THE IMPACT OF GENOME DEFENSE SYSTEMS ON ANTIBIOTIC RESISTANCE DISSEMINATION IN THE OPPORTUNISTIC PATHOGEN *ENTEROCOCCUS FAECALIS*

by

Valerie Jane Price

APPROVED BY SUPERVISORY COMMITTEE:

Kelli L. Palmer, Chair

Stephen Spiro

Lawrence J. Reitzer

Duane D. Winkler

This dissertation is dedicated to my parents, Jack and Mary, and my brother, Jackie.

THE IMPACT OF GENOME DEFENSE SYSTEMS ON ANTIBIOTIC RESISTANCE DISSEMINATION IN THE OPPORTUNISTIC PATHOGEN *ENTEROCOCCUS FAECALIS*

by

VALERIE JANE PRICE, BS, MS

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN

MOLECULAR AND CELL BIOLOGY

THE UNIVERSITY OF TEXAS AT DALLAS

December 2017

ACKNOWLEDGMENTS

First, I would like to thank my family for always being supportive and encouraging. To my parents, I would not be who I am today without your continued love and understanding. To my brother, thank you for being such a positive role model in my life and for showing me how to be a life-long student. I think we are up to around 20 years of college combined, hopefully it will pay off for me as well as it has for you.

Next, I would like to acknowledge the faculty that have helped me reach this point. Dr. Kelli Palmer, you have been the best doctoral advisor anyone could ask for. I am grateful for being given the opportunity to learn from you. A special thank you goes to Dr. Santosh D'Mello. I would not be receiving a PhD if you had not given me the chance to try out research in your lab. Thank you for giving me that chance and for continuing to be an active part of my academic career. To my committee, thank you all for dedicating your time to ensure my success. Whether it was writing letters or providing reagents for my research, nothing went unnoticed.

To the friends I have made in the Palmer and D'Mello labs over the years at UT Dallas, thank you all for making this journey so memorable. I would especially like to thank Wenwen, Hannah, Pooja, Marinelle and Farry; I am so happy we all started this journey together in the Palmer lab and thank you for all of the help and support over the years.

Finally, to my best friend: Simon, I will be forever grateful for your patience, support, and encouraging words, not only during graduate school but in life as well.

October 2017

THE IMPACT OF GENOME DEFENSE SYSTEMS ON ANTIBIOTIC RESISTANCE DISSEMINATION IN THE OPPORTUNISTIC PATHOGEN *ENTEROCOCCUS FAECALIS*

Valerie Jane Price, PhD
The University of Texas at Dallas, 2017

Supervising Professor: Dr. Kelli L. Palmer

Since the advent of the modern antibiotic era there has been an increasing number of bacteria that are resistant to antibiotics. The growing antibiotic resistance and the lack in development of new antibiotics poses a serious public health threat worldwide. One of the first steps to combat this problem is to understand the molecular mechanisms associated with antibiotic resistance. *Enterococcus faecalis* is one of the most common causes of healthcare-associated infections (HAI). *E. faecalis* is an opportunistic pathogen that normally resides in the gastrointestinal (GI) tracts of humans and other animals. This bacterium possesses intrinsic antibiotic resistance and can also acquire resistance to other antibiotics through horizontal gene transfer (HGT). Some multi-drug resistant (MDR) *E. faecalis* strain have acquired resistance to vancomycin, an antibiotic of last resort, leaving very few treatment options. It has been shown that MDR *E. faecalis* have expanded genomes enriched with mobile genetic elements (MGEs) that were acquired through horizontal gene transfer (HGT). Plasmid conjugation is one of the most common forms of HGT in this species. The pheromone-responsive plasmids (PRPs) are a group

of narrow host range plasmids that are rapid disseminators of antibiotic resistance in clinical isolates of *E. faecalis*. PRPs have high conjugation frequencies and often encode virulence factors that have been associated with enhanced pathogenicity. The acquisition of foreign DNA is usually associated with a fitness cost to the host cell, therefore prokaryotes encode defense systems that limit HGT. The two forms of genome defense studied here are restriction-modification (R-M) and CRISPR-Cas. Interestingly, MDR *E. faecalis* lack active CRISPR-Cas systems, leading to the hypothesis that MDR *E. faecalis* emerge due to the lack of barriers to HGT. The primary focus of the research presented here was to establish a role for R-M and CRISPR-Cas in providing genome defense in *E. faecalis* and to elucidate the interactions between CRISPR-Cas and MGEs. Conjugation assays were used to determine that R-M and CRISPR-Cas work cooperatively to significantly limit PRP acquisition in a natively drug-susceptible *E. faecalis* isolate, T11. Through *in vitro* evolution and deep sequencing analysis, it was concluded that under antibiotic selection for a PRP, the CRISPR-Cas system of T11 becomes compromised, providing a potential mechanism for the emergence of MDR *E. faecalis*. Finally, a mouse model of *E. faecalis* colonization was used to show that CRISPR-Cas provides defense against PRP acquisition *in vivo*. Overall, the data presented here demonstrates that R-M and CRISPR-Cas limit MGE acquisition in *E. faecalis*, supporting the overall hypothesis that MDR *E. faecalis* emerge due to the absence of genome defense. This coupled with the robust *in vivo* activity of CRISPR-Cas indicates that the CRISPR-Cas systems associated with *E. faecalis* strains could be harnessed for use as alternative therapies to treat *E. faecalis* infections.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
AS	Aggregation substance
BHI	Brain heart infusion
CFU	Colony forming unit
CRISPR-Cas	Clustered, regularly interspaced, short palindromic repeats with CRISPR-associated genes
crRNA	CRISPR RNA
DND	DNA phosphorothioation
EfCR3Cas9	*Enterococcus faecalis* CRISPR3 Cas9
GI	Gastrointestinal
HAI	Hospital associated infection
HGT	Horizontal gene transfer
LB	Lysogeny broth
MDR	Multi-drug resistant
m6A	6-methyladenine
m4C	4-methylcytosine
m5C	5-methylcytosine
MGE	Mobile genetic element
MTase	Methyltransferase
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
PNPase	Polynucleotide phosphorylase
PRP	Pheromone-responsive plasmid
PSK	Post-segregational killing
RAST	Rapid annotation using subsystems technology
REase	Restriction endonuclease
R-M	Restriction modification
SaCas9	*Staphylococcus aureus* Cas9
SMRT	Single molecule real time sequencing
SpCas9	*Streptococcus pyogenes* Cas9
tracrRNA	trans-activating CRISPR RNA
UV	Ultraviolet
VRE	Vancomycin resistant enterococci

CHAPTER 1

INTRODUCTION

1.1 Author contribution

The contents of this chapter were written and prepared by Valerie Price. This chapter was edited by Valerie Price and Kelli L. Palmer.

1.2 The threat of antibiotic resistance

The modern antibiotic era is commonly associated with the public use of penicillin that began in the 1940's (1). However, the use of antimicrobial agents has been traced back to ancient human history. Evidence for this comes from discovering tetracycline in the bones of human skeletal remains dating to the Roman period (2). In addition, traditional Chinese medicine remedies against malaria suggest that civilizations have known about and used antimicrobials to cure diseases throughout the course of human history (3). Despite the breakthroughs of the 1950's, which unearthed many novel classes of antibiotics, there has been a rapid decline in the discovery of new antimicrobial agents over the last fifty years (4). The slow discovery of new antimicrobial agents and the accelerated rate at which bacteria are acquiring resistance to existing antibiotics has resulted in increased mortality rates from bacterial infections that were once treatable. This has placed undue burden on healthcare systems around the world (5).

The human body is a complex organism that plays host to an entire ecosystem of commensal bacteria that act as symbionts and cause no harm to their human host. The trillions of bacteria

that make up the human microbiome consist of many different species each with their own unique genetic makeup, creating a very genetically diverse population of commensal organisms. The use of antibiotics disrupts the inherent balance within the microbiome by targeting certain populations of bacteria while allowing antibiotic-resistant populations to thrive. The disruption of the normal microbiota, at the hand of antibiotics, is the leading reason for the emergence of antibiotic-resistant organisms that are untreatable with our existing arsenal of antibiotic compounds (6). This emergence of antibiotic-resistant organisms has negatively impacted the overall health of the human population. The Centers for Disease Control and Prevention estimates that over two million illnesses and 23,000 deaths can be attributed to antibiotic-resistant organisms annually in the United States (7). The economic burden associated with antibiotic-resistant bacterial infections has an estimated annual cost of $20 billion in the United States alone (7).

Indeed, the threat of antibiotic resistance is a serious public health and economic threat that if left unresolved will have an increasingly negative impact on the health care industry. Current research on antibiotic-resistant organisms focuses on the mechanisms employed by bacteria to become resistant to available antibiotic treatments. The goal of understanding these details is to discover new cellular targets in order to generate novel therapies or to modify existing antimicrobial molecules such that they have improved efficiency. The focus of the research presented in this dissertation is centered on one bacterial species that is among the leading causes of drug resistant infections in clinical settings. The following sections will introduce this

bacterium and the current understanding of how it has developed resistance to most available treatment options.

1.3 *Enterococcus faecalis* and clinical significance

The enterococci are gram-positive bacteria that can be found in a wide variety of habitats ranging from soil to fermented foods (8). Species of enterococci are armed with metabolic features that allow them to withstand wide ranges of pH, temperature, and salt concentrations, contributing to their resilience in diverse habitats (9). A recent study traced the evolutionary origins of this genus and determined that it diverged from its last common ancestor concurrently with the terrestrialization of animals (10). Therefore, it is not surprising that certain species of enterococci are considered to be commensal organisms in the gastrointestinal (GI) tracts of humans, birds and other mammals (8). Within the last half century, two particular species of *Enterococcus*, *E. faecalis* and *E. faecium*, have gained notoriety as common hospital-associated pathogens despite the fact that they normally reside in the human gut at no apparent cost to their host (11, 12). It has been argued that some of the same metabolic features that drove the emergence of the *Enterococcus* genus also play roles in the ability of these bacteria to adapt to and thrive in a number of diverse environments, including hospitals (10).

E. faecalis and *E. faecium* are among the leading causes of device-associated and surgical site nosocomial infections in the United States (13). A steady rise in antibiotic resistance among these opportunistic pathogens began in the 1970s and continues today, which has led to their classification as a serious public health threat by the Centers for Disease Control and Prevention (14). A recent report on the number of hospital-associated infections attributed to *E. faecalis* and

E. faecium revealed that together they cause at least 66,000 infections per year (7). The most common types of infections caused by *E. faecalis* and *E. faecium* are those of the catheterized urinary tract, surgical site infections, and the life-threatening infections bacteremia and endocarditis secondary to intravascular catheter use (11). It has been determined that approximately 20-30% of hospital-associated *E. faecalis* and *E. faecium* infections are caused by vancomycin-resistant strains, or as they are often abbreviated, VRE (vancomycin resistant enterococci) (14-16). Vancomycin is an antibiotic that is used as a last resort in the treatment of certain bacterial infections after other viable treatment options have been pursued (17). Resistance to vancomycin, coupled with intrinsic antibiotic resistance to more commonly prescribed antibiotics, such as aminoglycosides and macrolides, leaves very few treatment options for infections caused by *E. faecalis* and *E. faecium* (18). Although both of the *Enterococcus* species described here are often grouped into the same category in terms of reporting hospital-associated infections (HAIs), the types of infections they cause and their associated mechanisms for antibiotic resistance differ between the two species (18). The studies conducted in this dissertation focused on the *faecalis* species, therefore all subsequent information will be discussed in the context of *E. faecalis*.

Although *E. faecalis* is a member of the normal gut flora, it is an opportunistic pathogen capable of causing life-threatening infections. The transition from commensal to pathogen is often afforded through the acquisition of antibiotic resistance and virulence-enhancing traits. However, the molecular mechanisms involved in this process have not been fully elucidated. It has been demonstrated that administration of antibiotics with activity against gram-negative bacteria leads

to an increase in *E. faecalis* density in the intestine (11, 12, 19). Now that *E. faecalis* is more abundant, it able to cross mucosal barriers and cause more serious infections. More commonly, hospital-associated infections result from the translocation and colonization of hospital-adapted strains of *E. faecalis* with enhanced pathogenicity (20). It has been documented that *E. faecalis* can persist on hospital surfaces for up to a week and can survive on the skin of health care workers and patients for up to an hour, presenting the opportunity for VRE outbreaks in clinical settings (21). Much of the focus of *E. faecalis* research has been aimed at understanding the mechanisms of colonization and resistance of hospital-adapted strains.

Advances in genome sequencing technologies have been at the forefront of understanding how multidrug resistant (MDR) *E. faecalis* have emerged in clinical settings. The genome sequence of *E. faecalis* strain V583, the first VRE *E. faecalis* strain isolated in the United States, became available in 2003, representing the first fully sequenced *E. faecalis* genome (22). The complete genome sequence of V583 revealed a genome that is enriched in mobile genetic elements (MGEs), including a pathogenicity island, multiple plasmids, integrated prophage, and transposons that encoded vancomycin resistance and other virulence traits (22). One of the next *E. faecalis* strains to be fully sequenced was OG1RF, a drug-susceptible oral isolate. Comparison of these two genomes revealed that the MDR strain, V583, had a genome at least 600 kilo-bases larger than OG1RF; this extra DNA was almost exclusively in the form of MGEs (23). Subsequent studies comparing the size and gene content of hospital-adapted strains to commensal isolates have come to the consensus that MDR strains of *E. faecalis* contain MGEs that contribute to their virulence and success as hospital-associated pathogens (24-26).

Mobile genetic elements, such as plasmids, bacteriophage, and transposons, are disseminated through the process of horizontal gene transfer (HGT). This process refers to the lateral transfer of genetic material between strains or species rather than vertical inheritance from parent to daughter cell. In *E. faecalis*, HGT is commonly mediated through plasmids, such as the pheromone-responsive plasmids and members of the broad host range Inc18 group, and conjugative transposons (27, 28). The prevalence of MGEs in MDR strains of *E. faecalis* is an indication that this species readily exchanges DNA with other bacteria in the human intestinal tract. This is a reason for concern because this species has the ability to act as a reservoir for antibiotic resistance determinants that can be disseminated to other bacterial genera. This illustrates the necessity of elucidating the mechanisms that govern HGT and the evolution of MDR *E. faecalis* in health care settings.

1.4 Pheromone-Responsive Plasmids

The pheromone-responsive plasmids (PRPs) represent a unique class of conjugative plasmids that are almost exclusive to the *E. faecalis* species. The PRPs are clinically relevant as they are often identified in clinical isolates of *E. faecalis* (27), where up to three PRPs have been identified in a single isolate (29, 30). There is evidence that these plasmids can be transferred to other genera of bacteria, but they are usually unable to replicate in non-*E. faecalis* hosts (31). The absence of PRPs in other bacterial genomes suggests that this class of plasmids and *E. faecalis* have co-evolved. This exclusive relationship and the evidence that hospital-adapted *E. faecalis* genomes are enriched with MGEs suggests that PRPs can shape the evolutionary trajectory of MDR *E. faecalis* and perhaps the species as a whole.

PRPs are relatively large (usually 60 kilobase pairs or greater) and encode a unique mechanism for conjugation among other traits such as antibiotic resistance, bacteriocin production, post-segregational killing mechanisms and resistance to ultraviolet (UV) light (32). The conjugation of these plasmids is afforded through sex pheromone peptides that are encoded on the chromosome of *E. faecalis* strains. At least nine chromosomally-encoded pheromone peptides have been identified and each has at least one cognate PRP that can be mobilized upon peptide recognition (27). The distribution of other traits that contribute to PRP maintenance and virulence in *E. faecalis* populations vary between individual plasmids. Much of our knowledge of PRP conjugation and the function of auxiliary traits encoded by them has come from the extensive study of two model PRPs, pCF10 and pAD1 (32, 33). The data presented in this dissertation is primarily focused on pAD1, a 60 kilobase pair PRP that encodes UV resistance, cytolysin, and aggregation substance. The following sections will highlight relevant aspects of PRPs using pAD1 as the model plasmid.

1.4.1 The pheromone response and conjugation

The use of peptide pheromones as a means to promote plasmid transfer between bacterial strains was first identified in *E. faecalis*. *E. faecalis* strains readily secrete multiple, unique short peptide sequences into the extracellular environment that act as signals for PRPs (34, 35). These short peptides are the result of cleavage events associated with chromosomally encoded lipoproteins for which cellular function have not been elucidated. Once in the extracellular environment, these peptides can be recognized by their cognate PRP; for example, plasmid pAD1 recognizes the peptide, cAD1. The overall result of the pheromone response is the transfer of a PRP to a

recipient cell in frequencies that are on the magnitude of 1 transconjugant per 10-100 donor cells (32, 33).

The pheromone response is tightly regulated, mostly by plasmid-encoded components, and will be briefly discussed here. The recognition of the pheromone cAD1 in the environment by pAD1 results in the activation of mating functions beginning with the transcription and export of aggregation substance (AS); it has been documented that AS enhances the pathogenicity of *E. faecalis*-associated endocarditis (36, 37). AS coats the surface of the pAD1-containing donor cell, which promotes the physical interaction between itself and the plasmid-free recipient cell that is producing cAD1 (38). This cell-to-cell contact can be observed by the naked eye in the laboratory in the form of cell aggregates and marks the initiation of plasmid transfer from donor to recipient cell (see Figure 1.1). The transfer of pAD1 from donor to recipient cell results in a transconjugant cell that now has at least one copy of the PRP. At this point, pAD1 in the transconjugant cell begins to express iAD1, an inhibitor of the cAD1 peptide that competes for binding to an external receptor protein (39). Expression of this inhibitor is thought to prevent the transconjugant from mounting a pheromone response to any residual cAD1.

Recent studies have looked more in depth at the dynamics between donor and recipient cell densities and how this affects the efficiency of plasmid transfer (40, 41). It is intuitive to reason that having a high density of plasmid donor cells and an equivalent or lesser amount of recipient cells would result in higher conjugation frequencies; however, the opposite is observed in PRP-mediated conjugation, where a lower donor density results in increased conjugation frequencies

(41). The overarching conclusions made from these recent studies implicate cAD1 as a measure of recipient cell density and iAD1 as a measure of donor cell density, whereby donor cells use the concentration of iAD1 to regulate the mating response. Due to their ability to prevent the rapid dissemination of PRPs by inhibiting the pheromone response, pheromone peptide inhibitors, such as iAD1, can be exploited as novel therapies against *E. faecalis* related infections.

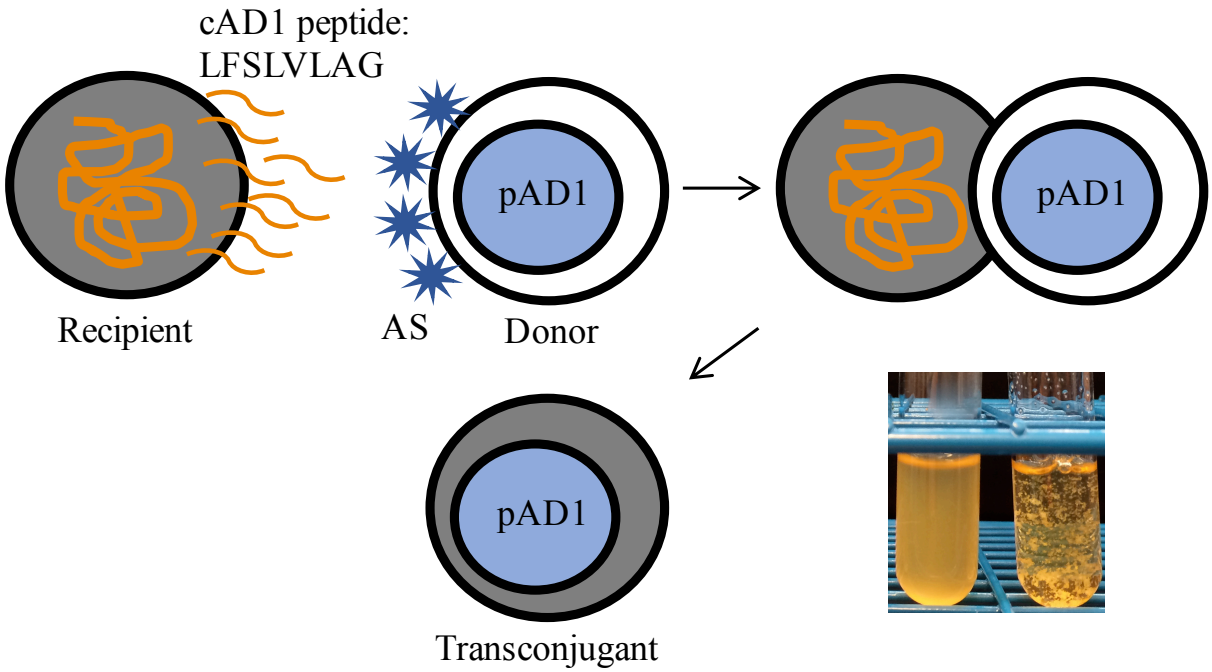


Figure 1.1. Model of pheromone-responsive plasmid conjugation. A plasmid-free cell (recipient) secretes a pheromone peptide, cAD1, into the extracellular environment that is recognized by the PRP, pAD1, in a donor cell. In response, pAD1 produces an aggregation substance (AS) that coats the outside of the donor cell, promoting cell-to-cell contact between donor and recipient cell. The physical interaction between the two cells allows for conjugation of pAD1 from donor to recipient, resulting in a transconjugant**.** (33). The picture shows aggregation of cells in a monoculture of a donor strain with (right) and without (left) synthetic cAD1 that induces the pheromone response.

1.4.2 Toxin-Antitoxin systems

Many bacterial plasmids possess mechanisms that promote their inheritance among a bacterial population, called post-segregational killing (PSK) mechanisms. PSK systems often come in the form of toxin and antitoxin molecules, where the antitoxin is produced at a concentration acceptable to keep the toxin molecule from killing the cell (42). The entire PSK system is reliant on discrepancies in stability of the toxin and antitoxin molecules, with the toxin possessing a longer half-life than the antitoxin. However, when there is improper segregation of the plasmid, the amount of antitoxin is depleted and can no longer protect the cell from the toxin, resulting in cell death. PSK mechanisms are advantageous to plasmids within a bacterial population as they ensure that every living daughter cell will have acquired at least one complete, functional plasmid copy.

pAD1 encodes a toxin-antitoxin system called *par*, representing a type I PSK system (43, 44); interestingly, the *par* system of pAD1 was the first reported instance of a type I PSK system occurring in gram-positive bacteria. The toxin and antitoxin molecules of pAD1 are convergently transcribed RNA molecules, called RNAI and RNAII, respectively (Figure 1.2). A portion of the RNAI transcript generates the peptide toxin, designated *fst*. RNAII does not code for a protein, but can form an extensive secondary structure with RNAI through shared direct repeat sequences, DRa and DRb. RNAI and RNAII also physically interact through complementary sequences in their respective terminators which block the ribosomal binding site of the RNAI transcript, thereby preventing production of the Fst toxin (45). Thus, as long as pAD1 is present, toxin expression is kept in check by RNAII; however, if a daughter cell fails to receive a

functional copy of the plasmid, the RNAII molecule is rapidly degraded allowing the stable Fst toxin to kill the cell. Toxin overexpression studies have shown that Fst alone is capable of killing *E. faecalis* and other gram-positive bacteria, including *Staphylococcus aureus* and *Bacillus subtilis* (46, 47). The mechanism by which the Fst toxin kills cells is currently unknown. The toxin is predicted to have a hydrophobic component that resembles a transmembrane domain and elicits a nucleoid condensation phenotype in target cells, thus providing evidence that its mechanism of action is likely intracellular (47, 48).

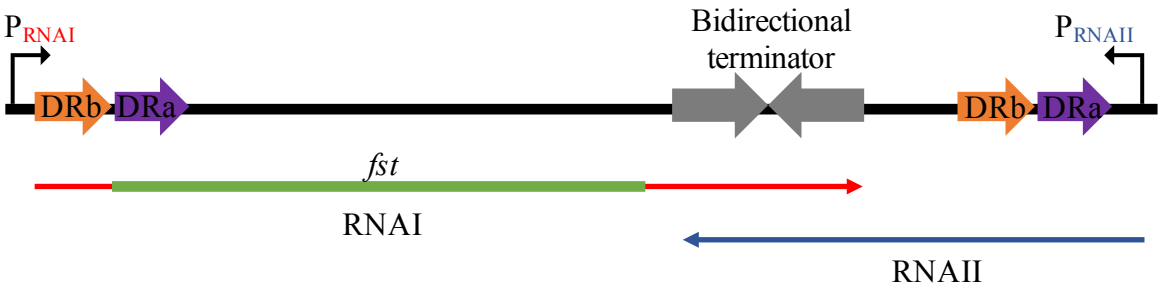


Figure 1.2. Genetic organization of the pAD1 plasmid addiction system, *par*. The RNA molecules that make up the toxin-antitoxin system of pAD1 are transcribed from convergent promoters, P_{RNAI} and P_{RNAII}, toward bidirectional terminator sequences (grey arrows). The transcripts, RNAI and RNAII, originating from this locus form significant RNA:RNA interactions mediated through direct repeat sequences, DRa and DRb (purple and orange arrows). The interaction between the direct repeats prevents *fst* translation (green box overlay on RNAI transcript) by occluding the ribosome binding site (47).

Interestingly, not all PSK systems are encoded on MGEs. In fact, *par* homologs not associated with integrated MGEs have been identified on the chromosomes of *E. faecalis* strains (27). Subtle differences in the RNA molecule sequences indicate that cross-talk between the *par* homologs is unlikely, allowing for the possibility of alternative functions unrelated to PSK (49). There is evidence that these chromosome-associated toxin-antitoxin systems of *E. faecalis* might contribute in some way to carbohydrate metabolism (27, 47). The discovery of *par* in pAD1 has

indeed opened up new avenues of research into the functional aspects of toxin-antitoxin systems in the context of both plasmid maintenance and host metabolism.

1.4.3 Bacteriocin production

Bacteriocins are molecules produced by bacteria that act as antimicrobial agents to inhibit the growth of other bacteria. These antimicrobial molecules are thought to help bacteria establish control in certain niches by eliminating any competitors that could be using up necessary resources. pAD1 encodes a bacteriocin called cytolysin. Cytolysin has activity against a number of gram-positive bacteria including members of the genera *Bacillus*, *Streptococcus*, *Lactobacillus* and *Clostridium* (50). Cytolysin kills cells by forming pores in the bacterial membrane, leading to cell lysis and death. In addition to killing bacterial cells, cytolysin also has hemolytic activity which can lyse human, horse, and bovine red blood cells (50). Finally, cytolysin has been implicated in numerous animal models of *E. faecalis* infections as an agent that promotes virulence and enhances the severity of infections (51-54). The broad-spectrum activity of pAD1-endoded cytolysin against multiple bacterial species and its capability of lysing red blood cells highlight its importance as a clinically relevant virulence factor that contributes to the pathogenicity of some hospital-associated *E. faecalis* strains.

The pAD1 cytolysin is produced from an operon of six genes encoding functions including cytolysin modification, translocation, and activation (see Figure 1.3) (55). Active cytolysin is composed of a pair of peptides, $CylL_L$ and $CylL_S$, that are initially post-translationally modified inside the cell by the product of the *cylM* gene (56). After this modification, CylB, an ATP-

binding cassette transporter (ABC transporter), secretes and trims the modified $CylL_L$ and $CylL_S$ peptide products, generating $CylL_L$' and $CylL_S$' (57, 58). CylA performs a final cleavage on $CylL_L$' and $CylL_S$' resulting in the active subunits, $CylL_L$" and $CylL_S$", that form the mature cytolysin molecule (59). These mature products are capable of forming extensive interactions with each other; however, the $CylL_L$" peptide preferentially binds target cells, whereas the $CylL_S$" peptide has an additional function as an autoinducer in a novel quorum sensing system that upregulates cytolysin production when an appropriate target is available (60). Active $CylL_S$" interacts with the $CylR_1$ and $CylR_2$ two-component system to induce the transcription of the entire cytolysin operon, resulting in an increased production of cytolysin molecules capable of interacting and killing more target cells.

The cytolysin operon of pAD1 also encodes an immunity factor, CylI, that protects the host cell from the activity of cytolysin (61). The mechanism of immunity has yet to be elucidated, but it is generally accepted that the immunity factor is a membrane-associated protein. Aside from the current animal models demonstrating pAD1-encoded cytolysin as a means of enhanced infection and the association of pAD1 with clinical isolates of *E. faecalis*, little is known about the impacts of cytolysin on pAD1 conjugation frequencies or the interactions, if any, between it and other virulence traits encoded by this PRP. It is of interest to investigate if donor densities could impact the efficacy of cytolysin activity against its targets.

1.5 Mechanisms of genome defense

Bacteria and archaea are among the most abundant organisms on Earth. The actual number is

largely underestimated as many genera and species remain unidentified due their diverse habitats and the inability to isolate them. Even more numerous than prokaryotic organisms are bacteriophage, or viruses that exclusively infect prokaryotes; like prokaryotes their abundance is underrepresented due to the limitations of culturing them (62). Bacteriophage and plasmids are selfish elements that are out to maintain their existence. Due to their ability to infect bacteria and archaea, bacteriophages and plasmids are constantly co-evolving with their hosts, thereby putting prokaryotes and MGEs in constant conflict, or an arms race (63). Although some MGEs can provide benefits to their hosts under certain conditions, this often comes with increased metabolic burden to the host. Therefore, prokaryotes often encode defense systems that protect themselves from infection by bacteriophage and plasmids.

Genome defense mechanisms can be classified into two broad groups, those that rely on DNA modifications and those that are related to cell death (64). The first group can be further classified into innate and adaptive immunity systems. Perhaps the most common form of innate immunity is restriction-modification (R-M), which relies on DNA methylation to distinguish self and non-self DNA. R-M systems will be discussed in more detail in the following section. Unlike R-M systems, DNA phosphorothioation, or DND systems, utilize S-modification of the DNA backbone to identify self and non-self DNA (65, 66). The most recent addition to the family of innate immunity systems are the Argonaut proteins which were originally thought to only occur in eukaryotic organisms (67, 68). It is known that Argonaut proteins often use single stranded DNA guides to recognize and target non-self DNA; however, what determines this

recognition is yet to be determined. The only known member of adaptive immunity systems is CRISPR-Cas, a form of RNA-guided adaptive immunity that will be discussed in detail later.

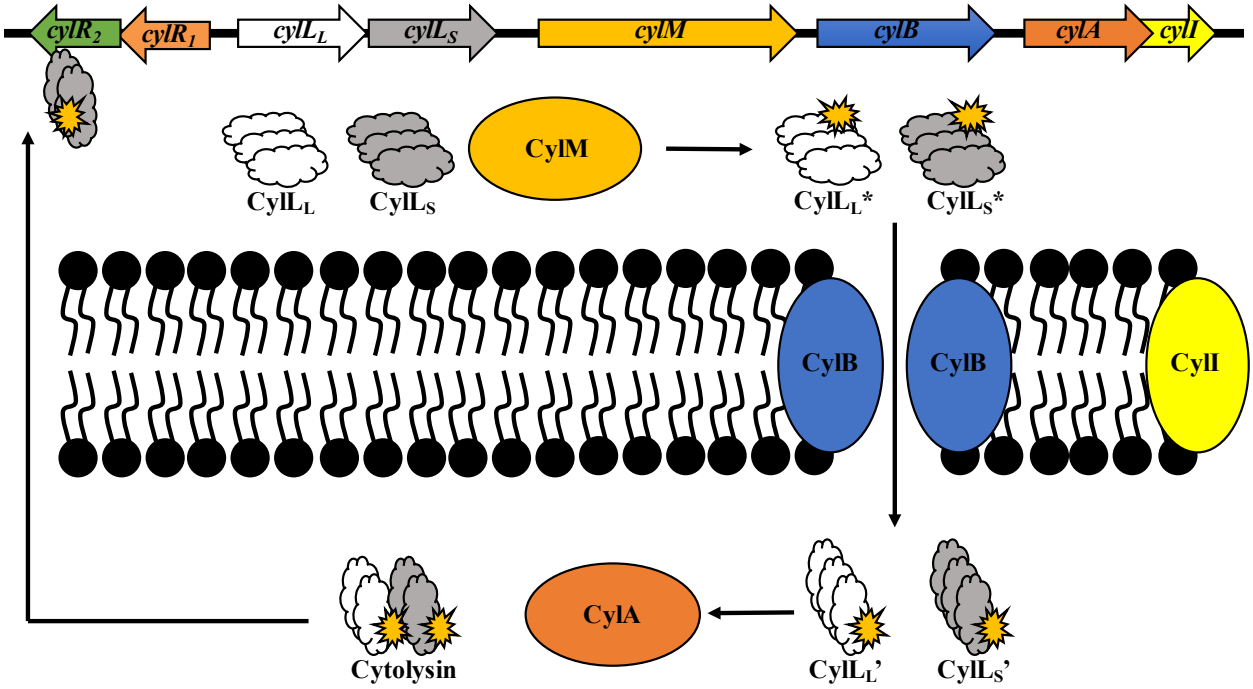


Figure 1.3. Cytolysin production from pAD1. The cytolysin operon of pAD1 encodes the two structural components of the cytolysin molecule, $CylL_L$ and $CylL_S$, and three proteins that are involved in the processing and export of these subunits (55). CylM modifies the $CylL_L$ and $CylL_S$ products with lanthionine residues followed by the export of the modified subunits through membrane associated CylB. Once outside of the cell, CylA performs a final processing of the cytolysin subunits resulting in fully active peptides that form the cytolysin molecule. After complete activation, the individual subunits can also play independent roles: $CylL_L$ physically binds target cells and $CylL_S$ provides a quorum sensing role by acting as the inducer molecule for the two-component regulatory system of *cylR1* and *cylR2*. Finally, the cytolysin operon also encodes the immunity factor, CylI, that protects the host cell from the action of cytolysin by an unknown mechanism (61).

The second group of genome defense mechanisms rely on cell death to prevent the propagation of MGEs. The first member of this group is toxin-antitoxin systems that are sometimes encoded on MGEs themselves, such as the toxin-antitoxin system of pAD1 that was previously discussed

(69). The other member of this group is abortive infection that functions by halting the propagation of bacteriophage through death of the cell that has become infected with the unwanted phage (70). All of these defense systems share a common goal which is to prevent the metabolic burden and decreased fitness associated with unwanted entry of MGEs into the bacterial genome.

Much is still unknown about the prevalence, distribution, and impact of the genome defense systems in the opportunistic pathogen *E. faecalis*. Until recently, it was unclear if *E. faecalis* encoded any active R-M systems or if CRISPR-Cas systems were present. New advances in the field have identified both innate (R-M) and adaptive (CRISPR-Cas) immunity systems in *E. faecalis;* the following chapters will present some of these findings.

1.5.1 Restriction-Modification (R-M)

As previously mentioned, R-M systems provide defense against MGEs through methylation signatures that allow host genomes to distinguish between self and non-self DNA. R-M systems are classified into four types (I-IV) based on their gene cohort, dependence on ATP, and mechanism used to cleave their target (71-73). At the most basic level, R-M systems carry out two functions: 1) they add a methyl group to adenine or cytosine residues at specific motifs on the host genome via a methyltransferase, and 2) they cleave non-methylated DNA using a cognate restriction endonuclease (74). Type II R-M systems are the most well characterized group and have been used in molecular cloning strategies (75). Type II R-M systems are the

most relevant to the work presented in this dissertation, therefore the specific mechanisms of the other types of R-M systems will not be discussed.

Type II R-M systems are composed of two proteins, a methyltransferase (MTase) and a restriction endonuclease (REase) that are usually encoded within the same operon on the bacterial chromosome. The MTase is responsible for methylating host DNA at a specific motif that is usually 4-8 base pairs long and palindromic. The REase recognizes the same motif as the methyltransferase and restricts DNA that is not methylated at a particular position (see Figure 1.4), thereby hindering the establishment of foreign DNA, such as bacteriophage or plasmid, in the host cell (76). The host chromosome becomes methylated following rounds of replication, therefore it is possible for some bacteriophage (and even plasmids) to become methylated and therefore are recognized as self DNA following their integration into the host chromosome (74). This provides bacteriophage and plasmids the opportunity to escape the action of the REase, highlighting an inherent leaky nature of R-M systems whereby they are not always perfect at providing defense against MGEs.

1.5.2 R-M in *E. faecalis*

There has been an overall lack of knowledge on the existence of R-M systems in *E. faecalis*; only a handful of studies have identified putative systems in this species. The majority of these systems belong to the type II group and appear to be derived from MGEs as opposed to the chromosome (77-79). Due to the ubiquitous distribution of R-M systems in bacteria, it is reasonable to believe that R-M systems are more widespread among *E. faecalis* strains. Recent

studies have added to our knowledge on the prevalence and function of R-M systems in *E. faecalis.*

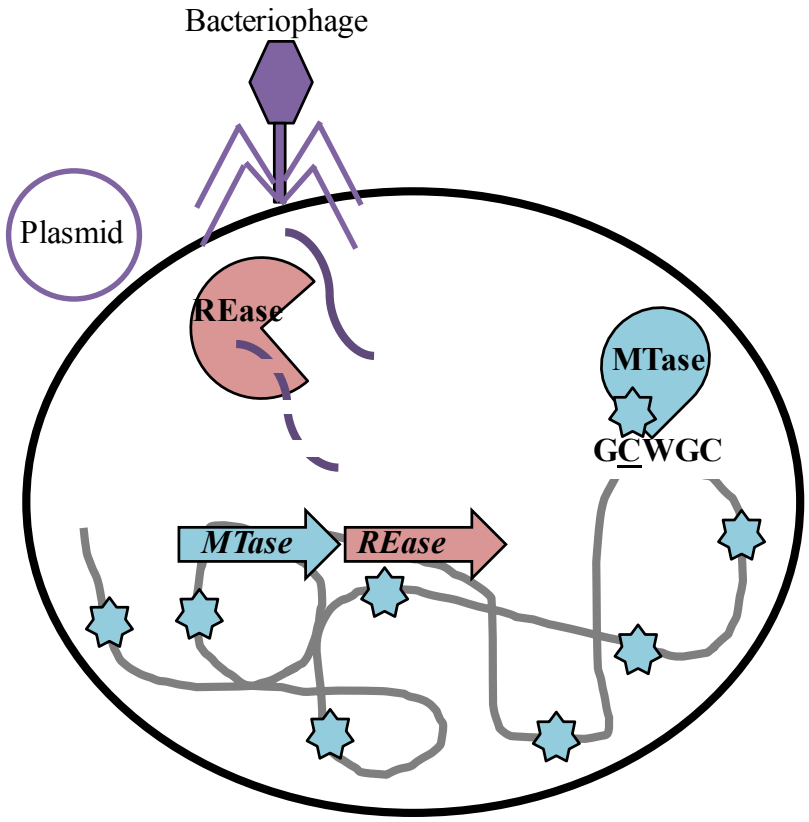


Figure 1.4. Model of function of type II restriction modification systems. Type II R-M systems generally encode a methyltransferase (*MTase*) and restriction endonuclease (*REase*), often in the form of an operon (71, 74). The MTase is responsible for the recognition and methylation of a particular motif on the host genome. In this model, the MTase recognizes 5'-GCWGC-3' motifs (where W can be either A or T) and specifically methylates (blue stars) the 5' cytosine base. There is a cognate REase that also recognizes 5'-GCWGC-3' motifs and restricts any DNA within the cell that does not have methylation at this motif. The model shows two MGEs, a plasmid and a bacteriophage, trying to enter a cell with 5'-GCWGC-3' methylation; the incoming DNA is un-methylated, allowing the REase to recognize the bacteriophage DNA as non-self, leading to cleavage of the foreign DNA and protection from infection by this MGE.

The approach taken to identify R-M systems in *E. faecalis* was a combination of bioinformatics and next-generation sequencing analysis. New England Biolabs maintains a database of

predicted and experimentally confirmed R-M systems called REBASE (80). This database was used to predict the presence of MTases in a model strain of *E. faecalis*, OG1RF (81). In conjunction with bioinformatic predictions, single molecule real time sequencing, or SMRT sequencing (82, 83), was also used to identify potential genome modifications. SMRT sequencing, a third-generation sequencing method, utilizes polymerase kinetics to monitor the addition of nucleotides to a growing chain of DNA. Pauses in the addition of new bases are interpreted as positions where the DNA has been modified by methylation, making it easy to identify the exact base and motifs that are methylated. The combination of these two methods resulted in the identification of one functional R-M system, EfaRFI, in *E. faecalis* OG1RF that recognizes and modifies host DNA at 5'-GCWGC-3' motifs (where W represents an A or T base) (81). Additional genomic analyses concluded that the distribution of EfaRFI is not well conserved among a collection of *E. faecalis* strains (81). Although the characterization of EfaRFI did not identify core R-M systems in the *faecalis* species, it did demonstrate that R-M systems and their recognition motifs can be accurately predicted using a combination of the methods described here.

1.5.3 CRISPR-Cas

CRISPR-Cas systems, or clustered, regularly interspaced short palindromic repeats and CRISPR-associated genes, are characterized by the presence of multiple direct DNA repeats separated by unique sequences of similar length. This configuration of repeats was initially identified in 1987 in a region flanking the *iap* gene of *E. coli* (84). It was more than a decade before questions about the functional aspect of these repeats were answered. Paramount to answering such

questions was the identification of genes that were associated with these repeats (85). Protein domain analysis of the predicted gene products revealed characteristics of nucleases and helicases, suggesting an interaction with DNA. Analysis of the sequences, or spacers, between direct repeats revealed identities to extrachromosomal DNA elements including plasmid and phage-derived sequences. This discovery was independently reported, almost concurrently, from three different groups analyzing spacer content of multiple species of bacteria (86-88). These systems have since been rapidly studied and many key characteristics have been elucidated, including the striking estimation that almost all archaea and at least 50% of bacterial genomes possess a CRISPR-Cas system (89).

Several seminal studies were published demonstrating that CRISPR-Cas systems can provide bacterial genomes with defense against phage and plasmid infection in a sequence-specific manner that is mediated through a non-coding RNA and *cas*-encoded nucleases (90-93). Since the initial discovery that CRISPR-Cas acts as a genome defense system and is thus a barrier to horizontal gene transfer, there has been an explosion of research surrounding these systems that has exponentially increased our knowledge about them. Using genome sequencing data, it was initially determined that there were three basic types (I-III) or classifications of CRISPR-Cas systems. Within the last five years, this classification has expanded into two broad classes (1 and 2) that include five types (I-V) of CRISPR-Cas systems (94, 95). The mechanisms of how these systems provide genome defense is generally analogous among all five types and will therefore be discussed in the context of type II systems, which are the focus of this dissertation.

1.5.4 Type II CRISPR-Cas in *E. faecalis*

Type II CRISPR-Cas systems have become the subject of intensive investigation as they have been adapted for biotechnological applications that have revolutionized the way scientists manipulate DNA sequences. The ability of this system to be used in such a way came about after the basic mechanistic details about the function of this system were elucidated. Type II CRISPR-Cas systems are defined by the type specific gene *cas9*. The Cas9 protein is a large protein and contains two endonuclease domains that mediate strand-specific DNA cleavage of foreign DNA elements (96). In addition to the *cas9* gene, type II loci also contain the conserved genes *cas1* and *cas2*; some type II systems encode a fourth *cas* gene, either *csn2* or *cas4* (97). Type II systems also employ a non-coding RNA called trans-activating CRISPR RNA (tracrRNA), usually encoded in the vicinity of the *cas9* gene, which is capable of forming an RNA duplex via complementary base pairing with the direct repeat sequence associated with the CRISPR locus (98). Downstream of the *cas* genes is the CRISPR array that is composed of direct repeats interspaced by short segments of unique sequence called spacers, which have identity to foreign DNA; a schematic representation of a type II CRISPR-Cas locus is presented in Figure 1.5.

As previously mentioned, the overall mechanism by which these systems provide genome defense is shared among the five types of CRISPR-Cas systems and will be presented here in the terms of a type II system. The process of CRISPR-Cas-mediated DNA interference can be broken down into three steps: adaptation, expression, and interference, that are depicted in Figure 1.5. Adaptation involves the incorporation of a new spacer into the leader end of the CRISPR array, thereby establishing a new ‘memory’ of an infection by a MGE. There is increasing

evidence that a complex of Cas proteins mediates the identification, excision and incorporation of new spacers into type II CRISPR loci (99, 100). The basis of the identification of putative spacer sequences is afforded through the recognition of a protospacer adjacent motif, or PAM; the PAM is a 2-8 base pair sequence located downstream of protospacer sequences (101). It has been shown that a particular domain of the Cas9 protein is responsible for the identification of the PAM motif and that Cas9 proteins derived from type II loci of other bacteria often recognize unique PAM sequences (102, 103). The second step in mediating genome defense is expression, wherein the entire set of repeats and spacers is transcribed into a pre-CRISPR RNA (pre-crRNA). The pre-crRNA transcript associates with tracrRNA molecules through complementary base pairing and is then processed by Cas9 and the host factor RNase III that is encoded by the *rnc* gene (98). The use of host-encoded factors not associated with the CRISPR-Cas locus is a unique feature of type II systems. These processing events culminate in the production of mature crRNA that are composed of partial spacer and adjacent repeat sequences; it is this form of crRNA that directs the cleavage event. Finally, interference with MGEs occurs when a mature crRNA:tracrRNA duplex guides Cas9 to the foreign DNA target by base-pairing with a sequence complementary to the spacer (104). This process also involves PAM recognition by Cas9 ultimately leading to double stranded DNA cleavage through the two endonuclease domains of Cas9; thus, immunizing the host genome from infection by the MGE (96, 105).

