intergenic_variant						86.21%
30	Bc10	1096083	G	A	missense_variant	MODERATE	CDR20291_0892	CDR20291_0892	287G>A	Gly96Glu	100%
30	Bc10	1096085	G	T	missense_variant	MODERATE	CDR20291_0892	CDR20291_0892	289G>T	Ala97Ser	100%
30	Bc10	1199231	T	A	missense_variant	MODERATE	CDR20291_0982	mreB2	333T>A	Ser111Arg	100%
30	Bc10	2028051	TA	T	frameshift_variant	HIGH	CDR20291_1730	aroC	478delA	Ile160fs	67.38%
30	Bc10	2290551	AT	A	frameshift_variant	HIGH	CDR20291_1962	CDR20291_1962	110delA	Asn37fs	39.04%
30	Bc10	2480295	AT	A	frameshift_variant	HIGH	CDR20291_2114	CDR20291_2114	251delA	Asn84fs	89.29%
30	Bc10	2564518	T	C	synonymous_variant	LOW	CDR20291_2181	CDR20291_2181	2466A>G	Pro822Pro	94.53%
30	Bc10	2578256	T	TA	frameshift_variant	HIGH	CDR20291_2191	CDR20291_2191	15dupA	Gly6fs	39.31%
30	Bc10	2674915	CT	C	intergenic_variant						33.01%
30	Bc10	2694070	T	C	synonymous_variant	LOW	CDR20291_2298	CDR20291_2298	192T>C	Ser64Ser	93.69%
30	Bc10	2703106	T	C	synonymous_variant	LOW	CDR20291_2304	CDR20291_2304	27A>G	Ala9Ala	42.59%
30	Bc10	2748910	T	C	missense_variant	MODERATE	CDR20291_2349	CDR20291_2349	473T>C	Met158Thr	94.04%
30	Bc10	2767834	C	CA	intergenic_variant						87.89%
30	Bc10	2963094	AT	A	frameshift_variant	HIGH	CDR20291_2526	CDR20291_2526	625delA	Ile209fs	20.38%
30	Bc10	3113632	AT	A	intergenic_variant						91.64%
30	Bc10	3362738	GA	G	frameshift_variant	HIGH	CDR20291_2841	CDR20291_2841	400delA	Ile134fs	52.27%

30	Bc10	3730065	A	AT	frameshift_variant	HIGH	CDR20291_3123	CDR20291_3123	33dupA	Phe12fs	90.82%
30	Bc10	3822970	AT	A	frameshift_variant	HIGH	CDR20291_3206	CDR20291_3206	11delA	Asn4fs	87.47%
30	Bc10	3964363	AT	A	frameshift_variant	HIGH	CDR20291_3330	CDR20291_3330	512delA	Asn171fs	18.40%
30	Bc10	4086199	T	C	missense_variant	MODERATE	CDR20291_3438	dacR	532A>G	Thr178Ala	100%
30	Bc10	4115049	C	T	synonymous_variant	LOW	CDR20291_3465	CDR20291_3465	1305C>T	Asn435Asn	96.79%
30	Bc10	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	98.30%
30	Bc11	1932986	CA	C	frameshift_variant	HIGH	CDR20291_1645	CDR20291_1645	24delA	Val9fs	99.16%
30	Bc11	3075606	AT	A	frameshift_variant	HIGH	CDR20291_2614	CDR20291_2614	581delA	Asn194fs	95.42%
30	Bc11	3120023	GT	G	frameshift_variant	HIGH	CDR20291_2649	CDR20291_2649	12delA	Lys4fs	97.80%
30	Bc11	3375243	AT	A	frameshift_variant	HIGH	CDR20291_2852	CDR20291_2852	1895delA	Asn632fs	96.17%
30	Bc11	128972	C	T	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	12897 2C>T		21.78%
30	Bc11	143974	G	A	intergenic_variant						26.44%
30	Bc11	434673	T	A	missense_variant	MODERATE	CDR20291_0362	CDR20291_0362	341T>A	Ile114Asn	98.87%
30	Bc11	488533	A	G	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	913T>C	Tyr305His	32.39%
30	Bc11	489013	G	T	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	433C>A	Gln145Lys	46.87%
30	Bc11	489163	C	T	missense_variant	MODERATE	CDR20291_0405	CDR20291_0405	283G>A	Glu95Lys	51.34%
30	Bc11	581487	T	TA	intergenic_variant						85.37%
30	Bc11	734846	C	CA	frameshift_variant	HIGH	CDR20291_0594	CDR20291_0594	54dupA	Phe19fs	94.33%
30	Bc11	1511820	AT	A	frameshift_variant	HIGH	CDR20291_1273	CDR20291_1273	311delA	Asn104fs	93.11%
30	Bc11	1711702	G	A	synonymous_variant	LOW	CDR20291_1449	CDR20291_1449	1833G>A	Lys611Lys	100%
30	Bc11	1798764	A	G	missense_variant	MODERATE	CDR20291_1526	vanT	359A>G	His120Arg	100%
30	Bc11	1806760	A	G	intergenic_variant						100%
30	Bc11	2025318	TA	T	frameshift_variant	HIGH	CDR20291_1728	aroB	97delA	Ile33fs	94.61%
30	Bc11	2084450	G	A	missense_variant	MODERATE	CDR20291_1786	CDR20291_1786	250G>A	Gly84Arg	18.41%

