THE ROLE OF Enterococcus faecalis BIOFILM FORMATION
IN THE REGULATION OF CONJUGATION

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

LAURA CAROL CASE COOK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DR. GARY M. DUNNY
APRIL 2012
Acknowledgements

So many people are responsible for helping me complete my thesis research. First, I want to thank my mentor, Gary Dunny, for his helpful insights and conversations. Throughout the last five years he has kept me excited about my project even when things weren’t working out. I hope that someday I can mentor people with the same mix of rational thought and passionate enthusiasm he has. Thank you to all of my committee members for your advice.

All of the members of the Dunny lab helped me in so many ways throughout my thesis research. Thanks to Heather Haemig, Katie Ballering, Thinh Le, Chris Johnson, Kristi Frank, Olivia Chuang-Smith, Frank Barr, Nick Dillon, Yuqing Chen and Anne-Marie Lueck. Thanks Aaron Barnes who spent countless hours showing me what I was doing wrong with the microscope and to Suzanne Grindle for RNAseq processing and analysis. I would like to especially thank Dawn Manias. She has been such an amazing and supportive mentor and always there to listen to my failures or celebrate my successes. I also want to thank all of my wonderful mentees throughout graduate school who have taught me an amazing amount about science and teaching: Francois, Chi, Ben, and Katie.

I appreciate the lessons I have learned from our collaborators Drs. Wei-Shou Hu and Anushree Chatterjee. They taught me more about mathematical modeling than I ever wanted to know but always provided the rationale and explanations to help me understand even the most difficult modeling concepts. I would also like to thank Dr. Jeremy Yarwood for giving me the opportunity to do
an internship in his laboratory at 3M. This was invaluable in my thesis work and a wonderful industry experience.

I would like to thank the members of the Schlievert lab: Tim Tripp, Sara Vetter, Mandy Brosnahan, Kristi Strandberg, Ted Sibley, John Mleziva, Petra Kohler, Evan Henke (E-Train), Jessica Rotschaffer (J-Dawg), Kyle Williams, Alexa Pragman and all of the others who went through the lab. Tim hired me as an undergraduate and Sara mentored me on my first project and without them I might not have ever joined the graduate program. They, along with Kristi and Mandy provided the fun and laughter that kept me going on days when research seemed overwhelming. I would specifically like to thank Pat Schlievert. That fact that he was usually right with all of his crazy ideas gave me the courage to try new things. I greatly appreciate his mentorship both before and during graduate school.

Last but not least, I would like to thank my friends and family. My friends were there to listen to practice talks about things they had never heard of and didn’t understand. They were a wonderful way to get away from it all after the science day ended. My family, especially my parents and siblings, are a wonderful source of support and love without which I never would have made it through. Lastly, I want to thank my husband, Gabe, for his patience, love, and understanding. He listened to my problems and helped me make sense of them when he didn’t even understand what I was talking about. During the stressful times, he was always there to give me a big hug to keep me going.
Dedication

This dissertation is dedicated to my husband, Gabriel Morton-Cook who kept me sane during graduate school.
Abstract

*Enterococcus faecalis* has recently emerged as an important nosocomial pathogen. Pathogenicity of these organisms depends greatly on a few important aspect of enterococcal physiology. The ability of enterococci to form biofilms greatly enhances their virulence. Their innate resistance to many antibiotics and their ability to transfer these resistance genes through conjugation heightens their threat to human health. The work described in this thesis attempts to explain the roles of biofilm growth, conjugation, and cell communication in *E. faecalis*. To examine the role of biofilm growth on the *E. faecalis* transcriptome, RNAseq analysis was undertaken. We found that over 100 genes were measurably upregulated during biofilm growth while approximately 26 genes were downregulated. These data gives us important insights into the biology of enterococcal biofilms. In clinical settings, biofilms are likely locations for antibiotic resistance transfer events involving nosocomial pathogens such as *E. faecalis*. Conjugation is an important mode of horizontal gene transfer in bacteria, enhancing the spread of antibiotic resistance. In this work, I demonstrate that growth in biofilms alters the induction of conjugation by a sex pheromone in *E. faecalis*. Mathematical modeling suggested that a higher plasmid copy number in biofilm cells would enhance a switch-like behavior in the pheromone response of donor cells with a delayed, but increased response to the mating signal. Alterations in plasmid copy number and a bimodal response to induction of conjugation in populations of plasmid-containing donor cells were both observed.
in biofilms, consistent with the predictions of the model. The pheromone system may have evolved such that donor cells in biofilms are only induced to transfer when they are in extremely close proximity to potential recipients in the biofilm community. These results have important implications for development of chemotherapeutic agents to block resistance transfer and treat biofilm-related clinical infections.
# Table of Contents

List of Tables................................................................................................................. vi

List of Figures................................................................................................................... vi

Introduction..................................................................................................................... 1

  Cell-cell communication in bacteria........................................................................... 1
  Bacterial biofilms......................................................................................................... 2
  RNAseq as a tool for examining the effects of the environment on bacterial gene expression........................................................................................................... 3
  The role of the biofilm environment on regulation of conjugation............................ 4

Materials and Methods................................................................................................. 12

  Bacterial strains and media......................................................................................... 12
  Cloning and DNA manipulation.................................................................................. 15
  Plasmid and strain construction and induction............................................................ 15
  Genomic DNA (gDNA) preparation and quantification.............................................. 20
  RNA extraction/cDNA preparation............................................................................. 21
  Quantitative real-time PCR (qRT-PCR) ...................................................................... 22
  Biofilm growth and induction..................................................................................... 22
  Serial biofilm growth................................................................................................. 24
  RNASeq....................................................................................................................... 24
  Conjugation experiments......................................................................................... 25
  Flow cytometry and fluorescent activated cell sorting (FACS)................................. 27
  Statistical analysis..................................................................................................... 28
Microscopy ........................................................................................................ 28
Pulse field gel electrophoresis (PFGE) ......................................................... 29
iCF10 donor density and de-induction cell preparation ......................... 29
Mathematical modeling ............................................................................... 30