Genome sequencing analysis of a collection of *E. faecalis* strains revealed the presence of Type II CRISPR-*cas* loci, of which the composition and distribution varied among the strains analyzed. This resulted in the definition of three type II CRISPR-Cas systems, CRISPR1-Cas,

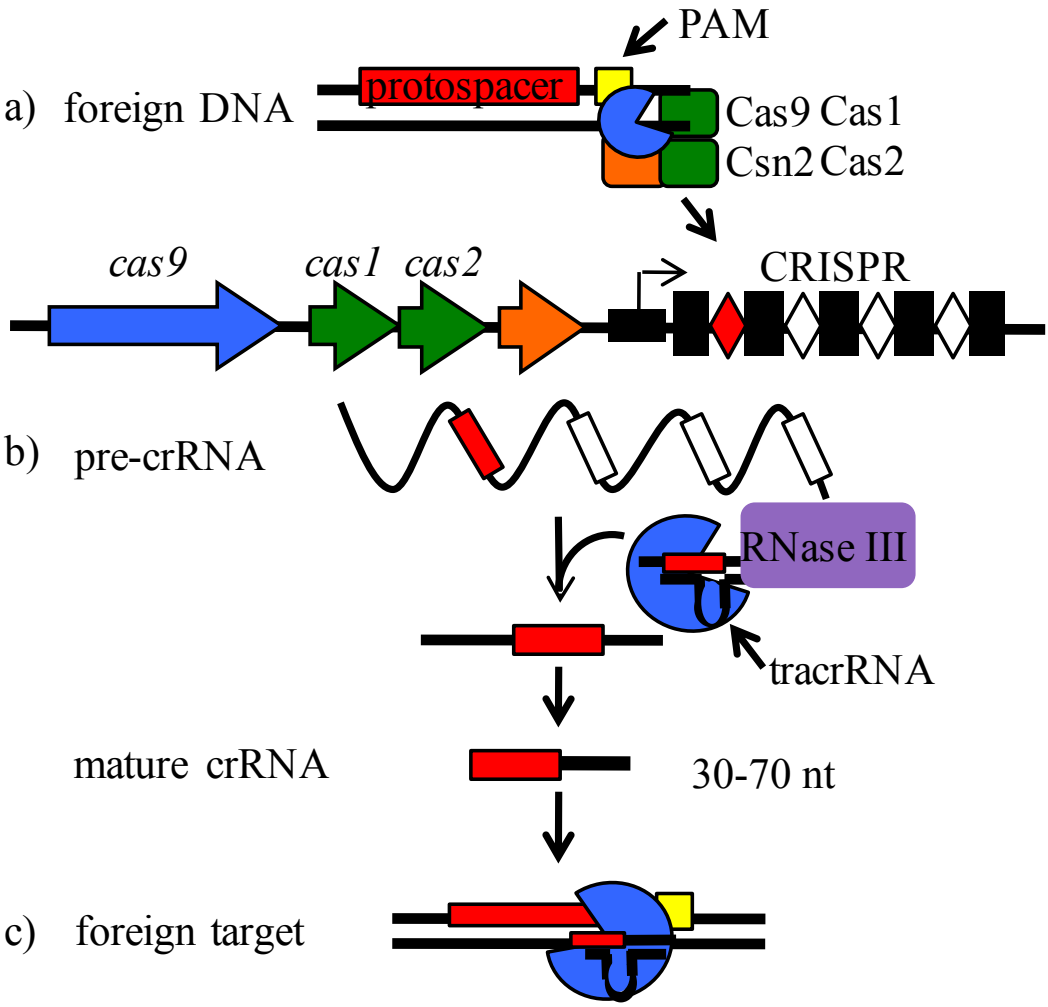


Figure 1.5. Type II CRISPR-Cas interference mechanism. The model for Type II CRISPR-Cas function shown here has been adapted from a previously proposed model where the locus depicted is a generalization of a basic type II CRISPR-Cas system (94, 95). The mechanism is described in three stages: a) adaptation, b) expression, and c) interference. During adaptation, a new spacer originating from foreign DNA (red protospacer sequence) is recognized in a PAM-dependent manner and subsequently integrated into the leader end of the array by a complex of Cas proteins. Next, the CRISPR array is transcribed into pre-crRNA; the pre-crRNA and tracrRNA form a complex that is processed by the host RNase III (98). Another processing event produces the mature crRNA consisting of parts of a spacer and adjacent repeat sequence. Finally, in interference, the mature crRNA:tracrRNA duplex guides Cas9 to the foreign DNA target by base-pairing with a sequence complementary to the spacer and PAM proximity, promoting cleavage of DNA through the two endonuclease domains of Cas9 (96). PAM, protospacer adjacent motif.

CRISPR2 and CRISPR3-Cas (23, 26); a general structure of these three loci is shown in Figure 1.6. Interestingly, all *E. faecalis* genomes analyzed to date possess CRISPR2, an orphan locus that lacks associated *cas* genes (106). In addition to the orphan CRISPR2 locus, *E. faecalis* strains can also possess either a CRISPR1-*cas* or CRISPR3-*cas* locus that vary in their *cas* gene content, location on the chromosome and direct repeat sequence (26). MDR strains of *E. faecalis* possess only the orphan CRISPR2 locus and therefore lack *cas* genes (26, 106). This observation in conjunction with the overall lack of knowledge about the distribution and function of R-M in this species led to the hypothesis that MDR *E. faecalis* emerge due to the lack of barriers to horizontal gene transfer.

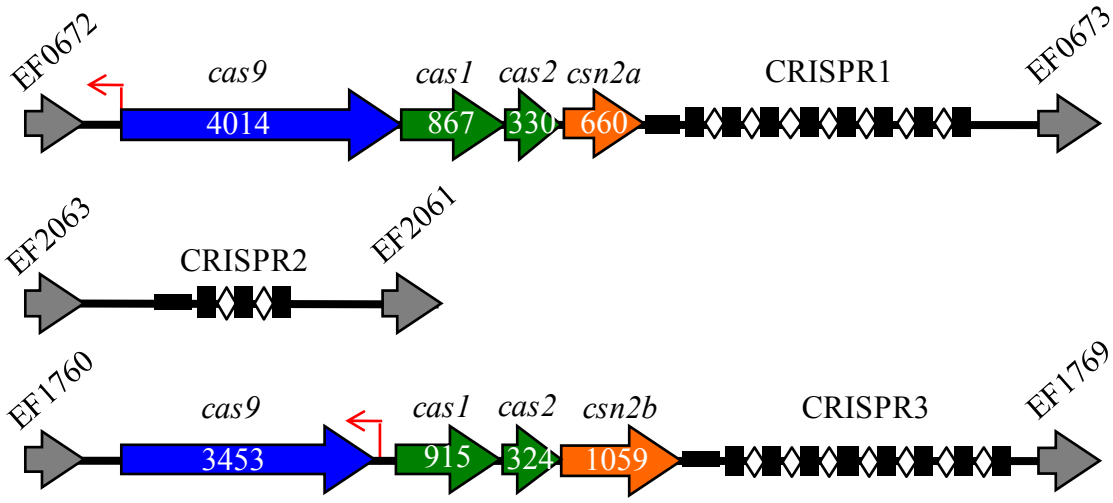


Figure 1.6. CRISPR-Cas systems in *E. faecalis*. Three type II CRISPR-Cas loci have been identified in *E. faecalis*. The organization of the systems depicted here are modeled from representative strains possessing these loci (26). The overall structure of CRISPR1-Cas and CRISPR3-Cas is similar, but the location of tracrRNA (red arrows) and the sizes of genes within the loci differ; nucleotide lengths of genes are given within arrows. The lengths of repeats (black rectangles) and spacers (white diamonds) are 36 and 30 base pairs, respectively; these sizes are conserved across the three systems although the individual sequences vary. Grey arrows represent the location of the CRISPR loci on the chromosome relative to orthologs of the V583 genome and the designation of *csn2a* or *csn2b* is based off of (97).

1.6 Justification of study

Due to the emergence of nosocomial *E. faecalis* infections, it is of great importance to understand the mechanisms associated with the acquisition of resistance to multiple antibiotics. *E. faecalis* readily engages in horizontal gene transfer, most often mediated through plasmids and to a lesser extent bacteriophage. It has been well established that MDR isolates of *E. faecalis* have genomes enriched with MGEs, providing a means of acquiring new traits (including antibiotic resistance) to help them thrive in specialized environments such as healthcare settings. The use of genome sequencing has established the existence of genome defense systems in *E. faecalis*, including R-M and CRISPR-Cas, but their presence is often limited to drug susceptible isolates whereas MDR strains appear to lack functional genome defense systems. However, there has been little in the way of experimental evidence exploring whether or not these systems actually provide genome defense in *E. faecalis* strains.

The focus of this dissertation is the study of type II R-M and CRISPR-Cas systems and their impact on HGT in the opportunistic pathogen *E. faecalis*. A combination of genetic, bioinformatics, experimental evolution, *in vitro* and *in vivo* approaches were taken to elucidate the impact of genome defense systems on the conjugative transfer of the model PRP, pAD1. The data presented in Chapters 2 and 3 will show that these genome defense systems impact plasmid transfer and that the stability of plasmid maintenance is altered by growth environment. Finally, Chapter 4 will provide a more clinical context to the relevance of genome defense in a mouse model of *E. faecalis* colonization. Ultimately, the knowledge gained from this work provides

novel insight into the interactions between MGEs and genome defense systems of *E. faecalis* that will impact the future development of alternative therapies to treat *E. faecalis* infections.

1.7 References

1. Flemming A (1929) On the Antibacterial Action of Cultures of a Penicillium, With Special Reference to their Use in the Isolation of *B. influenzae*. *British Journal of Experimental Pathology* X(3):226-236.

2. Cook M, Molto EL, & Anderson C (1989) Fluorochrome Labelling in Roman Period Skeletons From Dakhleh Oasis, Egypt. *American Journal of Physical Anthropology* 80:137-143.

3. Cui L & Su XZ (2009) Discovery, mechanisms of action and combination therapy of artemisinin. *Expert Rev Anti Infect Ther* 7(8):999-1013.

4. Aminov RI (2010) A brief history of the antibiotic era: lessons learned and challenges for the future. *Frontiers in microbiology* 1:134.

5. Gandra S, Barter DM, & Laxminarayan (2014) Economic burden of antibiotic resistance: how much do we really know? *Clin Microbiol Infect* 20:973-979.

6. Levy SB & Marshall B (2004) Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* 10(12 Suppl):S122-129.

7. Sievert DM*, et al.* (2013) Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. *Infection control and hospital epidemiology* 34(1):1-14.

8. Lebreton F, Willems RJL, & Gilmore MS (2014) *Enterococcus* Diversity, Origins in Nature, and Gut Colonization. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

9. Sherman JM (1938) The Enterococci and Related Streptococci. *J Bacteriol* 35(2):81-93.

10. Lebreton F*, et al.* (2017) Tracing the Enterococci from Paleozoic Origins to the Hospital. *Cell* 169(5):849-861 e813.

11. Agudelo Higuita N & Huycke MM (2014) Enterococcal Disease, Epidemeiology, and Implications for Treatment. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

12. Arias CA & Murray BE (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nature reviews. Microbiology* 10(4):266-278.

13. Weiner LM*, et al.* (2016) Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014. *Infection control and hospital epidemiology* 37(11):1288-1301.

14. Health UDo & Services H (2013) Antibiotic resistance threats in the United States, 2013. *Atlanta: CDC*.

15. Périchon B & Courvalin P (2009) Glycopeptide Resistance in Enterococci. 229-240.

16. Gilmore MS, Lebreton F, & van Schaik W (2013) Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. *Current opinion in microbiology* 16(1):10-16.

17. Barna JCJ & Williams DH (1984) The Structure and Mode of Action of Glycopeptide Anitibiotics of the Vancomycin Group. *Annual review of microbiology* 38:339-357.

18. Kristich CJ, Rice LB, & Arias C, A (2014) Enterococcal Infection-Treatment and Antibiotic Resistance. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

19. Ubeda C*, et al.* (2010) Vancomycin-resistant *Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *The Journal of clinical investigation* 120(12):4332-4341.

20. Mundy IM, Sahm DF, & Gilmore M (2000) Relationships between Enterococcal Virulence and Antimicrobial Resistance. *Clinical Microbiology Reviews* 13(4):513-522.

21. Noskin GA, Stosor V, Cooper I, & Peterson LR (1995) Recovery of Vancomycin-Resistant Enterococci on Fingertips and Environmental Surfaces. *Infection control and hospital epidemiology* 16(10):577-581.

22. Paulsen IT*, et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299(5615):2071-2074.

23. Bourgogne A*, et al.* (2008) Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome biology* 9(7):R110.

24. McBride SM, Fischetti VA, Leblanc DJ, Moellering RC, Jr., & Gilmore MS (2007) Genetic diversity among *Enterococcus faecalis*. *PloS one* 2(7):e582.

25. Solheim M, Aakra A, Vebo H, Snipen L, & Nes IF (2007) Transcriptional responses of *Enterococcus faecalis* V583 to bovine bile and sodium dodecyl sulfate. *Applied and environmental microbiology* 73(18):5767-5774.

26. Palmer KL & Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4).

27. Clewell DB*, et al.* (2014) Extrachromosomal and Mobile Elements in Enterococci: Transmission, Matintenance, and Epidemiology. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

28. Palmer KL, Kos VN, & Gilmore MS (2010) Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Current opinion in microbiology* 13(5):632-639.

29. Clewell D*, et al.* (1982) Sex pheromones in *Streptococcus faecalis*: multiple pheromone systems in strain DS5, similarities of pAD1 and pAMγ1, and mutants of pAD1 altered in conjugative properties. *Microbiology*:97-100.

30. Murray BE, An FY, & Clewell D (1988) Plasmids and pheromone response of the beta-lactamase producer *Streptococcus* (*Enterococcus*) *faecalis* HH22. *Antimicrobial agents and chemotherapy* 32(4):547-551.

31. Francia MV & Clewell DB (2002) Tansfer origins in the conjugative *Enterococcus faecalis* plasmids pAD1 and pAM373: identification of the pAD1 *nic* site, a specific relaxase and a possible TraG-like protein. *Molecular microbiology* 45(2):375-395.

32. Clewell DB (2007) Properties of *Enterococcus faecalis* plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin. *Plasmid* 58(3):205-227.

33. Dunny GM (2007) The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 362(1483):1185-1193.

34. Dunny GM, Brown BL, & Clewell DB (1978) Induced cell aggregation and mating in *Streptococcus faecalis*: Evidence for a bacterial sex pheromone. *Proceedings of the National Academy of Sciences of the United States of America* 75(7):3478-3483.

35. Dunny GM, Craig RA, Carron RL, & Clewell DB (1979) Plasmid Transfer in *Streptococcus faecalis*: Production of Multiple Sex Pheromones by Recipients. *Plasmid* 2(454-465).

36. Chow JW*, et al.* (1993) Plasmid-Associated Hemolysin and Aggregation Substance Production Contribute to Virulence in Experimental Enterococcal Endocarditis. *Antimicrobial agents and chemotherapy* 37(11):2474-2477.

37. Chuang ON*, et al.* (2009) Multiple functional domains of *Enterococcus faecalis* aggregation substance Asc10 contribute to endocarditis virulence. *Infection and immunity* 77(1):539-548.

38. Ike Y & Clewell DB (1985) Genetic Analysis of the pAD1 Pheromone Response in *Streptococcus faecalis,* Using Transposon Tn*917* as an Insertional Mutagen. *Journal of Bacteriology* 158(3):777-783.

39. Mori M*, et al.* (1987) Isolation and structure of the sex pheromone inhibitor, iPD1, excreted by *Streptococcus faecalis* donor strains harboring plasmid pPD1. *J Bacteriol* 169(4):1747-1749.

40. Chatterjee A*, et al.* (2013) Antagonistic self-sensing and mate-sensing signaling controls antibiotic-resistance transfer. *Proceedings of the National Academy of Sciences of the United States of America* 110:7086–7090.

41. Bandyopadhyay A, O'Brien S, Frank KL, Dunny GM, & Hu WS (2016) Antagonistic Donor Density Effect Conserved in Multiple Enterococcal Conjugative Plasmids. *Applied and environmental microbiology* 82(15):4537-4545.

42. Gerdes K, Rasmussen PB, & Molin S (1986) Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences of the United States of America* 83:3116-3120.

43. Weaver KE & Clewell DB (1989) Construction of *Enterococcus faecalis* pAD1 Miniplasmids: Identification of a Minimal Pheromone Response Regulatory Region and Evaluation of a Novel Pheromone-Dependent Growth Inhibition. *Plasmid* 22:106-119.

44. Weaver KE & Tritle DJ (1994) Identification and Characterization of an *Enterococcus faecalis* Plasmid pAD1-Encoded Stability Determinant Which Produces Two Small RNA Molecules Necessary for Its Function. *Plasmid* 32:168-181.

45. Greenfield TJ & Weaver KE (2000) Antisense RNA regulation of the pAD1 *par* post-segregational killing system requires interaction at the 5' and 3' ends of the RNAs. *Molecular microbiology* 37(3):661-670.

46. Patel S & Weaver KE (2006) Addiction toxin Fst has unique effects on chromosome segregation and cell division in *Enterococcus faecalis* and *Bacillus subtilis*. *J Bacteriol* 188(15):5374-5384.

47. Weaver KE*, et al.* (2009) Identification and characterization of a family of toxin-antitoxin systems related to the *Enterococcus faecalis* plasmid pAD1 *par* addiction module. *Microbiology* 155(Pt 9):2930-2940.

48. Weaver KE*, et al.* (2003) *Enterococcus faecalis* Plasmid pAD1-Encoded Fst Toxin Affects Membrane Permeability and Alters Cellular Responses to Lantibiotics. *Journal of Bacteriology* 185(7):2169-2177.

49. Van Melderen L (2010) Toxin-antitoxin systems: why so many, what for? *Current opinion in microbiology* 13(6):781-785.

50. Cox CR, Coburn PS, & Gilmore MS (2005) Enterococcal Cytolysin: A Novel Two Component Peptide System that Serves as a Bacterial Defense Against Eukaryotic and Prokaryotic Cells. *Current Protein and Peptide Science* 6(1):77-84.

51. Dupont H, Montravers P, Mohler J, & Carbon C (1998) Disparate Findings on the Role of Virulence Factors of *Enterococcus faecalis* in Mouse and Rat Models of Peritonitis. *Infection and immunity* 66(6):2570-2575.

52. Huycke MM, Gilmore MS, Jett BD, & Booth LJ (1992) Transfer of Pheromone-Inducible Plasmids between *Enterococcus faecalis* in the Syrian Hamster Gastrointestinal Tract. *Journal of Infectious Diseases* 166:1188-1191.

53. Huycke MM, Joyce WA, & Gilmore MS (1995) *Enterococus faecalis* Cytolysin without Effect on the Intestinal Growth of Susceptible Enterococci in Mice. *Journal of Infectious Diseases* 172:273-276.

54. Ike Y, Hashimoto H, & Clewell DB (1984) Hemolysin of *Streptococcus faecalis* Subspecies *zymogenes* Contributes to Virulence in Mice. *Infection and immunity* 45(2):528-530.

55. Gilmore MS, Coburn PS, Nallapareddy SR, & Murray BE (2002) Enterococcal virulence. *The Enterococci*, (American Society of Microbiology), pp 301-354.

56. Gilmore MS*, et al.* (1994) Genetic structure of the *Enterococcus faecalis* plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants. *J Bacteriol* 176(23):7335-7344.

57. Booth MC*, et al.* (1996) Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic. *Molecular microbiology* 21(6):1175-1184.

58. Gilmore M, Segarra R, & Booth M (1990) An HlyB-type function is required for expression of the *Enterococcus faecalis* hemolysin/bacteriocin. *Infection and immunity* 58(12):3914-3923.

59. Segarra R, Booth M, Morales D, Huycke M, & Gilmore M (1991) Molecular characterization of the *Enterococcus faecalis* cytolysin activator. *Infection and immunity* 59(4):1239-1246.

60. Haas W, Shepard BD, & Gilmore MS (2002) Two-component regulator of *Enterococcus faecalis* cytolysin responds to quorum-sensing autoinduction. *Letters to Nature* 415:84-87.

61. Coburn PS, Hancock LE, Booth MC, & Gilmore MS (1999) A Novel Means of Self-Protection, Unrelated to Toxin Activation, Confers Immunity to the Bactericidal Effects of the *Enterococcus faecalis* Cytolysin. *Infection and immunity* 67(7):3339-3347.

62. Labrie SJ, Samson JE, & Moineau S (2010) Bacteriophage resistance mechanisms. *Nature reviews. Microbiology* 8(5):317-327.

63. Weitz JS, Hartman H, & Levin SA (2005) Coevolutionary arms races between bacteria and bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* 102(27):9535-9540.

64. Makarova KS, Wolf YI, & Koonin EV (2013) Comparative genomics of defense systems in archaea and bacteria. *Nucleic acids research* 41(8):4360-4377.

65. Chen F*, et al.* (2012) Crystal Structure of the Cysteine Desulfurase DndA from *Streptomyces lividans* Which Is Involved in DNA Phosphorothioation. *PloS one* 7(5):e36635.

66. You D, Wang L, Yao F, Zhou X, & Deng Z (2007) A Novel DNA Modification by Sulfur: DndA Is a NifS-like Cysteine Desulfurase Capable of Assembling DndC as an Iron−Sulfur Cluster Protein in *Streptomyces lividans*. *Biochemistry* 46(20):6126-6133.

67. Koonin EV, Makarova KS, & Wolf YI (2017) Evolutionary Genomics of Defense Systems in Archaea and Bacteria. *Annual review of microbiology* (0).

68. Swarts DC*, et al.* (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* 507(7491):258-261.

69. Yamaguchi Y, Park JH, & Inouye M (2011) Toxin-antitoxin systems in bacteria and archaea. *Annual review of genetics* 45:61-79.

70. Chopin M-C, Chopin A, & Bidnenko E (2005) Phage abortive infection in lactococci: variations on a theme. *Current opinion in microbiology* 8(4):473-479.

71. Pingoud A, Wilson GG, & Wende W (2014) Type II restriction endonucleases—a historical perspective and more. *Nucleic acids research* 42(12):7489-7527.

72. Roberts RJ*, et al.* (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic acids research* 31(7):1805-1812.

73. Roberts RJ, Vincze T, Posfai J, & Macelis D (2007) REBASE—enzymes and genes for DNA restriction and modification. *Nucleic acids research* 35:D269-D270.

74. Tock MR & Dryden DT (2005) The biology of restriction and anti-restriction. *Current opinion in microbiology* 8(4):466-472.

75. Pingoud A & Jeltsch A (2001) Structure and function of type II restriction endonucleases. *Nucleic acids research* 29(18):3705-3727.

76. Pingoud A & Jeltsch A (1997) Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur J Biochem* 246:1-22.

77. Okhapkina S*, et al.* (2002) Comparison of the Homologous SfeI and LlaBI Restriction–Modification Systems. *Molecular Biology* 36(3):333-337.

78. Radliñska M, Piekarowicz A, Galimand M, & Bujnicki JM (2005) Cloning and Preliminary Characterization of a GATC-specific β2-class DNA:m6A Methyltransferase Encoded by Transposon Tn*1549* from *Enterococcus* spp. *Polish Journal of Microbiology* 54(3):249-252.

79. Wu R, King CT, & Jay E (1978) A new sequence-specific endonuclease from *Streptococcus faecalis* subsp. *zymogenes*. *Gene* 4(4):329-336.

80. Roberts RJ, Vincze T, Posfai J, & Macelis D (2009) REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic acids research* 38:D234-D236.

81. Huo W, Adams HM, Zhang MQ, & Palmer KL (2015) Genome Modification in *Enterococcus faecalis* OG1RF Assessed by Bisulfite Sequencing and Single-Molecule Real-Time Sequencing. *J Bacteriol* 197(11):1939-1951.

82. Clark TA*, et al.* (2011) Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. *Nucleic acids research* 40(4):e29-e29.

83. Flusberg BA*, et al.* (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature methods* 7(6):461-465.

84. Ishino Y, Shinagawa H, Makino K, Amemura M, & Nakata A (1987) Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of bacteriology* 169(12):5429-5433.

85. Jansen R, van Embden JDA, Gasstra W, & Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular microbiology* 43(6):1565-1575.

86. Bolotin A, Quinquis B, Sorokin A, & Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(Pt 8):2551-2561.

87. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, & Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of molecular evolution* 60(2):174-182.

88. Pourcel C, Salvignol G, & Vergnaud G (2005) CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151(Pt 3):653-663.

89. Grissa I, Vergnaud G, & Pourcel C (2007) The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC bioinformatics* 8:172.

90. Barrangou R*, et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709-1712.

91. Brouns SJJ*, et al.* (2008) Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* 321(5891):990-964.

92. Garneau JE*, et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468(7320):67-71.

93. Marraffini LA & Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322(5909):1843-1845.

94. Makarova KS*, et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nature reviews. Microbiology* 9(6):467.

95. Makarova KS*, et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nature reviews. Microbiology* 13(11):722-736.

96. Jinek M*, et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816-821.

97. Chylinski K, Le Rhun A, & Charpentier E (2013) The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA biology* 10(5):726-737.

98. Deltcheva E*, et al.* (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471(7340):602-607.

99. Heler R*, et al.* (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519(7542):199-202.

100. Wei Y, Terns RM, & Terns MP (2015) Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes & development* 29(4):356-361.

101. Mojica F, Diez-Villasenor C, Garcia-Martinez J, & Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155(3):733-740.

102. Anders C, Niewoehner O, Duerst A, & Jinek M (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513(7519):569-573.

103. Nishimasu H*, et al.* (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156(5):935-949.

104. Sapranauskas R*, et al.* (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic acids research* 39(21):9275-9282.

105. Sternberg SH, Redding S, Jinek M, Greene EC, & Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507(7490):62-67.

106. Hullahalli K*, et al.* (2015) Comparative Analysis of the Orphan CRISPR2 Locus in 242 *Enterococcus faecalis* Strains. *PloS one* 10(9):e0138890.

CHAPTER 2

CRISPR-CAS AND RESTRICTION MODIFICATION ACT ADDITIVELY AGAINST CONJUGATIVE ANTIBIOTIC RESISTANCE PLASMID TRANSFER IN *ENTEROCOCCUS FAECALIS*

2.1 Author contribution

This chapter was published as an original manuscript along with Wenwen Huo (WH), Ardalan Sharifi (AS) and Kelli L. Palmer (KP) in *mSphere* in 2016. The experiments for this work were conceived and designed by VP and KP. Data acquisition was carried out by VP, WH and AS. The analysis of results was performed by VP, WH and KP. The manuscript was written by VP and KP. The ASM Journals Statement of Authors' Rights allows the reuse of this full article as a part of the author's dissertation. Copyright © American Society for Microbiology, mSphere, volume 1, 2016, DOI: 10.1128/mSphere.00064-16.

2.2 Abstract

Enterococcus faecalis is an opportunistic pathogen and leading cause of nosocomial infections. Conjugative pheromone-responsive plasmids are narrow host range mobile genetic elements (MGEs) that are rapid disseminators of antibiotic resistance in the *faecalis* species. CRISPR-Cas and restriction-modification confer acquired and innate immunity, respectively, against MGE acquisition in bacteria. Most multidrug-resistant *E. faecalis* lack CRISPR-Cas and possess an orphan locus lacking *cas* genes, CRISPR2, that is of unknown function. Little is known about restriction-modification defense in *E. faecalis*. Here, we explore the hypothesis that multidrug-

resistant *E. faecalis* are immune-compromised. We assessed MGE acquisition by *E. faecalis* T11, a strain closely related to the multidrug-resistant hospital isolate V583, but which lacks the ~620 kb of horizontally acquired genome content that characterizes V583. T11 possesses the *E. faecalis* CRISPR3-*cas* locus and a predicted restriction-modification system, neither of which occur in V583. We demonstrate that CRISPR-Cas and restriction-modification together confer a four-log reduction in acquisition of the pheromone-responsive plasmid pAM714 in biofilm matings. Additionally, we show that the orphan CRISPR2 locus is functional for genome defense against another pheromone-responsive plasmid, pCF10, only in the presence of *cas9* derived from the *E. faecalis* CRISPR1-*cas* locus, which most multidrug-resistant *E. faecalis* lack. Overall, our work demonstrates that the loss of only two loci leads to a dramatic reduction in genome defense against a clinically relevant MGE, highlighting the critical importance of the *E. faecalis* accessory genome in modulating horizontal gene transfer. Our results rationalize the development of antimicrobial strategies that capitalize upon the immune-compromised status of multidrug-resistant *E. faecalis*.

2.3 Importance

Enterococcus faecalis is a bacterium that normally inhabits the gastrointestinal tracts of humans and other animals. Although these bacteria are members of our native gut flora, they can cause life-threatening infections in hospitalized patients. Antibiotic resistance genes appear to be readily shared among high-risk *E. faecalis* strains, and multidrug resistance in these bacteria limits treatment options for infections. Here, we find that CRISPR-Cas and restriction-modification systems, which function as adaptive and innate immune systems in bacteria,

significantly impact the spread of antibiotic resistance genes in *E. faecalis* populations. The loss of these systems in high-risk *E. faecalis* suggests that they are immune-compromised, a trade-off that allows them to readily acquire new genes and adapt to new antibiotics.

2.4 Introduction

Enterococcus faecalis is a Gram-positive bacterium that normally colonizes the gastrointestinal (GI) tracts of humans and other animals (1) and opportunistically colonizes wounds and the bloodstream, leading to the life-threatening infections bacteremia and endocarditis (1-3). Since the 1980s, *E. faecalis* have become increasingly associated with nosocomial (hospital-acquired) infections (4-6).

E. faecalis appear to have a remarkable propensity for acquisition of antibiotic resistance genes by horizontal gene transfer (HGT). Mobile genetic elements (MGEs) such as conjugative and mobilizable plasmids and transposons are common in *E. faecalis* clinical isolates. They encode resistance to vancomycin, aminoglycosides, tetracycline, chloramphenicol, ampicillin, linezolid and other antibiotics (7-13). Vancomycin-resistant *E. faecalis* are of particular concern and have been deemed serious public health threats by the United States Centers for Disease Control and Prevention (14). The emergence of HGT-acquired antibiotic resistance in *E. faecalis* is an ongoing problem that will limit the usefulness of future antibiotics. A unique group of narrow host range conjugative plasmids called the pheromone-responsive plasmids (PRPs) are rapid disseminators of antibiotic resistance, cytolytic toxin biosynthesis, and other virulence traits

among *E. faecalis* but cannot replicate outside the species (8, 15-17). The *in vivo* transfer frequency of PRPs occurs on the order of one transconjugant per 10-100 donor cells (18-20).

Genome analyses indicate that multidrug-resistant (MDR) *E. faecalis* strains are undergoing HGT-driven genome expansion (21-25). Exemplary of this, one fourth of the 3.36 Mb genome of *E. faecalis* V583, a hospital infection isolate from 1987 that was among the first vancomycin-resistant enterococci to be reported (26, 27), was acquired by HGT (23, 26). V583 originates from one of a group of high-risk enterococcal clonal complexes that are associated with nosocomial infections and are commonly resistant to multiple antibiotics (28, 29). In comparison to V583, the genome of the 1992 inherently drug-susceptible urinary tract isolate *E. faecalis* T11 is only 2.74 Mbp (21, 23). V583 and T11 share 99.5% average nucleotide sequence identity in their core genomes, thus these strains are very closely related. However, V583 has an additional ~620 kb of HGT-acquired content (21, 30). V583 and T11 are useful comparators for understanding the impacts of HGT on enterococcal biology.

In previous work, we proposed a model for the emergence of MDR, genome-expanded *E. faecalis* (30). Our hypothesis is that these strains lack or have lost endogenous barriers to HGT. Antibiotic use inadvertently selects for outgrowth of these immune-compromised strains with enhanced abilities to acquire MGEs, thereby assisting their rapid adaptation to the GI tracts of antibiotic-treated patients and the hospital environment.

CRISPR-Cas systems (clustered, regularly interspaced, short palindromic repeats with CRISPR-associated genes) are genome defense systems that are endogenous barriers to HGT in bacteria. CRISPR loci consist of short repeat sequences interspersed by unique spacer sequences (31, 32). A set of genes (*cas* genes) encoding nucleases are typically encoded near the CRISPR (33). Type II CRISPR-Cas loci consist of a CRISPR array, the type-specific gene *cas9*, and *cas1* and *cas2* genes (34, 35) (see Figure A.1 in Appendix A). The mechanism for Type II CRISPR-Cas genome defense has been recently reviewed (36) and is summarized here. When cells with Type II CRISPR-Cas are challenged with MGEs, some cells incorporate a short segment (protospacer) of the invading MGE genome into the CRISPR as a novel spacer; this is the adaptation phase. By this mechanism, the CRISPR serves as a heritable memory of MGE encounters. Short sequence motifs adjacent to protospacers, called protospacer adjacent motifs (PAMs), as well as the Cas nucleases are required for adaptation. To provide immunity to MGEs, the CRISPR is transcribed into a pre-crRNA and processed to mature crRNAs using RNase III, Cas9, and a trans-activating crRNA (tracrRNA) that has sequence complementarity to CRISPR repeats. This is the expression phase. If a MGE possessing the protospacer and PAM enters the cell, the Cas9 nuclease is directed to the MGE genome by a crRNA/tracrRNA complex with sequence complementarity to the protospacer. The HNH endonuclease domain of Cas9 cleaves the complementary protospacer strand, and the RuvC endonuclease domain of Cas9 cleaves the non-complementary protospacer strand, generating a double-strand DNA break in the invading MGE. This is the interference phase. In summary, Type II CRISPR-Cas systems provide adaptive immunity against MGEs.

Two Type II CRISPR-Cas systems, called CRISPR1-Cas and CRISPR3-Cas, occur with variable distribution across the *faecalis* species (22, 30, 37-39). There is an additional Type II locus, CRISPR2, that lacks associated *cas* genes but whose presence is conserved across the species (see Figure A.1 in Appendix A) (39). There is a striking relationship between HGT-acquired antibiotic resistance and CRISPR-Cas presence in *E. faecalis*. Specifically, most multidrug-resistant *E. faecalis* lack CRISPR-Cas and only possess the orphan CRISPR2 (30, 39). This suggests that CRISPR-Cas systems, by acting as barriers to MGE acquisition, are antagonistic to the evolution of multidrug resistance in *E. faecalis*. However, a role for CRISPR-Cas in *E. faecalis* genome defense has yet to be experimentally demonstrated.

Restriction-modification (R-M) systems provide another form of genome defense by acting as barriers to HGT through self vs. non-self recognition of methylation signatures. In R-M defense, a cell modifies its 'self' DNA at specific sequence motifs. Common modifications conferred by DNA methyltransferases (MTases) are 6-methyladenine (m6A), 4-methylcytosine (m4C), and 5-methylcytosine (m5C) (40). Restriction endonucleases (REases) recognize and degrade non-modified 'non-self' DNA (41, 42). In previous work, we studied R-M systems in the model *E. faecalis* strain OG1RF (43). We determined that *E. faecalis* OG1RF possesses a Type II R-M system, EfaRFI, that is capable of providing modest but significant defense against the PRP pCF10 (43). Additional analysis of 17 *E. faecalis* strains revealed that no core R-M systems occur in the species, signifying that these systems occur within the accessory genome of *E. faecalis*.

In this study, we used *E. faecalis* T11 as a model to assess roles of CRISPR3-Cas and the orphan CRISPR2 locus in genome defense against PRPs. We also evaluated synergism between two types of genome defense, R-M and CRISPR-Cas. By using conjugation assays and the model PRPs pAM714 and pCF10, we demonstrate that CRISPR3-Cas is active for sequence-specific genome defense. Our results also demonstrate that CRISPR-Cas and R-M, together, provide additive defense for the cell, with a striking 4-log difference in plasmid acquisition frequencies between strains equipped with or deficient for CRISPR-Cas and R-M defense. Our analysis of the orphan CRISPR2 locus revealed that this locus requires CRISPR1-Cas encoded factors in order to provide genome defense and cannot provide defense against MGEs on its own. Overall, our results are significant because they support the hypothesis that MDR hospital *E. faecalis* are immune-compromised.

2.5 Materials and Methods

Bacteria and reagents used.

Strains and plasmids used in this study are shown in Table 1. *E. faecalis* T11 is natively sensitive to antibiotics, therefore, a rifampicin- and fusidic acid-resistant derivative was isolated by sequential exposure to the antibiotics at 50 μg/mL and 25 μg/mL, respectively. *E. faecalis* strains were cultured in brain heart infusion (BHI) broth or agar at 37°C, unless otherwise stated. Antibiotic concentrations for *E. faecalis* were as follows: rifampicin, 50 μg/mL; fusidic acid, 25 μg/mL; spectinomycin, 500 μg/mL; streptomycin, 500 μg/mL; chloramphenicol, 15 μg/mL; tetracycline, 10 μg/mL; erythromycin, 50 μg/mL. *Escherichia coli* strains were cultured in lysogeny broth (LB) with aeration at 225 rpm or LB agar at 37°C, unless otherwise stated.

Antibiotic concentrations for *E. coli* are as follows: chloramphenicol, 15 μg/mL. Antibiotics were purchased from Sigma-Aldrich. Restriction enzymes were purchased from New England Biolabs and used according to manufacturer protocols. Routine PCR analysis was performed using *Taq* Polymerase (New England Biolabs). PCR for cloning procedures utilized Phusion Polymerase (Fisher Scientific). Plasmid isolation was performed using the GeneJET Plasmid Miniprep Kit (Thermo Scientific). PCR products and restriction digestion reactions were purified using the GeneJET PCR Purification Kit (Thermo Scientific). DNA sequencing was performed at the Massachusetts DNA Core Facility (Boston, MA). Primers used in this study are shown in Table A.1 of Appendix A.

Spacer analysis of *E. faecalis* T11.

The T11 CRISPR3-*cas* and CRISPR2 spacer sequences were used as queries in BLASTn analysis against the NCBI non-redundant nucleotide database. A significance threshold of 86% sequence identity, which allows four mismatches between the query and subject, was used to identify protospacer candidates.

Generation of T11RF strains used in this study.

In-frame deletions of CRISPR3 *cas9* and CRISPR3 spacer 6 were generated using a previously established protocol (44). Briefly, ~1 kb regions up- and downstream of *cas9* or CRISPR3 spacer 6 in *E. faecalis* T11RF were amplified, digested, and ligated into pLT06 (44) to generate pVP102 and pAS106, respectively. The resulting plasmids were transformed into competent T11RF cells via electroporation (45) and cultured at the permissive temperature of 30°C. Following

transformation, a shift to the non-permissive temperature of 42°C and counterselection on p-chloro-phenylalanine were performed to generate in-frame, markerless deletions. The predicted RuvC and HNH nuclease domain coding regions of CRISPR3 *cas9* were mutated such that residues D7 and H601 were changed to alanine. This was accomplished by amplifying ~1 kb arms up- and downstream of the codons for the 7th and 601st amino acids, but instead of using a restriction site to connect the two arms, overlapping sequence on the internal primers was used to generate the amino acid coding change (underlined in Table A.1 of Appendix A), generating T11RF*cas9*D7A and T11RF*cas9*H601A. Sequencing was used to confirm all modified regions.

Complementation of the *cas9* deletion was accomplished by integrating the gene into a neutral site on the T11 chromosome at a location between ORFs EFMG_00904 and EFMG_00905. pWH03, a derivative of pLT06 containing ~1 kb arms corresponding to the genes at this site, was used as the backbone vector for insertion of T11 CRISPR3 *cas9* (pVP301) as well as ATCC 4200 CRISPR1 *cas9* (pG19) into the T11RFΔ*cas9* strain. The putative promoter and predicted tracrRNA was included in the complementation constructs for both CRISPR3 *cas9* and CRISPR1 *cas9*, generating strains T11RFΔ*cas9*+CR3 and T11RFΔ*cas9*+CR1; the entire integrated region was confirmed by sequencing.

Generation of OG1SSp mutants.

The EfaRFI R-M system was deleted in OG1SSp pAM714 as in previous work using the pLT06 derivative, pWH01 (43); the deletion was confirmed by sequencing, resulting in strain VP701. Complementation was performed via knock-in of EfaRFI at the neutral locus. Briefly,

OG1RF_11622-OG1RF_11621 including its putative promoter region was ligated into pWH03, resulting in pWH43. pWH43 was electroporated into competent VP701 cells and temperature shift and counter-selection were used as described above to generate WH702; the insertion was confirmed by sequencing.

Generation of pCF10 mutants.

To insert the T11 CRISPR2 spacer 1 sequence into pCF10, 100 bp single stranded DNA oligos were annealed to each other to generate dsDNA. The 100 bp oligos included sequence from pCF10 *uvrB*, the spacer 1 sequence, and either a CRISPR1/2 PAM or CRISPR3 PAM. Annealed oligos were sub-cloned into the pGEM T-Easy vector (Promega) for amplification and ligation into pLT06 derivatives designed to insert these sequences into the *uvrB* gene of pCF10 by homologous recombination. See Figure A.5 in Appendix A for a schematic of constructs used to generate strains OG1SSp pVP501 and OG1SSp pVP502.

R-M system prediction in T11.

E. faecalis T11 contigs were downloaded from the Broad Institute (*Enterococcus* I initiative, broadinstitute.org) and annotated using RAST (46, 47). Protein sequences were blasted against the NEB REBASE gold standards list. Using a bit score cutoff of 60 for MTase identity to the gold standard list, we predicted only one MTase in T11 (EFMG_00924), which is also a homolog of M.EfaRFI (sequence identity 56%; query coverage 93%; e-value $2e^{-125}$).

Conjugation experiments.

For all conjugation reactions, donor and recipient strains were cultured overnight in BHI broth without antibiotic selection. The next day, cultures were diluted 1:10 into fresh BHI and incubated at 37°C for 1.5 hours. Next, 100 μL donor culture was mixed with 900 μL recipient culture and pelleted at 13, 000 rpm for 1 min. 100 μL of supernatant was used to resuspend the pellet, which was then plated on BHI agar and incubated at 37°C for 18 hours. Cells were collected from the plate with 2 mL 1X PBS supplemented with 2 mM EDTA. Dilutions were plated on BHI agars supplemented with antibiotics to quantify donor (spectinomycin and streptomycin with either erythromycin or tetracycline), recipient (rifampicin and fusidic acid), or transconjugant (rifampicin and fusidic acid with either erythromycin or tetracycline) populations. Plates were incubated for 36-48 hours at 37°C to allow colonies to develop. Plates with 30 to 300 colonies were used to calculate CFU/mL. Conjugation frequency was determined by dividing the number of transconjugants by the number of donors.

PAM identification.

Strains with complete CRISPR arrays (no sequence gaps) were used to identify putative PAMs for the three *E. faecalis* CRISPR loci. Protospacers were identified as described above. 15 nucleotides downstream of the protospacer sequence were extracted and subjected to motif detection using MEME (48). The same CRISPR2 spacer sequences often occur in multiple strains (39), therefore spacer hits to CRISPR2 loci were manually curated from the analysis so that a CRISPR2 spacer was not overrepresented.

2.6 Results

CRISPR3-Cas is a genome defense system in *E. faecalis*.

E. faecalis T11 is closely related to the hospital strain V583, but lacks the multidrug resistance and HGT-driven genome expansion that is characteristic of V583 (21). T11 possesses CRISPR3-Cas and the orphan CRISPR2 (30). Spacer 6 of the T11 CRISPR3 locus is identical to *repB* sequence from the model 60-kb pheromone-responsive plasmid pAD1 (30). The T11 CRISPR3 locus is shown in Figure 2.1 and an analysis of T11 CRISPR3 spacer identities is shown in Table A.2 of Appendix A. By aligning protospacers and adjacent sequences, the CRISPR3 PAM sequence was found to be NNRTA (Figure A.2 and Table A.2 of Appendix A).