30	Bc11	2408126	GT	G	frameshift_variant	HIGH	CDR20291_2055	CDR20291_2055	48delA	Lys16fs	68.20%
30	Bc11	2491603	T	C	synonymous_variant	LOW	CDR20291_2124	aroD	462A>G	Pro154Pro	100%
30	Bc11	2677841	CT	C	frameshift_variant	HIGH	CDR20291_2283	CDR20291_2283	298delA	Arg100fs	64.17%
30	Bc11	2678026	AT	A	frameshift_variant	HIGH	CDR20291_2283	CDR20291_2283	113delA	Asn38fs	28.97%
30	Bc11	3039561	AT	A	intergenic_variant						90.49%
30	Bc11	3070068	T	C	missense_variant	MODERATE	CDR20291_2609	CDR20291_2609	878A>G	Asn293Ser	100%
30	Bc11	3073126	A	AT	frameshift_variant	HIGH	CDR20291_2611	CDR20291_2611	25dupA	Ile9fs	20.06%
30	Bc11	3399917	T	C	intergenic_variant						100%
30	Bc11	3680127	ACTG	A	disruptive_inframe_deletion	MODERATE	CDR20291_3082	sdaB	887_889delCAG	Ala296del	92.06%
30	Bc11	3953261	AT	A	frameshift_variant	HIGH	CDR20291_3321	hemK	461delA	Asn154fs	22.63%
30	Bc11	3993945	CT	C	intergenic_variant						68.45%
30	Bc11	4047709	C	CA	intergenic_variant						87.66%
30	Bc11	4128000	T	C	synonymous_variant	LOW	CDR20291_3478	CDR20291_3478	1104A>G	Thr368Thr	99.51%
30	Bc11	4128108	C	A	synonymous_variant	LOW	CDR20291_3478	CDR20291_3478	996G>T	Val332Val	99.32%
30	Bc11	4129099	G	A	missense_variant	MODERATE	CDR20291_3478	CDR20291_3478	5C>T	Ala2Val	100%
30	Bc11	4190617	CT	C	frameshift_variant	HIGH	CDR20291_3541	rpmH	64delA	Arg22fs	93.55%
30	Bc2	31600	T	G	intergenic_variant						19.10%
30	Bc2	31602	T	A	intergenic_variant						17.60%
30	Bc2	31621	T	G	intergenic_variant						20.43%
30	Bc2	128972	C	T	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128972C>T		17.86%
30	Bc2	168011	G	A	missense_variant	MODERATE	CDR20291_0126	mreB1	604G>A	Gly202Arg	43.20%
30	Bc2	581487	T	TA	intergenic_variant						87.61%
30	Bc2	581494	T	TA	intergenic_variant						93.86%
30	Bc2	3399845	A	T	intergenic_variant						37.65%