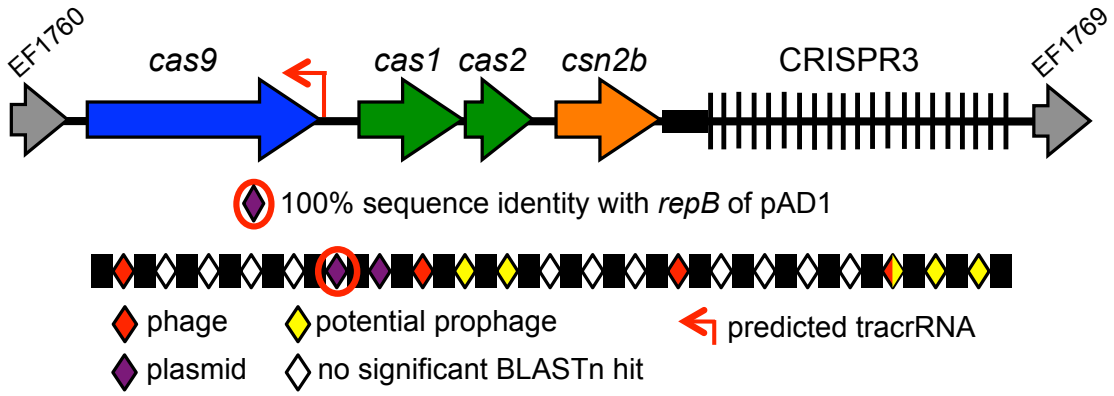


Figure 2.1. CRISPR3-*cas* locus of *E. faecalis* T11. The CRISPR3 locus of T11 consists of 21 unique spacer sequences of 30 nucleotides (diamonds) flanked by direct repeat sequences of 36 nucleotides each (rectangles); the entire set of repeats and spacers are expanded below the locus for clarity. Spacers having significant identity with MGEs (Table A.2 of Appendix A) are colored based on the type of genetic element they have identity with: red, phage; purple, plasmids; yellow, potential prophage. Grey arrows denote V583 gene orthologs. The red arrow between *cas9* and *cas1* represents the predicted location of the CRISPR3 tracrRNA. The black rectangle upstream of the CRISPR array represents the leader region. The designation of *csn2b* is based off of (49).

We tested the hypothesis that T11 CRISPR3-Cas interferes with pAD1 acquisition, using conjugation assays with *E. faecalis* OG1SSp as a plasmid donor and with T11 and its derivatives as plasmid recipients (see Table 2.1 for a list of plasmids and strains used in this study). T11 was passaged to create a rifampicin- and fusidic acid-resistant derivative for use in conjugation experiments (referred to as T11RF). Deletion of CRISPR3 *cas9* from T11RF resulted in a significant increase in acquisition of a pAD1 derivative conferring erythromycin resistance (pAM714 (50, 51)) in plate (biofilm) matings (Figure 2.2), providing evidence that CRISPR3-Cas is active for genome defense in this strain. This increase in conjugation frequency was not observed for the 67-kb pheromone-responsive plasmid pCF10, which is not targeted by CRISPR3 spacers (Figure 2.2). We complemented the T11RF CRISPR3 *cas9* deletion with T11 CRISPR3 *cas9* (Δ*cas9*+CR3) at a neutral site on the T11 chromosome. However, complementation was not observed upon integration of *cas9* derived from the *E. faecalis* ATCC 4200 CRISPR1-Cas locus (Δ*cas9*+CR1) (Figure 2.2). Deletion of CRISPR3 spacer 6 (ΔCR3S6) resulted in an increase in conjugation frequency similar to what was observed for the *cas9* deletion, confirming that the CRISPR is required for genome defense. Finally, alignment with the *Streptococcus pyogenes* Cas9 (SpCas9) and *S. aureus* Cas9 (SaCas9) sequences was used to predict the locations of the RuvC and HNH endonuclease domains of *E. faecalis* CRISPR3 Cas9 (EfCR3Cas9; Figure A.3 of Appendix A). Single amino acid substitutions were made in these two domains of EfCR3Cas9, generating a D7A substitution in the RuvC-I domain (*cas9*D7A) and a H601A substitution in the HNH domain (*cas9*H601A). These positions correspond to D10 and H557 in SaCas9, for which D10A and H557A substitutions result in a loss of DNA cleavage activity (52), and D10 and H840 in SpCas9, for which D10A substitution results in a loss of

protospacer non-complementary strand cleavage and H840A substitution results in a loss of protospacer complementary strand cleavage (53). A final strain, *cas9*DM, was generated that possessed both substitutions. Conjugation frequencies obtained with these strains as recipients were similar to the *cas9* deletion mutant (Figure 2.2), implicating these residues as active sites in EfCR3Cas9. Further, that the *E. faecalis* Cas9 D7A and H601A substitutions have equivalent impacts on pAD1 acquisition suggests that double-stranded pAD1 DNA is required for PRP interference by CRISPR3-Cas. These experiments establish that CRISPR3-Cas is a sequence-specific genome defense system in *E. faecalis* T11.

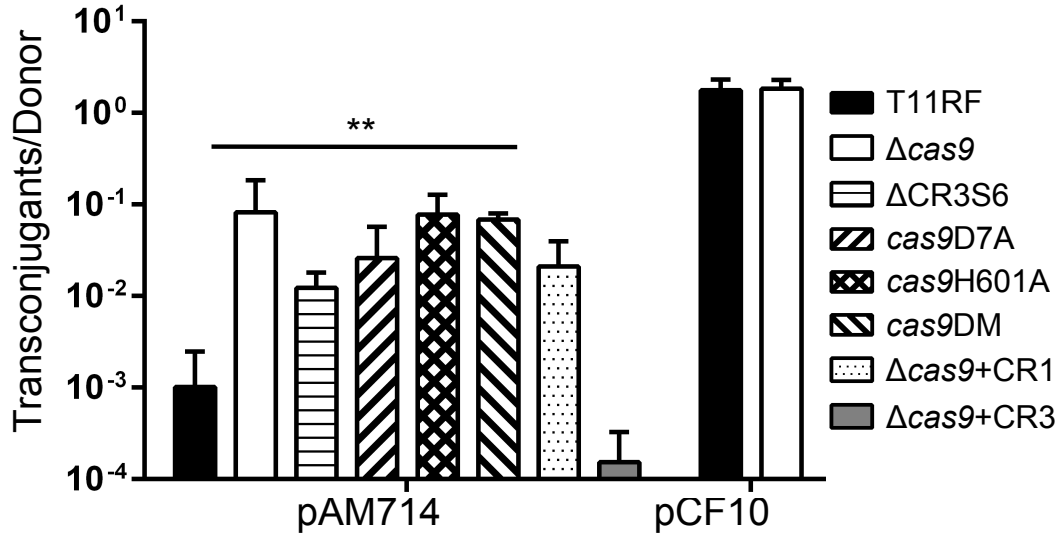


Figure 2.2. *E. faecalis* CRISPR3-*cas* provides sequence-specific defense against PRPs. Conjugation frequencies between *E. faecalis* OG1SSp harboring either pAM714 (left) or pCF10 (right) and T11RF and its derivatives. Conjugation frequency is represented by the ratio of transconjugants to donors in mating reactions. The pAM714 conjugation frequency is significantly higher for recipients that lack *cas9* (Δ*cas9*), lack CRISPR3 spacer 6 (ΔCR3S6), or have mutations in either (*cas9*D7A; *cas9*H601A) or both (*cas9*DM) RuvC and HNH endonuclease coding regions of *cas9*. Complementation was observed with CRISPR3 *cas9* (Δ*cas9*+CR3) but not with CRISPR1 *cas9* (Δ*cas9*+CR1). Data represent a minimum of 3 independent mating experiments. Significance was assessed using a one-tailed Student's t-Test; *P*-values are relative to T11RF: **, $P<0.005$.

Table 2.1. Plasmids and Strains.

Strain or Plasmid Name	Description	Reference
***E. coli* strains**		
EC1000	*E. coli* cloning host, providing *repA* in trans, for pLT06 and pGEM-T-Easy derived plasmids	(54)
***E. coli* plasmids**		
pGEM T-Easy	Plasmid containing T-overhangs in MCS, used for subcloning of DNA fragments for mutant generation in *E. faecalis*	Promega
pLT06	Markerless exchange plasmid; confers chloramphenicol resistance	(44)
pWH03	Derivative of pLT06 containing OG1RF_11778 and OG1RF_11789 for integration into neutral site on chromosome	(43)
pVP102	Derivative of pLT06 to create markerless, in-frame deletion of CRISPR3-*cas9* in T11RF	This study
pAS106	Derivative of pLT06 to create deletion of spacer 6 in CRISPR3 locus of T11RF	This study
pVP105	Derivative of pLT06 to change amino acid 7 of T11 CRISPR3 Cas9 from aspartic acid to alanine	This study
pG19	Derivative of pWH03 to integrate the CRISPR1-*cas9* gene, its native promoter, and predicted tracrRNA into the T11 chromosome between EFMG_00904 and EFMG_00905	This study
pVP301	Derivative of pWH03 to integrate the CRISPR3-*cas9* gene, its native promoter, and predicted tracrRNA into the T11 chromosome between EFMG_00904 and EFMG_00905	This study
pWH01	Derivative of pLT06 to create markerless, in-frame deletion of OG1RF_11621-OG1RF_11622 in OG1SSp	(43)
pWH43	Derivative of pWH03 to integrate OG1SSp OG1RF_11621-OG1RF_11622 and its native promoter into the chromosome between OG1RF_11778 and OG1RF_11789	This study
pVP401	Derivative of pGEM-T-Easy with 100 bp insert including T11 CRISPR2 spacer 1 and the consensus CRISPR2 PAM	This study

pVP107	Derivative of pLT06 to knock-in the T11 CRISPR2 spacer 1 sequence and consensus CRISPR2 PAM into the *uvrB* gene of pCF10	This study
pVP402	Derivative of pGEM-T-Easy with 100 bp insert including T11 CRISPR2 spacer 1 and the consensus CRISPR3 PAM	This study
pVP108	Derivative of pLT06 to knock-in T11 CRISPR2 spacer 1 and the consensus CRISPR3 PAM into the *uvrB* gene of pCF10	This study
pVP109	Derivative of pLT06 to change amino acid 601 of T11 CRISPR3 Cas9 from histidine to alanine	This study
***E. faecalis* strains**		
T11RF	Rifampicin-Fusidic acid resistant derivative of T11	(23) and this study
T11RFΔ*cas9*	T11RF CRISPR3-cas9 deletion mutant	This study
T11RFΔ*cas9*+CR3	T11RFΔ*cas9* with chromosomal integration of CRISPR3 *cas9* between EFMG_00904 and EFMG_00905	This study
T11RFΔ*cas9*+CR1	T11RFΔ*cas9* with chromosomal integration of CRISPR1 *cas9* and the predicted CRISPR1 tracrRNA between EFMG_00904 and EFMG_00905	This study
T11RFΔCR3S6	T11RF with a deletion of CRISPR3 spacer 6	This study
T11RF*cas9*D7A	T11RF with chromosomal mutation in the RuvC nuclease coding region of *cas9*	This study
T11RF*cas9*H601A	T11RF with chromosomal mutation in the HNH nuclease coding region of *cas9*	This study
T11RF*cas9*DM	T11RF with chromosomal mutations in the predicted RuvC and HNH nuclease coding regions of *cas9*	This study
OG1SSp pAM714	Spectinomycin-Streptomycin resistant derivative of OG1 harboring pAM714, an erythromycin (carried on Tn*917*) resistant derivative of pAD1	(50, 51)
VP701	OG1SSp pAM714 EfaRFI deletion mutant	This study
WH702	VP701 with chromosomal integration of EfaRFI (OG1RF_11621-OG1RF_11622) and its native promoter	This study

	between OG1RF_11778 and OG1RF_11789	
OG1SSp pCF10	Spectinomycin-Streptomycin resistant derivative of OG1 harboring pCF10 encoding tetracycline resistance on Tn*925*	(55)
VP703	OG1SSp pCF10 EfaRFI deletion mutant	This study
OG1SSp pVP501	OG1SSp pCF10 with insertion of T11 CRISPR2 spacer 1 and consensus CRISPR2 PAM into *uvrB* of pCF10	This study
OG1SSp pVP502	OG1SSp pCF10 with insertion of T11 CRISPR2 spacer1 and consensus CRISPR3 PAM into *uvrB* of pCF10	This study

The relative contributions of R-M and CRISPR-Cas defense in *E. faecalis* T11.

In a previous study, we determined that the genomes of *E. faecalis* OG1RF, OG1SSp, and T11 are modified by 5'-G^{m5}CWGC-3' (43). Deletion of EfaRFI, the R-M system responsible for 5'-G^{m5}CWGC-3' modification in OG1RF and OG1SSp, significantly but modestly (~3-fold) reduced pCF10 conjugation frequency between OG1RF mutant cells and OG1SSp (43). Using the same strategy for MTase identification as in our previous study, we predicted only one MTase in the T11 genome (EFMG_00924), and it has 56% amino acid sequence identity with M.EfaRFI (Table A.3 of Appendix A). We infer that this MTase is responsible for the 5'-G^{m5}CWGC-3' DNA modification observed for T11 (43). However, the prediction of the corresponding REase for the T11 M.EfaRFI homolog is not straightforward as there are four genes surrounding the MTase that have conserved endonuclease domains, three of which are predicted to recognize ^{m5}C signatures (Figure 2.3A and Table A.3 of Appendix A), only one of which has high amino acid sequence identity with R.EfaRFI (EFMG_00925, 43% identity). Analysis of the *faecalis* pan genome revealed that this region occurs in a subset of strains with available genome sequence (*E. faecalis* T11, B301, B345, B347 and T19). Synteny analyses with T11 and V583 suggest that these accessory genes were displaced in V583 by a transposon carrying the *vanB* vancomycin resistance cassette (23, 26).

The impact of DNA modification on plasmid transfer into T11 was assessed by conjugative transfer of pCF10 from OG1SSp donor strains with (OG1SSp pCF10) or without (VP703) EfaRFI. For OG1SSp pCF10 donors, the plasmid is modified by 5'-G^{m5}CWGC-3' and should be recognized as 'self' by the T11 R-M system. For VP703 donors, the plasmid is not modified by

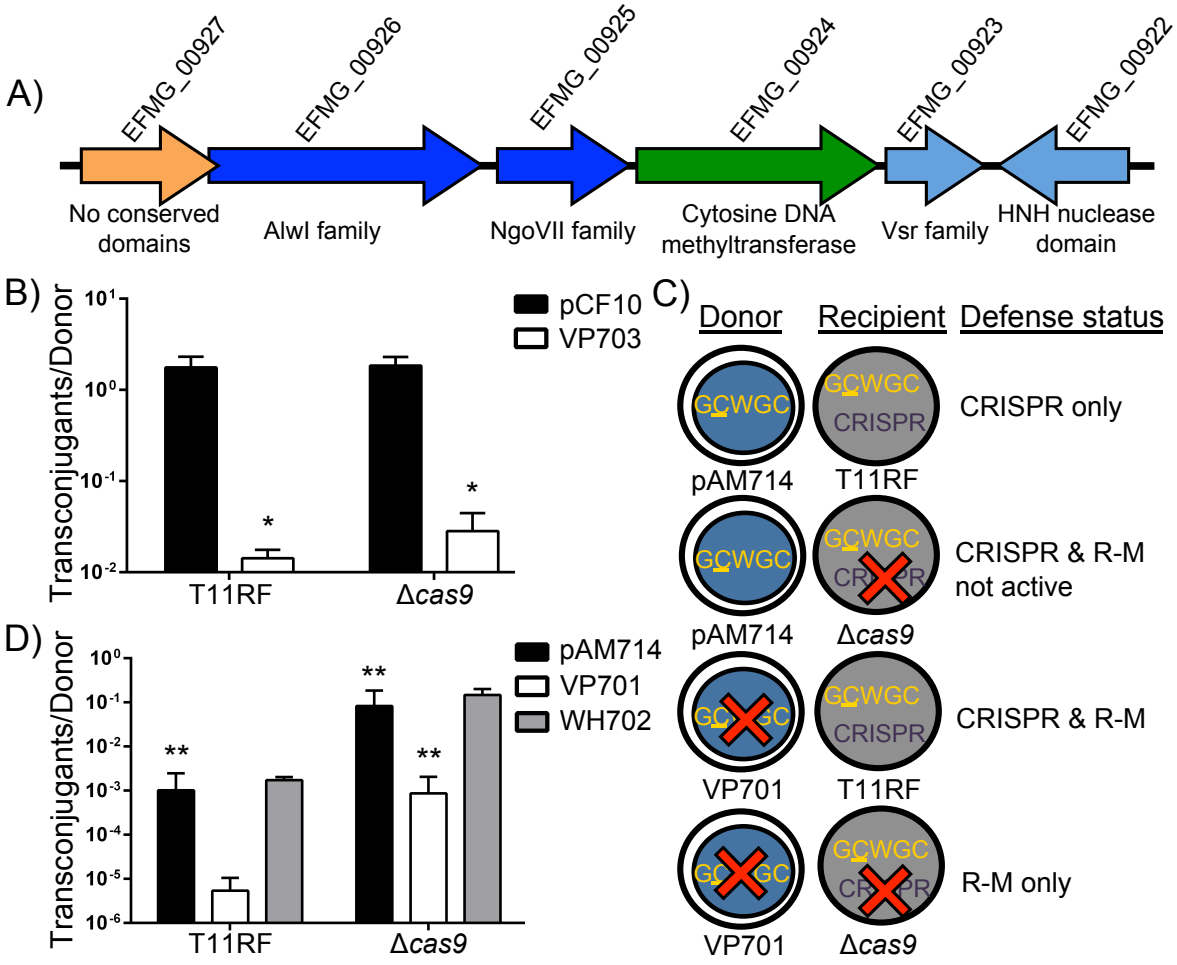


Figure 2.3. CRISPR-Cas and R-M provide additive defense against PRPs in *E. faecalis*. A) Organization of the predicted R-M locus of T11; multiple predicted REases are encoded near the MTase. B) Conjugation frequencies with T11RF and T11RFΔ*cas9* strains as recipients in mating reactions with OG1SSp pCF10 and VP703 as donors. P-value relative to transfer of OG1SSp pCF10 to T11RF: *, $P<0.05$. C) Schematic representing donor and recipient strains used to assess the individual and collective contributions of R-M and CRISPR-Cas on genome defense. D) Conjugation frequencies with T11RF and T11RFΔ*cas9* strains as recipients (x-axis) and with OG1SSp pAM714 (black columns), OG1SSp pAM714 ΔEfaRFI (VP701, white columns) and OG1SSp pAM714 ΔEfaRFI + EfaRFI (WH702, grey columns) as donors. Frequencies are shown as the ratio of transconjugants to donors. Results of these experiments show that the combined effect of CRISPR-Cas and R-M outweighs the effect of either system alone. Data represent a minimum of three independent conjugations for all experiments shown. P-values are relative to transfer of pAM714 from VP701 to T11RF: **, $P<0.005$. Significance in B) and D) was assessed using a one-tailed Student's t-Test.

5’-G^{m5}CWGC-3’ and should be recognized as 'non-self' by the T11 R-M system. Abolishment of DNA modification in the donor strain resulted in a 124-fold reduction in pCF10 plasmid transfer into T11 (Figure 2.3B). This effect is much more pronounced than the 3-fold decrease in pCF10 transfer observed in a previous study for the EfaRFI system (43), suggesting that the T11 R-M system possesses features that provide more robust genome defense than EfaRFI.

Next, we sought to determine whether CRISPR-Cas and R-M confer additive genome defense in *E. faecalis* T11. pAM714 possesses 59 GCWGC motifs, none of which overlap the protospacer and PAM sequences in *repB*. pAM714 is expected to be modified with 5’-G^{m5}CWGC-3’ by OG1SSp donor strains. For experiments shown in Figure 2.2, pAM714 transferred from OG1SSp to T11 was modified by 5’-G^{m5}CWGC-3' and recognized as 'self' DNA by the T11 R-M system. Therefore CRISPR3-Cas but not R-M defense was active in that condition. We modulated self versus non-self signals at 5’-GCWGC-3' motifs in the donor strain to determine individual and collective impacts of R-M and CRISPR-Cas defense on pAM714 acquisition. The design of these experiments is shown in Figure 2.3C. Donor strains used were OG1SSp pAM714, an OG1SSp pAM714 derivative with a deletion of EfaRFI (strain VP701), and a VP701 complement strain with EfaRFI genes integrated into a neutral site on the chromosome (WH702). When both CRISPR-Cas and R-M defense are active, the average conjugation frequency (expressed as transconjugants/donors) is 5.4 x 10^{-6}; we used this value as a reference for comparisons (Figure 2.3D). When CRISPR-Cas defense has been compromised by the loss of *cas9*, but R-M defense is active, the average conjugation frequency is 8.7 x 10^{-4}, a 160-fold increase in plasmid transfer. When R-M defense is not active due to the incoming plasmid being

modified as 'self,' but CRISPR-Cas defense is active, the average conjugation frequency is 1 x 10^{-3}, a 188-fold increase in plasmid transfer. When neither defense system is active, the average conjugation frequency is 8.25 x 10^{-2}, a 15,277-fold increase in plasmid transfer. Overall, we conclude that R-M and CRISPR-Cas, both individually and collectively, have significant impacts on conjugative plasmid transfer in *E. faecalis* T11.

T11 CRISPR2 does not provide genome defense unless CRISPR1 Cas9 is present.

An orphan CRISPR locus lacking *cas* genes and of varying spacer composition, called CRISPR2, occurs in all *E. faecalis* genomes, including multidrug-resistant strains (30, 39). The consensus repeats of CRISPR2 and CRISPR1-Cas are identical, suggesting they are functionally linked (Figure A.4 of Appendix A). The repeat sequences of CRISPR3 are only 58% identical to those of CRISPR1/CRISPR2 (Figure A.4 of Appendix A). In previous work, we hypothesized that CRISPR2 is inactive for genome defense in strains lacking CRISPR1-Cas; i.e., high-risk lineages (30). An alternative hypothesis that would explain the conservation of CRISPR2 is that CRISPR2 confers genome defense by a Cas-independent mechanism. We used T11 as a model strain to determine whether CRISPR2 can confer genome defense alone or in conjunction with CRISPR-Cas-encoded factors.

The spacer content of CRISPR1 and CRISPR2 loci of six *E. faecalis* strains was used to determine their respective PAM sequences, which are predicted to be identical (NGG, Figure A.2 and Table A.2 of Appendix A). The CRISPR2 of T11 possesses 4 spacers that lack identity to known MGEs, but are identical to spacers that occur in CRISPR2 loci of other *E. faecalis* strains,

two of which are present in the CRISPR2 of V583 (39). We inserted a protospacer identical to T11 CRISPR2 spacer 1, along with a NGG PAM sequence (for CR1 and CR2) or a NNRTA PAM sequence (for CRISPR3; Figure A.2 and Table A.2 of Appendix A) into pCF10, generating pVP501 and pVP502, respectively (Figure 2.4A and Figure A.5 of Appendix A). The integration of the same protospacer with either of two different PAM sequences is to assess Cas9 specificity to its cognate target recognition motif. We then evaluated conjugative transfer of these two plasmids and wild-type pCF10 from OG1SSp to T11RF and its derivatives (Figure 2.4B).

As previously shown (Figures 2.2 and 2.3), there is no significant change in conjugation frequency of pCF10 between T11RF and T11RFΔ*cas9* (Figure 2.4C). Moreover, the addition of the T11 CRISPR2 spacer 1 and PAM sequences into pCF10 had no effect on conjugation frequency in T11RF or T11RFΔ*cas9.* We conclude that under these conditions, CRISPR2 alone cannot provide defense in the presence of a protospacer target and the predicted PAM. We then set out to determine if the presence of the *E. faecalis* CRISPR1 *cas9* and its predicted tracrRNA would impact plasmid transfer. To test this, we integrated the CRISPR1 *cas9* and predicted tracrRNA coding regions into a neutral site on the T11RFΔ*cas9* chromosome. A 6-fold reduction in conjugation frequency was observed between mating of pVP501 to T11RF Δ*cas9* and pVP501 to T11RF Δ*cas9*+CR1, revealing that CRISPR2 requires CRISPR1-Cas factors to provide genome defense. Finally, no change in conjugation frequency was observed when using pVP502. This result, in conjunction with observing a similar conjugation frequency of pVP502 into T11RF, provides experimental evidence that support the predicted PAM for CRISPR1/CRISPR2.

These results demonstrate a functional linkage between CRISPR1-Cas and CRISPR2 through CRISPR1-Cas encoded factors.

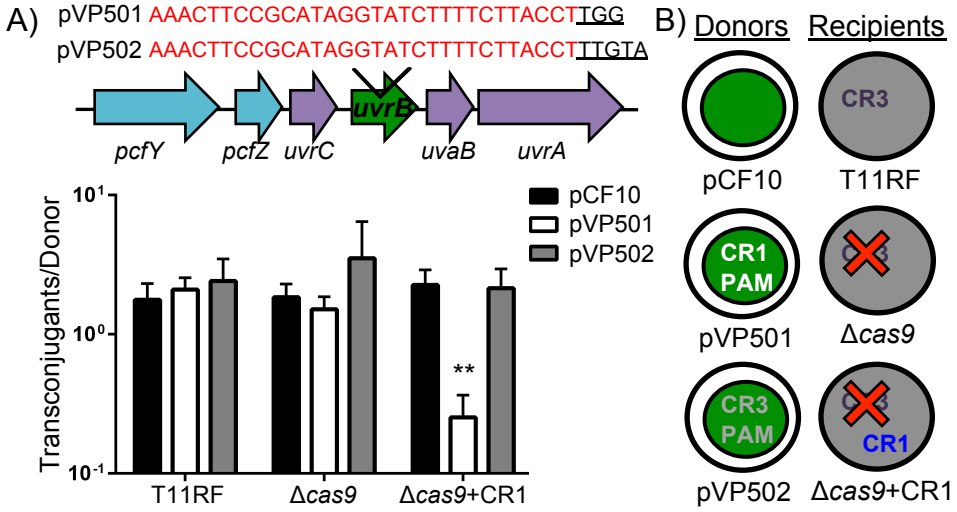


Figure 2.4. Orphan CRISPR2 provides defense against PRPs in the presence of CRISPR1 *cas9*. A) Schematic of how the T11 CRISPR2 spacer 1 sequence and corresponding PAM sequences (underlined) were introduced into pCF10. B) Outline of donor and recipient strains used for assessing the function of CRISPR2. C) Conjugation frequency of T11RF and its derivatives as recipient strains in conjugation with OG1SSp harboring pCF10, pVP501 and pVP502. The T11 CRISPR2 locus provides genome defense against pVP501 in the presence of CRISPR1 *cas9* and its predicted tracrRNA. Results also demonstrate CRISPR1 *cas9* PAM specificity to the NGG sequence. A minimum of three independent conjugation reactions are represented. Significance was assessed using a one-tailed Student's t-Test; P-values relative to pVP501 transfer to Δ*cas9*+CR1: **, $P<0.005$.

2.7 Discussion

A correlation between the lack of CRISPR-Cas and multidrug resistance in *E. faecalis* has been previously established using genome analysis (30). The aim of the current work was to experimentally assess genome defense strategies in *E. faecalis* using clinically relevant conjugative plasmids as model MGEs. Broadly, our study exemplifies the importance of the

variable genome of *E. faecalis.* We explored genome defense in *E. faecalis* T11, a strain closely related to the high-risk MDR strain V583. Two components of the *faecalis* variable genome that occur in T11 but are absent from V583, CRISPR3-Cas and a predicted R-M system, have a combined four-log impact on the conjugative transfer of the pheromone-responsive plasmid pAM714 in biofilm settings. These results substantiate our hypothesis that high-risk *E. faecalis* have readily acquired resistance to antibiotics due to their lack of genome defense. In future work, it will be of interest to assess the kinetics of CRISPR-Cas and R-M defense against antibiotic resistance plasmids, as well as their comparative efficiencies in providing genome defense in biofilm, planktonic, and polymicrobial settings.

Our work demonstrates that the orphan CRISPR2 locus in T11 does not confer genome defense in the absence of CRISPR1-Cas-encoded factors. This is significant because all high-risk, MDR *E. faecalis* possess orphan CRISPR2 loci. The conservation of CRISPR2 among *E. faecalis* strains lacking CRISPR1-Cas remains to be explained. CRISPR2 may be maintained in the species by providing another function for the cell, perhaps by acting as a non-coding regulatory RNA. Indeed, both CRISPR2 and a transcript antisense to CRISPR2 have been detected in transcriptome studies of V583 (56, 57), demonstrating that this region is transcriptionally active in the absence of CRISPR1-Cas. There is a precedent for a role for orphan CRISPR loci in regulation of gene expression; the orphan CRISPR *rliB* in *Listeria monocytogenes* regulates expression of *feoAB* (ferrous iron acquisition genes) and impacts virulence (58, 59). This locus undergoes an alternative processing pathway involving polynucleotide phosphorylase (PNPase) (60), therefore requirement of host-encoded factors beyond RNase III in *E. faecalis* CRISPR2

function cannot be ruled out. Studies of the V583 CRISPR2 locus are of interest for future work. Of particular interest is testing whether the reintroduction of CRISPR1-Cas into high-risk MDR *E. faecalis* leads to CRISPR adaptation against endogenous MGEs and genome reduction when antibiotic selection is absent.

Although CRISPR3-Cas had a significant impact on conjugation frequency, it was not a perfect barrier to plasmid transfer as some transconjugants were obtained in every mating reaction. This suggests that a subset of recipient cells have mutations in CRISPR3-Cas that inactivate defense, or that pAM714 plasmids have mutations in the *repB* protospacer or PAM, or perhaps that pAD1 has a mechanism for actively evading CRISPR-Cas defense in a subset of cells. Whether CRISPR-Cas is equally expressed in every recipient cell, and how the system is regulated, are also unknown. Interestingly, high frequencies of CRISPR-Cas mutations have been observed in other Type II CRISPR systems (61, 62). Further analysis of these 'escaper' transconjugants will be the focus of future work. Importantly, R-M defense can still impede plasmid transfer in CRISPR-Cas mutant cells. Our observation that CRISPR-Cas and R-M defense each contribute significantly to anti-plasmid genome defense is consistent with a previous report that the two defenses work additively against phage infection in *Streptococcus thermophilus* (63).

How can this information be applied? Our work supports the development of antimicrobial strategies that monopolize the immune-compromised status of high-risk, MDR *E. faecalis*. These applications include phage therapy, and pre-programmed CRISPR-Cas9 systems, introduced by phagemids, that target the bacterial chromosome for destruction (64, 65). These strategies could

be used for surface and gastrointestinal tract decolonization of problematic *E. faecalis*. Critical to the success of these strategies will be a greater understanding of *E. faecalis* phage biology, about which little is known, as well as the potential for Cas9-directed chromosome cleavage in *E. faecalis*.

2.8 Acknowledgments

We thank Karthik Hullahalli for plasmid pG19, which was used to generate strain T11RFΔ*cas9*+CR1. We also thank Dr. Zhenyu Xuan for his help with statistical analyses and Dr. Michael Gilmore and Dr. Gary Dunny for providing enterococcal strains used in this study.

This work was supported by grants K22AI099088 and R01AI116610 from the National Institutes of Health to KLP. The authors declare that the funders of this work had no role in the design of experiments, interpretation of data, or decision to publish this work.

2.9 References

1. Lebreton F, Willems RJL, & Gilmore MS (2014) Enterococcus Diversity, Origins in Nature, and Gut Colonization. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*, eds Gilmore MS, Clewell DB, Ike Y, & Shankar N. Boston.

2. Agudelo Higuita NI & Huycke MM (2014) Enterococcal Disease, Epidemiology, and Implications for Treatment. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*, eds Gilmore MS, Clewell DB, Ike Y, & Shankar N. Boston.

3. Arias CA & Murray BE (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nature reviews. Microbiology* 10(4):266-278.

4. George RC & Uttley AH (1989) Susceptibility of enterococci and epidemiology of enterococcal infection in the 1980s. *Epidemiology and infection* 103(3):403-413.

5. Mundy LM, Sahm DF, & Gilmore M (2000) Relationships between enterococcal virulence and antimicrobial resistance. *Clin Microbiol Rev* 13(4):513-522.

6. Sievert DM*, et al.* (2013) Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. *Infection control and hospital epidemiology* 34(1):1-14.

7. Clewell DB*, et al.* (2014) Extrachromosomal and Mobile Elements in Enterococci: Transmission, Maintenance, and Epidemiology. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*, eds Gilmore MS, Clewell DB, Ike Y, & Shankar N. Boston.

8. Palmer KL, Kos VN, & Gilmore MS (2010) Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Current opinion in microbiology* 13(5):632-639.

9. Arthur M, Molinas C, Depardieu F, & Courvalin P (1993) Characterization of Tn*1546*, a Tn*3*-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. *Journal of bacteriology* 175(1):117-127.

10. Leclercq R, Derlot E, Duval J, & Courvalin P (1988) Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *The New England journal of medicine* 319(3):157-161.

11. Murray BE, An FY, & Clewell DB (1988) Plasmids and pheromone response of the beta-lactamase producer *Streptococcus* (*Enterococcus*) *faecalis* HH22. *Antimicrobial agents and chemotherapy* 32(4):547-551.

12. Uttley AH*, et al.* (1989) High-level vancomycin-resistant enterococci causing hospital infections. *Epidemiology and infection* 103(1):173-181.

13. Liu Y*, et al.* (2013) Transferable multiresistance plasmids carrying *cfr* in *Enterococcus* spp. from swine and farm environment. *Antimicrobial agents and chemotherapy* 57(1):42-48.

14. Health UDo & Services H (2013) Antibiotic resistance threats in the United States, 2013. *Atlanta: CDC*.

15. Dunny GM (2007) The peptide pheromone-inducible conjugation system of *Enterococcus faecali*s plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 362(1483):1185-1193.

16. Clewell DB (2007) Properties of *Enterococcus faecalis* plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin. *Plasmid* 58(3):205-227.

17. Jensen LB, *et al.* (2010) A classification system for plasmids from enterococci and other Gram-positive bacteria. *Journal of microbiological methods* 80(1):25-43.

18. Hirt H, *et al.* (2005) Characterization of the pheromone response of the *Enterococcus faecalis* conjugative plasmid pCF10: complete sequence and comparative analysis of the transcriptional and phenotypic responses of pCF10-containing cells to pheromone induction. *J Bacteriol* 187(3):1044-1054.

19. Licht TR, Laugesen D, Jensen LB, & Jacobsen BL (2002) Transfer of the Pheromone-Inducible Plasmid pCF10 among *Enterococcus faecalis* Microorganisms Colonizing the Intestine of Mini-Pigs. *Applied and environmental microbiology* 68(1):187-193.

20. Huycke MM, Gilmore MS, Jett BD, & Booth JL (1992) Transfer of pheromone-inducible plasmids between *Enterococcus faecalis* in the Syrian hamster gastrointestinal tract. *The Journal of infectious diseases* 166(5):1188-1191.

21. Palmer KL, *et al.* (2012) Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. *mBio* 3(1):e00318-00311.

22. Bourgogne A, *et al.* (2008) Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome biology* 9(7):R110.

23. McBride SM, Fischetti VA, Leblanc DJ, Moellering RC, Jr., & Gilmore MS (2007) Genetic diversity among *Enterococcus faecalis*. *PloS one* 2(7):e582.

24. Solheim M, *et al.* (2011) Comparative genomic analysis reveals significant enrichment of mobile genetic elements and genes encoding surface structure-proteins in hospital-associated clonal complex 2 *Enterococcus faecalis*. *BMC microbiology* 11:3.

25. Shankar N, Baghdayan AS, & Gilmore MS (2002) Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. *Nature* 417(6890):746-750.

26. Paulsen IT, *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299(5615):2071-2074.

27. Sahm DF, *et al.* (1989) In vitro susceptibility studies of vancomycin-resistant *Enterococcus faecalis*. *Antimicrobial agents and chemotherapy* 33(9):1588-1591.

28. Willems RJ, Hanage WP, Bessen DE, & Feil EJ (2011) Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. *FEMS microbiology reviews* 35(5):872-900.

29. Leavis HL, Bonten MJ, & Willems RJ (2006) Identification of high-risk enterococcal clonal complexes: global dispersion and antibiotic resistance. *Current opinion in microbiology* 9(5):454-460.

30. Palmer KL & Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4).

31. Jansen R, van Embden JD, Gaastra W, & Schouls LM (2002) Identification of a novel family of sequence repeats among prokaryotes. *Omics : a journal of integrative biology* 6(1):23-33.

32. Mojica FJ, Diez-Villasenor C, Soria E, & Juez G (2000) Biological significance of a family of regularly spaced repeats in the genomes of *Archaea*, *Bacteria* and mitochondria. *Mol Microbiol* 36(1):244-246.

33. Jansen R, Embden JD, Gaastra W, & Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43(6):1565-1575.

34. Makarova KS*, et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nature reviews. Microbiology* 9(6):467-477.

35. Makarova KS*, et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nature reviews. Microbiology* 13(11):722-736.

36. Marraffini LA (2015) CRISPR-Cas immunity in prokaryotes. *Nature* 526(7571):55-61.

37. Lindenstrauss AG*, et al.* (2011) Comparison of genotypic and phenotypic cluster analyses of virulence determinants and possible role of CRISPR elements towards their incidence in *Enterococcus faecalis* and *Enterococcus faecium*. *Systematic and applied microbiology* 34(8):553-560.

38. Burley KM & Sedgley CM (2012) CRISPR-Cas, a prokaryotic adaptive immune system, in endodontic, oral, and multidrug-resistant hospital-acquired *Enterococcus faecalis*. *Journal of endodontics* 38(11):1511-1515.

39. Hullahalli K*, et al.* (2015) Comparative Analysis of the Orphan CRISPR2 Locus in 242 *Enterococcus faecalis* Strains. *PloS one* 10(9):e0138890.

40. Davis BM, Chao MC, & Waldor MK (2013) Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Current opinion in microbiology* 16(2):192-198.

41. Tock MR & Dryden DT (2005) The biology of restriction and anti-restriction. *Current opinion in microbiology* 8(4):466-472.

42. Murray NE (2002) 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. *Microbiology* 148(Pt 1):3-20.

43. Huo W, Adams HM, Zhang MQ, & Palmer KL (2015) Genome Modification in *Enterococcus faecalis* OG1RF Assessed by Bisulfite Sequencing and Single-Molecule Real-Time Sequencing. *J Bacteriol* 197(11):1939-1951.

44. Thurlow LR, Thomas VC, & Hancock LE (2009) Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *Journal of bacteriology* 191(20):6203-6210.

45. Bae T, Kozlowicz B, & Dunny GM (2002) Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and *prgX* mRNA and in regulation of pheromone-inducible conjugation. *Journal of molecular biology* 315(5):995-1007.

46. Aziz RK*, et al.* (2008) The RAST Server: rapid annotations using subsystems technology. *BMC genomics* 9:75.

47. Overbeek R*, et al.* (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic acids research* 42(Database issue):D206-214.

48. Bailey TL*, et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic acids research* 37(Web Server issue):W202-208.

49. Chylinski K, Le Rhun A, & Charpentier E (2013) The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA biology* 10(5):726-737.

50. Clewell DB*, et al.* (1982) Mapping of *Streptococcus faecalis* plasmids pAD1 and pAD2 and studies relating to transposition of Tn*917*. *Journal of bacteriology* 152(3):1220-1230.

51. Ike Y, Clewell DB, Segarra RA, & Gilmore MS (1990) Genetic analysis of the pAD1 hemolysin/bacteriocin determinant in *Enterococcus faecalis*: Tn*917* insertional mutagenesis and cloning. *Journal of bacteriology* 172(1):155-163.

52. Nishimasu H*, et al.* (2015) Crystal Structure of *Staphylococcus aureus* Cas9. *Cell* 162(5):1113-1126.

53. Jinek M*, et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816-821.

54. Leenhouts K*, et al.* (1996) A general system for generating unlabelled gene replacements in bacterial chromosomes. *Molecular & general genetics : MGG* 253(1-2):217-224.

55. Dunny G, Funk C, & Adsit J (1981) Direct stimulation of the transfer of antibiotic resistance by sex pheromones in *Streptococcus faecalis*. *Plasmid* 6(3):270-278.

56. Fouquier d'Herouel A*, et al.* (2011) A simple and efficient method to search for selected primary transcripts: non-coding and antisense RNAs in the human pathogen *Enterococcus faecalis*. *Nucleic acids research* 39(7):e46.

57. Innocenti N*, et al.* (2015) Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*. *RNA* 21(5):1018-1030.

58. Toledo-Arana A*, et al.* (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459(7249):950-956.

59. Mandin P, Repoila F, Vergassola M, Geissmann T, & Cossart P (2007) Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic acids research* 35(3):962-974.

60. Sesto N*, et al.* (2014) A PNPase dependent CRISPR System in *Listeria*. *PLoS genetics* 10(1):e1004065.

61. Lopez-Sanchez MJ*, et al.* (2012) The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Mol Microbiol* 85(6):1057-1071.

62. Jiang W*, et al.* (2013) Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS genetics* 9(9):e1003844.

63. Dupuis ME, Villion M, Magadan AH, & Moineau S (2013) CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nature communications* 4:2087.

64. Bikard D*, et al.* (2014) Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nature biotechnology* 32(11):1146-1150.

65. Citorik RJ, Mimee M, & Lu TK (2014) Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nature biotechnology* 32(11):1141-1145.

CHAPTER 3

SELECTION FOR ANTIBIOTIC RESISTANCE PLASMIDS RESULTS IN COMPROMISED CRISPR-CAS DEFENSE IN *ENTEROCOCCUS FAECALIS*

3.1 Author contribution

The work presented in this chapter was conducted in collaboration with Wenwen Huo (WH), Ardalan Sharifi (AS), Michael Zhang (MZ) and Kelli L. Palmer (KP). Valerie Price (VP) and WH contributed equally to this work. The serial passaging experiments were conceived and designed by VP, WH, and KP. The data for serial passaging experiments was collected by VP, WH, and AS; analysis was performed by VP and WH. Deep sequencing experiments were designed by VP, WH, MZ and KP. WH performed the analysis of deep sequencing data and VP and WH analyzed the results. This chapter was written as an original manuscript by VP, WH and KP and is expected to be submitted for publication in 2017.

3.2 Abstract

Enterococcus faecalis is a Gram-positive bacterium that normally colonizes the human gastrointestinal tract and opportunistically causes life-threatening infections. Multidrug-resistant (MDR) *E. faecalis* strains have emerged that are replete with mobile genetic elements (MGEs). Some *E. faecalis* possess CRISPR-Cas systems, which reduce the conjugation frequency of pheromone-responsive plasmids (PRPs), a group of narrow host range plasmids that are highly efficient gene disseminators. However, despite CRISPR-Cas defense, PRP transconjugants still

arise. Using serial passage experiments in the presence and absence of antibiotic selection for the plasmids, we found that there is a cost to maintain both the plasmid and a functional CRISPR-Cas system in these transconjugants. Without antibiotic selection, CRISPR-Cas eliminated the target plasmid. Under antibiotic selection, compromised CRISPR-Cas systems emerged. We conclude CRISPR-Cas defense can be compromised when antibiotic selection for MGEs is present.

3.3 Introduction

Enterococcus faecalis is a gram-positive bacterium that normally colonizes the gastrointestinal tracts of humans and other animals (1). *E. faecalis* is also an opportunistic pathogen that causes serious infections in hospitalized patients (2-6). *E. faecalis* is intrinsically resistant to certain classes of antibiotics (7-11). Some *E. faecalis* strains have acquired additional resistances through horizontal gene transfer (HGT), mediated primarily by plasmids and integrative conjugative elements (12-14). Plasmid-mediated dissemination of antibiotic resistance contributes to the rapid emergence of multidrug-resistant bacterial pathogens, one of the most challenging problems facing health care today (13, 15-18).

Much research on antibiotic resistance focuses on the mechanisms underlying the persistence of antibiotic resistance plasmids in populations where there is no antibiotic selection (19-22). The persistence of these plasmids allows them to act as reservoirs for accessory genes that can be readily shared and utilized for rapid adaptation to new environments (19, 21, 23). Overall, this speeds up the evolution of bacteria by negating the reliance on adaptation by mutation (24, 25).

There is growing consensus in the field that the persistence of resistance plasmids in the absence of antibiotic selection is due to compensatory mutations made in the host, plasmid or both (21, 25, 26). These mutations reduce the metabolic burden of plasmid carriage and reduce the rate at which the plasmid can engage in horizontal transfer while stabilizing the vertical inheritance of the plasmid, ultimately leading to plasmid persistence (27).

Pheromone-responsive plasmids (PRPs) are a unique class of narrow host range plasmids that disseminate antibiotic resistance and virulence genes among *E. faecalis* strains and mobilize resistance genes to other pathogens (28-30). PRP-mediated conjugation is highly regulated and is initiated by plasmid recognition of short peptide sequences produced by plasmid-free cells. PRPs and *E. faecalis* are highly co-evolved because these plasmids usually cannot replicate outside of the species (31, 32). PRPs shape the evolutionary trajectory of the *faecalis* species.

A facet of host-plasmid co-evolution that has yet to be extensively studied is the impact of genome defense systems that block the acquisition of potentially beneficial mobile genetic elements (MGEs). CRISPR-Cas systems confer programmable genome defense against plasmids and phage. CRISPR-Cas systems consist of *cas* genes and a CRISPR array composed of spacers interspersed by direct and partially palindromic repeats (33, 34). The spacers bear identity to foreign DNA elements, usually of plasmid or phage origin (35-37); each spacer is a memory of a previously encountered MGE. CRISPR-Cas defense is afforded in three stages, adaptation, expression and interference, that ultimately result in sequence-specific cleavage of MGEs by a *cas*-encoded endonuclease. *E. faecalis* possesses Type II CRISPR-Cas systems that encode the

endonuclease Cas9 (12). A Cas protein complex, including Cas9, recognizes a protospacer from a newly encountered MGE in a PAM (Protospacer Adjacent Motif)-dependent manner, after which the protospacer is incorporated into the leader end of the CRISPR array (38, 39). During expression, the CRISPR array is transcribed into a pre-crRNA that is processed by Cas9, RNase III, and tracrRNA (trans-activating crRNA) generating a mature crRNA (40, 41). Each mature crRNA is bound by Cas9 and tracrRNA to form an active targeting complex. When the bacterial host is invaded by a MGE that has complementarity to a crRNA, the active targeting complex is able to recognize the target in a PAM-dependent manner and create a double-stranded DNA break to prevent MGE invasion (42-44).

Interestingly, there is a strong correlation between the absence of complete CRISPR-Cas systems and the presence of horizontally acquired antibiotic resistance in *E. faecalis* clinical isolates (45). This suggests that the absence of barriers to horizontal gene transfer allows the rapid dissemination of MGEs among hospital-adapted strains of *E. faecalis*, transforming these MDR strains into reservoirs for antibiotic resistance and other virulence traits. On the other hand, drug-susceptible isolates of *E. faecalis* possess complete type II CRISPR-Cas systems (45, 46). We reported in previous studies that Type II CRISPR-Cas systems can reduce PRP dissemination in *E. faecalis* colony biofilms by 80-fold (47). These results are significant because they demonstrate that some *E. faecalis* strains possess endogenous barriers to horizontal antibiotic resistance acquisition while adding credence to the hypothesis that compromised genome defense systems may lead to the acquisition of MGEs and emergence of MDR strains of *E. faecalis*.

Much is now known about how CRISPR/Cas9 can be harnessed for genome editing applications (48-55). However, there is much to learn about how interactions between MGEs and CRISPR can shape the evolutionary trajectory of a species. It has been shown that CRISPR array regions, the set of repeats and spacers, undergo dynamic evolution (56-61). The addition of new spacers into the leader end provides fresh immunity to newly evolved MGEs, while leader distal spacers act as molecular 'fossils' to track the evolutionary history of strains (46, 60). At the same time, CRISPR array expansion is not unlimited as internal spacers can be deleted, providing a basis for diversification and the emergence of heterogeneous bacterial populations with dynamic CRISPR array-allele variations (62, 63). Previously, researchers used mathematical modeling to estimate the rate of these deletion events to be around e^{-4} for a type III-A CRISPR-Cas system in *Staphylococcus epidermidis* (64). This study also concluded that the ability of a MGE to escape CRISPR-Cas defense was dependent on the existence of pre-existing CRISPR mutants in recipient populations.

Although the CRISPR3-Cas system of *E. faecalis* had a significant impact on PRP transfer, we still observed a high number (10^5) transconjugants (47). This observation suggests that unique interactions occur under these mating conditions that allow the PRP to escape genome defense. Moreover, a conflict is potentially established between the *E. faecalis* CRISPR-Cas system and one of its targets. These transconjugants present a unique opportunity to study the role of CRISPR-Cas systems in plasmid-host interactions. In this study, we used a combination of *in vitro* evolution and deep sequencing analysis to investigate how *E. faecalis* resolves conflicts between CRISPR-Cas and antibiotic resistance plasmids, and the role that antibiotic selection

plays in this process. We conclude that antibiotic-driven PRP maintenance in *E. faecalis* can lead to compromised genome defense and enhanced susceptibility to other MGEs.

3.4 Materials and Methods

Strains, reagents, and routine molecular biology procedures.

Bacterial strains and plasmids used in this study are listed in Table 3.1. *E. faecalis* strains were grown in Brain Heart Infusion (BHI) broth or on agar plates at 37°C unless otherwise noted. Antibiotics were used for *E. faecalis* at the following concentrations: erythromycin, 50 μg/mL; chloramphenicol, 15 μg/mL; streptomycin, 500 μg/mL; spectinomycin, 500 μg/mL; rifampicin, 50 μg/mL; fusidic acid, 25 μg/mL. *Escherichia coli* strains used for plasmid propagation and were grown in lysogeny broth (LB) broth or on agar plates at 37°C. Chloramphenicol was used at 15 μg/mL for *E. coli*. PCR was performed using *Taq* (New England Biolabs) or Phusion (Fisher Scientific) polymerases. Primer sequences used are in Table B.1 in Appendix B. Routine DNA sequencing was carried out at the Massachusetts General Hospital DNA core facility (Boston, MA). *E. faecalis* electrocompetent cells were made using the lysozyme method as previously described (65).

Generation of mutant *E. faecalis* strains and plasmids.

In-frame deletion of *recA* in T11RF was generated using a previously established protocol (66). Briefly, ~750 bp regions up- and downstream of *recA* in *E. faecalis* T11RF were amplified, digested, and ligated into pLT06 (66) to generate pWH*recA*. The resulting plasmid was transformed into competent T11RF cells via electroporation (65). Following transformation at

30°C, a shift to the non-permissive temperature of 42°C and counterselection on p-chloro-phenylalanine were performed to generate an in-frame, markerless deletion.

To insert the T11 CRISPR3 S_1, S_6, and S_7 sequences and CRISPR3 PAM (TTGTA) into pCF10, 47 bp and 39 bp single stranded DNA oligos were annealed to each other to generate dsDNA with restriction enzyme overhangs for *BamHI* and *PstI*. The annealed oligos were ligated into the pLT06 derivative pWH107 that includes sequence from pCF10 *uvrB*, to insert these sequences into the *uvrB* gene of pCF10 by homologous recombination. A knock-in protocol was performed as previously described (47).

Conjugation experiments.

E. faecalis donor and recipient strains were grown in BHI overnight to stationary phase. A 1:10 dilution was made for both donor and recipient in fresh BHI broth and allowed to grow for 1.5 hr to reach mid-exponential phase. A mixture of 100 µL donor cells and 900 µL recipient cells was pelleted and plated on BHI agar to allow conjugation. After 18 h incubation, the conjugation mixture was scraped from the plate using 2 mL 1X PBS supplemented with 2 mM EDTA. Serial dilutions were prepared from the conjugation mixture and plated on selective BHI agars. After 24-48 h incubation, colony forming units per milliliter (CFU/mL) was determined using plates with 30 - 300 colonies. The conjugation frequency was calculated as the CFU/mL of transconjugants divided by the CFU/mL of donors.

Serial passage.

Transconjugant or transformant colonies were suspended in 50 μL BHI broth. The 50 μL suspension was used as follows: 3 μL was used for PCR to confirm the integrity of the CRISPR array, 10 μL was inoculated into plain BHI broth, another 10 μL was inoculated into selective BHI broth for plasmid selection, and another 10 μL was used for serial dilution and plating on selective medium to enumerate the initial number of plasmid-containing cells in the transconjugant colonies. Broth cultures were incubated for 24 h, followed by 1:1000 dilution into either fresh plain BHI or fresh selective BHI. At each 24 h interval, 3 μL of each culture from the previous incubation was used for PCR to check CRISPR array integrity, and 10 μL was used for serial dilution and plating on agars to determine CFU/mL for total viable cells and plasmid-containing cells. The cultures were passaged in this manner for 14 days; cryopreserved culture stocks were made daily in glycerol. To use the Day 14 transconjugant populations in conjugation reactions, the glycerol stocks were completely thawed on ice, and 20 μL was inoculated into plain BHI broth. The cultures were incubated for 6-8 h to allow them to reach mid-exponential phase ($OD_{600nm} \approx 0.5–0.7$), and 900 μL was used as recipient in conjugation reactions as described above.

Deep sequencing of CRISPR3 amplicons and genomic DNA.

For CRISPR3 amplicon sequencing, 3 μL from a broth culture was used as template in PCR using Phusion Polymerase with CR3_seq_F/R primers (Table B.1 in Appendix B). The PCR products were purified using the Thermo Scientific PCR purification kit (Thermo Scientific). Genomic DNA was isolated using the phenol-chloroform method (67). The purified PCR

amplicons and genomic DNA samples were sequenced using 2 x 150 bp paired end sequencing chemistry by Molecular Research LP (MR DNA; Texas).

Whole genome sequencing analysis.

T11 supercontig and pAD1 plasmid contig references were downloaded from NCBI (accession numbers: T11: NZ_GG688637.1-NZ_GG688649; pAD1: AB007844, AF394225, AH011360, L01794, L19532, L37110, M84374, M87836, U00681, X17214, X62657, X62658). Reads were aligned to these references using default parameters in CLC Genomics Workbench (Qiagen) where ≥50% of each mapped read has ≥80% sequence identity to the reference. Variations occurring with ≥35% frequency at positions with ≥10X coverage between our samples and the reference contigs were detected using the Basic Variant Detector. At the same time, local realignment was performed, followed by Fixed Ploidy variant detection using default parameters and variants probability ≥90% in CLC Genomics Workbench. The basic variants and fixed ploidy variants were combined for each sequencing sample and subjected to manual inspection. The variants that were detected in the T11 genome from all samples are possibly pre-existing variants in our parent T11 stock, hence were manually removed. The variants that were detected in pAD1 genome from all transconjugant samples are possibly pre-existing variants in our pAM714 stock, hence were also manually removed. Next, variants within the CRISPR3 array were removed as we analyzed CRISPR3 alleles using a different approach (amplicon deep sequencing). The variants with 0% detection rates were manually checked for coverage depth to eliminate the detection bias. The variants detected in all samples are shown in Table B.2 in Appendix B.

Analysis of CRISPR3 amplicon sequencing.

Reads from the 1,763 bp CRISPR3 amplicon were mapped to the T11 CRISPR3 reference (NZ_GG688647.1: 646834 - 648596) using stringent mapping conditions in CLC Genomics Workbench. The stringent mapping conditions require 100% of each mapped read to have ≥95% identity to the reference. The percent mapped reads were calculated by dividing the number of reads mapped by the total number of reads, these percentages are listed in Table B.3 in Appendix B, step 1. The coverage depth was then calculated for each position within the PCR amplicon region using CLC Genomics Workbench, normalized using reads per million, and plotted against reference positions (Figure 3.4).

To further analyze CRISPR3 spacer deletions and rearrangements, we manually created CRISPR3 references. The CRISPR3 amplicon references contain two spacers connected by a T11 CRISPR3-Cas repeat: 5'-spacer[x]-repeat-spacer[y]-3' (5'-SxRSy-3'), where spacer[x] and spacer[y] could be 30 bp upstream of the first repeat (leader end; or S_0 hereafter; Figure 3.1a), or any internal spacer within the CRISPR3 array (from spacer 1 to spacer 21; or S_1 to S_{21}; Figure 3.1a). Each manually generated CRISPR3 amplicon reference is 96 bp in length. The references where y=x+1 represent wild-type alleles. The terminal repeat following S_{21} in the CRISPR3 array is divergent from the regular direct repeat sequence, so references containing 5'-spacer[x]-TerminalRepeat-S_T-3' (5'-S_xTRS_T-3') were constructed, where spacer[x] ranges from S_0 to S_{21} and spacer S_T represents the sequence 30 bp downstream of the terminal repeat (Figure 3.1a). The 5'-S_{21}TRS_T-3' reference represents the wild-type. In total, 484 references with length of 96 bp were generated for the CRISPR3 amplicon. Considering that the read length is 150 bp, we

manually split the reads into two shorter subsequences to enhance mapping efficiency. The subsequences are of similar length to the references allowing for retrieval of maximal sequence information. The split amplicon sequencing reads were mapped to the 5'-SxRSy-3' and 5'-$SxTRS_T$-3' references using stringent mapping parameters in CLC Genomics Workbench (Qiagen). The stringent mapping parameters require 100% of each mapped read to be ≥95% identical to one unique reference. Thus, the sequencing reads from different CRISPR alleles will be distinguished. These amplicon mapping results were applied to the and forward spacer deletion and backward spacer rearrangement calculations.

To determine the mapping efficiency, the unmapped reads from initial mapping to the T11 CRISPR3 reference (Table B.3 in Appendix B, step 1) were subjected to additional quality control analysis. The unmapped reads were mapped to the 484 manually created spacer[x]-repeat-spacer[y] references using the same mapping parameters in CLC as above (Table B.3 in Appendix B, step 2 mapping; ignore unspecific mapping). The unmapped reads from step 2 were subjected to mapping to all possible references (CRISPR3 region plus manually created references) using default mapping parameters, ignoring unspecific mapping (80% of each mapped read has at least 50% identity to the reference sequence; Table B.3 in Appendix B, step 3 mapping). The unmapped reads from step 3 were mapped to all possible references using the default mapping parameters and randomly map unspecific matching reads (Table B.3 in Appendix B, step 4 mapping).

Percentage of S_6-containing cells.

To evaluate the percentage of bacterial cells containing S_6, we calculated the percentage of S_6-containing reads relative to S_1-containing reads. Assuming all bacteria containing intact CRISPR3 possess S_1, the percentage P of S_6-containing reads relative to S_1-containing reads represents the percentage of bacterial cells maintaining S_6, hence 1-P represents the percentage of bacterial cells without S_6 (Table 3.4).

Forward spacer deletion and backward spacer rearrangement.

We observed two categories of mutant CRISPR3 alleles: 5'-S_xRS_y-3' (y > x+1) and 5'-S_xRS_y-3' (y < x). The forward deletion mutants with 5'-S_xRS_y-3' (y > x+1) are the result of spacer deletions, with spacers from S_{x+1} to S_{y-1} deleted; while the backward rearrangement mutants with 5'-S_xRS_y-3' (y < x) are the result of spacer rearrangement, where a downstream spacer S_y flips to become upstream of an upstream spacer S_x. To study if there were positional preferences, the average forward spacer deletion rate and backward spacer rearrangement rate was calculated for each 5'-S_x (0 < x < 21) within the CRISPR3 array. For each 5'-S_x, the average forward deletion and backward rearrangement rate are calculated as:

$$\text{P(5'-Sx Forward)} = \frac{\#\text{ mapped reads to the reference of 5'-SxRSy-3'}}{\sum_{y=x+1}^{n} \#\text{ mapped reads to the references of 5'-SxRSy-3' and 5'-SnTRST-3'}}$$

$$\text{P(5'-Sx Backward)} = \frac{\#\text{ mapped reads to the reference of 5'-SxRSy-3'}}{\sum_{y=0}^{x-1} \#\text{ mapped reads to the references of 5'-SxRSy-3'}}$$

where n is the total number of spacers within a CRISPR array, hence S_n represents terminal spacer, as described above (Figure 3.1a).

3.5 Results

Design of *in vitro* evolution assay to study CRISPR-Cas and plasmid dynamics.

E. faecalis T11RF possesses a Type II CRISPR-Cas system, CRISPR3-Cas, that has 21 unique spacers (Figure 3.1) (47). Spacer 6 has 100% sequence identity to the *repB* gene of the PRP pAD1 (68, 69). Previous research from our lab demonstrated that T11RF CRISPR3-Cas significantly reduces the conjugation frequency of pAM714, a derivative of pAD1 conferring erythromycin resistance via *ermB* (47). CRISPR3-Cas genome defense against pAM714 required both *cas9* and spacer 6 sequences. However, despite the activity of CRISPR3-Cas, a large number (~10^5) of T11RF pAM714 transconjugants were obtained from these conjugation reactions. We hypothesized that the T11RF pAM714 transconjugants were subject to intracellular conflict between the endogenous CRISPR3-Cas and its pAM714 target, and that antibiotic selection for pAM714 could impact the outcome of this conflict.

To understand how T11RF transconjugants resolve conflicts between active CRISPR3-Cas defense and a CRISPR-Cas target, we utilized an *in vitro* evolution assay. We randomly selected transconjugant colonies from two mating schemes, T11RF pAM714 and T11RF Δ*cas9* pAM714; the Δ*cas9* strain was included as a control for the condition where CRISPR-Cas is inactive. Next, the colonies were split equally into two growth media, BHI medium and BHI medium with erythromycin to maintain selection for pAM714; see Figure 3.2a (a detailed explanation of the assay conditions can be found in Materials and Methods). These populations were then passaged daily for 14 days. We performed *in vitro* evolution on a total of six T11RF pAM714 (referred to

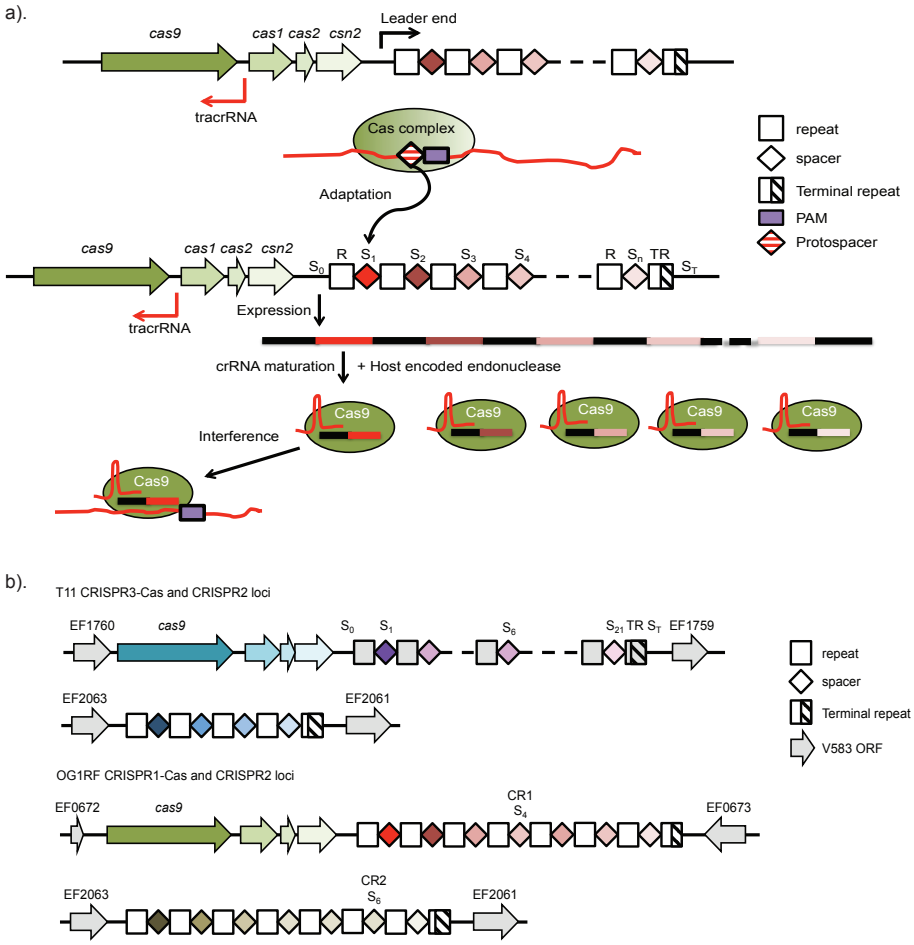


Figure 3.1. Type II-A CRISPR-Cas system in *E. faecalis*. a) Schematic mechanism of Type II CRISPR-Cas defense in bacteria. Upon MGE invasion, CRISPR-Cas acts as a genome defense system. When a new MGE is encountered, the protospacer is recognized based on Protospacer Adjacent Motif (PAM). A complex of Cas proteins incorporates the protospacer into the leader end of CRISPR array to form a new spacer (Adaptation). During the expression stage, the CRISPR array is transcribed into pre-crRNA, which is further processed into mature crRNA by Cas9, tracrRNA and a host-encoded endonuclease. The mature crRNA consists of part of a repeat and part of a spacer, which is bound to a Cas9:tracrRNA complex to form an effector complex. When the previously encountered MGE invades again, the effector complex recognizes the target by sequence complementary and the presence of a PAM. Upon recognition, the target is cleaved and thus invasion by the MGE is blocked. The definition of R, TR and S_n is described in Material and Methods. b). CRISPR-Cas loci occurring in *E. faecalis* T11RF and OG1RF. *E. faecalis* T11RF contains a CRISPR3-Cas system and a CRISPR2 array. S_6 within the CRISPR3 array targets pAM714 and pCF10+Sp6, while S_1 and S_7 within CRISPR3 array target pCF10+Sp1 and pCF10+Sp7, respectively, as described in Material and Methods. *E. faecalis* OG1RF contains a CRISPR1-Cas system and a CRISPR2 array. S_4 within CRISPR1 array targets pKHS96, while S_6 within CRISPR2 array targets pKHS5, as described in Material and Methods.

as WT1-WT6) and six T11RF Δ*cas9* pAM714 (referred to as Δ1- Δ6) transconjugants originating from two independent conjugation experiments each. Every 24 h during the course of the passage, the proportion of cells within the population that maintained pAM714 was enumerated by determining the percentage of erythromycin-resistant cells relative to the total viable population.

We established the frequency of pAM714 carriage in our transconjugant colonies at Day 0, prior to serial passage (Figure 3.2b). As expected, pAM714 was detected at ~100% frequency for the T11RF Δ*cas9* pAM714 transconjugant colonies. The frequency of plasmid carriage in the T11RF pAM714 transconjugant colonies varied greatly and was <25% for five of the six transconjugant colonies evaluated. We attribute the variability of plasmid carriage in the T11RF pAM714 transconjugants to the biofilm-like mode of colony growth, where different cells may be exposed to different antibiotic concentrations as a result of spatial heterogeneity (discussed further later in this work).

In addition to determining the frequency of plasmid carriage over the course of the *in vitro* evolution experiments, we also selected a genetic locus to assay for variation. We chose the CRISPR3 array, which is required to produce an active Cas9-crRNA targeting complex and houses a molecular memory of previous interactions with MGEs. To assess our transconjugants for pre-existing deletions in the CRISPR3 array, we performed PCR on the individual colonies, prior to initiating the passage experiments. We observed no fixed, pre-existing CRISPR3 array

deletions that would allow stable maintenance of pAM714 prior to the evolution assay (Figure 3.2c), an observation confirmed by Sanger sequencing of CRISPR3 amplicons.

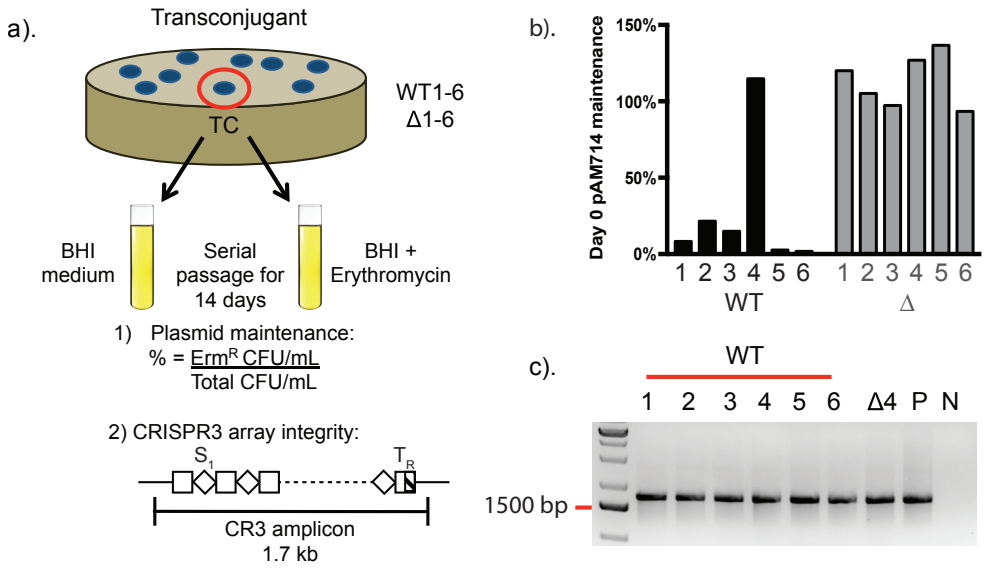


Figure 3.2. Experimental evolution design and analysis of pAM714 transconjugant colonies at Day 0. a) Design of *in vitro* evolution assay. Randomly selected T11RF pAM714 and T11RFΔ*cas9* pAM714 transconjugants were passaged for 14 days in the presence and absence of antibiotic selection for pAM714. These populations were monitored daily for: 1) pAM714 maintenance by determining the percentage of the population that was erythromycin-resistant, and 2) deviations in the CRISPR3 array by amplifying the 1.7 kb region encompassing the CRISPR3 array. b) Frequency of pAM714 carriage in transconjugant colonies used to initiate serial passage experiments. c) CRISPR3 amplicon PCR results for transconjugant colonies used to initiate serial passage experiments. Shown are CRISPR3 amplicon sizes for six T11RF pAM714 transconjugants (WT 1-6), a representative T11RFΔ*cas9* pAM714 transconjugant (Δ4), T11RF genomic DNA as a positive control (P), and a reagent control (N).

CRISPR3-Cas eliminates its target during passage in non-selective medium.

For passage without erythromycin selection, a gradual decrease in frequency of pAM714-containing cells was observed for five out of the six WT transconjugants (Figure 3.3a); a discussion of WT4 will be provided later. These data are consistent with CRISPR3-Cas

eliminating its target, pAM714, via Cas9 programmed with a crRNA derived from the spacer 6 (S_6) sequence. In contrast, pAM714 was stably maintained at high frequencies in all of the T11RFΔ*cas9* pAM714 transconjugant populations (Figure 3.3a). CRISPR3 array integrity was maintained over the course of serial passage for both T11RF pAM714 and T11RFΔ*cas9* pAM714 transconjugant populations (Figure 3.3c). Overall, these data demonstrate that *cas9*-dependent pAM714 loss occurs in *E. faecalis* when passaged in the absence of antibiotic selection for pAM714. By extension, for 5 of 6 WT transconjugants, the progenitor recipient cells for these lineages must have had functional CRISPR-Cas defense.

As stated above, the WT4 population did not exhibit plasmid loss in the absence of antibiotic selection. We reasoned that this transconjugant may have been CRISPR-Cas-deficient prior to serial passage. We sequenced the *cas9* coding region from passage Day 1 of the WT4 population and identified a mutation resulting in an Ala749Thr substitution. Ala749 occurs within the RuvC nuclease domain in T11RF Cas9 and is conserved in the model *Streptococcus pyogenes* Cas9 (47). Due to the critical catalytic function of the RuvC domain, we hypothesize that the Ala749Thr substitution confers a loss of Cas9 function. We describe the inability of the WT4 population to interfere with plasmid targets in a later section.

Under continuous antibiotic selection, conflicts can be resolved by CRISPR memory loss.

Although pAM714 initially escapes CRISPR-Cas defense in some cells, when antibiotic selection for the plasmid is absent, over time, CRISPR-Cas depletes pAM714 from transconjugant populations (Figure 3.3a). When passaging the same original transconjugant

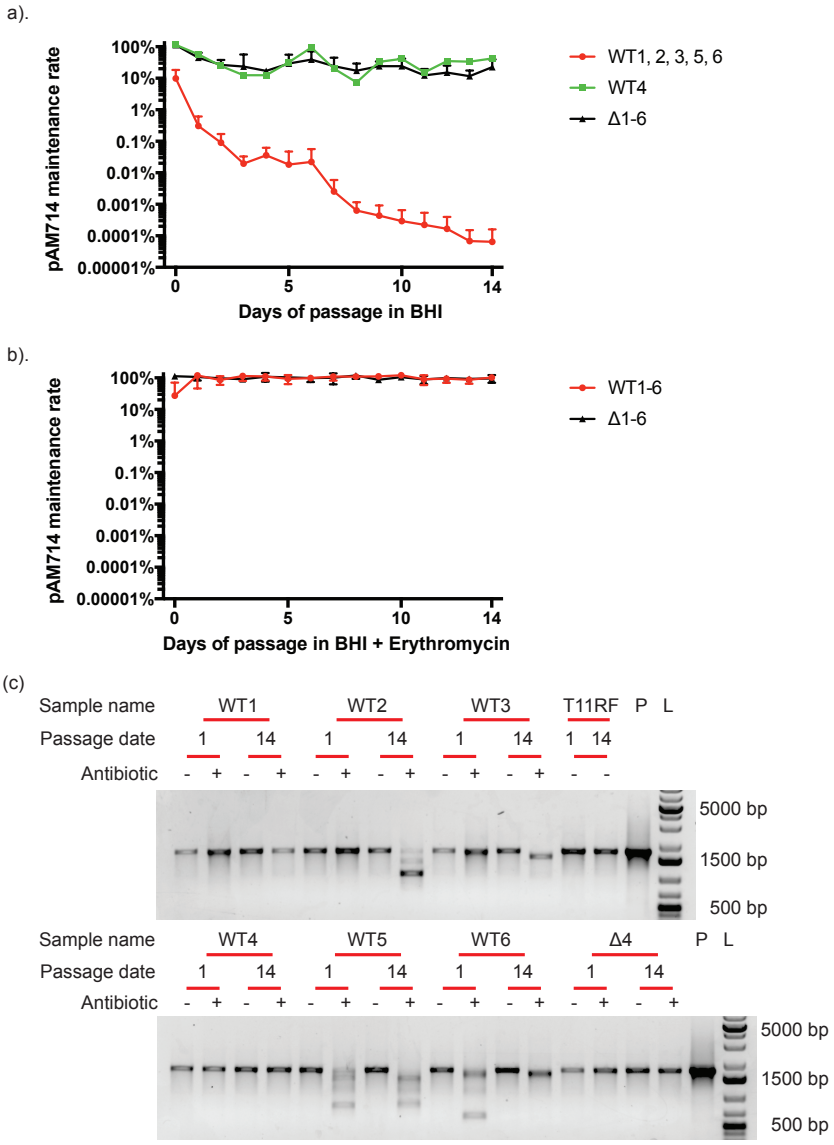


Figure 3.3. ***In vitro*** **evolution experiments.** a-b) pAM714 maintenance over the course of passage without (a) and with (b) antibiotic selection. Plasmid maintenance is expressed as percentage of bacterial cells conferring erythromycin resistance. WT populations are shown in green or red and Δ*cas9* populations are shown in black. c) CRISPR3 amplicon size from early (Day 1) and late (Day 14) passage dates for six WT transconjugant populations and a representative Δ*cas9* transconjugant population (Δ4). As a control, T11RF without pAM714 was passaged for 14 days and the CRISPR3 locus was queried (T11RF). P: positive control, T11RF genomic DNA. L, DNA ladder.

populations with erythromycin selection, we observed stable maintenance of pAM714 in both WT and Δ*cas9* transconjugants (Figure 3.3b). Knowing that the 5 of the 6 WT passage experiments each initiated with at least some cells in the population having active CRISPR-Cas defense, we wondered how the conflict between CRISPR-Cas and pAM714 was resolved in these populations under antibiotic selection.

We probed the CRISPR3 region of erythromycin-passaged transconjugants and observed significant heterogeneity in the CRISPR3 array for only the WT transconjugant populations (Figure 3.3c). By Day 14, four of the six T11RF pAM714 transconjugants had visibly reduced CRISPR3 arrays (Figure 3.2c); the variation in array size initiated sporadically over the 14 days and was unique in pattern of emergence for each transconjugant (Figure B.1 in Appendix B). We utilized Sanger sequencing as a first-line assessment of CRISPR3 allele composition present in Day 1 and Day 14 erythromycin-passaged populations. The results showed that, for transconjugant populations where CRISPR array reduction was observed, S_6 was either deleted from the array or had low sequencing quality (Table 3.2). Low Sanger sequencing quality likely resulted from mixed populations with different deletion events, each resulting in S_6 deletion. In contrast to the WT populations, CRISPR3 arrays for the T11RF Δ*cas9* pAM714 transconjugants were unchanged (Figure 3.3c and Table 3.2). We chose the Δ4 population as a representative of the T11RFΔ*cas9* pAM714 transconjugant populations for future analyses as all data indicated that the six populations were equivalent. In summary, S_6 was poorly tolerated in four WT transconjugants whereas no diversification of the CRISPR3 array occurred in Δ*cas9* transconjugants. This suggests that under antibiotic selection, the conflict between CRISPR3-Cas

and its target can be resolved by compromising the CRISPR3-Cas system, in the form of either S_6 loss (populations WT2, WT3, WT5, WT6) or *cas9* mutation (population WT4).

Table 3.1. Bacterial strains and plasmids used.

Name	Description	Reference
***E. faecalis* strains**		
T11RF	Human urine isolate	(47, 70)
T11RFΔ*cas9*	Derivative of T11RF with *cas9* deleted	(47)
T11RFΔ*recA*	Derivative of T11RF with *recA* deleted	This study
OG1RF	Human oral isolate	(12, 71)
OG1SSp	Donor strain for conjugation assay	(68, 69, 72)
Plasmids		
pAM714	65 kb PRP encoding erythromycin on Tn*917*, derivative of pAD1	(69, 73)
pCF10	67 kb PRP encoding tetracycline resistance on Tn*925*	(72)
pLZ12	broad host range shuttle vector encoding chloramphenicol resistance	(74)
pKH12	pLZ12 with *oriT*	(62)
pKHS5	pKH12 with CRISPR2protospacerS5 and CRISPR1/2 PAM	This study
pKHS96	pKH12 with CRISPR1protospacerS96 and CRISPR1/2 PAM	(62)
pWH*recA*	pLT06 with ~750 bp up- and downstream of *recA* from T11RF	
pVP107	pLT06 with T11CR2protospacer and CRISPR1/2 PAM	(47)
pWH107	pVP107 digested with XbaI/SphI and re-ligated with primers pVP107_ XbaI _For/pVP107_SphI_Rev to remove PstI enzyme site	This study
pWH107.S1	pWH107 with T11CR3protospacerS1 and CRISPR3 PAM inserted between BamHI/PstI	This study
pWH107.S6	pWH107 with T11CR3protospacerS6 and CRISPR3 PAM inserted between BamHI/PstI	This study
pWH107.S7	pWH107 with T11CR3protospacerS10 and CRISPR3 PAM inserted between BamHI/PstI	This study

The WT1 population is discussed further here. PCR analysis of transconjugant WT1 indicated that the wild-type CRISPR3 allele was present after 14 days of passage with erythromycin (Figure 3.3c). However, Sanger sequencing detected a mixed population in the region of S_6 and S_7 after passage Day 1, which was not detected after passage Day 14 (Table 3.2). Therefore, a S_6 deletion arose during the passage experiment but did not become fixed. Mutations in other CRISPR-associated factors could have arisen in the WT1 population.

To investigate this possibility, we performed whole genome Illumina deep sequencing on five Day 14 erythromycin-passaged T11RF pAM714 transconjugant populations and a control population, Δ4 (see Table 3.2). We observed variation in *cas9* sequence in the WT1, WT2, and WT3 populations (Table 3.3). All of the mutations led to nonsynonymous changes and are predicted to result in Cas9 loss of function (Table 3.3). In addition to *cas9* mutations, we observed variation in six other genes in some of the populations (Table B.2 in Appendix B). No variations were identified in the S_6 protospacer or PAM region of pAM714, although one variation was identified elsewhere in *repB* in the WT2 population (Table B.2 in Appendix B).

Reduced tolerance of S_6 in T11RF pAM714 transconjugant populations.

To attain greater resolution of CRISPR3 alleles beyond what Sanger sequencing could achieve, we deep-sequenced CRISPR3 amplicons from populations of interest, beginning with BHI-passaged T11RF (lacking pAM714) as a control. We first mapped CRISPR3 amplicon reads to the T11 reference sequence and calculated coverage depth to analyze mapping efficiency (Figure 3.4a). Serial passaging in BHI medium for 14 days slightly altered the distribution of reads

Table 3.2. CRISPR alleles.

Sample name	Day 1 Sanger[c]	Day 14 Sanger[c]	Day 1 Amplicon[d]	Day 14 Amplicon[d]
T11RF control[a]	WT	WT	WT	WT
Δ4[b]	WT	WT	WT	WT
WT1[b]	Poor quality at S_6-S_7	WT	ΔS_6 ΔS_6-S_7	WT
WT2[b]	WT	ΔS_2-S_{11}	WT	ΔS_2-S_{11}
WT3[b]	Poor quality at S_6-S_7	ΔS_5-S_7	ΔS_6 ΔS_5-S_7	ΔS_5-S_7 ΔS_6
WT5[b]	Poor quality at S_3-S_8	Poor quality at S_3-S_9	ΔS_5-S_8 ΔS_3-S_{16} ΔS_6 ΔS_4-S_9 ΔS_1-S_{14}	ΔS_5-S_8 ΔS_4-S_9 ΔS_6 ΔS_6-S_{18} ΔS_3-S_{16}
WT6[b]	Poor quality at S_1-S_7	Poor quality at S_5-S_6	ΔS_5-S_6 ΔS_1-S_{17} ΔS_5-S_7 ΔS_6 ΔS_6-S_9	ΔS_6 ΔS_5-S_6 ΔS_5-S_7 ΔS_6-S_{10}

[a]T11RF without pAM714 passaged for 14 days in BHI medium.
[b]pAM714 transconjugants passaged for 14 days in BHI medium with erythromycin.
[c]CRISPR3 alleles detected by Sanger sequencing.
[d]CRISPR3 alleles detected by Illumina amplicon deep sequencing. Mutant alleles with >0.3% abundance are shown for each population and are listed from highest to lowest abundance. If no mutant alleles were detected above this threshold, "WT" is stated

Table 3.3. Nonsynonymous *cas9* mutations detected by whole genome sequencing.

Contig #	Position	Ref	Allele	Amino acid change	WT1[a]	WT2[a]	WT3[a]
1.11	652983	G	A	Gln506*	31.6% (689x)	0% (590x)	0% (577x)
1.11	653184	C	T	Glu439Lys	0% (640x)	0% (536x)	24.2% (594x)
1.11	653180	AG	T	Leu440fs	0% (640x)	0% (528x)	23.5% (590x)
1.11	654165	G	-	Arg112fs	0% (676x)	48.4% (659x)	0% (772x)

[a]Variation frequency and coverage depth at the indicated nucleotide position are shown.

across the amplicon but did not result in a strong preference for the abundance or absence of any spacer. A representative *cas9* pAM714 transconjugant passage experiment was also analyzed. As expected, the coverage depth within the CRISPR3 amplicon of Day 1 and Day 14 erythromycin-passaged Δ4 populations showed a uniform distribution (Figure 3.4b).

We then expanded this analysis to the T11RF pAM714 transconjugants, excepting WT4. As expected, depletion of S_6 was detected for WT2, WT3, WT5, and WT6 populations after 14 days of passage with antibiotic selection (Figure 3.4d-g). For WT3, WT5, and WT6 populations, S_6 depletion was evident after one day of passage with selection (Figure 3.4e-g). For WT1, depletion of S_6 was not detected after 14 days passage with selection (Figure 3.4c), consistent with our Sanger sequencing results (Table 3.2).

To identify specific mutant CRISPR alleles in the amplicon deep sequencing, we manually constructed artificial CRISPR reference sequences for every possible spacer deletion event (see

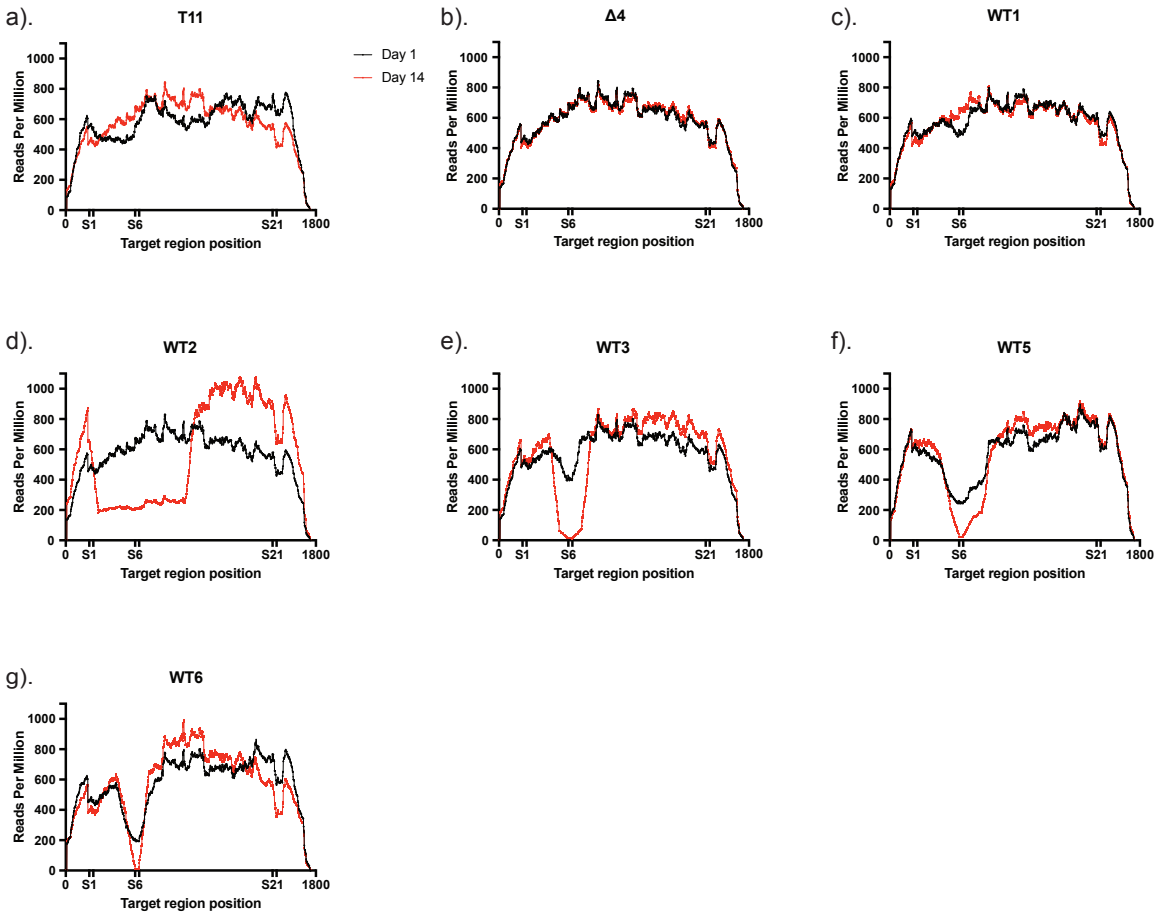


Figure 3.4. Amplicon sequencing revealed the depletion of S_6 antibiotic-passaged WT transconjugants. The coverage depth is calculated for each position within the amplicon and normalized using reads per million, which is then plotted against the genomic position. For each sample, the results for Day 1 and Day 14 are represented in black and red lines, respectively. The beginning and end of the regions along the amplicon corresponding to S_1, S_6 and S_{21} within the CRISPR3 array are labeled with vertical hash marks on the x-axis. a) BHI passaged T11RF parent strain. b) Erythromycin passaged Δ4 transconjugant. c-g) Erythromycin passaged WT transconjugants. Here, the WT4 population is not included due to the inactivating *cas9* mutation, as discussed in the main text.

Materials and Methods for more information). In total, 484 references were constructed, where wild type CRISPR alleles were represented by the wild type references: 5'-$S_xRS_{(x+1)}$-3' ($0 \leq x$

<21) and 5'-$S_{21}TRS_T$-3'. Mutant alleles were represented by 5'-S_xRS_y-3' ($y \neq x+1$). For control T11RF passaged for 1 or 14 days in plain BHI medium, only 28 (Day 1) and 4 (Day 14) mutant alleles out of 484 possible alleles were not detected. We conclude that CRISPR3 heterogeneity naturally occurs in T11RF populations, possibly as a result of slippage during DNA replication and/or recombination between CRISPR repeat sequences. This is consistent with previous research that proposed that heterogeneity exists within CRISPR arrays in bacterial populations (63, 64, 75).

The electrophoresis analysis shown in Figure 3.3 revealed that some T11RF pAM714 transconjugant populations passaged in erythromycin possessed multiple CRISPR3 alleles. We resolved the most abundant mutant CRISPR3 alleles in transconjugant populations by mapping amplicon reads to wild-type and mutant CRISPR3 references. The amplicon sequencing provided greater resolution than Sanger sequencing (Table 3.2). Moreover, all T11RF transconjugant populations, other than WT1, possessed multiple co-existing CRISPR3 alleles after 14 days of passage with antibiotic selection, and each of those alleles lacked S_6.