30	Bc2	3730727	C	A	missense_variant	MODERATE	CDR20291_3124	CDR20291_3124	692G>T	Arg231Leu	100%
30	Bc2	3808824	CT	C	frameshift_variant	HIGH	CDR20291_3193	bclA3	642delA	Ile214fs	87.80%
30	Bc2	4085224	T	A	missense_variant	MODERATE	CDR20291_3437	dacS	798A>T	Lys266Asn	100%
30	Bc2	3808827	T	TCTGC	frameshift_variant	HIGH	CDR20291_3193	bclA3	639_640insGCAG	Ile214fs	99.18%
30	Bc3	29915	A	G	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29915A>G		28.57%
30	Bc3	1384166	AGAAGAAACTTTAGCTCACTTT	A	disruptive_inframe_deletion	MODERATE	CDR20291_1159	comR	836_856delCTCACTTTGAAGAAACTTAG	Ala279_Leu285del	92.31%
30	Bc3	1799093	G	A	missense_variant	MODERATE	CDR20291_1526	vanT	688G>A	Glu230Lys	100%
30	Bc3	3873178	C	T	intragenic_variant	MODIFIER	Gene_3873065_3874689	16S_rRNA	3873178C>T		40%
30	Bc4	581494	T	TA	intergenic_variant						98.02%
30	Bc4	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960T>C		23.98%
30	Bc4	31600	T	G	intergenic_variant						17.51%
30	Bc4	128090	G	A	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128090G>A		20.59%
30	Bc4	128972	C	T	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128972C>T		11.58%
30	Bc4	132468	C	T	intergenic_variant						14.09%
30	Bc4	143974	G	A	intergenic_variant						25.77%
30	Bc4	149615	C	G	missense_variant	MODERATE	CDR20291_0108	CDR20291_0108	173C>G	Ala58Gly	100%
30	Bc4	525201	C	A	intergenic_variant						27.56%
30	Bc4	525407	T	G	intergenic_variant						24.50%
30	Bc4	581480	T	TA	intergenic_variant						68.88%
30	Bc4	581487	T	TA	intergenic_variant						80%
30	Bc4	1384681	G	T	missense_variant	MODERATE	CDR20291_1159	comR	1337G>T	Cys446Phe	100%

30	Bc4	1798676	C	A	missense_variant	MODERATE	CDR20291_1526	vanT	271C>A	Pro91Thr	100%
30	Bc5	29960	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29960T>C		26.36%
30	Bc5	31600	T	G	intergenic_variant						22.49%
30	Bc5	31602	T	A	intergenic_variant						20.50%
30	Bc5	973962	G	A	missense_variant	MODERATE	CDR20291_0801	maa	71G>A	Arg24Lys	100%
30	Bc5	1722891	A	G	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	246A>G	Glu82Glu	11.11%
30	Bc5	1722892	C	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	247C>T	Leu83Leu	10.49%
30	Bc5	1722898	C	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	253C>T	Leu85Leu	11.54%
30	Bc5	1722910	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	265G>A	Asp89Asn	11.62%
30	Bc5	1722912	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	267T>C	Asp89Asp	11.58%
30	Bc5	1722921	A	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	276A>T	Gly92Gly	11.76%
30	Bc5	1798752	G	A	missense_variant	MODERATE	CDR20291_1526	vanT	347G>A	Gly116Glu	100%
30	Bc7	1479264	GA	G	frameshift_variant	HIGH	CDR20291_1249	CDR20291_1249	489delA	Gly164fs	95.61%
30	Bc7	2436981	AT	A	frameshift_variant	HIGH	CDR20291_2081	CDR20291_2081	636delA	Lys212fs	96.62%
30	Bc7	3024436	CT	C	frameshift_variant	HIGH	CDR20291_2579	hpt	51delA	Val18fs	95.99%
30	Bc7	3728936	GT	G	intergenic_variant						96.24%
30	Bc7	8347	T	C	intergenic_variant						99.75%
30	Bc7	128972	C	T	intragenic_variant	MODIFIER	Gene_127804_129434	16S_rRNA	128972C>T		14.17%
30	Bc7	143974	G	A	intergenic_variant						22.73%
30	Bc7	251284	GA	G	frameshift_variant	HIGH	CDR20291_0197	CDR20291_0197	725delA	Lys242fs	46.36%
30	Bc7	451877	T	C	synonymous_variant	LOW	CDR20291_0377	CDR20291_0377	564T>C	Ile188Ile	24.37%
30	Bc7	525201	C	A	intergenic_variant						22.02%
30	Bc7	525407	T	G	intergenic_variant						17.94%
30	Bc7	933007	C	CT	intergenic_variant						93.68%