To further evaluate the level of heterogeneity, we calculated the average spacer rearrangement rate, which denotes the average percentage of total mapped reads that mapped to mutant alleles (Figure 3.5 and Table 3.4). Day 1 and Day 14 BHI-passaged T11RF had an average spacer rearrangement rate of $4e^{-4}$ to $8e^{-4}$, which is consistent with Day 1 and Day 14 erythromycin-passaged Δ4. This indicates the natural heterogeneity was maintained during the course of passaging. The WT1 population experienced reduced heterogeneity within the CRISPR3 array

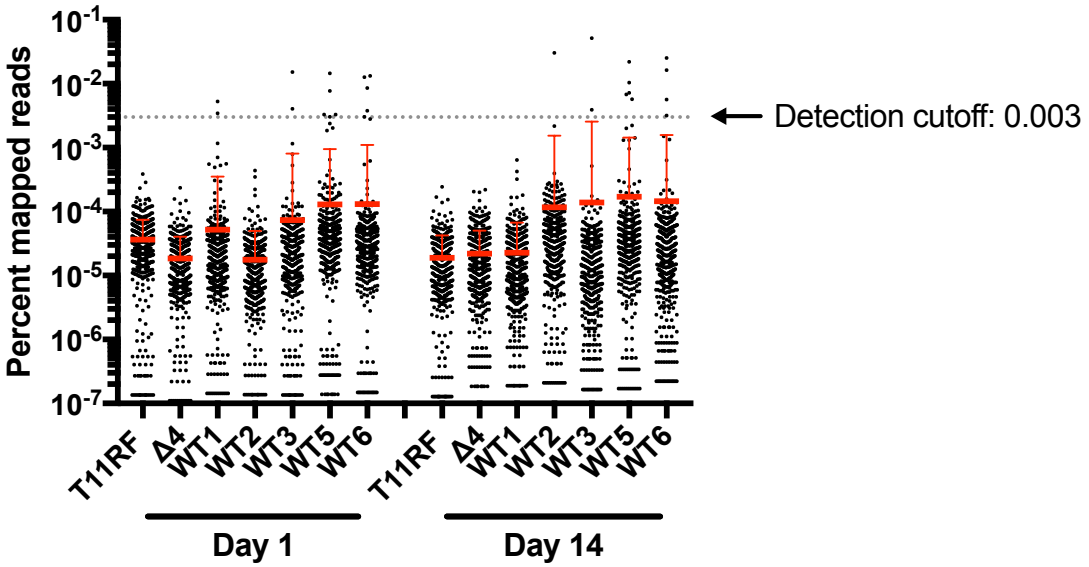


Figure 3.5. Percent mapped reads to mutant alleles. Percent mapped reads were calculated for each artificial reference by dividing mapped reads at each position to the total number of mapped reads. The percent mapped reads to mutant alleles (dots) are shown here with average (thick red bar) and standard deviation (thin red bar). A detection cutoff value was applied so that mutant alleles with high abundances can be detected.

region, which is possibly due to a *cas9* mutation as previously mentioned. WT2, WT3, WT5 and WT6 showed variable increases in the heterogeneity. The WT2 population experienced the most change in heterogeneity with an 8-fold increase, which is also explained by the increased number of PCR bands (Figure 3.3) and the change of dominant CRISPR3 alleles (Table 3.2). Day 1 erythromycin-passaged WT5 and WT6 started with a high level of heterogeneity and this heterogeneity was well-maintained during the course of passaging, hence the WT5 and WT6 populations had the least increase in the level of heterogeneity. Overall, we conclude that natural heterogeneity in CRISPR3 exists at a low frequency in BHI-passaged T11RF and erythromycin-passaged Δ4 populations, and that this heterogeneity provides a genetic basis for the survival of erythromycin-passaged WT transconjugant populations. In other words, mutant CRISPR alleles

occur at a low frequency in T11RF populations, and cells with mutant alleles lacking spacer 6 have a competitive advantage when antibiotic selection for pAM714 is present.

Preference for forward spacer deletion events.

We categorized mutant CRISPR3 alleles into two groups of events: forward spacer deletion and backward spacer rearrangement. The forward mutant group represents the mutant alleles with spacer deletion (5'-S_xRS_y-3' where $y > x+1$ and 5'-S_xTRS_T-3' where $x < 21$), while the backward mutant group represents the mutant alleles with a terminal spacer becoming more leader-proximal (5'-S_xRS_y-3' where $y < x$). We calculated forward spacer deletion and backward spacer rearrangement rates for each 5' spacer as described in Materials and Methods.

We observed slightly higher forward spacer deletion rates than backward spacer rearrangement rates for leader end spacers in Day 1 and Day 14 BHI-passaged T11RF, suggesting that spacers at the leader end are more readily deleted than flipped (Figure 3.6a). The forward deletion and backward rearrangement rates are similar at spacers S_9-S_{15} for Day 1 and Day 14 T11RF populations, indicating an equal chance of spacer deletion and flip (Figure 3.6a). As the 5' spacer reaches the terminal end of the array, the forward deletion and backward rearrangement rates decrease, indicating a dormant activity of spacer rearrangement near the terminal end. On average, Day 14 T11RF showed slightly lower rates for forward deletion and backward rearrangement than Day 1 T11RF.

Table 3.4. Wild-type CRISPR3 allele frequencies before and after passage.

Sample name	Total mapped reads	# mapped to WT references	# mapped to mutant references	Percent S_6-containing reads	Average spacer rearrangement rate
Day 1 T11RF control[a]	7398495	7273291	125204	96.11%	$8.7e^{-4} \pm 10.0e^{-4}$
Day 1 Δ4[b]	9073461	8995910	77551	141.03%	$4.1e^{-4} \pm 4.8e^{-4}$
Day 1 WT1[b]	6976811	6808760	168051	97.24%	$1.2e^{-3} \pm 6.5e^{-3}$
Day 1 WT2[b]	7281948	7222339	59609	133.12%	$3.9e^{-4} \pm 7.0e^{-4}$
Day 1 WT3[b]	7387447	7136664	250783	79.41%	$1.7e^{-3} \pm 1.6e^{-2}$
Day 1 WT5[b]	7215381	6784881	430500	39.95%	$3.3e^{-3} \pm 2.1e^{-2}$
Day 1 WT6[b]	6744053	6337706	406347	43.02%	$3.4e^{-3} \pm 2.4e^{-2}$
Day 14 T11RF control[a]	7829494	7760594	68900	142.59%	$4.2e^{-4} \pm 5.1e^{-4}$
Day 14 Δ4[b]	5424100	5369063	55037	155.27%	$5.1e^{-4} \pm 6.8e^{-4}$
Day 14 WT1[b]	5275288	5219558	55730	142.87%	$5.2e^{-4} \pm 8.9e^{-4}$
Day 14 WT2[b]	4744257	4489642	254615	30.89%	$3.3e^{-3} \pm 3.2e^{-2}$
Day 14 WT3[b]	6039118	5653212	385906	2.36%	$4.3e^{-3} \pm 5.5e^{-2}$
Day 14 WT5[b]	5844417	5384356	460061	4.03%	$5.2e^{-3} \pm 3.7e^{-2}$
Day 14 WT6[b]	4505777	4201804	303973	2.01%	$4.1e^{-3} \pm 4.3e^{-2}$

[a]T11RF without pAM714 passaged for 14 days in BHI medium.
[b]pAM714 transconjugants passaged for 14 days in BHI medium with erythromycin.
[c]Relative WT frequency is the fold change of WT frequency relative to Day 1 T11RF WT frequency.

The population diversity of the CRISPR3 array was evaluated in the same manner by analyzing the forward deletion and backward rearrangement rates in five T11RF pAM714 transconjugants and the Δ4 populations from both Day 1 and Day 14 erythromycin passages. The distribution of forward deletion and backward rearrangement rates in the Δ4 and WT1 populations (Figure 3.6b and c) were similar to T11RF (Figure 3.6a), except that we observed an elevated forward deletion rate at S_5 for WT1, demonstrating that mutant CRISPR3 alleles with at least a S_6 deletion have increased. This indicates that internal spacer deletion occurs in the Day 1 and Day 14 WT1 populations even through these events were not abundant enough to be detected by PCR. The rate of forward spacer deletion at S_5 for the WT1 Day 1 (black dot) population had decreased by Day 14 (red dot) of the passage resulting in the absence of mutant alleles within the detection limit (see Table 3.2). Consequently, we conclude that the erythromycin-passaged WT1 population is a heterogeneous mixture of CRISPR3 memory loss and *cas9* mutations, both of which resolve the conflict between pAM714 and CRISPR-Cas under antibiotic selection.

The forward spacer deletion events in T11RF pAM714 erythromycin-passaged transconjugants, WT2, WT3, WT5 and WT6, have unique positional preferences based on an increase in the average number of reads mapped to the mutant allele references containing spacer deletions (Figure 3.6 c-g; red or black dots). The elevated forward deletion rates were often observed for spacers upstream of S_6, indicating a positional preference for forward deletion events upstream of S_6. We speculate that this is because internal spacer deletions upstream of S_6 provide a selective advantage under these conditions. Finally, we did not observe significant fluctuation of

backward rearrangements in erythromycin-passaged transconjugants (Figure 3.6 c-g), supporting our claim that spacers are more readily deleted than flipped.

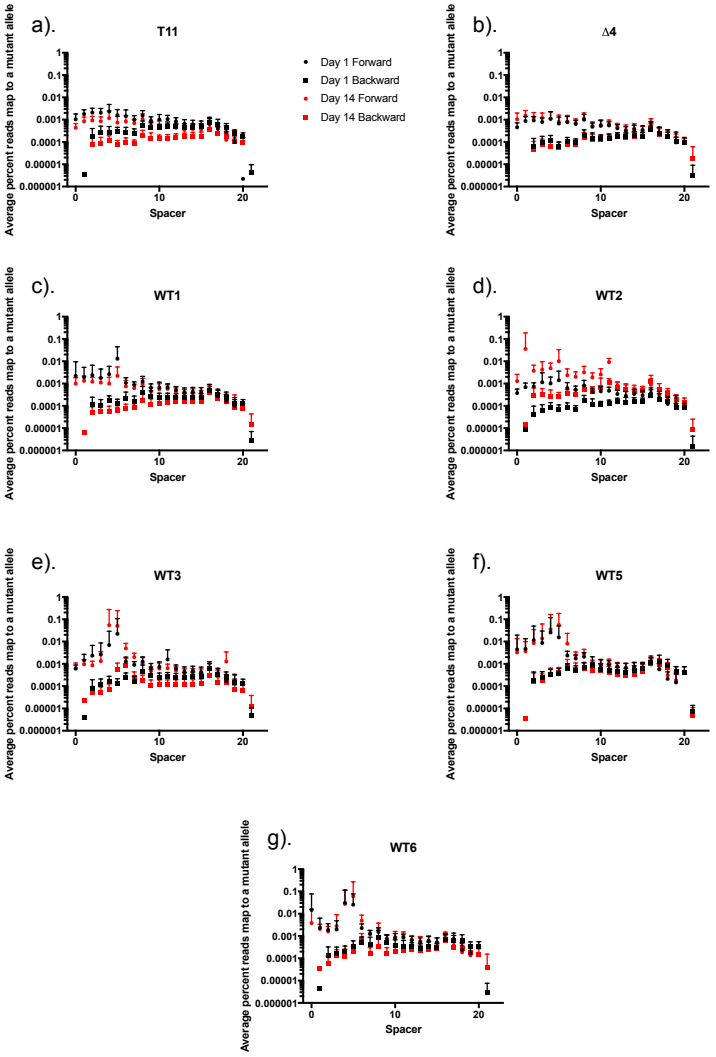


Figure 3.6. Amplicon sequencing revealed elevated forward spacer deletion rates at spacers upstream of S_6 in WT transconjugants. The forward spacer deletion and backward rearrangement rates (y-axis) were calculated for the CRISPR3 amplicon of each passaged population and are plotted against each spacer occurring in the CRIPSR3 array shown on the x-axis. For each sample, the forward deletion (dots) and backward rearrangement (squares) rates for Day 1 and Day 14 of the passage are shown in black and red, respectively. a) BHI passaged T11RF parent strain. b) Erythromycin passaged Δ4 transconjugant. c-g) Erythromycin passaged WT transconjugants.

Compromised CRISPR3-Cas resulting from pAM714 conflict benefits other MGEs.

Under antibiotic selection for pAM714, the CRISPR3-Cas system was compromised by S_6 deletion or *cas9* mutation. Considering that each spacer bears a unique memory of a previously encountered MGE, the loss of spacers surrounding S_6 could lead to compromised defense against multiple MGEs. To investigate this, we engineered pCF10, a PRP conferring tetracycline resistance (72), to encode different T11RF CRISPR3 protospacer targets along with the consensus CRISPR3 PAM sequence (47) (Table 3.1). We generated three pCF10 derivatives that would be targets for CRISPR3 S_1, S_6, and S_7, generating plasmids pWH107.S1, pWH107.S6, and pWH107.S7, respectively. Mutant CRISPR3 alleles with S_6 and S_7 deletions arose in all transconjugant populations that experienced array degeneracy after antibiotic passage (Table 3.2) and would therefore allow us to make conclusions about compromised defense. S_1 was maintained in all transconjugant populations (Table 3.2) and serves as a test for intact CRISPR-Cas function in passaged populations. We used wild type pCF10 as a control for baseline conjugation frequency as pCF10 is not targeted by T11RF CRISPR3-Cas (47).

We performed conjugation using Day 14 BHI-passaged T11RF as a recipient and *E. faecalis* OG1SSp bearing pCF10 and its derivatives as donors to ascertain the impact of CRISPR3-Cas on conjugation frequency of the plasmid constructs. All protospacers were targeted, resulting in significant reductions in conjugation frequencies relative to wild-type pCF10 (Figure 3.7). However, the degree of interference with plasmid transfer was different for each target; CRISPR-Cas defense against a MGE bearing a target for S_7 was weak compared to S_1 and S_6.

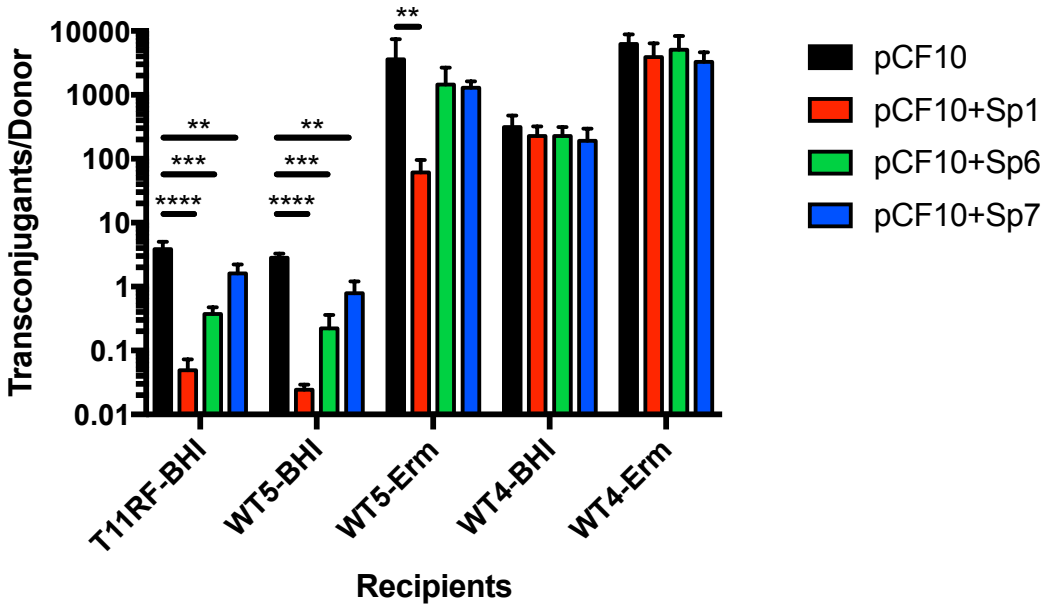


Figure 3.7. Compromised CRISPR-Cas primes populations for MGE acquisition. Day 14 transconjugant populations passaged in BHI and erythromycin were used as recipients in conjugation with OG1SSp:pCF10 and derivatives with protospacers corresponding to spacers 1, 6 and 7 of the T11RF CRISPR3 array. Day 14 BHI passaged T11RF parent strain was used as recipient in conjugation, serving as positive control. The graph shows the conjugation frequency or ratio of transconjugants to donors from mating reactions. Statistical significance was determined using a Student's T-test; P-values: ** ≤ 0.01; *** ≤ 0.001; **** ≤ 0.0001.

To determine the impact of serial passage on CRISPR-Cas defense in our transconjugant populations, we assessed conjugation of pCF10 and its derivatives to Day 14 BHI- and erythromycin-passaged populations of WT5. The erythromycin-passaged WT5 transconjugant population experienced the greatest diversification within the CRISPR3 array (Table 3.2), and we expected it to be compromised in terms of its ability to defend against other CRISPR targets. Day 14 BHI-passaged WT5 had conjugation frequencies very similar to BHI-passaged T11RF for all pCF10 plasmids (Figure 3.7). This was expected because the BHI-passaged WT5 population became depleted for pAM714 over time (Figure 3.3a), and did not have any changes in the CRISPR3 array (Table 3.2), therefore it was expected to maintain CRISPR-Cas activity

against all CRISPR3 targets. In contrast, Day 14 erythromycin-passaged WT5 exhibited defense only against pCF10 bearing a target for S_1 (Figure 3.7). This is consistent with the amplicon analysis that identified multiple CRISPR3 alleles with deletions of S_6 and S_7 in the erythromycin-passaged WT5 population (Table 3.2).

We also tested the WT4 transconjugant populations for CRISPR-Cas activity. We detected a mutation within the RuvC catalytic domain coding region of *cas9* in WT4 after passage Day 1, and WT4 failed to deplete pAM714 when passaged without erythromycin selection (Figure 3.3a). We expected both the BHI- and erythromycin-passaged populations of WT4 to be completely deficient for CRISPR-Cas activity if the observed mutation conferred loss of Cas9 function. CRISPR-Cas activity against S_1, S_6, and S_7 targets was in fact absent in these populations (Figure 3.7).

We observed that the transfer frequencies of pCF10 and its derivatives were higher all populations containing pAM714 (WT5-Erm, WT4-BHI, and WT4-Erm in Figure 3.7). We infer that pAM714 enhances pCF10 conjugation frequency via an unknown mechanism.

Spacer deletion is not exclusively RecA-dependent.

Upon antibiotic selection, the T11RF transconjugants lost S_6 to resolve the conflict between CRISPR-Cas and its target. The loss of S_6 was often coupled with the loss of surrounding spacers, ranging from S_1 to S_{18} (Table 3.2). The rearrangements associated with shortened CRISPR3-Cas arrays occurred between repeat-spacer junctions leaving behind perfectly intact

repeat-spacer-repeat sequences that are still of use as guides for CRISPR interference. This phenomenon led us to hypothesize that either homologous recombination or DNA replication slippage plays a role in eliminating S_6 from the array. To study if homologous recombination had an impact on spacer loss, we constructed an in-frame deletion of *recA* in T11RF, generating strain T11RF Δ*recA*. The pAM714 plasmid was introduced into T11RF *ΔrecA* through the same conjugation procedures described previously and two select transconjugants (recA.TC1 and recA.TC2) were serially passaged for 14 days with continuous erythromycin selection. The PCR analysis for the select transconjugant colonies indicated that recA.TC1 had a wild type CRISPR3-Cas array size while recA.TC2 had a shortened CRISPR3-Cas array (Figure 3.8). Using Sanger sequencing, we observed that recA.TC1 lost S_6 after one day of passage in erythromycin, while the initial recA.TC2 colony had a deletion of S_6-S_7. The same CRISPR3 alleles were detected by Sanger sequencing from the Day 14 erythromycin-passaged transconjugants. These data demonstrate that spacer deletion can occur in the absence of *recA*, and implicates DNA replication slippage in the emergence of mutant CRISPR alleles.

Mechanism of conflict resolution is not specific to *E. faecalis* T11RF or pAM714.

We wanted to determine if our observations were limited to one host-plasmid pair. Therefore, we expanded our analysis to include *E. faecalis* OG1RF, which possesses a Type II CRISPR-Cas system, CRISPR1-Cas, that is related to but distinct from CRISPR3-Cas of T11RF (12, 45) Like all sequenced *E. faecalis* strains, OG1RF also possesses the orphan CRISPR2 array (Figure 3.1). Previous research demonstrated that the T11RF CRISPR2 array is active for genome defense in the presence of CRISPR1 *cas9* (47). We interpret this to mean that the orphan CRISPR2 locus

can be used as a native genome defense system in *E. faecalis* OG1RF, due to the presence of endogenous CRISPR1-Cas which was recently demonstrated to provide genome defense (62).

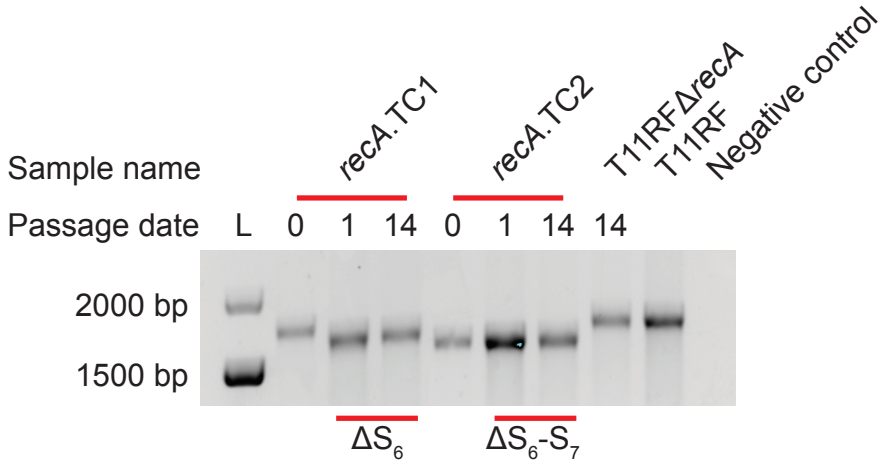


Figure 3.8. T11RFΔ*recA* transconjugants containing pAM714 experienced reduction in CRISPR3 size. Two randomly selected transconjugants were passaged *in vitro* with erythromycin selection and the CRISPR3 amplicon sizes were monitored using PCR and gel electrophoresis. As a control, the T11RFΔ*recA* parent strain was passaged in BHI and used as a control in PCR analysis. L, DNA ladder.

We utilized the shuttle vector pLZ12, which confers chloramphenicol resistance, as a backbone for the generation of artificial OG1RF CRISPR1-Cas and CRISPR2 protospacer targets (Table 3.1). The pKH12 plasmid does not natively contain a protospacer that would be targeted by either CRISPR1-Cas or CRISPR2 spacers (62). pKHS96 is a pKH12 derivative with an engineered CRISPR1-Cas protospacer that is targeted by OG1RF CRISPR1-Cas S_4. pKHS5 is a pKH12 derivative with an engineered CRISPR2 protospacer that is targeted by OG1RF CRISPR2 S_6. The consensus PAM sequence for both CRISPR1-Cas and CRISPR2 is NGG (47) and was included adjacent to the engineered protospacers. pKH12, pKHS96 and pKHS5 were each transformed into electrocompetent OG1RF. Twenty random transformants for each plasmid

were selected as templates for PCR to determine the initial integrity of the CRISPR1-Cas and CRISPR2 arrays using Sanger sequencing. We determined that the CRISPR1-Cas and CRISPR2 arrays were intact in all selected transformants, regardless of the plasmid that was transformed.

We randomly selected three transformants for each plasmid to be used for *in vitro* evolution experiments. Each transformant was passaged in plain BHI medium and BHI medium supplemented with chloramphenicol for a period of 14 days. Similar to our observations for T11RF pAM714 transconjugants, we observed loss of pKHS5 and pKHS96 over the course of passaging without antibiotic selection (Figure 3.9a). On the other hand, pKH12 was stably maintained during the passage in the absence of chloramphenicol.

CRISPR1 and CRISPR2 integrity was assessed for transformants on passage Day 14 (Figure 3.9b). All three pKHS96 transformants had reduced CRISPR1 arrays and two of three pKHS5 transformants had reduced CRISPR2 arrays after 14 days passage with chloramphenicol selection. Using Sanger sequencing, we confirmed that three pKHS96 transformants lost CRISPR1 S_4 while CRISPR2 remained intact, and two pKHS5 transformants lost CRISPR2 S_6 while CRISPR1 remained intact. The chloramphenicol-passaged pKHS5 transformant without a visible reduction in the CRISPR2 amplicon size was confirmed to have a mixed spacer population. Sanger sequencing revealed mixed nucleotides with low sequencing quality overlapping S_5 and S_6. This confirms that both OG1RF CRISPR1-Cas and CRISPR2 can become compromised when selection for CRISPR-targeted MGEs is present, which is consistent with what we observed in T11RF CRISPR3-Cas. Overall, we conclude that regardless of the CRISPR

subtype involved or the nature of the plasmid (naturally occurring PRP or shuttle vector), spacer loss events occur under antibiotic selection for maintenance of CRISPR-Cas targets in *E. faecalis*.

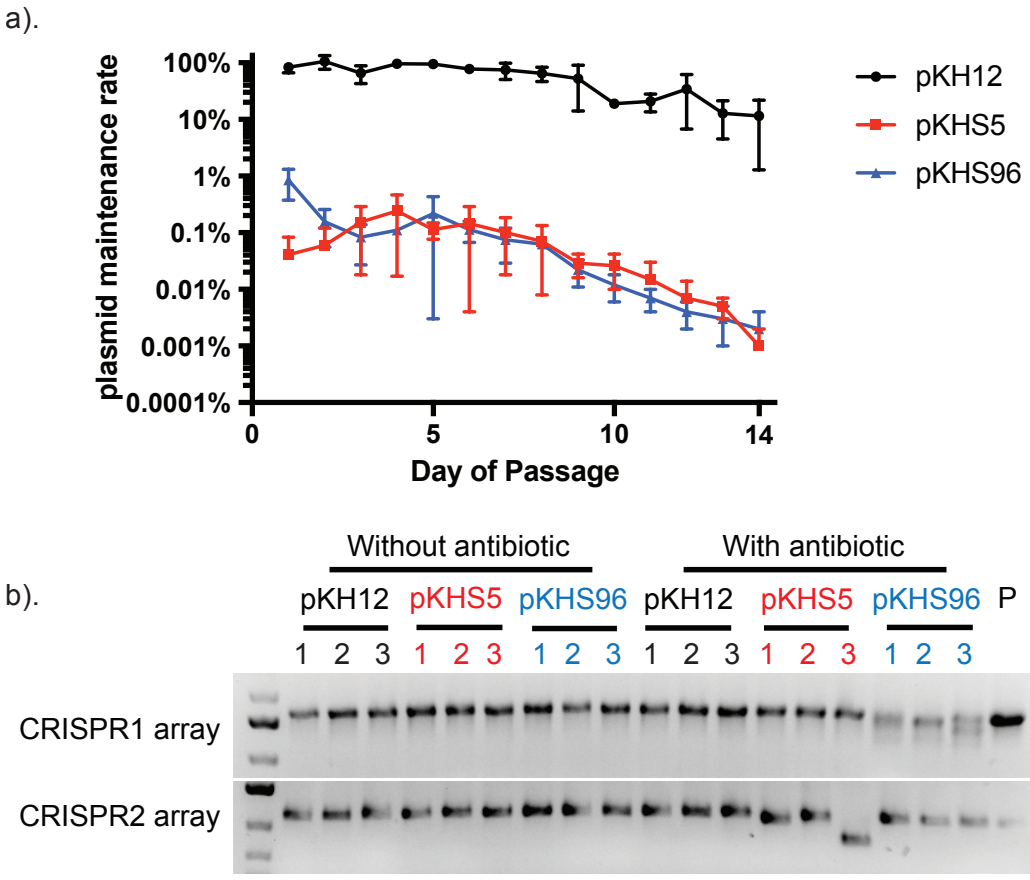


Figure 3.9. The phenomenon of plasmid elimination and spacer deletion upon different selection environment is conserved in CRISPR1-Cas and CRISPR2 of *E. faecalis* OG1RF. a) Plasmid maintenance rates of OG1RF transformants passaged in the absence of chloramphenicol are calculated as percentage of chloramphenicol resistant bacterial cells and plotted against passage days. Each dot represents the average rate from three transformants along with the standard deviation. b) CRISPR1 and CRISPR2 amplicon PCR results from Day 14 transformant populations passaged without antibiotic (left) and with antibiotic (right). P: positive control.

3.6 Discussion

It has been well documented that the misuse of antibiotics contributes to the emergence of MDR organisms. This is of particular concern in the opportunistic pathogen *E. faecalis* due to their intrinsic antibiotic resistance and their propensity to engage in HGT events that allow them to acquire genes providing resistance to other antibiotics. It is possible that antibiotic treatment helps to drive the evolution of MDR strains of *E. faecalis*. Usually, bacteria encode genome defense systems, such as CRISPR-Cas, to prevent HGT. Compromised CRISPR-Cas systems have been observed in MDR strains of *E. faecalis* (45) substantiating the claim that compromised genome defense leads to the evolution of MDR *E. faecalis*.

In our study, we used *in vitro* passaging experiments and deep sequencing analysis of CRISPR3 amplicons to study the dynamics of CRISPR-Cas and its plasmid target, pAM714, in transconjugants where these systems are in conflict. We find that the CRISPR3 array of T11RF populations is naturally heterogeneous in allelic structure, with most possible spacer deletion alleles occurring at low frequencies. When a CRISPR target is present, CRISPR-Cas eliminates its target from the population over time. However, when antibiotic selection for the target is present, CRISPR-Cas mutants arise that allow the plasmid to be maintained and promote bacterial survival. One would reason that the heterogeneity of a CRISPR array could be a result of either homologous recombination or slippage during DNA replication. Our results demonstrate that *recA* is not required for CRISPR compromisation by spacer deletion. However, the fact that we observed flipped spacers, where x > y in 5'-S_xRS_y-3', indicates that homologous recombination does play a role. It is likely that both mechanisms contribute to the emergence of

heterogeneous CRISPR alleles; we do not have an estimate of which process has a greater effect, nor whether additional stresses beyond antibiotic selection could influence rates for each. Moreover, we do not know whether sub-inhibitory antibiotic concentrations or fluctuating selection could alter outcomes of these conflicts.

Our studies utilized pAM714, which encodes a toxin-antitoxin system (76-78). The system encodes a stable toxin that will kill daughter cells that have not inherited a plasmid copy; an unstable antitoxin is encoded from the same locus that blocks toxin translation in cells with proper plasmid segregation. However, in our study, we observed a gradual decrease of erythromycin-resistant cells when we passaged T11RF pAM714 in BHI for 14 days. In this case, the toxin-antitoxin system in pAM714 seems to lose its efficiency over time, or its effect is overwhelmed by the active CRISPR-Cas system. The fact that pAM714, pKHS67 and pKHS5 were also eliminated gradually over the passage and not immediately suggests that *E. faecalis* CRISPR-Cas systems either act slowly or are not very effective under native conditions. An initial lag in Cas9 activation would explain the ability of pAM714 to become established in a subpopulation of CRISPR-Cas active cells. It is of interest to study the efficiency of plasmid elimination in T11RF that overexpresses Cas9. It is also of interest to identify mechanisms of *cas9* expression regulation.

The work presented here contributes to the understanding of different qualities of CRISPR-Cas systems from different species and how CRISPR-Cas reacts to its target during *in vitro* evolution. Our work stresses the importance of antibiotic usage on shaping the immune systems

of bacteria and its potential contribution to the emergence of MDR *E. faecalis*. We suspect that the phenomenon described in this study can be applied to other type II CRISPR-Cas systems as a function of the kinetics of Cas9 activity. Future studies will focus on what happens to the erythromycin-passaged T11RF pAM714 populations if they are maintained in medium without antibiotics. We would expect a subpopulation of the cells to acquire a new spacer targeting pAM714 while the rest of the subpopulation will maintain pAM714 at some level.

3.7 Acknowledgments

This work was supported by Public Health Service grant K22AI099088 to K.L.P and the Cecil H. and Ida Green Chairship to M.Q.Z. We thank Karthik Hullahalli for plasmid construction of pKH12 and its derivatives.

3.8 References

1. Lebreton F, Willems RJL, & Gilmore MS (2014) *Enterococcus* Diversity, Origins in Nature, and Gut Colonization. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

2. Arias CA & Murray BE (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nature reviews. Microbiology* 10(4):266-278.

3. Agudelo Higuita N & Huycke MM (2014) Enterococcal Disease, Epidemeiology, and Implications for Treatment. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

4. Jett BD, Huycke MM, & Gilmore MS (1994) Virulence of enterococci. *Clinical microbiology reviews* 7(4):462-478.

5. Maki DG & Agger WA (1988) Enterococcal bacteremia: clinical features, the risk of endocarditis, and management. *Medicine* 67(4):248.

6. Richards MJ, Edwards JR, Culver DH, & Gaynes RP (2000) Nosocomial infections in combined medical-surgical intensive care units in the United States. *Infection Control & Hospital Epidemiology* 21(8):510-515.

7. Arias CA, Contreras GA, & Murray BE (2010) Management of multidrug-resistant enterococcal infections. *Clinical microbiology and infection* 16(6):555-562.

8. Murray B (1992) Beta-lactamase-producing enterococci. *Antimicrobial agents and chemotherapy* 36(11):2355.

9. Portillo A*, et al.* (2000) Macrolide Resistance Genes in *Enterococcus* spp. *Antimicrobial agents and chemotherapy* 44(4):967-971.

10. Zervos M & Schaberg D (1985) Reversal of the *in vitro* susceptibility of enterococci to trimethoprim-sulfamethoxazole by folinic acid. *Antimicrobial agents and chemotherapy* 28(3):446-448.

11. Zimmermann RA, Moellering RC, & Weinberg AN (1971) Mechanism of resistance to antibiotic synergism in enterococci. *Journal of bacteriology* 105(3):873-879.

12. Bourgogne A*, et al.* (2008) Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome biology* 9(7):R110.

13. Palmer KL, Kos VN, & Gilmore MS (2010) Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Current opinion in microbiology* 13(5):632-639.

14. Paulsen IT*, et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299(5615):2071-2074.

15. Hegstad K, Mikalsen T, Coque T, Werner G, & Sundsfjord A (2010) Mobile genetic elements and their contribution to the emergence of antimicrobial resistant *Enterococcus faecalis* and *Enterococcus faecium*. *Clinical microbiology and infection* 16(6):541-554.

16. Huycke MM, Sahm DF, & Gilmore MS (1998) Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future. *Emerging infectious diseases* 4(2):239.

17. Kristich CJ, Rice LB, & Arias C, A (2014) Enterococcal Infection-Treatment and Antibiotic Resistance. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

18. Uttley A*, et al.* (1989) High-level vancomycin-resistant enterococci causing hospital infections. *Epidemiology & Infection* 103(1):173-181.

19. Andersson DI & Levin BR (1999) The biological cost of antibiotic resistance. *Current opinion in microbiology* 2:489-493.

20. De Gelder L, Ponciano JM, Joyce P, & Top EM (2007) Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology* 153(2):452-463.

21. Harrison E & Brockhurst MA (2012) Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in microbiology* 20(6):262-267.

22. Sota M*, et al.* (2010) Shifts in Host Range of a Promiscuous Plasmid through Parallel Evolution of its Replication Initiation Protein. *The ISME journal* 4(12):1568.

23. Hulter N*, et al.* (2017) An evolutionary perspective on plasmid lifestyle modes. *Current opinion in microbiology* 38:74-80.

24. Gandon S & Vale PF (2014) The evolution of resistance against good and bad infections. *J Evol Biol* 27(2):303-312.

25. Stalder T*, et al.* (2017) Emerging patterns of plasmid-host coevolution that stabilize antibiotic resistance. *Scientific reports* 7(1):4853.

26. Harrison E, Guymer D, Spiers AJ, Paterson S, & Brockhurst MA (2015) Parallel compensatory evolution stabilizes plasmids across the parasitism-mutualism continuum. *Current biology* 25(15):2034-2039.

27. Hughes JM, *et al.* (2012) The role of clonal interference in the evolutionary dynamics of plasmid-host adaptation. *mBio* 3(4):e00077-00012.

28. Clewell DB, *et al.* (2014) Extrachromosomal and Mobile Elements in Enterococci: Transmission, Matintenance, and Epidemiology. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

29. Dunny GM (2007) The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 362(1483):1185-1193.

30. Gilmore MS, *et al.* (1994) Genetic structure of the *Enterococcus faecalis* plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants. *J Bacteriol* 176(23):7335-7344.

31. Clewell DB (2007) Properties of *Enterococcus faecalis* plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin. *Plasmid* 58(3):205-227.

32. Dunny GM (2013) Enterococcal sex pheromones: signaling, social behavior, and evolution. *Annual review of genetics* 47:457-482.

33. Makarova KS, *et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nature reviews. Microbiology* 13(11):722-736.

34. Marraffini LA (2015) CRISPR-Cas immunity in prokaryotes. *Nature* 526(7571):55-61.

35. Bolotin A, Quinquis B, Sorokin A, & Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(Pt 8):2551-2561.

36. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, & Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of molecular evolution* 60(2):174-182.

37. Pourcel C, Salvignol G, & Vergnaud G (2005) CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151(Pt 3):653-663.

38. Heler R*, et al.* (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519(7542):199-202.

39. Wei Y, Terns RM, & Terns MP (2015) Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes & development* 29(4):356-361.

40. Chylinski K, Le Rhun A, & Charpentier E (2013) The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA biology* 10(5):726-737.

41. Deltcheva E*, et al.* (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471(7340):602-607.

42. Anders C, Niewoehner O, Duerst A, & Jinek M (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513(7519):569-573.

43. Nishimasu H*, et al.* (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156(5):935-949.

44. Sternberg SH, Redding S, Jinek M, Greene EC, & Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507(7490):62-67.

45. Palmer KL & Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4).

46. Hullahalli K*, et al.* (2015) Comparative Analysis of the Orphan CRISPR2 Locus in 242 *Enterococcus faecalis* Strains. *PloS one* 10(9):e0138890.

47. Price VJ, Huo W, Sharifi A, Palmer KL, & Fey PD (2016) CRISPR-Cas and Restriction-Modification Act Additively against Conjugative Antibiotic Resistance Plasmid Transfer in *Enterococcus faecalis*. *mSphere* 1(3):e00064-00016.

48. Barrangou R, *et al.* (2015) Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. *Nucleic acids research* 43(7):3407-3419.

49. Barrangou R & van Pijkeren JP (2015) Exploiting CRISPR-Cas immune systems for genome editing in bacteria. *Current opinion in biotechnology* 37:61-68.

50. Ji W, *et al.* (2014) Specific gene repression by CRISPRi system transferred through bacterial conjugation. *ACS synthetic biology* 3(12):929-931.

51. Jiang Y, *et al.* (2015) Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Applied and environmental microbiology* 81(7):2506-2514.

52. Kleinstiver BP, *et al.* (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529(7587):490-495.

53. Hsu PD, Lander ES, & Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157(6):1262-1278.

54. Jinek M, *et al.* (2013) RNA-programmed genome editing in human cells. *eLife* 2:e00471.

55. Ran FA, *et al.* (2013) Genome engineering using the CRISPR-Cas9 system. *Nature protocols* 8(11):2281-2308.

56. Andersson AF & Banfield JF (2008) Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320(5879):1047-1050.

57. Bzymek M & Lovett ST (2001) Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences* 98(15):8319-8325.

58. Erdmann S & Garrett RA (2012) Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Molecular microbiology* 85(6):1044-1056.

59. Iranzo J, Lobkovsky AE, Wolf YI, & Koonin EV (2013) Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *Journal of bacteriology* 195(17):3834-3844.

60. Louwen R*, et al.* (2013) A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain-Barre syndrome. *European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology* 32(2):207-226.

61. Tyson GW & Banfield JF (2008) Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environmental microbiology* 10(1):200-207.

62. Hullahalli K, Rodrigues M, & Palmer KL (2017) Exploiting CRISPR-Cas to manipulate *Enterococcus faecalis* populations. *eLife* 6.

63. Lopez-Sanchez MJ*, et al.* (2012) The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Molecular microbiology* 85(6):1057-1071.

64. Jiang W*, et al.* (2013) Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS genetics* 9(9):e1003844.

65. Bae T, Kozlowicz B, & Dunny GM (2002) Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and *prgX* mRNA and in regulation of pheromone-inducible conjugation. *Journal of molecular biology* 315(5):995-1007.

66. Thurlow LR, Thomas VC, & Hancock LE (2009) Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *Journal of bacteriology* 191(20):6203-6210.

67. Manson JM, Keis S, Smith JM, & Cook GM (2003) A clonal lineage of VanA-type *Enterococcus faecalis* predominates in vancomycin-resistant enterococci isolated in New Zealand. *Antimicrobial agents and chemotherapy* 47(1):204-210.

68. Clewell D*, et al.* (1982) Mapping of *Streptococcus faecalis* plasmids pAD1 and pAD2 and studies relating to transposition of Tn*917*. *Journal of Bacteriology* 152(3):1220-1230.

69. Ike Y, Clewell D, Segarra R, & Gilmore M (1990) Genetic analysis of the pAD1 hemolysin/bacteriocin determinant in *Enterococcus faecalis*: Tn*917* insertional mutagenesis and cloning. *Journal of bacteriology* 172(1):155-163.

70. McBride SM, Fischetti VA, Leblanc DJ, Moellering RC, Jr., & Gilmore MS (2007) Genetic diversity among *Enterococcus faecalis*. *PloS one* 2(7):e582.

71. Gold OG, Jordan H, & Van Houte J (1975) The prevalence of enterococci in the human mouth and their pathogenicity in animal models. *Archives of Oral Biology* 20(7):473IN415-477.

72. Dunny G, Yuhasz M, & Ehrenfeld E (1982) Genetic and physiological analysis of conjugation in *Streptococcus faecalis*. *Journal of Bacteriology* 151(2):855-859.

73. Jett BD, Jensen HG, Nordquist RE, & Gilmore MS (1992) Contribution if the pAD1-Encoded Cytolysin to the Severity of Experimental *Enterococcus faecalis* Endophthalmitis. *Infection and immunity* 60(6):2445-2452.

74. Perez-Casal J, Caparon M, & Scott J (1991) Mry, a trans-acting positive regulator of the M protein gene of *Streptococcus pyogenes* with similarity to the receptor proteins of two-component regulatory systems. *Journal of Bacteriology* 173(8):2617-2624.

75. Gudbergsdottir S*, et al.* (2011) Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Molecular microbiology* 79(1):35-49.

76. Greenfield TJ & Weaver KE (2000) Antisense RNA regulation of the pAD1 *par* post-segregational killing system requires interaction at the 5' and 3' ends of the RNAs. *Molecular microbiology* 37(3):661-670.

77. Weaver KE & Clewell DB (1989) Construction of *Enterococcus faecalis* pAD1 Miniplasmids: Identification of a Minimal Pheromone Response Regulatory Region and Evaluation of a Novel Pheromone-Dependent Growth Inhibition. *Plasmid* 22:106-119.

78. Weaver KE & Tritle DJ (1994) Identification and Characterization of an *Enterococcus faecalis* Plasmid pAD1-Encoded Stability Determinant Which Produces Two Small RNA Molecules Necessary for Its Function. *Plasmid* 32:168-181.

CHAPTER 4

ANTIBIOTIC RESISTANCE PLASMID TRANSFER KINETICS AND THE IMPACT OF CRISPR-CAS IN A MOUSE MODEL OF *ENTEROCOCCUS FAECALIS* COLONIZATION

4.1 Author contribution

This work was conducted in collaboration with Sara McBride (SM), Breck Duerkop (BD) and Kelli L. Palmer (KP). The conception and design of experiments was carried out by Valerie Price (VP), SM, BD and KP. VP collected and analyzed the data from *in vitro* experiments. The *in vivo* studies were performed by VP, SM and BD at the University of Colorado Denver Anschutz Medical Center in Aurora, CO. The analysis of *in vivo* data was performed by VP and SM. This chapter was written by VP and was edited by VP and KP. The contents of this chapter are being prepared as an original manuscript that will be submitted for publication in 2017.

4.2 Abstract

Enterococcus faecalis is one of the leading causes of health-care associated infections in the United States. This bacterium readily acquires antibiotic resistance through horizontal gene transfer mediated through plasmids and conjugative transposons. Prokaryotes encode genome defense systems, such as CRISPR-Cas, that limit the acquisition of foreign DNA. Some *E. faecalis* possesses type II CRISPR-Cas systems, where a correlation between the lack of CRISPR-Cas and the presence of multiple antibiotic resistance among clinical isolates of *E. faecalis* has been observed. This suggests that the absence of genome defense leads to the

emergence of hospital-adapted *E. faecalis.* Previous studies demonstrated that CRISPR-Cas in *E. faecalis* is an effective barrier to plasmid acquisition *in vitro.* Here, we use a clinically relevant plasmid conferring antibiotic resistance to monitor the kinetics of plasmid transfer under planktonic and biofilm mating conditions in the presence and absence of CRISPR-Cas. We reveal that a plasmid-encoded virulence factor, cytolysin, impacts transfer efficiency over time, and this effect is dependent on planktonic conjugation conditions. We also expand our knowledge of *E. faecalis* CRISPR-Cas activity to *in vivo* studies by assessing conjugation in a mouse model of *E. faecalis* gastrointestinal colonization. We find that CRISPR-Cas prevents plasmid dissemination in 17 out of 20 mice. Overall, our findings reveal that the efficiency of CRISPR-Cas defense varies between *in vitro* and *in vivo* studies. These data warrant more in-depth exploration of the impact of CRISPR-Cas *in vivo* and the application of this knowledge toward developing alternative treatments.

4.3 Introduction

Enterococcus faecalis is a gram-positive bacterium that natively inhabits the gastrointestinal tracts (GI) of humans and other animals (1). *E. faecalis* is also an opportunistic pathogen that is among the leading causes of hospital acquired infections in the United States (2-4). *E. faecalis* is commonly associated with the life-threatening infections, endocarditis and bacteremia (2). Patients with suppressed or compromised immune systems are at a higher risk for contracting *E. faecalis* infections. It has been reported that *E. faecalis* is able to persist on the skin of patients and health care workers as well as on hospital surfaces, adding to the threat for the dissemination of *E. faecalis* infections in health care settings (5). Moreover, hospital-associated *E. faecalis*

strains are often multidrug-resistant (MDR) and encode resistance to the drug-of-last-resort, vancomycin; this leaves very few, if any, treatment options for *E. faecalis*-associated infections (3, 6).

The genes encoding antibiotic resistance are often acquired by *E. faecalis* through the process of horizontal gene transfer (HGT) (7-9). This process involves the lateral exchange of mobile genetic elements (MGEs), such as bacteriophage, conjugative transposons, and plasmids, between strains or species. Genomic analysis of two model *E. faecalis* strains, V583, the first vancomycin-resistant strain isolated in the United States, and OG1RF, a vancomycin-susceptible oral isolate, revealed that the hospital-adapted strain V583 possessed roughly 600-kb additional DNA sequence compared to OG1RF (7). Further inspection of the V583 genome uncovered that this DNA almost exclusively corresponded to MGEs, including multiple plasmids, integrated prophage, transposons, and a pathogenicity island (9). The phenotype of an expanded genome is not limited to V583 and holds true for many hospital-adapted *E. faecalis* isolates (7, 9-11). The ability of *E. faecalis* to engage in HGT in concurrence with the intrinsic antibiotic resistance inherent to this species make *E. faecalis*-associated infections very difficult to treat. This exemplifies the importance of understanding the mechanisms of antibiotic resistance acquisition and the development of alternative therapeutic strategies.

One of the most relevant means of HGT in this species is mediated through plasmids belonging to the narrow host range pheromone-responsive plasmids (PRPs). The PRPs are a unique class of plasmids that are rarely found outside of the *faecalis* species and are highly co-evolved with *E.*

faecalis (10, 12). These plasmids are large (~60-kb) and are very efficient at conjugation, with frequencies reaching 1 transconjugant for every 10-100 donor cells (13). PRPs are often identified in the genomes of clinical isolates of *E. faecalis* where they provide accessory antibiotic resistance genes (9, 14-18). In addition to antibiotic resistance, some PRPs also encode other traits that add to their virulence; one example of this is bacteriocin production. Bacteriocins are a class of antimicrobials produced by bacteria that confer competitive advantages in polymicrobial environments. The PRP pAD1 encodes a bacteriocin called cytolysin (19). Cytolysin is a lantibiotic-like antimicrobial peptide with activity against a number of gram-positive bacteria and certain mammalian cell types (20, 21). Cytolysin has both bactericidal and hemolytic properties which have been linked to enhancing the virulence of *E. faecalis* in some infections (22-25).

Many bacteria employ genome defense strategies to prevent the acquisition of selfish genetic elements such as bacteriophage and plasmids (26). One benefit to encoding such systems is to avoid metabolic burden associated with MGE carriage. A form of adaptive immunity is conferred by CRISPR-Cas systems. CRISPR-Cas systems contain short segments of DNA, called spacers, that have been acquired from MGEs; these sequences serve as molecular memories of past infection by MGEs (27, 28). Spacers are transcribed into RNA molecules that, when bound to an effector nuclease, identify foreign DNA and target it for cleavage (29-32). Interestingly, there is a strong correlation between the absence of CRISPR-Cas systems and the presence of multidrug-resistance associated with hospital-adapted *E. faecalis* (33). This, in association with the MGE-rich nature of hospital-associated strains, led to the hypothesis that

MDR *E. faecalis* emerge due to the absence of barriers to HGT. Until recently, it was still unknown whether or not genome defense systems, if present, were functional in this species.

E. faecalis possesses three type II CRISPR-Cas systems, CRISPR1-Cas, CRISPR2 and CRISPR3-Cas, that have a unique distribution among strains of this species. The CRISPR2 locus is found in all *E. faecalis* isolates and is not associated with *cas* genes (33, 34). This is the only CRISPR locus present in most hospital-adapted strains of *E. faecalis*, providing evidence that the absence of genome defense contributes to the rapid acquisition of MGEs. Commensal *E. faecalis* isolates possess the orphan CRISPR2 locus as well as one of the other two systems, CRISPR1-Cas or CRISPR3-Cas. We have recently demonstrated that all three systems are active for genome defense when they are in the presence of their cognate Cas9 enzyme (35, 36). One of these studies focused on the ability of CRISPR3-Cas to block the acquisition of pAD1, where we determined that pAD1 conjugation frequencies were 80-fold higher in an *E. faecalis* T11 *cas9* deletion mutant compared to the isogenic wild-type strain with active CRISPR-Cas (36). This study was significant to the field as it was one of the first demonstrations of active genome defense in this species, and added to existing evidence that hospital-adapted *E. faecalis* readily acquire antibiotic resistance due to the absence of barriers to HGT.

Although it has been demonstrated that CRISPR-Cas is functional for genome defense in *E. faecalis,* little is known about how growth conditions might affect its efficiency. Moreover, no studies have investigated the impact of CRISPR-Cas on plasmid transfer *in vivo*. In this study, we monitor the kinetics of PRP transfer over time in both planktonic and biofilm mating

conditions, mimicking the variation in potential *E. faecalis* infection sites. These experiments allow us to assess the activity of CRISPR-Cas defense over time under different conjugation circumstances. We also use a mouse gut colonization model to ascertain the impact of CRISPR-Cas on plasmid transfer *in vivo*. Our studies utilize two derivatives of pAD1, pAM714 and pAM771, in order to explore the effect of bacteriocin (cytolysin) production on CRISPR-Cas activity and colonization in the mouse gut. Overall, we conclude that mating condition and PRP-encoded virulence traits influence the efficiency of CRISPR-Cas to limit plasmid transfer *in vitro*, but not in our mouse model of *E. faecalis* colonization where we observe a cytolysin-independent enhanced colonization phenotype of *E. faecalis* strains bearing PRPs in the mouse gut.

4.4 Materials and Methods

Bacteria and reagents used.

Strains used in this study are shown in Table 4.1. All *E. faecalis* strains were cultured in brain heart infusion (BHI) broth or agar at 37°C. Antibiotic concentrations used were as follows: rifampicin, 50 μg/mL; fusidic acid, 25 μg/mL; spectinomycin, 500 μg/mL; streptomycin, 500 μg/mL; erythromycin, 50 μg/mL. Antibiotics were purchased from Sigma-Aldrich.

Conjugation experiments.

The same initial procedures were followed for both planktonic and biofilm conjugation reactions, as follows. Donor and recipient strains were cultured overnight in BHI broth in the absence of antibiotic selection. The following day, cultures were diluted 1:10 into fresh BHI and incubated

at 37°C for 1.5 hours. For planktonic conjugations, 2 mL of donor and 18 mL of recipient were mixed in a flask and incubated without agitation at 37°C for 30 min to 18 h. At each time point, 1 mL of the mating reaction was removed and used for serial dilutions and plating on selective media. For biofilm mating reactions, 100 μL of donor was mixed with 900 μL of recipient. The mixture was centrifuged for 1 min at 16,000 x g. After centrifugation, 100 μL supernatant was used to resuspend the pellet, which was then plated on non-selective BHI agar. To allow for sampling of multiple time points, multiple identical conjugation reactions were generated using the same donor and recipient inocula. The conjugation reactions were then incubated at 37°C for 30 min to 18 h. At each time point, cells were collected from a plate using 2 mL 1X PBS supplemented with 2 mM EDTA, and serial dilutions were plated on selective media. BHI agar supplemented with antibiotics was used to quantify the donor (spectinomycin, streptomycin and erythromycin), recipient (rifampicin and fusidic acid), and transconjugant (rifampicin, fusidic acid and erythromycin) populations in our conjugation experiments. Plates were incubated for 36-48 h at 37°C. Plates with 30 to 300 colonies were used to calculate CFU/mL.

Cytolysin immunity assay.

Recipient strain susceptibility to cytolysin was assessed by a previously described method (37), with slight modifications. To eliminate residual cytolysin activity due to the presence of the donor strain, 50 μL of mating reactions from each time point during conjugation were incubated overnight in BHI broth supplemented with rifampicin and fusidic acid to select for recipient and transconjugant strains. The following day, these donor-deficient cultures were used in a soft agar overlay on normal BHI agar. After allowing the overlay to dry, 5 μL of OG1SSp pAM714 was

spotted on the overlay as an indicator for cytolysin susceptibility. Plates were incubated overnight at 37°C and were inspected for zones of inhibition the following day.

Mouse model of *E. faecalis* colonization.

7 days prior to bacterial inoculation, 6-8 week old Jax B6 mice were gavaged with 100 μL of an antibiotic cocktail (streptomycin 1 mg/mL, gentamicin 1 mg/mL, erythromycin 200 μg/mL), and given a water bottle *ad libitum* with the same antibiotic cocktail for 6 days following gavage. 24 h prior to bacterial inoculation, antibiotic water was removed and replaced with standard sterile RO water. Bacteria were grown overnight and mice were gavaged with $1e^9$ CFU/mL in PBS of each bacterial strain as experimental groups indicated. Cultures used for gavage were also plated on BHI to confirm that inocula were equal across strains. Fecal samples from mice were collected at 0 h, 24 h, 48 h and 96 h. Fecal samples were resuspended in 1 mL of PBS and dilutions were plated on BHI agar supplemented with rifampicin, fusidic acid, erythromycin, streptomycin and spectinomycin in combinations that would select for desired *E. faecalis* populations. Plates were incubated for 36-48 h at 37°C. Plates with 30 to 300 colonies were used to calculate CFU/gram of feces. All animal protocols were approved by the Institutional Animal Care and Use Committee of the University of Colorado Anschutz Medical Campus (protocol number B-113916(09)1E).

4.5 Results

Degree of plasmid acquisition differs between mating condition.

Previous work on the impact of CRISPR-Cas on horizontal transfer of pAM714 revealed an ~80-fold increase in plasmid acquisition in a *cas9*-null derivative of *E. faecalis* T11RF after 18 hours of biofilm mating on an agar surface (36). To gain insight into the kinetics of plasmid transfer and whether or not the type of mating condition impacts the efficiency of CRISPR-Cas defense, we sampled an 18-hour mating reaction over six time points under planktonic and biofilm mating conditions at a donor to recipient ratio of 1:9. The donor strain used in these experiments was *E. faecalis* OG1SSp bearing the PRP pAM714, a derivative of pAD1 encoding erythromycin resistance. Our recipient strains were T11RF and a T11RF derivative with an in-frame deletion of *cas9*, T11RFΔ*cas9* (see Table 4.1 for a list of strains). Spacer 6 of the CRISPR3-Cas locus of T11RF has 100% sequence identity with the *repB* gene of pAM714, therefore, CRISPR3-Cas activity reduces pAM714 plasmid acquisition in wild-type T11 populations (36).

Table 4.1. *E. faecalis* strains used in this study.

Strain Name	Description	Reference
T11RF	Rifampicin-fusidic acid resistant derivative of strain T11	(11, 36)
T11RFΔ*cas9*	T11RF with an in-frame deletion of *cas9*	(36)
OG1SSp pAM714	Spectinomycin-streptomycin resistant derivative of strain OG1 harboring pAM714, conferring erythromycin resistance via Tn*917* insertion upstream of the *par* locus; *cyl*$^+$	(38, 39)
OG1SSp pAM771	Spectinomycin-streptomycin resistant derivative of strain OG1 harboring pAM771, conferring erythromycin resistance via Tn*917* insertion disrupting *cylL* of the cytolysin operon; *cyl*$^-$	(40, 41)

The number of pAM714 transconjugants (i.e., the number of T11RF or T11RFΔ*cas9* cells that acquired pAM714) was used to compare the kinetics of conjugative plasmid transfer in the presence and absence of CRISPR-Cas defense. After only 30 minutes of mating, we observed at least 10^3 transconjugants for both recipient strains (T11RF and Δ*cas9*) under both mating conditions; however, the overall trend for plasmid acquisition over time differed between the two conditions (Figure 4.1 A and B). Under planktonic mating conditions in broth, the number of T11Δ*cas9* pAM714 transconjugants remained stable over all time points sampled while the amount of T11RF pAM714 transconjugants decreased steadily (Figure 4.1A). This suggests that CRISPR-Cas in T11RF is actively targeting pAM714 in planktonic conjugation either by preventing plasmid acquisition or by causing plasmid loss after initial conjugation success in some cells. Indeed, we observe a statistically significant difference in the number of pAM714 transconjugants obtained in CRISPR-active and CRISPR-deficient recipient strains after 1 hour of conjugation; this significance was maintained for the duration of the experiment. On the other hand, the transfer kinetics of pAM714 in a biofilm mating resembles a linear trend for both recipient strains that achieved transconjugant yields previously reported for an 18-hour conjugation reaction in a biofilm (36) (Figure 4.1B). Consistent with planktonic mating, we observed that the number of T11Δ*cas9* pAM714 transconjugants was higher at all time points in a biofilm mating, but the difference was not statistically significant at all time points, demonstrating that CRISPR-Cas defense does not have an immediate impact on conjugation in a biofilm.

When comparing the number of transconjugants obtained between planktonic and biofilm mating conditions, without considering the impact of CRISPR-Cas, we observe a one to two log difference in the amount of pAM714 transconjugants obtained. Due to available data showing that PRP-mediated conjugation is more efficient on solid surfaces and that PRPs often encode auxiliary functions, we hypothesized that the discrepancy in transconjugant numbers could be due to pAM714-specific traits that influence conjugation efficiency under the conditions tested.

PRPs encode not only antibiotic resistance genes, but also virulence factors that promote their survival in polymicrobial communities. pAM714 encodes one of these factors, cytolysin, a bacteriocin that is bactericidal to some gram-positive bacteria including *E. faecalis* (21). The operon encoding the cytolysin molecule also produces a membrane-associated immunity factor that protects the host cell from the activity of cytolysin (37); therefore, pAM714-containing cells (and by extension, newly generated transconjugants) are immune to the bacteriocin. To determine if cytolysin production could impact pAM714 transfer kinetics, we utilized another derivative of pAD1, pAM771. In pAM771, the cytolysin synthesis genes are disrupted by a Tn*917* insertion. Experiments described above for pAM714 were also performed with OG1SSp pAM771 as a plasmid donor. Under planktonic conditions, we obtained a similar number of T11RF and T11RFΔ*cas9* pAM771 transconjugants across all time points (Figure 4.1C). These data are in contrast to the same experiments performed with pAM714 (Figure 4.1A), suggesting that cytolysin production affects the population dynamics of donor and recipient strains under planktonic mating conditions. This is further supported by a 1-log increase in T11RF pAM771

transconjugants and a 3-log increase in T11RFΔ*cas9* pAM771 transconjugants after 18 hours of conjugation compared to pAM714 transconjugant values at the same time point.

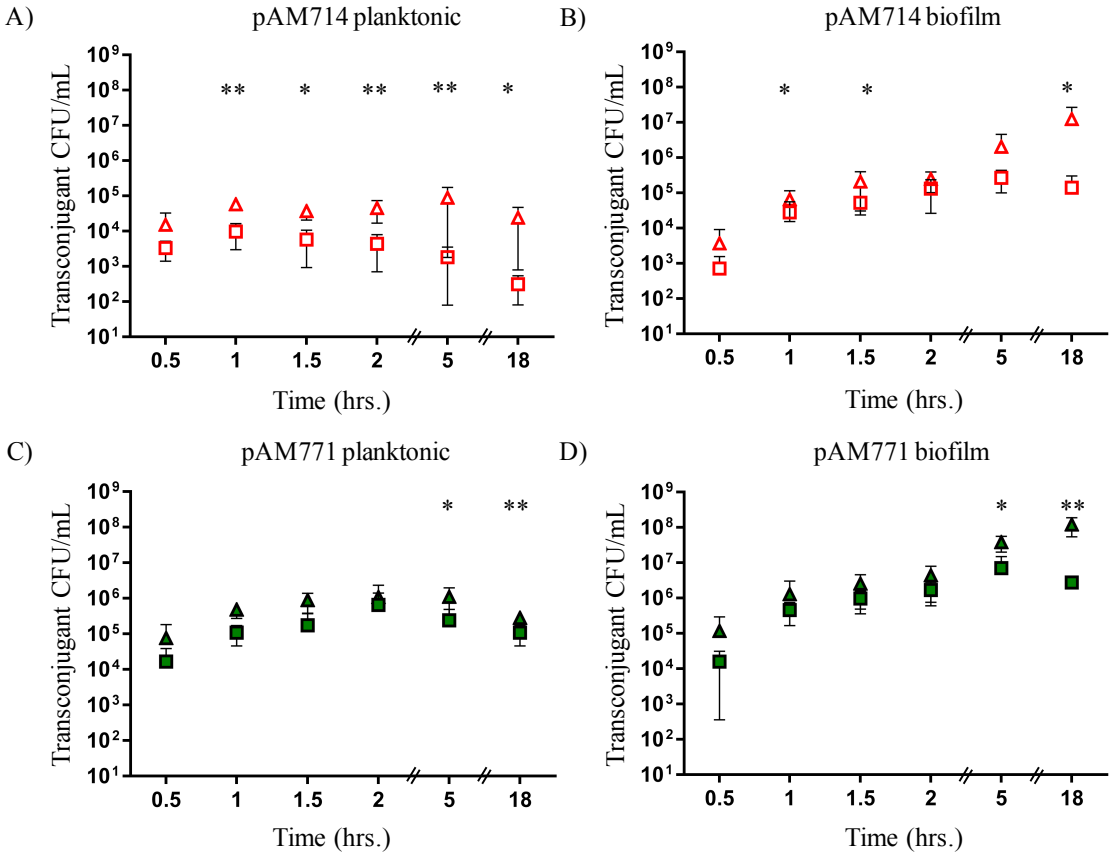


Figure 4.1. Mating condition impacts effectiveness of CRISPR-Cas on plasmid transfer *in vitro*. The CFU/mL of transconjugants obtained in mating reactions sampled over an 18-hour period is shown for T11RF (squares) and T11RFΔ*cas9* (triangles) recipient strains. Conjugation was performed under planktonic conditions in broth (A and C) and biofilm conditions on an agar plate (B and D) utilizing OG1SSp as a donor strain for plasmids pAM714 (open, red symbols) and pAM771 (closed, green symbols). These plasmids are isogenic except for the location of Tn*917* insertion; in pAM771, bacteriocin production is disrupted by Tn*917* insertion into *cylL*, whereas in pAM714 the Tn*917* occurs in an intergenic region upstream of the *par* locus. Data shown is the average and standard deviation from a minimum of three independent trials for each time point for both mating conditions. Statistical significance was assessed using a one-tailed Student's t-Test; *P*-values, *<0.05 and **<0.005.

Consistent with our results for pAM714 conjugation in a biofilm, the kinetics of pAM771 biofilm transfer also displayed a linear trend under these conditions (Figure 4.1D). In addition, we also detect a greater amount of transconjugants in biofilm mating with pAM771 than with pAM714. The impact of CRISPR-Cas on pAM771 conjugation achieves statistical significance in the later time points sampled (5 h and 18 h).

The data here have been presented as transconjugant CFU/mL, however the traditional way to express plasmid transfer is by calculating a conjugation frequency. This value is expressed as the CFU/mL of transconjugants divided by the CFU/mL of donors in a mating reaction. The conjugation frequencies from the planktonic and biofilm conjugation reactions were used to determine the impact of CRISPR-Cas on plasmid transfer by calculating a fold change that represents how much more plasmid is acquired in the absence of CRISPR-Cas defense (Δ*cas9* recipients). The fold change values are shown in Table 4.2. From these values, we conclude that the impact of CRISPR3-Cas on blocking the transfer of pAM714 or pAM771 is more robust under biofilm mating conditions. We also surmise that the PRP-encoded cytolysin impacts the kinetics of plasmid transfer and CRISPR-Cas efficiency under planktonic conditions.

Table 4.2. Fold Change of conjugation frequency

	Planktonic		Biofilm	
Time (hrs.)	pAM714	pAM771	pAM714	pAM771
0.5	4.56	4.12	11.26	0.73
1	4.26	5.54	4.90	3.76
1.5	6.74	5.99	1.64	2.94
2	10.56	1.71	2.42	2.55
5	50.12	10.12	3.49	6.96
18	37.84	3.38	45.80	53.24

Recipient strain susceptibility to cytolysin is enhanced in planktonic mating reactions.

Under planktonic mating conditions, we observed a decrease in pAM714 transconjugants over time (Figure 4.1A) and a strong difference in the overall number of T11RF pAM714 and pAM771 transconjugants after 5 and 18 h of conjugation (Figure 4.1 A and C). The only difference between the two plasmids used was the ability to produce cytolysin, therefore we hypothesized that pAM714-encoded cytolysin was having an impact on plasmid transfer kinetics under planktonic mating conditions. Cytolysin production is regulated by a quorum sensing mechanism in which one component of the active cytolysin molecule acts as an auto-inducer for a two-component regulatory system associated with the cytolysin operon (42, 43). Activation of the two-component system by cytolysin results in induction of cytolysin expression, thereby increasing the amount of cytolysin produced. The operon also encodes the immunity factor that provides protection from cytolysin activity, thereby immunizing the cell harboring pAM714, and allowing it to persist in a population (37).

When putting cytolysin production in the context of a mating reaction, there would need to be a sufficient amount of plasmid transfer into the recipient population such that enough immunity factor could be produced to protect new transconjugants cell from cytolysin activity. We monitored recipient cell density over time in the mating reactions described in the previous section using selective media that would only allow growth for T11RF or T11RFΔ*cas9* strains. Indeed, this particular antibiotic selection scheme would also allow the growth of cells that have acquired pAM714 or pAM771, but considering that CRISPR-Cas is active in T11RF recipient strains there will be a portion of the population that does not contain the plasmid justifying the

use of this selection as a measure of recipient cell density. Under planktonic mating conditions, T11RF recipient viability remained stable when the donor strain harbored pAM771; however, in the presence of pAM714 there was a linear decrease in recipient viability that initiated after only 30 minutes of conjugation (Figure 4.2A). The difference in T11RF density under these conditions was statistically significant at all six time points that were assessed. A similar trend was also observed for the Δ*cas9* recipient strain (Figure 4.2B), where there was a statistically significant difference in the number of viable cells at all time points when comparing pAM714 and pAM771 mating reactions. We conclude that under planktonic mating conditions, cytolysin production from pAM714 continuously depletes the recipient cell population. We performed the same experiments for biofilm conjugation reactions and found that the degree of cell death attributed to cytolysin production from pAM714 was not as robust (Figure 4.2 C and D). Significant changes in recipient cell viability did not occur until the later time points of conjugation. From this, we conclude that cytolysin activity has a greater impact on the population dynamics of conjugation reactions under planktonic mating conditions as compared to biofilm mating conditions.

Based on the differences in recipient cell density between planktonic and biofilm mating conditions, we hypothesized that the ability of cytolysin to kill cells is dependent on mating condition. To obtain a better resolution of the recipient cell sensitivity to cytolysin activity, we utilized a cytolysin immunity assay (37). This assay allowed us to determine the time point at which the recipient populations became immune to cytolysin activity, which can then be used as a means to confirm that the reduced recipient cell density observed in pAM714 planktonic

conjugations can be attributed to cytolysin. The cytolysin immunity assay was performed on the recipient cell populations from planktonic and biofilm mating reactions when pAM714 was used as a donor. The was not performed on conjugation reactions with pAM771 as there is no cytolysin production from that plasmid. The results of the cytolysin immunity assay are shown in Figure 4.3, where a zone of inhibition demonstrates cytolysin susceptibility and the absence of a zone of inhibition indicates cytolysin immunity.

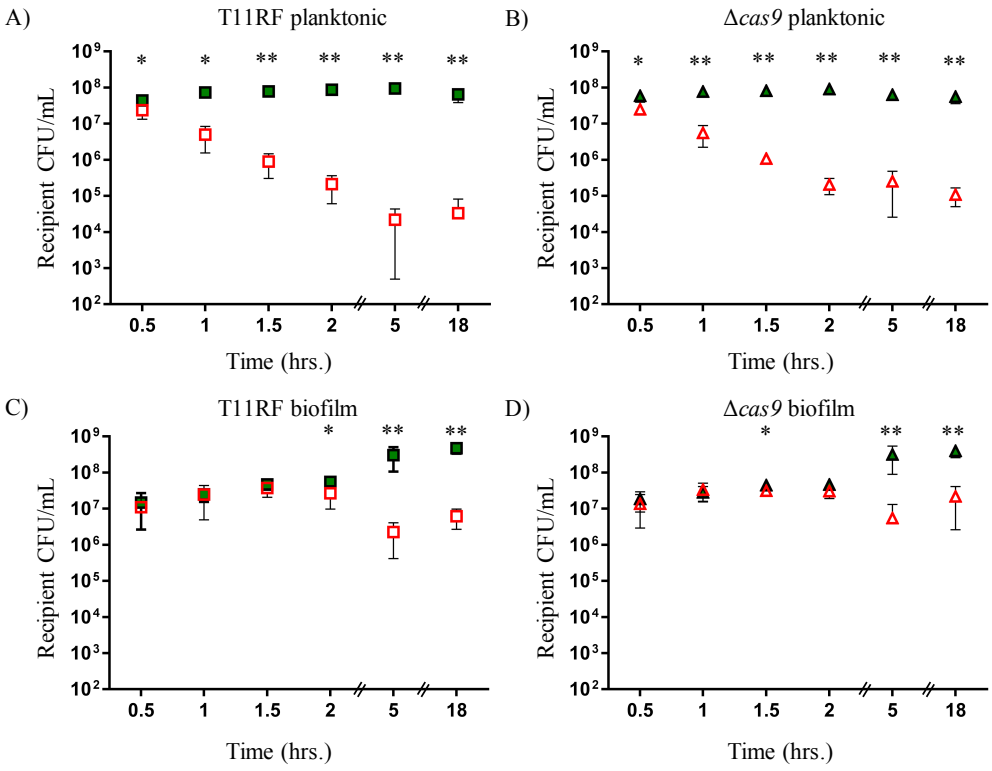


Figure 4.2. Production of cytolysin from pAM714 results in a reduction of recipient cell density. The portion of mating reactions corresponding to recipient cells was determined by plating conjugation reactions on selective media for the recipients, T11RF (squares) and T11RFΔ*cas9* (triangles). The recipient CFU/mL was determined for both planktonic (A and B) and biofilm (C and D) mating conditions utilizing plasmids pAM714 (open, red symbols) and pAM771(closed green symbols). Data shown is the average and standard deviation from a minimum of three independent trials for each time point for both mating conditions. Statistical significance was assessed using a one-tailed Student's t-Test; *P*-values, *<0.05 and **<0.005.

Under planktonic mating conditions (Figure 4.3, top panel), the T11RF population was susceptible to cytolysin for the duration of the experiment. Therefore, the activity of cytolysin reduces the number of available recipient cells (Figure 4.2A) while the activity of CRISPR-Cas, prevents pAM714 acquisition (Figure 4.1A). This results in sustained cytolysin susceptibility because in planktonic conjugation reactions approximately 1 out of 100 recipient cells have acquired pAM714. On the other hand, T11RFΔ*cas9* was immune to the action of cytolysin after one hour of conjugation (Figure 4.3, top panel). These results suggest that in the absence of active CRISPR-Cas, pAM714 is transferred into the recipient population at a level sufficient to produce enough immunity factor rendering the population immune to the activity of cytolysin. This is further substantiated by the observation that all available T11RFΔ*cas9* recipients have received pAM714 (Figure 4.1A and 4.2B).

In conjugation reactions performed in a biofilm, the T11RF strain becomes immune to cytolysin after five hours of mating and the Δ*cas9* strain is at least partially immune after one hour of mating on a solid surface (Figure 4.3, bottom panel). These data support the idea that the absence of CRISPR-Cas leads to more rapid acquisition of cytolysin immunity at a population level, presumably through unrestricted plasmid acquisition. Overall, we conclude that cytolysin has a greater impact on recipient cell viability under planktonic mating conditions compared to conjugation reactions under biofilm conditions. We also demonstrate that the presence of active CRISPR-Cas defense against pAM714 results in prolonged susceptibility to the activity of cytolysin.

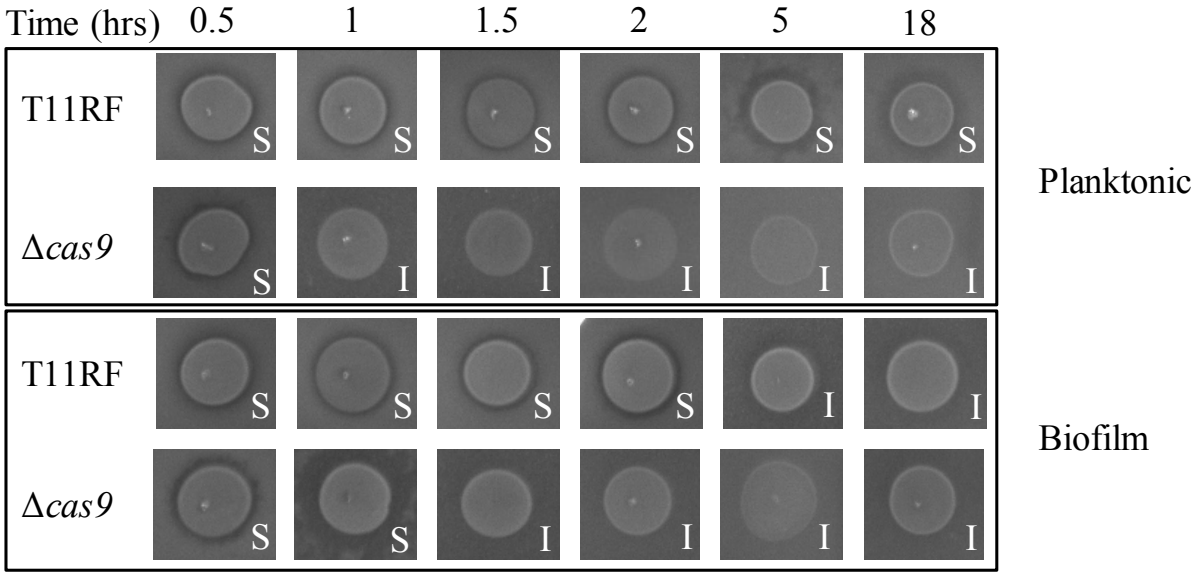


Figure 4.3. Cytolysin immunity assay shows increased susceptibility of T11RF to cytolysin in planktonic and biofilm conjugation reactions. A cytolysin immunity assay was performed as previously described (37) on planktonic and biofilm mating reactions where OG1SSp bearing pAM714 was used as a donor. 5 μL of an overnight culture of strain OG1SSp pAM714 was spotted onto an overlay of recipient strains to assay for cytolysin susceptibility. A zone of inhibition indicates susceptibility to cytolysin; S, susceptible and I, immune. Pictures presented are representative of results obtained from at least two independent trials of the assay.

CRISPR-Cas has robust activity against PRPs *in vivo*.

The function of the CRISPR-Cas systems of many bacterial species have been studied extensively *in vitro* over the last decade. However, there is little research on the impact of these systems against MGE acquisition in native habitats. Elucidating the role of CRISPR-Cas in preventing HGT *in vivo* is especially important in the opportunistic pathogen *E. faecalis*, where it is well established that the lack of barriers to horizontal gene transfer influence the degree of genome plasticity, specifically through the acquisition and maintenance of MGEs. Here, we apply our *in vitro* experimental design to a model for *E. faecalis* invasion in the mouse gut; a

detailed explanation of the procedure can be found in Materials and Methods. Briefly, mice were administered a cocktail of antibiotics in their water for seven days in order to deplete their normal microbiota. After antibiotic treatment, the mice went 24 hours on normal water and were then sequentially colonized by gavage with equal concentrations of a donor and recipient *E. faecalis* strain. Fecal pellets were collected at 24, 48, and 96 hours post co-colonization, homogenized, and cultured and selective and non-selective media. The number of transconjugants obtained at each time point was normalized to the weight of fecal pellets for each mouse so that data could be compared across groups. The mouse colonization experiments included three experimental groups with different combinations of donor and recipient strains: OG1SSp with T11RF served as a plasmid-free control, OG1SSp pAM714 (or pAM771) with T11RF was our test group for CRISPR-Cas activity, and OG1SSp pAM714 (or pAM771) with T11RFΔ*cas9* was our control group for the absence of CRISPR-Cas defense.

After 24 hours post co-colonization, we observed pAM714 transfer in only one out of ten mice when T11RF was the recipient (Figure 4.4A). Transconjugants were not recovered in this mouse after 48 hours; we speculate that active CRISPR-Cas was responsible for eliminating pAM714-containing cells in this mouse. Conversely, pAM714 transconjugants at densities up to 10^6 CFU/g of feces were observed for eight of ten mice colonized with T11RFΔ*cas9* recipients (Figure 4.4A). Transconjugant colonization was retained through the 96-hour time point for the eight mice.

We also performed these experiments with OG1SSp pAM771 as donor to assess the potential for cytolysin to influence *in vivo* plasmid transfer. At the 24 and 48-hour post co-colonization time points, one out of ten mice with T11RF as a recipient experienced pAM771 plasmid transfer (Figure 4.4B). As was observed with pAM714, transconjugants were not recovered in these two mice at the next time point sampled, presumably due to the action of CRISPR-Cas. All eleven mice in the group with Δ*cas9* as the recipient strain acquired pAM771 and retained the plasmid at a stable density over 96 hours (Figure 4.4B). These data show that there is a significant impact of CRISPR-Cas on plasmid transfer between *E. faecalis* strains in the mouse gut. We also determined that the presence of cytolysin does not impact the ability of plasmid transfer to occur in the absence of CRISPR-Cas activity (T11RFΔ*cas9* mice).

One caveat to these data, is that some of the no plasmid control mice had colony growth on media with selection for transconjugants. Growth did not occur until 48 or 96 hours post-co-colonization (Figure 4.4 A and B) meaning that the growth could be attributed to residual bacteria in the mouse gut that possessed resistance to that particular selection. We do not anticipate this confounding factor to greatly impact the conclusions made here for two reasons: 1) the majority of the test group mice (T11RF as recipient) did not experience any plasmid transfer, and if they did it was not sustained; and, 2) if the transconjugant data are normalized using donor density, only one control mouse had a conjugation frequency that was calculable (Figure C.1 in Appendix C). Therefore, we stand by our conclusion that the presence of CRISPR-Cas has a profound impact on preventing HGT of an *E. faecalis* PRP in the mouse gut.

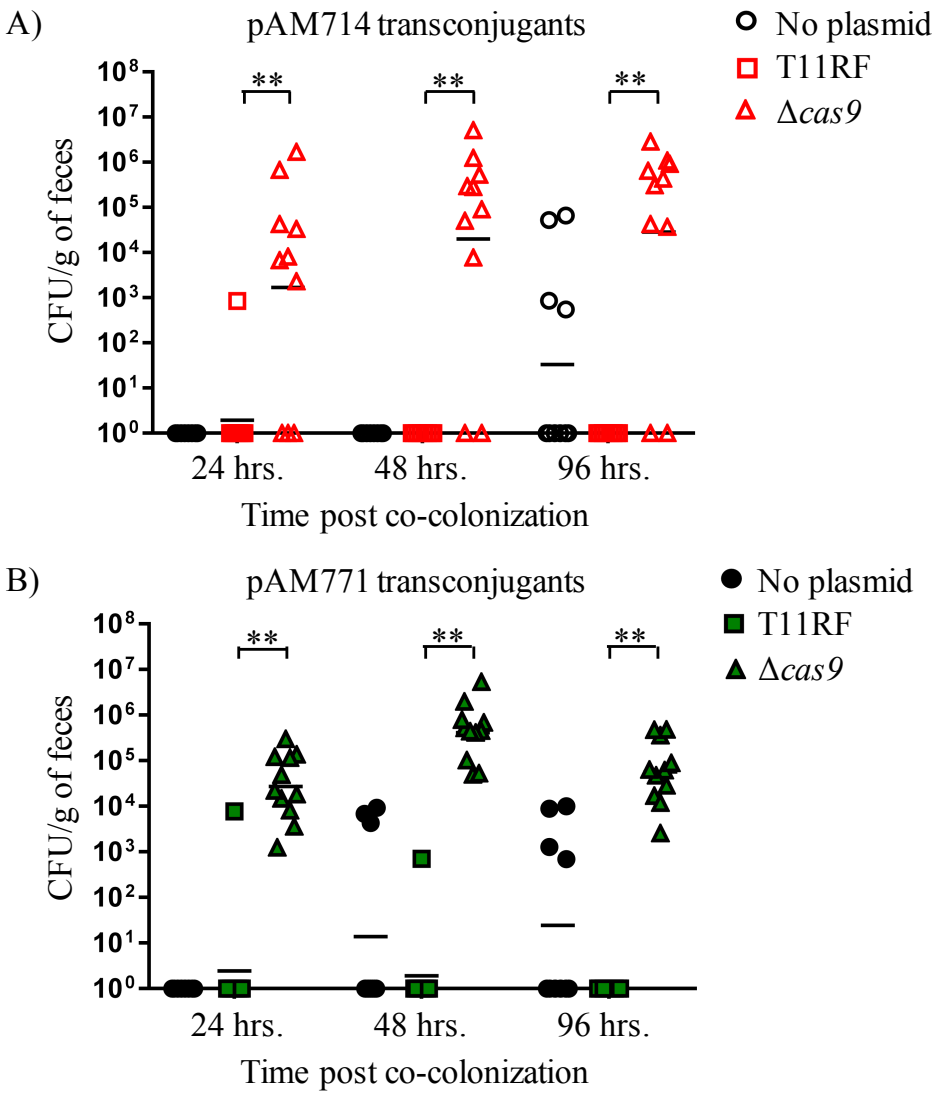


Figure 4.4. CRISPR-Cas has a strong impact on plasmid transfer in the mouse gut. The number of transconjugants obtained for each experimental group was determined by calculating the CFU/g of feces for each individual mouse; one symbol represents one mouse on the graph. Black horizontal bars represent the geometric mean of data in each group. For experiments with pAM714 as the plasmid donor (A), ten mice were used in each group. For experiments with pAM771 as the plasmid donor (B), ten mice were used for the no plasmid control and T11RF recipient groups, whereas the Δ*cas9* group had 11 mice. Statistical significance was assessed using a one-tailed Student's t-Test; *P*-value, **<0.005.

Cytolysin-independent colonization benefit to strains possessing a PRP.

Based on our *in vitro* studies, we were interested in exploring the ability of cytolysin to impact the colonization of *E. faecalis* in the mouse gut. To assess recipient cell density in the mice, fecal samples from the same time points analyzed in Figure 4.4 were plated on media to select for recipient populations only (similar to the rationale of Figure 4.2). If pAM714 had an impact on the viability of recipient cell populations *in vivo*, we would expect to see a decline in T11RF and Δ*cas9* density in mice where pAM714 was the donor and a stable colonization phenotype of the recipient strains when pAM771 was used as a donor. At 24 hours post co-colonization, the density of recipient strains from all three test groups was relatively equal in the experiments assessing pAM714 conjugation (Figure 4.5A). By 48 hours post co-colonization, the density of T11RF from the plasmid-free control group had increased by approximately one log and was significantly more than the T11RF strain in the test group where pAM714 was present. At the 96-hour time point, the difference in T11RF density between the no plasmid and pAM714 groups was more pronounced; the observed difference was due to a decrease in density of T11RF in the presence of pAM714. The total density of T11RF when co-colonized with OG1SSp pAM714 decreased by more than one log between the 24 and 96-hour time points. Similar results were obtained in our experimental groups where pAM771 was used as a donor (Figure 4.5B). After 48 hours of co-colonization, there was a significant decrease in T11RF density in the presence of pAM771, whereas T11RF in the absence of plasmid experienced an increase in density. This association remained the same at the 96-hour time point, where T11RF density in the presence of plasmid decreased further. It is also important to note that there were no obvious differences in density between the two recipient strains, T11RF and T11RFΔ*cas9*, suggesting

that the presence or absence of *cas9* does not impact strain colonization. Therefore, we conclude that the cytolysin-dependent impact on recipient cell viability observed in our *in vitro* studies did not translate to the mouse intestine.

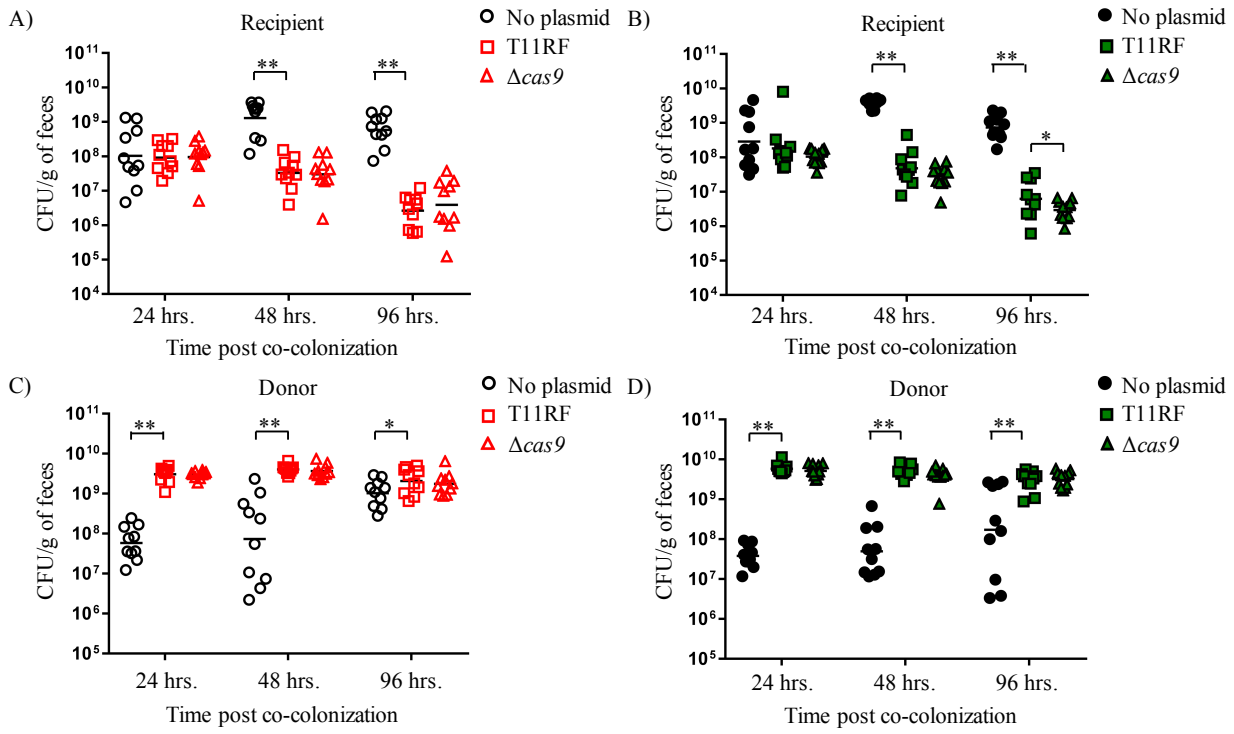


Figure 4.5. Cytolysin-independent colonization phenotypes are observed in the mouse gut. The density of recipient (A and B) and donor (C and D) strains are shown as the CFU/g of feces on the appropriate antibiotic selection. These experiments were performed on the same test groups described in Figure 4.4, where pAM714 (A and C) is represented with open, red symbols and pAM771 (B and D) is denoted by closed, green symbols. Statistical significance was assessed using a one-tailed Student's t-Test; *P*-values, *<0.05 and **<0.005.