30	Bc7	1093457	G	A	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	192G>A	Lys64Lys	100%
30	Bc7	1202386	G	A	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	1009G>A	Ala337Thr	99.48%
30	Bc7	1595613	T	TA	frameshift_variant	HIGH	CDR20291_1349	CDR20291_1349	616dupA	Thr206fs	43.15%
30	Bc7	1619675	TA	T	intergenic_variant						93.73%
30	Bc7	2771469	TA	T	frameshift_variant	HIGH	CDR20291_2368	CDR20291_2368	1213delT	Tyr405fs	91.14%
30	Bc7	2830047	T	C	synonymous_variant	LOW	CDR20291_2415	gbeA	1020T>C	His340His	49%
30	Bc7	3336373	AC	A	intergenic_variant						92.39%
30	Bc7	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	98.25%
30	Bc7	4086199	T	C	missense_variant	MODERATE	CDR20291_3438	dacR	532A>G	Thr178Ala	100%
30	Bc7	3750208	CT	C	frameshift_variant	HIGH	CDR20291_3141	CDR20291_3141	244delA	Ser82fs	95.25%
30	Bc8	864888	G	A	synonymous_variant	LOW	CDR20291_0697	CDR20291_0697	54G>A	Gln18Gln	11.11%
30	Bc8	1201871	T	A	missense_variant	MODERATE	CDR20291_0985	CDR20291_0985	494T>A	Ile165Lys	100%
30	Bc9	132923	C	CA	intergenic_variant						100%
30	Bc9	2571555	AT	A	intergenic_variant						95.65%
30	Bc9	2946680	CA	C	frameshift_variant	HIGH	CDR20291_2512	CDR20291_2512	1209delT	Phe403fs	95.16%
30	Bc9	3731131	CA	C	frameshift_variant	HIGH	CDR20291_3124	CDR20291_3124	287delT	Leu96fs	98.25%
30	Bc9	29988	T	C	intragenic_variant	MODIFIER	Gene_29898_30014	5S_rRNA	29988T>C		10.61%
30	Bc9	38555	C	T	synonymous_variant	LOW	CDR20291_0011	CDR20291_0011	231C>T	Asp77Asp	38.75%
30	Bc9	132914	N	G	intergenic_variant						100%
30	Bc9	132915	N	C	intergenic_variant						100%
30	Bc9	132939	G	T	intergenic_variant						100%
30	Bc9	132955	C	A	intergenic_variant						100%
30	Bc9	132958	T	G	intergenic_variant						100%
30	Bc9	132959	T	C	intergenic_variant						100%

30	Bc9	143507	N	A	intergenic_variant						100%
30	Bc9	143508	N	T	intergenic_variant						100%
30	Bc9	143974	G	A	intergenic_variant						42.11%
30	Bc9	525201	C	A	intergenic_variant						11.94%
30	Bc9	772099	GA	G	frameshift_variant	HIGH	CDR20291_0624	CDR20291_0624	324delA	Lys108fs	11.29%
30	Bc9	1093913	A	G	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	648A>G	Pro216Pro	100%
30	Bc9	1093916	A	G	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	651A>G	Leu217Leu	100%
30	Bc9	1093922	T	C	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	657T>C	Thr219Thr	100%
30	Bc9	1093958	T	C	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	693T>C	Asn231Asn	100%
30	Bc9	1093964	A	G	synonymous_variant	LOW	CDR20291_0891	CDR20291_0891	699A>G	Val233Val	100%
30	Bc9	1094468	T	G	missense_variant	MODERATE	CDR20291_0891	CDR20291_0891	1203T>G	Asp401Glu	100%
30	Bc9	1094475	T	A	missense_variant	MODERATE	CDR20291_0891	CDR20291_0891	1210T>A	Leu404Ile	100%
30	Bc9	1199434	T	C	missense_variant	MODERATE	CDR20291_0982	mreB2	536T>C	Val179Ala	100%
30	Bc9	1591582	AT	A	frameshift_variant	HIGH	CDR20291_1346	CDR20291_1346	1058delA	Asn353fs	23.88%
30	Bc9	1595965	AT	A	frameshift_variant	HIGH	CDR20291_1349	CDR20291_1349	969delT	Phe323fs	16.67%
30	Bc9	1722735	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	90T>C	Asn30Asn	10.34%
30	Bc9	1722804	T	C	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	159T>C	Cys53Cys	10.26%
30	Bc9	1722807	T	A	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	162T>A	Pro54Pro	10.13%
30	Bc9	1722828	A	G	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	183A>G	Lys61Lys	10.53%
30	Bc9	1722829	G	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	184G>A	Ala62Thr	10.67%
30	Bc9	1722830	C	A	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	185C>A	Ala62Glu	10.53%
30	Bc9	1722831	A	T	synonymous_variant	LOW	CDR20291_1463	CDR20291_1463	186A>T	Ala62Ala	10.96%
30	Bc9	1722841	A	T	missense_variant	MODERATE	CDR20291_1463	CDR20291_1463	196A>T	Thr66Ser	10.53%
30	Bc9	2027817	C	T	missense_variant	MODERATE	CDR20291_1730	aroC	236C>T	Thr79Met	41.79%