Despite the observation that cytolysin did not have a significant impact on recipient cell density *in vivo*, we wanted to explore the impact of cytolysin production on the colonization of the donor strains. We followed the same procedures as previously described except this time we plated the fecal samples on media that would enumerate the donor cell density. The density of OG1SSp

strains with and without plasmid at 24 hours post co-colonization differed by almost two logs regardless of the plasmid present (Figure 4.5 C and D). Additionally, the control group with no plasmid had a significantly reduced donor density compared to plasmid-bearing donors at all time points assessed in experiments utilizing both pAM714 and pAM771. However, we observed a gradual increase in OG1SSp density in the no plasmid control groups over the 96 hours, but this was accompanied by a high degree of variability between the mice within these groups. On the other hand, the density of donor strains harboring either plasmid remained stable throughout the experiment and was consistent across all mice in each group. Based on these data, we conclude that there was a cytolysin-independent colonization benefit for OG1SSp strains harboring pAM714 or pAM771.

4.6 Discussion

Prior to this study, little was known about how the environment in which conjugation occurs impacts plasmid transfer kinetics in *E. faecalis*. Here, we explored the differences in PRP conjugation over time under biofilm and planktonic mating conditions. In all of our *in vitro* mating reactions, we are able to detect transconjugants as early as 30 minutes post-conjugation initiation. We observed a linear trend to plasmid acquisition over time that was more prolonged for conjugations that took place in a biofilm. This result was not surprising based on current knowledge on the efficiency of PRP-mediated conjugation on solid surfaces. The increase in plasmid transfer over time, regardless of mating condition, is relevant to the rapid dissemination of PRPs in hospital settings. It would be of interest to understand how a mixed microbial community would impact the plasmid transfer kinetics seen here.

Previous research from our lab has shown that the CRISPR3-Cas system of T11RF has a significant impact on plasmid transfer in *E. faecalis* (36). Here, we studied the impact of mating condition on the efficiency of genome defense. From our data, we conclude that under mating conditions on a solid surface, CRISPR3-Cas functions to significantly impact the transfer of both pAM714 and pAM771. However, this impact is not immediate as it takes at least one and a half of conjugation to produce a significant difference between the number of transconjugants obtained in CRISPR-active and -deficient strains of *E. faecalis*. Despite the lag in time when the impact of CRISPR-Cas becomes significant, we always observe more transconjugants in *cas9*-deficient T11RF suggesting that CRISPR-Cas activity in *E. faecalis* is initiated early but is overall a slow process that requires several hours to reach its maximal efficiency. Interestingly, under planktonic mating conditions with pAM771, CRISPR3-Cas has little if any impact on plasmid transfer. We speculate the reason for the lack in genome defense is related to gene expression changes that occur in pAM771. It has been shown that expression of aggregation substance from PRPs is only necessary under broth mating conditions (13). We reason that aggregation substance production enhances the cell-to-cell interactions in planktonic conjugation and increases the efficiency of plasmid transfer. The enhanced plasmid transfer rates under these conditions likely result in the inability of the CRISPR-Cas system to produce enough effector complexes to match incoming plasmid targets. We have evaluated the expression of *cas9* after 1, 2 and 18 hours of conjugation under planktonic and biofilm conditions and determined that there is no change in expression (data not shown). This suggests that changes in crRNA expression might be the limiting factor to CRISPR-Cas efficiency. Future studies will focus on the expression profile of CRISPR-Cas components under different mating conditions.

The comparison between CRISPR3-Cas activity under biofilm and planktonic conditions is complicated by the action of cytolysin. We utilized two pAD1 derivatives, pAM714 and pAM771, that differ only in cytolysin production to assess the influence of bacteriocin production on plasmid transfer kinetics and CRISPR-Cas function. We observe a significant reduction in recipient cell density under planktonic mating conditions when using OG1SSp pAM714 as a donor, thus reducing the number of available cells to act as plasmid recipients. The cytolysin-dependent decrease in recipient cell viability coupled with active CRISPR-Cas against plasmid transfer resolves the discrepancies observed for CRISPR-Cas activity against pAM714 and pAM771 under planktonic mating conditions. Although the impact of CRISPR-Cas on pAM771 transfer in planktonic conjugation is not as drastic as conjugation with pAM714, we do notice a 1-3 log increase in the total number of transconjugants obtained compared to pAM714. This result can be explained in part by the findings of a recent study showing that high PRP donor density relative to recipient density resulted in decreased *in vitro* conjugation frequencies (44). Extrapolating this to our experiments, we observe that cytolysin-mediated cell death disrupts donor and recipient ratios resulting in the shutdown of the mating response, leading to decreased plasmid transfer rates for pAM714. The impact of cytolysin on recipient cell density observed here is not insignificant because it highlights the importance of understanding plasmid-host interactions and how those interactions can influence the population structure of a microbial community. This is even more applicable in the context of *E. faecalis* and PRPs as they are highly co-evolved and are rapid disseminators of antibiotic resistance in hospital settings.

We used a mouse model of *E. faecalis* invasion to study the impact of CRISPR-Cas on HGT *in vivo*. Our data reveal that CRISPR-Cas has a profound impact on preventing horizontal gene transfer in the mouse gut, where CRISPR-Cas activity blocked PRP dissemination in all but three of twenty mice. This is an important result in the context of *E. faecalis* as it has been well established that the lack of genome defense systems leads to the emergence of MDR strains. One caveat to our data is that we obtain growth on media selecting for plasmid transfer events in some control mice. We reason that the growth is not likely due to contamination, but is rather the result of interactions between our recipient strains and the residual microbiota of the mice. The fact that we did not observe comparable growth on selection for plasmid transfer in our test group with active CRISPR-Cas further supports that the aberrant growth in control mice is a byproduct of the microbiota and is not associated with our experimental design.

A recent study investigated the impact of *E. faecalis*-encoded bacteriocins on colonization efficiency in mice that did not have their microbiota disrupted by antibiotic treatment (45). The findings of the study showed that the bacteriocin produced by the PRP pPD1 enhances strain colonization and enables pPD1-encoding strains to out compete other *E. faecalis* strains in the mouse GI tract. These findings are in contrast to the data presented here where we observed no difference in the ability of our pAM714 (cytolysin-producing) or pAM771 donor strains to colonize the mouse GI tract. The absence of a colonization benefit to cytolysin production in our experiments can likely be attributed to the use of antibiotics to deplete the normal microbiota. Even though cytolysin did not enhance the colonization of our donor strain, there was an overall benefit to colonization when mice were inoculated with donor strains harboring a PRP. It is

possible that the differences in colonization density are due to other PRP-encoded factors that were not assessed here or have yet to be characterized.

Overall, our study found that mating condition impacts the efficiency of CRISPR-Cas activity *in vitro* and that PRP-encoded virulence factors, such as cytolysin, can impact plasmid transfer kinetics under certain *in vitro* conditions by altering recipient population density. These data illustrate the intricate interactions that exist among plasmids and their hosts. We also demonstrate for the first time that CRISPR-Cas prevents HGT between *E. faecalis* strains in the mouse gut. The efficiency of CRISPR-Cas *in vivo* has the potential to be exploited as an alternative therapy where the development of therapeutic *E. faecalis* strains that harbor CRISPR loci against known hospital-associated MGEs could be used to stop the dissemination of those elements.

4.7 Acknowledgments

The authors would like to thank Dr. Michael Gilmore for providing enterococcal strains used in this study. This work was supported by grants R01AI116610 to KLP and K01DK102436 to BD from the National Institutes of Health.

4.8 References

1. Lebreton F, Willems RJL, & Gilmore MS (2014) *Enterococcus* Diversity, Origins in Nature, and Gut Colonization. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

2. Agudelo Higuita N & Huycke MM (2014) Enterococcal Disease, Epidemeiology, and Implications for Treatment. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

3. Arias CA & Murray BE (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nature reviews. Microbiology* 10(4):266-278.

4. Sievert DM*, et al.* (2013) Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. *Infection control and hospital epidemiology* 34(1):1-14.

5. Noskin GA, Stosor V, Cooper I, & Peterson LR (1995) Recovery of Vancomycin-Resistant Enterococci on Fingertips and Environmental Surfaces. *Infection control and hospital epidemiology* 16(10):577-581.

6. Kristich CJ, Rice LB, & Arias C, A (2014) Enterococcal Infection-Treatment and Antibiotic Resistance. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

7. Bourgogne A*, et al.* (2008) Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome biology* 9(7):R110.

8. Palmer KL, Kos VN, & Gilmore MS (2010) Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Current opinion in microbiology* 13(5):632-639.

9. Paulsen IT*, et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299(5615):2071-2074.

10. Clewell DB*, et al.* (2014) Extrachromosomal and Mobile Elements in Enterococci: Transmission, Matintenance, and Epidemiology. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

11. McBride SM, Fischetti VA, Leblanc DJ, Moellering RC, Jr., & Gilmore MS (2007) Genetic diversity among *Enterococcus faecalis*. *PloS one* 2(7):e582.

12. Dunny GM (2013) Enterococcal sex pheromones: signaling, social behavior, and evolution. *Annual review of genetics* 47:457-482.

13. Dunny GM (2007) The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 362(1483):1185-1193.

14. Clewell D*, et al.* (1982) Sex pheromones in *Streptococcus faecalis*: multiple pheromone systems in strain DS5, similarities of pAD1 and pAMγ1, and mutants of pAD1 altered in conjugative properties. *Microbiology*:97-100.

15. Dunny GM & Clewell DB (1975) Transmissible Toxin (Hemolysin) Plasmid in *Streptococcus faecalis* and Its Mobilization of a Noninfectious Drug Resistance Plasmid. *Journal of Bacteriology* 124(2):784-790.

16. Freitas AR*, et al.* (2013) Microevolutionary events involving narrow host plasmids influences local fixation of vancomycin-resistance in *Enterococcus* populations. *PloS one* 8(3):e60589.

17. Mikalsen T*, et al.* (2015) Investigating the mobilome in clinically important lineages of *Enterococcus faecium* and *Enterococcus faecalis*. *BMC genomics* 16:282.

18. Wardal E, Gawryszewska I, Hryniewicz W, & Sadowy E (2013) Abundance and diversity of plasmid-associated genes among clinical isolates of *Enterococcus faecalis*. *Plasmid* 70(3):329-342.

19. Clewell DB (2007) Properties of *Enterococcus faecalis* plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin. *Plasmid* 58(3):205-227.

20. Booth MC*, et al.* (1996) Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic. *Molecular microbiology* 21(6):1175-1184.

21. Cox CR, Coburn PS, & Gilmore MS (2005) Enterococcal Cytolysin: A Novel Two Component Peptide System that Serves as a Bacterial Defense Against Eukaryotic and Prokaryotic Cells. *Current Protein and Peptide Science* 6(1):77-84.

22. Dupont H, Montravers P, Mohler J, & Carbon C (1998) Disparate Findings on the Role of Virulence Factors of *Enterococcus faecalis* in Mouse and Rat Models of Peritonitis. *Infection and immunity* 66(6):2570-2575.

23. Huycke MM, Gilmore MS, Jett BD, & Booth LJ (1992) Transfer of Pheromone-Inducible Plasmids between *Enterococcus faecalis* in the Syrian Hamster Gastrointestinal Tract. *Journal of Infectious Diseases* 166:1188-1191.

24. Huycke MM, Joyce WA, & Gilmore MS (1995) *Enterococus faecalis* Cytolysin without Effect on the Intestinal Growth of Susceptible Enterococci in Mice. *Journal of Infectious Diseases* 172:273-276.

25. Ike Y, Hashimoto H, & Clewell DB (1984) Hemolysin of *Streptococcus faecalis* Subspecies *zymogenes* Contributes to Virulence in Mice. *Infection and immunity* 45(2):528-530.

26. Makarova KS, Wolf YI, & Koonin EV (2013) Comparative genomics of defense systems in archaea and bacteria. *Nucleic acids research* 41(8):4360-4377.

27. Makarova KS*, et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nature reviews. Microbiology* 13(11):722-736.

28. Marraffini LA (2015) CRISPR-Cas immunity in prokaryotes. *Nature* 526(7571):55-61.

29. Barrangou R*, et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709-1712.

30. Brouns SJJ*, et al.* (2008) Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* 321(5891):990-964.

31. Garneau JE*, et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468(7320):67-71.

32. Marraffini LA & Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322(5909):1843-1845.

33. Palmer KL & Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4).

34. Hullahalli K*, et al.* (2015) Comparative Analysis of the Orphan CRISPR2 Locus in 242 *Enterococcus faecalis* Strains. *PloS one* 10(9):e0138890.

35. Hullahalli K, Rodrigues M, & Palmer KL (2017) Exploiting CRISPR-Cas to manipulate *Enterococcus faecalis* populations. *eLife* 6.

36. Price VJ, Huo W, Sharifi A, Palmer KL, & Fey PD (2016) CRISPR-Cas and Restriction-Modification Act Additively against Conjugative Antibiotic Resistance Plasmid Transfer in *Enterococcus faecalis*. *mSphere* 1(3):e00064-00016.

37. Coburn PS, Hancock LE, Booth MC, & Gilmore MS (1999) A Novel Means of Self-Protection, Unrelated to Toxin Activation, Confers Immunity to the Bactericidal Effects of the *Enterococcus faecalis* Cytolysin. *Infection and immunity* 67(7):3339-3347.

38. Clewell D*, et al.* (1982) Mapping of *Streptococcus faecalis* plasmids pAD1 and pAD2 and studies relating to transposition of Tn*917*. *Journal of Bacteriology* 152(3):1220-1230.

39. Ike Y, Clewell D, Segarra R, & Gilmore M (1990) Genetic analysis of the pAD1 hemolysin/bacteriocin determinant in *Enterococcus faecalis*: Tn*917* insertional mutagenesis and cloning. *Journal of bacteriology* 172(1):155-163.

40. Jett BD, Jensen HG, Nordquist RE, & Gilmore MS (1992) Contribution if the pAD1-Encoded Cytolysin to the Severity of Experimental *Enterococcus faecalis* Endophthalmitis. *Infection and immunity* 60(6):2445-2452.

41. Gilmore M, Segarra R, & Booth M (1990) An HlyB-type function is required for expression of the *Enterococcus faecalis* hemolysin/bacteriocin. *Infection and immunity* 58(12):3914-3923.

42. Gilmore MS, Coburn PS, Nallapareddy SR, & Murray BE (2002) Enterococcal virulence. *The Enterococci*, (American Society of Microbiology), pp 301-354.

43. Haas W, Shepard BD, & Gilmore MS (2002) Two-component regulator of *Enterococcus faecalis* cytolysin responds to quorum-sensing autoinduction. *Letters to Nature* 415:84-87.

44. Bandyopadhyay A, O'Brien S, Frank KL, Dunny GM, & Hu WS (2016) Antagonistic Donor Density Effect Conserved in Multiple Enterococcal Conjugative Plasmids. *Applied and environmental microbiology* 82(15):4537-4545.

45. Kommineni S*, et al.* (2015) Bacteriocin production augments niche competition by enterococci in the mammalian gastrointestinal tract. *Nature* 526(7575):719-722.

CHAPTER 5

SUMMARY AND DISCUSSION

5.1 Summary of findings

Enterococcus faecalis is an opportunistic pathogen that is among the leading causes of hospital-associated infections in the United States (1-3). *E. faecalis* strains are often resistant to multiple antibiotics making the infections attributed to this bacterium challenging to treat (1, 4, 5). Genomic analyses of *E. faecalis* strains, both commensal and hospital isolates, revealed that hospital-associated strains have expanded genomes enriched with mobile genetic elements (6-9). In addition, it has been determined that hospital-adapted, multidrug-resistant strains were devoid of genome defense systems that normally prevent horizontal gene transfer (HGT) (6, 10). These findings led to the hypothesis that the absence of endogenous barriers to HGT leads to the rampant dissemination of mobile DNA among this species and ultimately, the evolution of multidrug-resistant *E. faecalis* strains. The work presented in this dissertation aimed to elucidate the interactions between genome defense systems and mobile genetic elements in *E. faecalis*.

The studies performed here analyzed the interaction between strain T11RF and a derivative of the model pheromone-responsive plasmid, pAD1. Several considerations were made to arrive on studying this specific strain-plasmid pair. First, pAD1 is a clinically relevant pheromone-responsive plasmid that is often found in clinical isolates of *E. faecalis* (11). Second, *E. faecalis* strain T11RF possesses two genome defense systems, a type II restriction modification system and a type II CRISPR-Cas system which enabled us to assess multiple aspects of *E. faecalis*

genome defense in a single strain (7, 12). Third, the CRISPR3-Cas locus of T11RF natively targets pAD1 (12). Finally, T11RF and the MDR strain V583 share greater than 99% nucleotide sequence identity in their core genomes (13). This allows us to interpret the implications of the presence or absence of genome defense in closely related strains, adding a clinical importance to this work. All of the factors mentioned here allowed us to draw conclusions from these studies that could be directly applied to the native condition in which T11RF and pAD1 interact with one another.

The first study presented in this dissertation determined that the genome defense systems of strain T11RF, R-M and CRISPR-Cas, collectively provide up to a four-log impact on preventing the acquisition of plasmid pAD1 (12). This was a novel finding in *E. faecalis* because it was the first demonstration of active CRISPR-Cas activity in this species. Cooperation between R-M and CRISPR-Cas had previously been identified in one other study using bacteriophage (14); therefore, our results proved that the combined effort of these two genome defense systems can also be applied to plasmid biology. Furthermore, we determined that the orphan CRISPR2 locus can be activated for genome defense in the presence of its cognate Cas9 enzyme, identifying the use of CRISPR-Cas as a potential means to correct the genome defense deficiencies that are associated with hospital-adapted *E. faecalis*. Overall, it was concluded that R-M and CRISPR-Cas provide a significant impact on preventing HGT in *E. faecalis*, providing evidence to support the general claim that MDR *E. faecalis* emerge due to the lack of barriers to HGT.

It was determined from the aforementioned study that CRISPR-Cas alone provided an 80-fold impact on pAD1 acquisition in T11RF. However, significant numbers of transconjugants were nevertheless observed for T11RF, indicating that CRISPR3-Cas was not a perfect barrier against HGT. This observation led to a more in-depth analysis of plasmid-host interactions and how pAD1 was maintained in T11RF despite the presence of CRISPR-Cas. The second study described here used a combination of *in vitro* evolution and deep-sequencing analyses to uncover the interactions between T11RF and pAD1 post-conjugation under with and without selection for pAD1. It was established that the CRISPR3-Cas system becomes compromised, either through spacer deletion or *cas9* mutation, under selection for plasmid maintenance. These mutations were identified days after conjugation occurred, therefore it was concluded that host mutations in CRISPR-Cas were necessary for fitness under conditions that were favorable for the plasmid. These results were shown to be applicable to another *E. faecalis* strain with a different CRISPR-Cas subtype implicating this as a general mechanism for plasmid maintenance in this species. Broadly speaking, this suggests that the use of antibiotics generates immune-compromised *E. faecalis* through forced selection of MGEs. These strains have either reduced or lost the ability to block acquisition of foreign DNA, providing evidence for another mechanism leading to the emergence of MDR *E. faecalis*.

The final set of work described in this dissertation took a more clinical approach and looked at the impact of CRISPR-Cas on plasmid transfer over time and ascertained its relevance *in vivo*. *E. faecalis* infections can be identified at multiple sites in the human body including the bloodstream and on the surfaces of in-dwelling medical devices. This raised the question of how

growth conditions can impact conjugation rates. In the final study, plasmid transfer kinetics were monitored over an 18-hour period under planktonic and biofilm conditions. It was determined that conjugation occurs more efficiently in a biofilm than under planktonic conditions. Additionally, the combination of a PRP-encoded virulence factor and CRISPR-Cas activity significantly reduced the amount of plasmid acquisition in planktonic conjugation reactions. This finding is significant because it highlights the importance of PRPs in *E. faecalis* infections and further warrants investigations into the mechanisms mediating plasmid-host interactions. Finally, it was shown that *E. faecalis* CRISPR3-Cas has a profound impact on preventing HGT in a mouse model of *E. faecalis* invasion, where pAD1 acquisition was substantially reduced in mice colonized with T11RF. This is the first assessment of *E. faecalis* CRISPR-Cas activity *in vivo*. The strong phenotype observed here suggests that *E. faecalis* strains with active CRISPR-Cas have the potential to be exploited as an alternative treatment against MDR *E. faecalis* infections.

5.2 Implications on *E. faecalis* biology

The research presented in this dissertation has uncovered previously unknown aspects of the interactions between *E. faecalis* and its MGEs. More specifically, this work has provided evidence substantiating the claim that MDR *E. faecalis* emerge due to the absence of barriers to HGT. The demonstration that CRISPR-Cas is active for genome defense in an animal model of *E. faecalis* invasion will open up avenues for the development of alternative therapies. Even though this work has advanced our knowledge, many unanswered questions remain.

Perhaps one of the more obvious questions is how can these findings be applied to human infections? *E. faecalis* infections can reside in various locations in or on the human body. Often, *E. faecalis* is only one member of a complex microbial community consisting of a number of bacterial species. Therefore, future research should focus on HGT in the context of a mixed microbial community that includes MDR and commensal *E. faecalis* strains as well as other bacterial species of the normal microbiota. The presence of more than one bacterial species would more accurately resemble a human infection and could be more readily applicable to the treatment of *E. faecalis* infections *in vivo*. These mixed communities could be studied under various conditions (such as biofilm and planktonic growth, anaerobic growth and growth in a chemostat) that will allow for the exploration of the diverse metabolic features possessed by this bacterium. The use of mixed communities under various growth conditions would not only be useful for elucidating the mechanisms of HGT in *E. faecalis* under more relevant circumstances, but can be used as a model to test alternative therapies.

Along the same lines, there is limited data available about the impact that a course of antibiotic treatment has on a polymicrobial community. Here, evidence was shown that antibiotic selection for a MGE leads to host mutations that stabilized plasmid maintenance in a population containing only one strain of *E. faecalis*. It is of interest to determine if similar mutations occur in *E. faecalis* strains that are present in a more complex environment where selection may be more beneficial to other members of the population. Gaining insight into these situations will vastly improve our understanding of how MDR *E. faecalis* emerge in clinical settings. The

treatment of *E. faecalis* infections will become more successful once the mechanisms of how these opportunistic pathogens become MDR are realized.

The narrow-host range PRPs present a unique host-plasmid relationship. Due to the inability of PRPs to replicate in species other than *E. faecalis*, it is likely that these two entities are highly co-evolved. Much is known about the mechanism of the pheromone response and the regulation of PRP-encoded toxin-antitoxin systems, but there is little known about the long-range costs and benefits associated with *E. faecalis* maintaining a PRP as part of its genome. It is well accepted in the plasmid-host co-evolution field that compensatory mutations are made under purifying selection that allow the stable maintenance of plasmids for prolonged periods of time (15-18). It would be interesting to determine what these stabilizing mutations are in *E. faecalis* and if these mutations differ between MDR and commensal strains. Finding differences in the mutations that allow stable maintenance could provide more information on how MDR strains arise. Finally, it has recently been determined that some bacteriophage encode anti-CRISPR proteins that make the phage immune to targeting by CRISPR-Cas systems (19, 20). The available DNA sequence of pAD1 contains many uncharacterized open reading frames, therefore it would be interesting to explore the possibility that PRPs encode some kind of anti-CRISPR-like system. Understanding the short and long-term relationships between PRPs and their *E. faecalis* host will provide context to the role of PRPs in shaping the evolutionary trajectory of this species.

5.3 CRISPR-Cas function and applications in *E. faecalis*

Although the data presented here established a role for CRISPR-Cas in *E. faecalis,* there are aspects of CRISPR-Cas function in this species that have not been fully elucidated. One of these aspects is the efficiency of genome defense. It was shown that CRISPR-Cas activity is not 100% efficient at preventing plasmid transfer *in vitro*. This suggests that there is an unknown regulatory mechanism at work or that the conditions for optimal genome defense were not met. This second claim is somewhat substantiated by the data showing that CRISPR-Cas had increased efficiency against plasmid acquisition *in vivo*. In order to elucidate these differences, more in-depth expression studies would need to be performed. Preliminary data suggests that the expression of *cas9* does not change over time and that its promoter is constitutively active. This indicates one of the other CRISPR-Cas components, crRNA or tracrRNA, is the limiting factor for providing efficient genome defense. These mechanistic details are important to understand if CRISPR-Cas is to be exploited as an alternative therapy against MDR *E. faecalis*.

Another characteristic of CRISPR-Cas systems is their ability to update their molecular memory. This is the process by which the CRISPR locus is able to identify and incorporate a new spacer sequence into its array through the process of adaptation. Attempts have been made to detect adaptation events in the *E. faecalis* T11RF CRISPR3-Cas locus; however, there has been no evidence to suggest that adaptation occurs. This indicates that the conditions required for adaptation have not been met. It is likely that the same conditions required for optimal CRISPR defense would also favor adaptation to occur. The over-expression of Cas1 and Cas2 has proven beneficial in observing CRISPR adaptation in other species and could be explored in *E. faecalis*

(21-24). A more relevant means to elucidating these details would be to utilize next-generation sequencing methods on *E. faecalis* strains that are challenged with novel MGEs under conditions that mimic their natural habitat. Gaining an understanding of how likely it is for *E. faecalis* to acquire new spacers is important to understanding the interactions between *E. faecalis* and MGEs on an evolutionary time scale as the order of spacers in an array can be loosely used to track lineages.

The majority of the CRISPR-Cas literature within the past five years has focused on the application of type II CRISPR-Cas systems as tools for genome engineering. Many advancements have been made on this front which has increased the ease and efficiency of genetic manipulation in the laboratory and in animal models of genetic disorders. (25-31). Indeed, CRISPR/Cas9 genome editing technology has revolutionized the way we approach the ongoing efforts to correct and eradicate human diseases. The Cas9 protein studied in this dissertation, EfCR3Cas9, is among the smaller orthologs of the well-characterized Cas9 proteins that are commercially used for genome editing strategies. Its size and unique PAM requirement makes EfCR3Cas9 an attractive candidate to add to the list of available Cas9 enzymes. More characterization of this protein and its related crRNA and tracrRNA components are required before its use could be broadly applicable. However, this enzyme could be of more immediate use to the *E. faecalis* field through the development of a plasmid that would allow for Cas9-mediated control of gene expression.

5.4 Concluding remarks

Within the last fifty years, there has been an increase in the number of antibiotic resistant bacterial infections. If this trend continues, antibiotic resistant organisms will become one of the leading causes of death worldwide. The focus of this dissertation was to determine mechanistic aspects of antibiotic resistance acquisition in the nosocomial pathogen, *Enterococcus faecalis*. In summary, the work conducted here has established a role for genome defense systems in preventing the acquisition of foreign DNA. Specifically, the finding that CRISPR-Cas is able to prevent HGT *in vivo* will open up many new avenues for the development of alternative treatments for *E. faecalis*-associated infections. The ultimate goal is to identify a broad-spectrum application for CRISPR-based therapies that could be used against a number of hospital-associated pathogens. The successful development of CRISPR-based treatments would be a significant advancement in combating the ever-growing problem of antibiotic resistance.

5.5 References

1. Agudelo Higuita N & Huycke MM (2014) Enterococcal Disease, Epidemeiology, and Implications for Treatment. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

2. Arias CA & Murray BE (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nature reviews. Microbiology* 10(4):266-278.

3. Sievert DM*, et al.* (2013) Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. *Infection control and hospital epidemiology* 34(1):1-14.

4. Kristich CJ, Rice LB, & Arias C, A (2014) Enterococcal Infection-Treatment and Antibiotic Resistance. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*.

5. Périchon B & Courvalin P (2009) Glycopeptide Resistance in Enterococci. 229-240.

6. Bourgogne A*, et al.* (2008) Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome biology* 9(7):R110.

7. McBride SM, Fischetti VA, Leblanc DJ, Moellering RC, Jr., & Gilmore MS (2007) Genetic diversity among *Enterococcus faecalis*. *PloS one* 2(7):e582.

8. Paulsen IT*, et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299(5615):2071-2074.

9. Solheim M, Aakra A, Vebo H, Snipen L, & Nes IF (2007) Transcriptional responses of *Enterococcus faecalis* V583 to bovine bile and sodium dodecyl sulfate. *Applied and environmental microbiology* 73(18):5767-5774.

10. Palmer KL & Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4).

11. Clewell DB (2007) Properties of *Enterococcus faecalis* plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin. *Plasmid* 58(3):205-227.

12. Price VJ, Huo W, Sharifi A, Palmer KL, & Fey PD (2016) CRISPR-Cas and Restriction-Modification Act Additively against Conjugative Antibiotic Resistance Plasmid Transfer in *Enterococcus faecalis*. *mSphere* 1(3):e00064-00016.

13. Palmer KL*, et al.* (2012) Comparative Genomics of Enterococci: Variation in *Enterococcus faecalis,* Clade Structure in *E. faecium*, and Defining Characteristics of *E. gallinarum* and *E. casseliflavus. mBio* 3(1):e00318-00311.

14. Dupuis ME, Villion M, Magadan AH, & Moineau S (2013) CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nature communications* 4:2087.

15. Harrison E & Brockhurst MA (2012) Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in microbiology* 20(6):262-267.

16. Harrison E, Guymer D, Spiers AJ, Paterson S, & Brockhurst MA (2015) Parallel compensatory evolution stabilizes plasmids across the parasitism-mutualism continuum. *Current biology* 25(15):2034-2039.

17. Hughes JM*, et al.* (2012) The role of clonal interference in the evolutionary dynamics of plasmid-host adaptation. *mBio* 3(4):e00077-00012.

18. Stalder T*, et al.* (2017) Emerging patterns of plasmid-host coevolution that stabilize antibiotic resistance. *Scientific reports* 7(1):4853.

19. Bondy-Denomy J*, et al.* (2015) Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* 526(7571):136-139.

20. Bondy-Denomy J, Pawluk A, Maxwell KL, & Davidson AR (2013) Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493(7432):429-432.

21. Datsenko KA*, et al.* (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature communications* 3:945.

22. Fineran PC & Charpentier E (2012) Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. *Virology* 434(2):202-209.

23. Heler R*, et al.* (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519(7542):199-202.

24. Wei Y, Terns RM, & Terns MP (2015) Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes & development* 29(4):356-361.

25. Barrangou R, *et al.* (2015) Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. *Nucleic acids research* 43(7):3407-3419.

26. Barrangou R & van Pijkeren JP (2015) Exploiting CRISPR-Cas immune systems for genome editing in bacteria. *Current opinion in biotechnology* 37:61-68.

27. Jiang Y, *et al.* (2015) Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Applied and environmental microbiology* 81(7):2506-2514.

28. Hsu PD, Lander ES, & Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157(6):1262-1278.

29. Jinek M, *et al.* (2013) RNA-programmed genome editing in human cells. *eLife* 2:e00471.

30. Ran FA, *et al.* (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154(6):1380-1389.

31. Ran FA, *et al.* (2013) Genome engineering using the CRISPR-Cas9 system. *Nature protocols* 8(11):2281-2308.

APPENDIX A

CHAPTER 2 SUPPLEMENTAL FIGURES AND TABLES

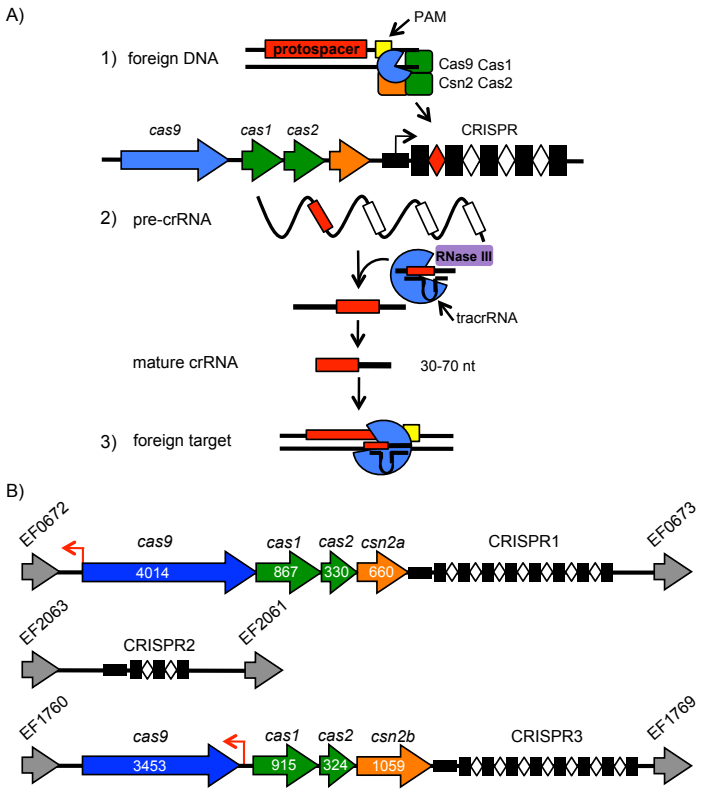


Figure A.1. Representative organization of the Type II CRISPR loci found in *E. faecalis*. A) Type II CRISPR-Cas interference mechanism. The model for Type II CRISPR-Cas function shown here has been adapted from a previously proposed model (36). The mechanism is described in three stages: 1) adaptation, 2) expression, and 3) interference. During adaptation, a new spacer originating from foreign DNA (red protospacer sequence) is integrated into the leader end of the array. Subsequently, the CRISPR array is transcribed into pre-crRNA; the pre-crRNA and tracrRNA form a complex that is processed by the host RNase III. Another processing event produces the mature crRNA consisting of parts of a spacer and adjacent repeat sequence. Finally, in interference, the mature tracrRNA:crRNA duplex guides Cas9 to the foreign DNA target by base-pairing with a sequence complementary to the spacer and PAM proximity, promoting cleavage of DNA through the two endonuclease domains of Cas9. PAM, protospacer adjacent motif. B) Three CRISPR loci identified in *E. faecalis*. Grey arrows represent the location of the CRISPR loci relative to orthologs of the V583 genome. The structure of CRISPR1-*cas* and CRISPR3-*cas* is similar, but the location of tracrRNA and the sizes of genes within the loci differs; nucleotide length of genes are given within arrows. Designation of *csn2a* or *csn2b* is based on (49).

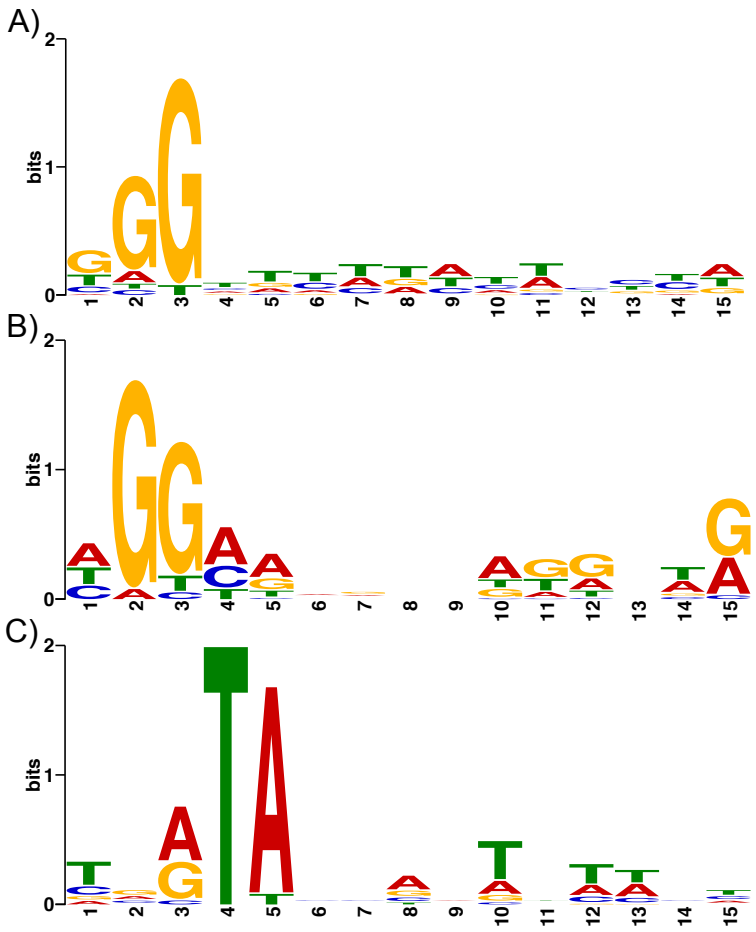
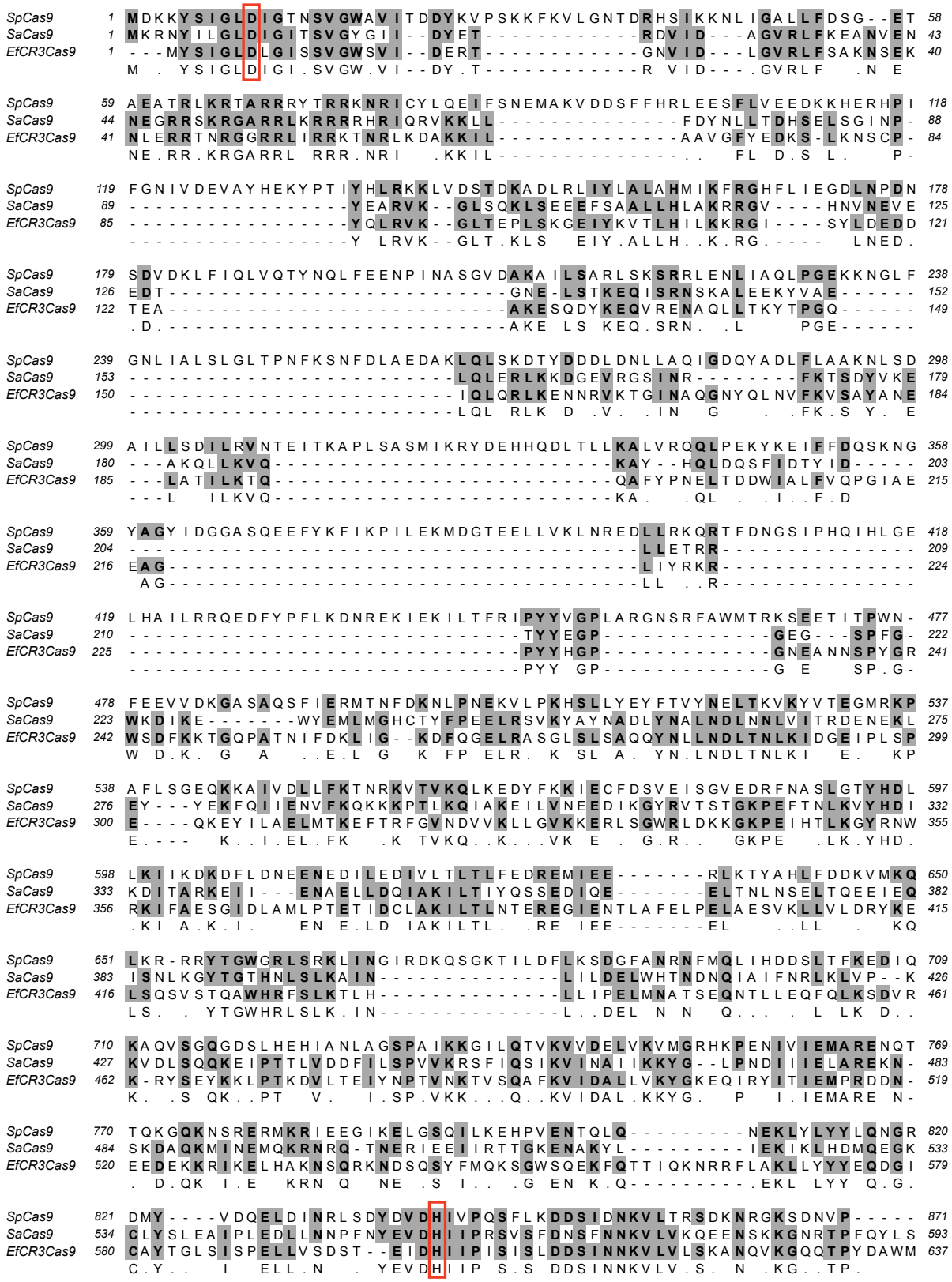


Figure A.2. Predicted PAM sequences for the CRISPR loci in *E. faecalis*. Motifs were determined utilizing the MEME motif alignment web server (62) for CRISPR1-Cas (A), CRISPR2 (B) and CRISPR3-Cas (C). Similarity in CRISPR1-Cas and CRISPR2 motifs and a unique CRISPR3-Cas motif are consistent with the differences in the consensus repeat sequences of the three loci.


SpCas9
SaCas9
EfCR3Cas9

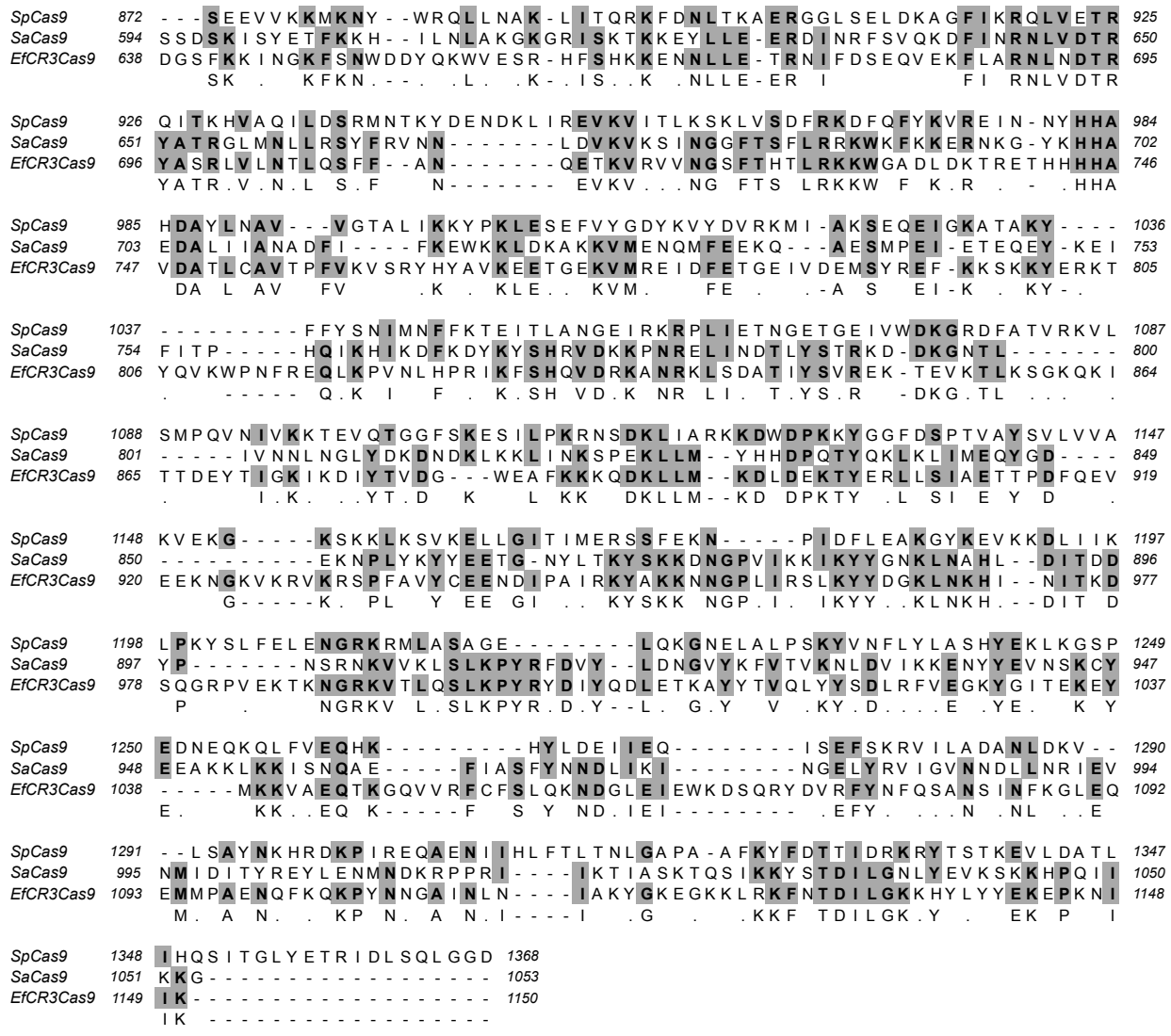


Figure A.3. *Streptococcus pyogenes* (Sp) Cas9 was used as the reference sequence in an alignment with *Staphylococcus aureus* (Sa) and *E. faecalis* CRISPR3 (EfCR3) Cas9 proteins. The MUSCLE alignment software was used with default parameters. Active site residues (D10 and H601 for T11) are boxed.

Figure A.4. Alignment of direct repeat sequences of the CRISPR loci found in *E. faecalis*. Consensus repeat sequences from each of the CRISPRs were aligned using Geneious. CRISPR1-Cas and CRISPR2 repeats are identical, whereas the CRISPR3 repeat only shares 58% identity with them. Repeats were derived from the following genomes: CRISPR1-Cas, OG1RF; CRISPR2, OG1RF and CRISPR3-Cas, T11.

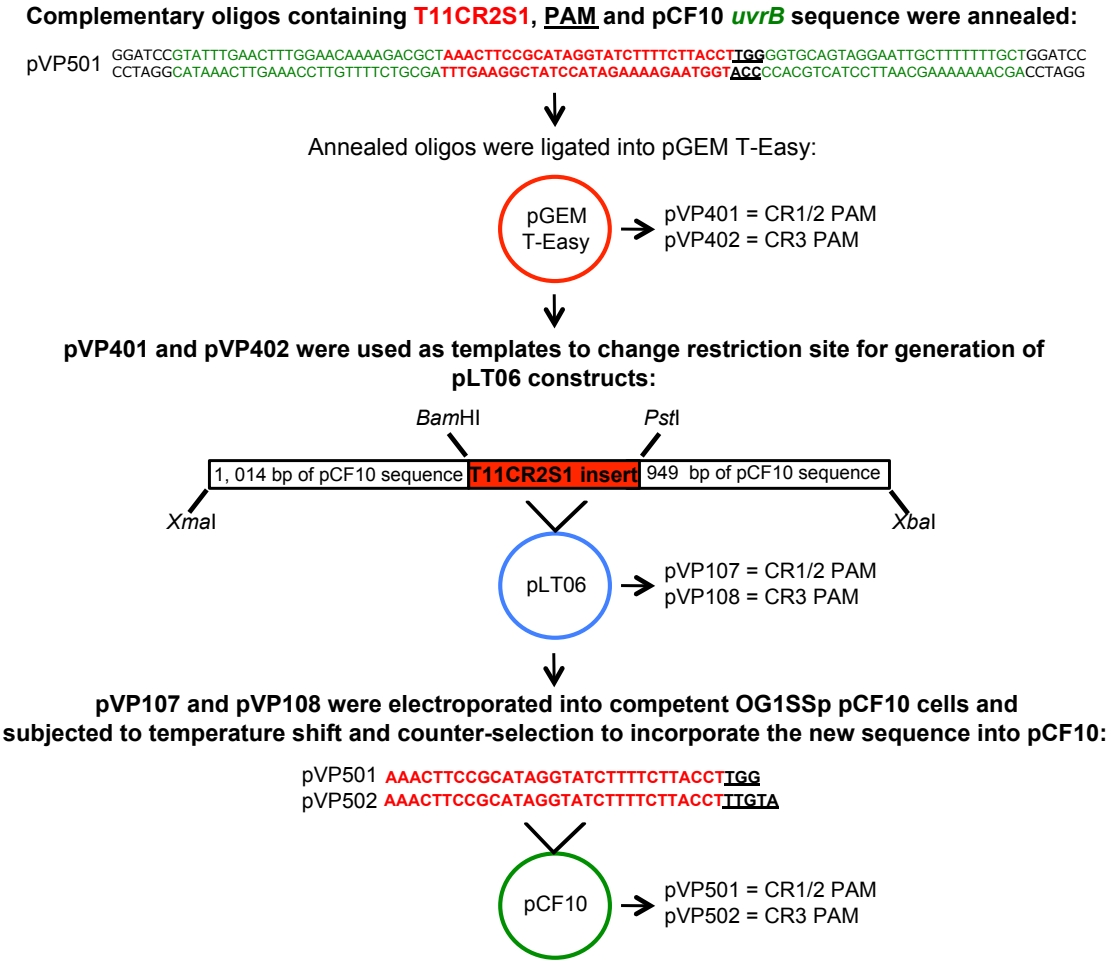


Figure A.5. Construction of pCF10 derivatives, pVP501 and pVP502.

Table A.1. Primers used in this study.

Primer name	Sequence
pLT06	
pLT06 oriF	CAATAATCGCATCCGATTGCA
pLT06 KS05	CCTATTATACCATATTTTGGAC
pWH03	
pWH03 F	CGAATTCAAATGCCTAGCATT
pWH03 R	GCAGATGAACCATCACCATT
pGEM-T-Easy	
T7	TAATACGACTCACTATAGGG
SP6	TATTTAGGTGACACTATAG
pVP102	
T11delcas9: arm1 F	GAGCAT*cccggg*CGAAGGCGACAGTAACCCAT
T11delcas9: arm1 R	CCAGTA*ggatcc*ACTGTACATGAGAACCTCCTTT
T11delcas9: arm2 F	TGCTAC*ggatcc*GGAAAGAAACATTATCTGTATTA
T11delcas9: arm2 R	TTCGAC*tctaga*GCAGCTACATATTCAGACATT
pAS106	
T11del6: arm1 F	GCGCAT*cccggg*CGAAGAGAACAGCCTGTCT
T11del6: arm1 R	GCTCGT*ctgcag*AACTCGTCTGTGCGCCGA
T11del6: arm2 F	GGATATctgcagACAGTTTTTGTACTCTCAAT
T11del6: arm2 R	TTCGAC*tctaga*GCTAGTTCATAGGTAATGGT
pVP105	
T11cas9 D7A arm1 F	GATCAT*cccggg*CGAAGGCGACAGTAACCCAT
T11cas9 D7A arm1 R	TAATCCCTAATGCTAAACCTATACTGTACATGAGAACCTCCTT
T11cas9 D7A arm2 F	TATAGGTTTAGCATTAGGGATTAGTAGTGTTGGTTGGT
T11cas9 D7A arm2 R	TTCGACtctagaGCCAATTCCGATAGCCTT
pG19	
CR1*cas9* SSI F	ATCACG*gcggccgc*GCAAACACAGTTAACCACG
CR1*cas9* SSI R	ACGTAC*gcggccgc*GCAAACGCTTATCATCGCAA
CR1*cas9* Seq F5	CGATTCATTGGATAACCTAGT
CR1cas9 Seq R5	GCATTGTAGCGTTGATCTAAG
CR1cas9 Seq F12	CAAGGTGGCTTGACTCTCGA

CR1cas9 Seq R7	GCATCGTGGTTAACGAAATG
CR1cas9 Seq F7	GCTACACGTCAATCAAAGAAA
CR1cas9 Seq F11	TCAAAGTGAGGAACCAATTC
CR1cas9 Seq F3	GCGAAATTATCTTCTAGCATGA
CR1cas9 Seq R3	GCTCCAGCAATATCTTGAAT
CR1cas9 Seq F10	TATGACAGAAGACTATCAGT
CR1cas9 Seq F1	GCGATTACTTTCTGGTTGTAA
CR1cas9 Seq R14	CCTTGGTCAGACGTTGGAAT
CR1cas9 Seq F14	AGCACAGTCCAGTGAACATG
pVP301	
CR3*cas9* SSI F	GATCAT*gcggccgc*CGGACAAAAACCTTGCCATC
CR3*cas9* SSI R	TCGTAC*gcggccgc*CGAAAACTGTTCGAAATCCC
CR3cas9 Seq F0	TCAACACACCAATCCATCG
CR3cas9 Seq F1	CGGACAGGCAATGTTATTGA
CR3cas9 Seq R3	GGCAGTTCAAAAGCTAACGTA
CR3cas9 Seq F11	ACTACGAGCAAGCGGCTTAA
CR3cas9 Seq R4	CGACTTGCTCACTGTCAAAA
CR3cas9 Seq R5	CGTTCATAGGTCTTTTCGTC
CR3cas9 Seq F4	TTTGACAGTGAGCAAGTCGA
CR3cas9 Seq F10	GCTAACCGCAAACTCAGTGA
CR3cas9 Seq F5	CGCAAAGTTACACTACAAAGT
CR3cas9 Seq F6	GGAAAGAAACATTATCTGTA
CR3cas9 Seq R1	CCTACTGAACAAAAACATTTCTC
CR3cas9 Seq R11	AAACGACGTCCACCTCGATT
pWH01	
1F_BamHI	AATGAA*ggatcc*ACCTAAAGCTATTCCTCCTGGA
1R_XbaI	AGTACA*tctaga*TCAATCTTAAAAGGTCGTGGCT
2F_PstI	TCATCT*ctgcag*ACTCAACAGATAAAGCATCCCC
2R_BamHI	TTATCT*ggatcc*GGAAGATTGGATGTAGAGATAACA
pWH43	
11621-2GISE_For	ATAAGAAT*gcggccgc*TTCATTTAATGAATAACGCTTAAAGGGAC
11621-2GISE_Rev	ATAAGAAT*gcggccgc*TTAACCAAAAGGATTAAAATCTAAAT
pVP107	

pCF10 uvrB arm1 F	GCGCAT*cccggg*GCCAGAAGTATCGTTGACTT
pCF10 uvrB arm1 R	ATCACG*ggatcc*CCAAAATTCATTAAACTCTCC
pCF10 uvrB arm2 F	ATCACG*ctgcag*GGAAATATGAAACAAGAAAAAAGG
pCF10 uvrB arm2 R	TTCGAC*tctaga*CATTGTCTTCAATGGATCAAG
T11CR2S1 CR1 PAM REase +	*ggatcc*GTATTTGAACTTTGGAACAAAAGACGCTAAACTTCCGCATAGG TATCTTTTCTTACCTTGGGGTGCAGTAGGAATTGCTTTTTTTGCT*ggatcc*
T11CR2S1 CR1 PAM REase -	*ggatcc*AGCAAAAAAAGCAATTCCTACTGCACCCCAAGGTAAGAAAAGA TACCTATGCGGAAGTTTAGCGTCTTTTGTTCCAAAGTTCAAATAC*ggatcc*
CR1 PAM from pGEM Pst F	GTGATT*ctgcag*AGCAAAAAAAGCAATTCCTA
CR1 PAM from pGEM Bam R	TCGATT*ggatcc*GTATTTGAACTTTGGAACAA
pVP108	
T11CR2S1 CR3 PAM REase +	*ggatcc*TATTTGAACTTTGGAACAAAAGACGCTAAACTTCCGCATAGGT ATCTTTTCTTACCTTTGTAGGTGCAGTAGGAATTGCTTTTTTTGC*ggatcc*
T11CR2S1 CR3 PAM REase -	*ggatcc*GCAAAAAAAGCAATTCCTACTGCACCTACAAAGGTAAGAAAAG ATACCTATGCGGAAGTTTAGCGTCTTTTGTTCCAAAGTTCAAATA*ggatcc*
CR3 PAM from pGEM Pst F	GTGATT*ctgcag*GCAAAAAAAGCAATTCCTAC
CR3 PAM from pGEM Bam R	TCGATT*ggatcc*TATTTGAACTTTGGAACAAA
pVP109	
T11cas9 H601A arm1 F	GATCAT*cccggg*GCAAAGACTTTCAAGGTG
T11cas9 H601A arm1 R	AAATTGGAATAATTGCATCAATTTCTGTTGAATCACTAACCAAT
T11cas9 H601A arm2 F	AGAAATTGATGCAATTATTCCAATTTCGATTAGCTTAGATGATT
T11cas9 H601A arm2 R	TTCGAC*tctaga*GCAAACGGTGAACGTTTA

Italicized bases represent restriction enzyme sites used for cloning. Underlined bases show the amino acid change for generation of T11RF*cas9*D7A and T11RF*cas9*H601A as well as the CRISPR1/CRISPR2 PAM or CRISPR3 PAM sequence used for studying CRISPR2 in genome defense.

Table A.2. MGE identities of spacers used to determine PAMs for the three CRISPR loci of *E. faecalis*.

Strain	Spacer[a]	Sequence Identity	Representative BLASTn hit	Genetic Element	PAM query sequence[b]
			CRISPR1-Cas		
ATCC 29200	6	27/30	DENG_01063, hypothetical protein	*E. faecalis* DENG1	AGGCATTTATACAAA
D6	1	30/30	EFS1_1522-EFS1_1523; integrase and conserved hypothetical protein (overlaps both genes)	*E. faecalis* Symbioflor1	GGGACTTTCGATGTG
	2	30/30	ENT_04360; phage major tail protein, ϕ13 family	*E. faecalis* 7L76	GGGAGCTGTTACTCC
	8	29/30	DR75_2165-DR75_2166; hypothetical proteins (sequence located between the two genes)	*E. faecalis* ATCC 29212	CGGTGCTACCGTTCG
			EF0133-EF0134; hypothetical proteins (sequence located between the two genes)	*E. faecalis* V583	CGGCGCTACCGTTCG
DS5	1	28/30	DR75_1628; hypothetical protein	*E. faecalis* ATCC 29212	GGGTTCCTATTGGCA
		27/30	EFS1_2343; conserved hypothetical protein	*E. faecalis* Symbioflor1	GGGTTCGTATTGGCA
			EF2810; hypothetical protein	*E. faecalis* V583	GGGGTCATATTGGCA
	3	30/30	DR75_1520-DR75_1521; tRNA-glu and ArpU family, phage transcriptional regulator (sequence located between the two genes)	*E. faecalis* ATCC 29212	TGGTTTAGTCACCGT
		28/30	gp34-gp35; predicted ArpU transcriptional regulator and conserved hypothetical protein	*Enterococcus* phage ϕFL1A	TCGTTTAGTCACCGT

			(sequence located between the two genes)		
ATCC 4200	3	30/30	DR75_1503; tape measure domain protein	*E. faecalis* ATCC 29212	TGGCTGATTTACCTA
	4	29/30	Genome sequence is not annotated	*Enterococcus* phage ϕEF24C-P2	CGGTTTACCATTTTA
	7	29/30	EFD32_2263-EFD32_2264; hypothetical proteins (sequence located between the two genes)	*E. faecalis* D32	GGTCGACTAGAAAGA
OG1RF	1	30/30	DR75_1516; phage portal, SPP1 Gp6-like family protein	*E. faecalis* ATCC 29212	TAGACTTAAATACTT
		27/30	EFAU085_02222; phage portal protein	*E. faecium* Aus085	TTGATTTAAACACCT
	3	28/30	EF2528-EAF2529; Cro/CI family transcriptional regulator and hypothetical protein (sequence located between the two genes)	*E. faecalis* V583	GGGTAGAAACTTTTG
		26/30	DR75_2173-DR75_2174; hypothetical protein and helix-turn-helix family protein (sequence located between the two genes)	*E. faecalis* ATCC 29212	GGGTTTTTGACGTAA
E1Sol	2	29/30	EF62_pC0030-EF62_pC0031; conserved hypothetical protein and toxin-antitoxin system, toxin component, RelE family (sequence located between the two genes)	*E. faecalis* 62 plasmid EF62pC	GAGGAACGTTTCCTT
			DR75_2995-DR75_2996; hypothetical protein and *pcfJ*-like family protein (sequence located between the two genes)	*E. faecalis* ATCC 29212 plasmid 2	GGGGAACGTTTCCTT

			CRISPR2		
ATCC 29200	39	30/30	LPS glycosyl transferase	*Enterococcus* phage IME-EF1	AAGAGCGCATTGCCG
		29/30	VD13_058; LPS glycosyl transferase	*Enterococcus* phage VD13	AGCAACGCATTGCCG
			LPS glycosyl transferase	*Enterococcus* phage SAP6	AGGAACGCATTGCTG
D6	21	29/30	EF1486; prophage 3, ply-2, predicted endolysin	*E. faecalis* V583	TGGAAGTGGAGTTTG
		28/30	EFD32_2243; N-acetylmuramoyl-L-alanine amidase family protein	*E. faecalis* D32	TGGAAGTCCAGTTTG
	22	30/30	EFS1_2364; hypothetical protein	*E. faecalis* Symbioflor1	TGGTTATAAAAAAGA
	25	30/30	EF62_pB0048; *pcfV*, hypothetical protein	*E. faecalis* 62 plasmid EF62pB	AGGCGTCTTCGGGGA
DS5	74	30/30	Tape measure domain-containing protein	*Enterococcus* phage EFC-1	CGGTATCGAACAAAA
ATCC 4200	7	30/30	IME-EF4_48-IME-EF4_49; hypothetical protein and major tail protein (sequence located between the two genes)	*Enterococcus* phage IME-EF4	TGGCGTGCTAGGGTG
			EfaCPT1_gp11-EfaCPT1_gp12; putative major tail tube protein and putative tail tape measure chaperone protein (sequence located between the two genes)	*Enterococcus* phage EFaCPT1	TGGCGTACTAGGGTG
	8	28/30	IME-EF4_27; prim-pol domain protein	*Enterococcus* phage IME-EF4	AGGAAGTTCAAGGAG
	9	29/30	EF2813; prophage 6, tail tape measure protein	*E. faecalis* V583	AGGTCAAATGGAAAA
	10	30/30	EF2836-EF2837; prophage 6,	*E. faecalis*	AGGAAAGGGAGTTA

			hypothetical proteins (sequence located between the two genes)	V583	A
	11	29/30	locus ID?; *pcfS*, putative ssb protein	*E. faecalis* plasmid pCF10	CGTCAAGTGGTGTAG
OG1RF	34	30/30	ENT_20330; phage terminase, large subunit, PBSX family	*E. faecalis* 7L76	CGGATCAGTAGCATA
	37	27/30	M7W_1349; hypothetical protein	*E. faecium* NRRL B-2354	TGGATAAACGGAGTC
	38	29/30	pLG2-0017; hypothetical protein	*E. faecalis* plasmid pGL2	CGGCAGCAGAAGCTA
	5	29/30	EFD32_2253; major tail protein	*E. faecalis* D32	CGTCAAGTGGTGTAG
E1Sol	91	30/30	EF2535; nucleotidyltransferase domain-containing protein	*E. faecalis* V583	AGGAGAAAGTTAAAG
			CRISPR3-Cas		
T11	1	30/30	EfaCPT1_gp29; putative metallo-beta-lactamase domain protein	*Enterococcus* phage EFaCPT1	TGGTACCACAATTGA
	6	30/30	*repB,* replication-associated protein	*E. faecalis* plasmid pAD1	TTGTAATTGTTTTGC
	7	30/30	EF62_pC0029; conserved hypothetical protein	*E. faecalis* 62 plasmid EF62pC	AAGTAATACGACAAT
	8	27/30	gp31, conserved hypothetical protein	*Enterococcus* phage φFL1A	TGATAGAAATTTAGT
	9	29/30	EFD32_2249, Hypothetical protein	*E. faecalis* D32[c]	AAGTACCTGTTAACT
	10	29/30	EFD32_2266, ArpU family transcriptional regulator	*E. faecalis* D32[c]	TAATATCATTCTCAC
	14	30/30	hypothetical protein	*Enterococcus* phage EFC1	TGTAAAACGTTCTTC
	19	27/30	PHIEF11_0021; putative phage structural protein	*Enterococcus* phage φEF11	TGTATAAGTTGATTG
	20	29/30	EFD32_2276, N-6 DNA	*E. faecalis* D32[c]	TTGTATACTTTTTTC

			methylase		
	21	29/30	ENT_02110; DNA segregation ATPase FtsK/SpoIIIE and related proteins	*E. faecalis* 7L76	CAATAAGCGTATTTG
		27/30	EF2533; FtsK/SpoIIIE family protein	*E. faecalis* V583	GTTATTCGCATAAAC
F1	3	30/30	VD13_034; putative structural protein	*Enterococcus* phage VD13	TGGTACTCCTTCGCT
	4	30/30	DR75_1629; hypothetical protein	*E. faecalis* ATCC 29212	CCATAGAATCGATAC
	5	30/30	EFD32_2267; hypothetical protein	*E. faecalis* D32	CAATAGCGCTCATCA
	6	30/30	DR75_1503; tape measure domain protein	*E. faecalis* ATCC 29212	TGATACGGCGGTTGG
Fly1	2	28/30	EFD32_2260; phage portal protein	*E. faecalis* D32	GCCTATCAAAAACTT
	3	28/30	ENT_04320; phage tail tape measure protein, TP901 family	*E. faecalis* 7L76	CGATATGAAACTACT
	4	30/30	EFS1_2367; DNA replication protein DnaC	*E. faecalis* Symbioflor1	GCATAAAGATTTCCT
	5	30/30	DR75_2999-DR75_3000; hypothetical protein and addiction molecule toxin, RelE/StbE family protein (sequence located between the two genes)	*E. faecalis* ATCC 29212 plasmid 2	TGGTACTGTAGCAAA
	6	29/30	EF2825; prophage 6, hypothetical protein	*E. faecalis* V583	TGATTCCAATCAAGA
	8	29/30	DENG_01652; hypothetical protein	*E. faecalis* DENG1	TCATAAAAATCGACC

[a]Spacer numbers correspond to a previously published *E. faecalis* CRISPR2 spacer dictionary (39).

[b]Protospacer-adjacent sequence extracted from representative hits used to create putative PAM for each CRISPR locus; Figure A2.
[c]These protospacers align within 18 kb of each other on the D32 genome; this region could be a prophage or pathogenicity island.

Table A.3. Protein features around the predicted methyltransferase in the T11 R-M region.

Locus ID	Annotation	Pfam hit (e-value)	Notes
EFMG_00927	Predicted protein	No conserved domains	
EFMG_00926	AlwI family type II restriction endonuclease	RE_AlwI; AlwI family restriction endonuclease ($1.2e^{-32}$)	60% identity (99% coverage) with LmoJ2R of *L. monocytogenes* J2749 (AlwI family Type II restriction endonuclease) [e-value, 0.0]
EFMG_00925	NgoFVII family restriction endonuclease	RE_NgoFVII; NgoFVII restriction endonuclease ($5.9e^{-26}$)	43% identity with OG1RF R.EfaRFI [e-value, $1e^{-62}$]
EFMG_00924	DNA cytosine methyltransferase	DNA_methylase; C-5 cytosine specific DNA methylase ($5.1e^{-71}$)	56% identity to OG1RF M.EfaRFI [e-value, $2e^{-125}$]
EFMG_00923	Very short patch repair endonuclease	Vsr; DNA mismatch endonuclease Vsr ($2.2e^{-25}$)	73% identity with endonuclease of *Staphylococcus carnosus* subsp. *utilis* [e-value,e^{-70}]
EFMG_00922	Predicted protein	HNH; HNH endonuclease ($5.8e^{-7}$)	35% identity with 5-methylcytosine-specific restriction enzyme A of *Bacillus* sp. GeD10 [e-value, $2e^{-24}$]

In addition to T11, this configuration of genes occurs in the following *E. faecalis* strains: B301, B345, B347 and T19.

APPENDIX B

CHAPTER 3 SUPPLEMENTAL FIGURE AND TABLES

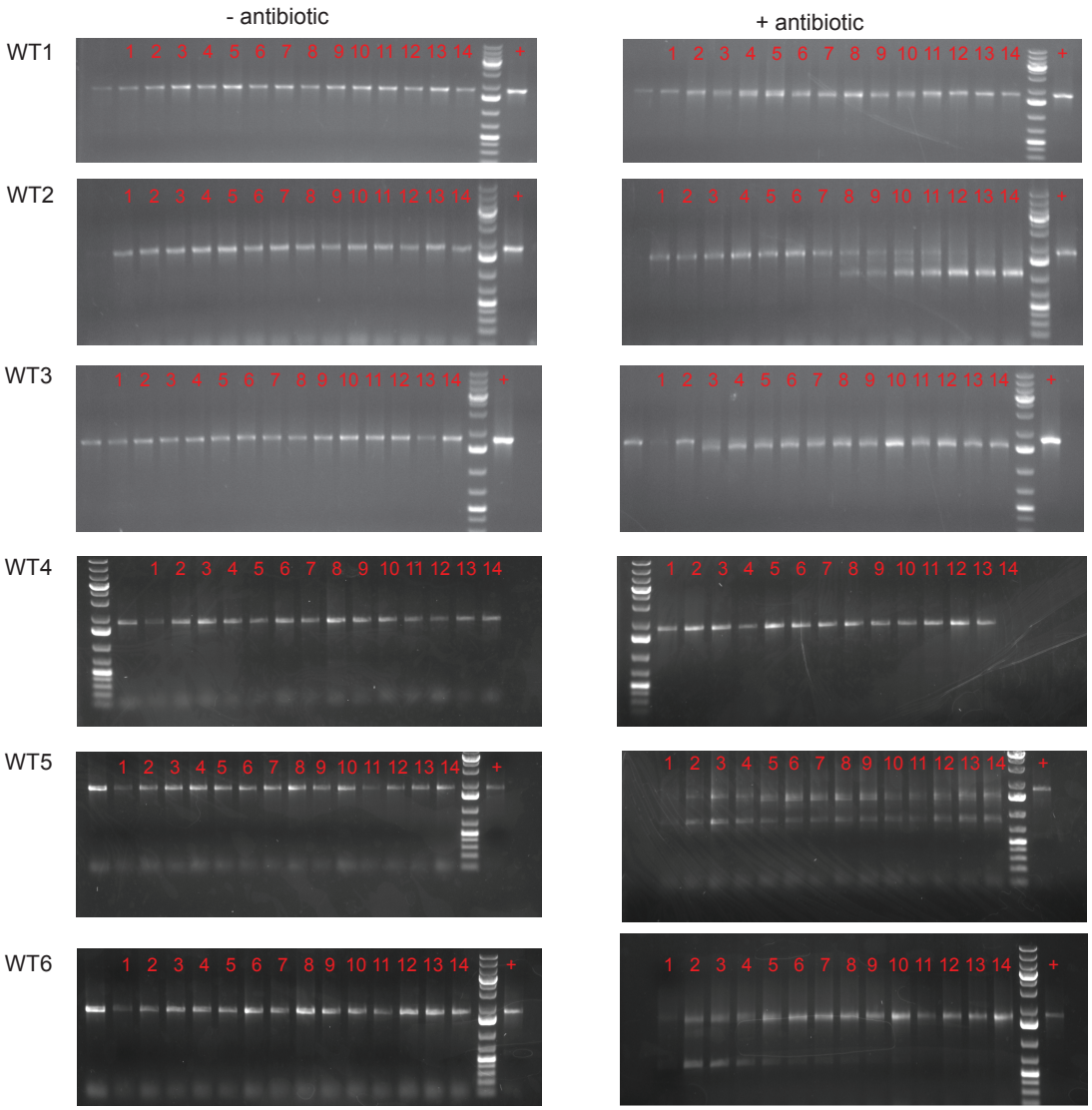


Figure B.1. Gel electrophoresis of CRISPR3 amplicons reveal spacer loss in T11RF pAM714 transconjugants passaged with antibiotic. Six T11RF pAM714 transconjugants were serially passaged for 14 days in BHI (left panel) or BHI with erythromycin (right panel). The size of the CRISPR3 array was monitored using PCR and gel electrophoresis on each passage day.

Table B.1. Primers used in this study.

Name	Primer sequence
CRISPR seq primers	
CRISPR1 seq For	CGTATTTGACAGAGGATGAAG
CRISPR1 seq Rev	CGAATATGCCTGTGGTGAAA
CRISPR2 seq For	TGCTGTTACAGCTACTAAA
CRISPR2 seq Rev	GCCAATGTTACAATATCAAACA
CRISPR3 seq For	GCTCACTGTATTGGAAGAAC
CRISPR3 seq Rev	CATCGATTCATTATTCCTCCAA
pWH107 and derivatives	
pVP107_XbaI_For	CTAGAGATAATATATCTTTTATATAGAAGATGGGTACCAT GGCATG
pVP107_SphI_Rev	CCATGGTACCCATCTTCTATATAAAAGATATATTATCT
T11CR3sp6_BamPst_F	GATCCTCCCGATACAGCTCTTTATTCTTCTAATTACATTGT ACTGCA
T11CR3sp6_BamPst_R	GTACAATGTAATTAGAAGAATAAAGAGCTGTATCGGGAG
T11CR3sp1_BamPst_F	GATCCTCAAAAGTTGAATATGTTTCGCTTTGGTGTAATTGT ACTGCA
T11CR3sp1_BamPst_R	GTACAATTACACCAAAGCGAAACATATTCAACTTTTGAG
T11CR3sp7_BamPst_F	GATCCTCTGCTTTTGGAGGAATACAAATGAGAAGATTTTG TACTGCA
T11CR3sp7_BamPst_R	GTACAAAATCTTCTCATTTGTATTCCTCCAAAAGCAGAG
T11CR2S1 seq arm1 F	CGAAATCAGCACATGGAACA
T11CR2S1 seq arm2 R	CCAGTAACTGTATCAACTAC

Table B.2. SNPs detection in all gDNA sequencing samples.

Supercontig	Pos.	Ref	Allele	Annotation	AA change	T11RF[a]	WT3[b]	WT1[c]	WT2[c]	WT3[c]	WT5[c]	WT6[c]	Δ4[c]
pAD1repabc	1904	C	T	Replication protein	AAB00504.1:p.Pro100Leu				45.49				
1.11	51613	C	T	Collagen adhesin								25.00	
1.11	51674	G	A	Collagen adhesin	WP_002379313.1:p.Gly392Ser						33.90	33.39	
1.11	51745	A	G	Collagen adhesin							33.02		
1.11	337315	G	T	Helicase	WP_002382285.1:p.Glu425*							37.16	
1.11	652983	G	A	Cas9	WP_002379510.1:p.Gln506*			31.64					
1.11	653180	AG	T	Cas9	WP_002379510.1:p.Leu440fs		22.74			23.52			
1.11	653184	C	T	Cas9	WP_002379510.1:p.Glu439Lys		23.21			24.17			
1.11	654165	G	-	Cas9	WP_002379510.1:p.Arg112fs				48.41				
1.11	685523	A	T	Conserved hypothetical protein	WP_002382413.1:p.Glu5Val		23.24						
1.11	685938	A	T	Conserved hypothetical protein	WP_002382413.1:p.Glu143Asp		26.00						
1.11	974866	AT	GC	Cell wall surface anchor family protein	WP_002379620.1:p.Ile972Ala							23.41	
1.11	974991	A	G	Cell wall surface anchor								22.67	

				family protein		
1.2	2903	C	T	Peptidylprolyl isomerase	WP_002355150.1:p.Val167Ile	46.03

[a]Day 14 BHI-passaged
[b]Day 1 Erythromycin-passaged population WT3 to confirm presence of *cas9* mutations prior to Day 14
[c]Day 14 Erythromycin-passaged

Table B.3. Quality control of the amplicon sequencing reads.

Sample name	# of Total reads	% reads mapped in Step 1	% reads mapped in Step 2	% reads mapped in Step 3	% reads mapped in Step 4	% reads left
Day 1						
T11RF CR3	19,082,652	97.543%	0.642%	0.543%	0.865%	0.408%
Δ4	20,808,404	98.155%	0.362%	0.616%	0.702%	0.165%
WT1	16,736,162	97.629%	0.985%	0.544%	0.709%	0.132%
WT2	16,859,152	98.292%	0.344%	0.562%	0.664%	0.139%
WT3	17,357,970	97.141%	1.416%	0.626%	0.672%	0.145%
WT5	17,744,230	96.024%	2.393%	0.699%	0.747%	0.137%
WT6	16,179,400	95.871%	2.488%	0.814%	0.672%	0.154%
Day 14						
T11RF CR3	17,973,946	98.233%	0.373%	0.601%	0.649%	0.143%
Δ4	14,403,868	98.173%	0.445%	0.625%	0.529%	0.227%
WT1	14,158,158	98.136%	0.454%	0.624%	0.545%	0.240%
WT2	14,538,836	96.340%	2.166%	0.707%	0.527%	0.260%
WT3	17,174,946	96.018%	2.538%	0.738%	0.514%	0.192%
WT5	15,200,646	95.193%	2.998%	0.885%	0.738%	0.187%
WT6	12,325,042	94.419%	2.735%	1.501%	1.005%	0.340%

APPENDIX C

CHAPTER 4 SUPPLEMENTAL FIGURE

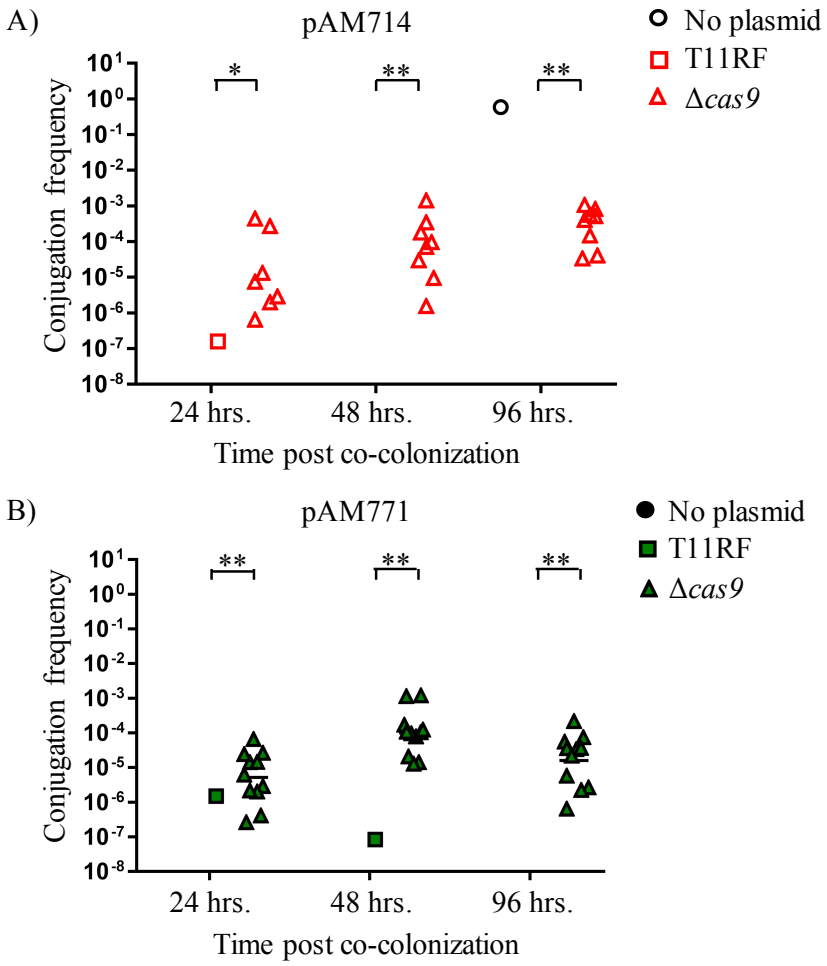


Figure C.1. Frequency of conjugation in the mouse gut. The conjugation frequency for each mouse was determined by dividing the transconjugant CFU/g by the donor CFU/g; one symbol represents one mouse on the graph. Black horizontal bars represent the geometric mean of data in each group. No symbol means that a frequency could not be calculated because one or both of the values (donor CFU/g or transconjugant CFU/g) were zero. For experiments with pAM714 as the plasmid donor (A), ten mice were used in each group. For experiments with pAM771 as the plasmid donor (B), ten mice were used for the no plasmid control and T11RF recipient groups, whereas the Δ*cas9* group had 11 mice. Statistical significance was assessed using a one-tailed Student's t-Test; *P*-values, *<0.05 and **<0.005.

BIOGRAPHICAL SKETCH

Valerie Jane Price was born in Fort Worth, Texas on November 29, 1988. She graduated from Castleberry High School in 2007. Valerie began her undergraduate studies at UT Dallas in 2007. During this time, she participated in research in the lab of Dr. Santosh D'Mello studying the process of neuronal programmed cell death. Valerie graduated in 2011 with a BS in Molecular Biology. After graduation, Valerie worked as a research technician in the D'Mello lab until entering the doctoral program at UT Dallas in 2012. She changed research fields for her doctoral training and joined the lab of Dr. Kelli L. Palmer where she studied genome defense systems in the opportunistic pathogen, *Enterococcus faecalis.* Valerie received her MS in Molecular and Cell Biology in 2014. She was recognized by The University of Texas System as a Jess Hay Chancellor's Graduate Research Fellow in 2015.

CURRICULUM VITAE

Valerie Price
Department of Biological Sciences, The University of Texas at Dallas,
800 W. Campbell Rd. BSB12, Richardson, TX 75080 / valerie.price@utdallas.edu

Education

2014-Present	The University of Texas at Dallas Ph.D. Candidate, Molecular and Cell Biology Advisor: Dr. Kelli Palmer	Richardson, TX
2012-2014	The University of Texas at Dallas M.S. in Molecular and Cell Biology Advisor: Dr. Kelli Palmer	Richardson, TX
2007-2011	The University of Texas at Dallas B.S. in Molecular Biology	Richardson, TX

Research Experience

The University of Texas at Dallas
Teaching Assistant: August 2012-May 2014
Research Assistant: June 2014-Present
Advisor: Dr. Kelli Palmer

Duties:
• Primary independent research project is to understand the role of CRISPR-Cas, a bacterial genome defense system, in the emergence of antibiotic resistant *Enterococcus faecalis* endemic to clinical settings. A combination of microbiology techniques and genome sequencing analysis are being used to complete work for this project.

• Mentoring: Supervised seven undergraduate students on individual research projects.

The University of Texas at Dallas
Technician: June 2011-August 2012
Student Worker: August 2009-May 2011
Advisor: Dr. Santosh R. D'Mello

Duties:
• Primary independent research project during employment was studying histone deacetlyase-4 and its role in neuronal survival in the central nervous system by using two conditional knock-out mice lines. Responsibilities and data generated in this study were: colony management (including breeding and genotyping), performing phenotypic analyses and establishing expression data.

• Aside from the primary research project, breeding and expression analysis were performed on multiple neurodegenerative disease model mice lines, which produced data for several publications.

• Mentoring: Trained undergraduate students in molecular biology techniques including PCR, western blot, cell culture, immunocytochemistry, genotyping and mouse colony breeding and management.

Teaching Experience

The University of Texas at Dallas, Department of Biological Sciences

- Molecular and Cell Biology Lab. Teaching Assistant (January 2014-May 2014 and August 2012 – May 2013)
- Molecular and Cell Biology. Teaching Assistant (August 2013 – December 2013)
- Biochemistry Lab. Teaching Assistant (May 2013-August 2013)

Publications

Price VJ, Huo W, Sharifi A, Palmer KL. 2016. CRISPR-Cas and restriction-modification act additively against conjugative antibiotic resistance plasmid transfer in *Enterococcus faecalis*. mSphere. 1(3):e00064-16.

Price V, Wang L, D'Mello SR. 2013. Conditional deletion of histone deacetylase-4 in the central nervous system has no major effect on brain architecture or neuronal viability. J. Neurosci. Res. **91**:407–15.

Bardai FH, **Price V**, Zaayman M, Wang L, D'Mello SR. 2012. Histone deacetylase-1 (HDAC1) is a molecular switch between neuronal survival and death. J. Biol. Chem. **287**:35444–53.

Dastidar SG, Bardai FH, Ma C, **Price V**, Rawat V, Verma P, Narayanan V, D'Mello SR. 2012. Isoform-specific toxicity of Mecp2 in postmitotic neurons: Suppression of neurotoxicity by FoxG1. J. Neurosci. **32**:2846–2855.

Zhao K, Ippolito G, Wang L, **Price V**, Kim MH, Cornwell G, Fulenchek S, Breen GA, Goux WJ, D'Mello SR. 2010. Neuron-selective toxicity of tau peptide in a cell culture model of neurodegenerative tauopathy: essential role for aggregation in neurotoxicity. J. Neurosci. Res. **88**:3399–413.

Wang L, Ankati H, Akubathini SK, Balderamos M, Storey CA, Patel A V, **Price V**, Kretzschmar D, Biehl ER, D'Mello SR. 2010. Identification of novel 1,4-benzoxazine compounds that are protective in tissue culture and in vivo models of neurodegeneration. J. Neurosci. Res. **88**:1970–84.

Poster Presentations – Presenting Author

Price, VJ., Huo, W., Sharifi, A., Palmer, KL. Plasmid maintenance results in CRISPR memory loss in *E. faecalis*. American Society for Microbiology Texas Branch Fall Meeting, Richardson, TX, November 10th-12th 2016.

Price, VJ., Huo, W., Sharifi, A., Palmer, KL. Plasmid maintenance results in CRISPR memory loss in *E. faecalis*. Evolution General Meeting. Austin, Texas, June 17st-21st 2016.

Price, VJ., Sharifi, A. Huo, W. and Palmer, KL. Evolutionary studies of CRISPR-Cas defense in *Enterococcus faecalis*. Molecular Genetics of Bacteria and Phages, Madison, WI, August 4th-8th 2015.

Price, VJ., Sharifi, A., Huo, W. and Palmer, KL. Evolutionary studies of CRISPR-Cas defense in *Enterococcus faecalis.* Gordon Research Conference on Microbial Population Biology, Andover, NH, July 19th-24th 2015.

Price, VJ., Sharifi, A., Huo, W. and Palmer, KL. Evolutionary studies of CRISPR-Cas defense in *Enterococcus faecalis.* Gordon Research Seminar on Microbial Population Biology, Andover, NH, July 18th-19th 2015.

Price, VJ., Sharifi, A. and Palmer, KL. CRISPR-Cas defense and acquired antibiotic resistance in *Enterococcus faecalis*. Molecular Genetics of Bacteria and Phages, Madison, WI, August 5th-9th 2014.

Price, VJ., Sharifi, A. and Palmer, KL. CRISPR-Cas defense and acquired antibiotic resistance in *Enterococcus faecalis*. American Society for Microbiology Texas Branch Annual Meeting, New Orleans, LA, October 31st-November 2nd 2013.

Poster Presentations – Another Author Presenting

Sharifi, A., **Price, V.**, and Palmer, K. Antibiotic Resistance and the Emergence of Immune-Deficient Bacteria. American Society of Microbiology General Meeting, New Orleans, LA, May 30th-June 2nd 2015.

Singaravelu, S., **Price, V.**, and Palmer, K. CRISPR-Cas9 Antimicrobials & Bacterial "Auto-Immunity." Green Fellowship Presentation, UT Southwestern Medical Center, Dallas, TX, May 7th 2015.

Sharifi, A., **Price, V.**, and Palmer, K. Antibiotic Resistance and the Emergence of Immune-Deficient Bacteria. American Society of Microbiology Texas Branch Fall Meeting, Houston, TX, November 6th-8th 2014.

Sharifi, A., **Price, V.**, and Palmer, K. Antibiotic Resistance and the Emergence of Immune-Deficient Bacteria. UT Dallas Undergraduate Research Presentation, Richardson, TX April 2014.

Oral Presentations – Presenting Author

Price, VJ., Huo, W., Sharifi, A., Palmer, KL. Plasmid maintenance results in CRISPR memory loss in *E. faecalis*. Evolution General Meeting. Austin, Texas, June 17st-21st 2016.

Price, VJ., Huo, W., Sharifi, A., Palmer, KL. Plasmid maintenance results in CRISPR memory loss in *E. faecalis*. 60th Wind River Conference on Prokaryotic Biology. Estes Park, Colorado, June 1st-5th 2016.

Oral Presentations – Another Author Presenting

Huo, W., **Price, V.**, Zhang, M., Palmer, K. Genome Defense Systems Compromise for Beneficial MGE Acquisition in *E. faecalis*. Molecular Genetics of Bacteria and Phages, Madison, WI, August 8th-12th 2016.

Huo, W., **Price, V.**, Zhang, M., Palmer, K. Genome Defense Systems Compromise for Beneficial MGE Acquisition in *E. faecalis*. American Society for Microbiology General Meeting: Microbe 2016, Boston, MA, June 16th-20th 2016.

Certifications

Institutional Responsible Conduct of Research training; specific topics covered: Research Misconduct & Scientific Fraud, Responsible Authorship, Publication and Peer Review, Mentor/Mentee Responsibilities and Relationships, Human Subjects Research.

Certified by the Collaborative Institutional Training Initiative (CITI) in the area of Biomedical Responsible Conduct of Research.

Professional Development and Other Training

Online Scientific Writing and Publishing Workshop. Presented by the American Society for Microbiology, January-April 2017.

Teaching Workshop: Assessing Learning Outcomes. Presented by, UTD Office of Assessment and Gloria Shenoy, PhD, February 2016.

Teaching Development: Crafting a Teaching Philosophy. Presented by, UTD Center for Teaching and Learning, February 2016.

Teaching Lecture and Workshop: The Reader Expectation Approach. Presented by, Dr. George Gopen, January 2016.

Graduate Student Workshop Series: Writing a Literature Review/Going Beyond Undergraduate Research. Presented by, the Eugene McDermott Library and the SSC Writing Center, January 2016.

Awards

Samuel Kaplan Award 2016

- In recognition of: Outstanding Scientific Achievement in Molecular Microbiology for a Poster given by a Graduate Student at the Fall Meeting of the Texas Branch of the American Society for Microbiology

Jess Hay Chancellor's Graduate Research Fellowship 2015

- $10,000 fellowship awarded by The University of Texas System recognizing excellence in research benefiting the state of Texas and biomedical science.

Harris Scholarship 2014

- $500 award from the Department of Biological Sciences at UT Dallas for the best research discussion presented by a graduate student; selection committee consisted of peers and faculty members.

Service to the field

Royston Clowes Memorial Lecture Coordinator 2017

- Invited guest speaker on behalf of graduate student association and was responsible for organizing the schedule of events during the speaker's visit to UT Dallas.

Member of the American Society for Cell Biology 2017

Guest Lecturer at STEAM camp 2016

- Mentored a group of high school girls on the topic of global access to clean water as part of a week-long camp sponsored by the UT Dallas Science and Engineering Education Center.

Member of the American Society for Microbiology 2013-Present