30	Bc9	2910395	A	AT	frameshift_variant	HIGH	CDR20291_2482	pyrR	452dupA	Asn151fs	27.42%
30	Bc9	4079340	GT	G	intergenic_variant						97.18%
30	Bc9	3561410	G	A	synonymous_variant	LOW	CDR20291_2987	CDR20291_2987	688C>T	Leu230Leu	46.15%
30	Bc9	3680965	G	A	missense_variant	MODERATE	CDR20291_3082	sdaB	52C>T	His18Tyr	24.05%
30	Bc9	4085474	A	G	missense_variant	MODERATE	CDR20291_3437	dacS	548T>C	Val183Ala	100%
30	Bc9	4087103	T	C	missense_variant	MODERATE	CDR20291_3439	dacJ	478A>G	Thr160Ala	100%
30	Bc9	4146836	A	G	synonymous_variant	LOW	CDR20291_3496	CDR20291_3496	864T>C	Arg288Arg	100%

Appendix VII – Nanopore Variants

Table A7 - Nanopore Variants

BAR CODE	POS	REF	ALT	TYPE	CONSEQUENCE	LOCUS TAG	GENE NAME	GENE FUNCTION	NT	AA
Bc1	1197338	ATATGGAATAATTTCCATATCTCTAATGAGTACATAAAACAGAGG	A	del	bidirectional_gene_fusion	CDR20291_0979	CDR20291_0979			
Bc1	1795507	T	TGAACAACGTAAATTTGAGACACAAATGGCA	ins	disruptive_inframe_insertion	CDR20291_1523	vanS	two-component sensor histidine kinase	378_407dupATTTGAGACACAAATGGCAGAACAACGTAA	Lys136_Asn137insPheGluThrGlnMetAlaGluGlnArgLys
Bc1	4085308	C	A	snp	missense_variant	CDR20291_3437	dacS	two-component sensor histidine kinase	714G>T	Glu238Asp
Bc2	967909	CA	C	del	frameshift_variant	CDR20291_0794	CDR20291_0794	conserved hypothetical protein	732delA	Glu245fs
Bc2	2182601	GT	G	del	frameshift_variant	CDR20291_1871	CDR20291_1871	putative ABC transporter, permease protein	2178delA	Lys726fs
Bc2	3730727	C	A	snp	missense_variant	CDR20291_3124	CDR20291_3124	sensor protein	692G>T	Arg231Leu
Bc2	3808825	TAT	ATCTGC	complex	missense_variant&conservative_inframe_insertion	CDR20291_3193	bclA3	putative exosporium glycoprotein	640_642delATAinsGCAGAT	Ile214delinsAlaAsp

Bc2	4085224	T	A	snp	missense_variant	CDR20291_3437	dacS	two-component sensor histidine kinase	798A>T	Lys266Asn
Bc3	1384166	AGAAGAAACTTTAGCTCACTTT	A	del	disruptive_inframe_deletion	CDR20291_1159	comR	polyribonucleotide nucleotidyltransferase	836_856delCTCACTTTGAAGAAACTTTAG	Ala279_Leu285del
Bc3	1799093	G	A	snp	missense_variant	CDR20291_1526	vanT	serine/alanine racemase	688G>A	Glu230Lys
Bc4	149615	C	G	snp	missense_variant	CDR20291_0108	CDR20291_0108	putative membrane protein	173C>G	Ala58Gly
Bc4	1384681	G	T	snp	missense_variant	CDR20291_1159	comR	polyribonucleotide nucleotidyltransferase	1337G>T	Cys446Phe
Bc4	1798676	C	A	snp	missense_variant	CDR20291_1526	vanT	serine/alanine racemase	271C>A	Pro91Thr
Bc5	973962	G	A	snp	missense_variant	CDR20291_0801	maa	maltose O-acetyltransferase	71G>A	Arg24Lys
Bc5	1798752	G	A	snp	missense_variant	CDR20291_1526	vanT	serine/alanine racemase	347G>A	Gly116Glu
Bc5	4190468	ATTACAAAAGGGCCGCTGCAAAGGTGGCCTTTTTGTGTAAATAATTACAATTCATAGCTGTGTTTATGGAGGTTAAC	A	del	intergenic_variant					