Results ............................................................................................................ 32

I. Development of a reproducible system to examine differential gene
expression in Enterococcus faecalis biofilms........................................ 32

  E. faecalis biofilms are actively growing and there are no rapid
mutagenic changes selecting for biofilm growth................................. 32

  Whole transcriptome RNA sequencing shows differential gene
expression between biofilm and planktonic cells.............................. 36

II. Biofilm growth alters regulation of conjugation by a bacterial
pheromone ................................................................................................... 51

  The pattern of conjugation induction differs between cells growing in
liquid culture and those grown in a biofilm......................................... 51

  The biofilm matrix is not responsible for differences between biofilm
and planktonic conjugation regulation .............................................. 53

  RNaseIII and endogenous pheromone production do not promote
bistability in planktonic cells................................................................. 62

  pCF10 copy number and heterogeneity are increased in biofilm
cells........................................................................................................... 64
The copy number of various enterococcal plasmids is altered by growth in a biofilm................................................................. 73

III. The dual role of iCF10 as inhibitor and quorum sensing molecule........... 82
    Donor density affects $Q_L$ expression........................................ 82
    Donor density affects mating frequencies................................. 86
    iCF10 acts to repress induction of $Q_L$ by cCF10......................... 88

Discussion.............................................................................................. 91
    The \textit{E. faecalis} biofilm transcriptome................................. 91
    Effects of the environment on conjugation in pCF10.................... 94
    Plasmid copy number is altered by changing environments.......... 96
    Donor density and iCF10: Quorum sensing in \textit{E. faecalis} conjugation.. 101

References......................................................................................... 104

Appendices........................................................................................... 111
    I. Unstable GFP variants do not function properly in \textit{E. faecalis} cells.... 111
    II. Tdtomato does currently fluoresce optimally in \textit{E. faecalis}.......... 114
    III. Mathematical modeling design, equations, and parameters.......... 116
LIST OF TABLES

1. Bacterial strains ............................................................................................................. 13
2. Cloning vectors and plasmids....................................................................................... 14
3. Oligonucleotides........................................................................................................... 19
4. RNAseq analysis genes down-regulated in biofilm formation................................... 40
5. RNAseq analysis genes up-regulated in biofilm formation........................................ 41
6. ABC transporters up-regulated in biofilms................................................................. 48
7. Biofilm and planktonic mating rates............................................................................ 61
8. Enterococcal plasmids for copy number measure..................................................... 75
9. Mathematical modeling equations................................................................................. 118
10. List of variables and parameters used in the mathematical model......................... 119
LIST OF FIGURES

1. Induction of pCF10.............................................................................................................. 10
2. Cloning of gfp into pCF10.............................................................................................. 11
3. Growth of biofilm and planktonic cells in the CDC biofilm reactor.................. 33
4. Serial growth of biofilm cells for five days does not significantly increase biofilm cell yield.................................................................................................................. 35
5. Gene categories found in RNAseq analysis................................................................. 47
6. Growth in a biofilm alters the induction pattern of pCF10 conjugation............. 55
7. Growth in a biofilm alters the time course induction pattern of pCF10 conjugation......................................................................................................................................... 57
8. Growth pattern changes are seen as early as 4 hours post biofilm inoculation.................................................................................................................................................. 58
9. Bimodal induction is not an artifact of GFP expression from the prgQ promoter.................................................................................................................................................... 60
10. Knockouts of RNaseIII, eep, and ccfA do not affect the induction pattern biofilm and planktonic cells...................................................................................................................... 63
11. Modeling the effect of pCF10 plasmid copy number (N) on the cCF10 pheromone response..................................................................................................................................... 66
12. The copy number of pCF10 as well as the population heterogeneity of copy number is increased in biofilm cells compared to planktonic........... 70
13. Pulse field gel electrophoresis (PFGE) to examine cell copy number...... 72
14. The copy number of four different plasmids is increased during biofilm growth versus growth in liquid culture ............................................................... 76
15. The amount of total gDNA/CFU extracted is the same for biofilm and planktonic cells ........................................................................................................ 77
16. Copy number of pMSP3535VAX is increased as early as 4 hours post-inoculation of the biofilm ............................................................................. 78
17. The prgX-$Q_L$ portion of the pCF10 plasmid is sufficient to produce bimodal induction of biofilm cells and sorting of induced cells demonstrates copy number heterogeneity in biofilms that is not present in planktonic populations ............................................................................... 80
18. iCF10 and cCF10 do not act at a 1:1 ratio in terms of induction independently of donor density ................................................................................................ 84
19. Donor density affects induction and continued expression of $Q_L$ transcripts .............................................................................................................. 85
20. Mating efficiency is affected by donor density .................................................. 87
21. iCF10 is responsible for shutdown of $Q_L$ expression and abolishing endogenous iCF10 abrogates donor density effects on $Q_L$ ......................... 90
22. Model of donor density effect on conjugation .................................................. 103
23. Unstable GFP variants either remain stable or are not expressed properly in *E. faecalis* cells ........................................................................................................ 113
24. Tdtomato expression in *E. faecalis* cells is inconsistent .............................. 115
INTRODUCTION

*Enterococcus faecalis*, a gram positive bacterium, is a member of the normal flora of the human intestinal tract. Although often non-pathogenic, *E. faecalis* has recently emerged as an important nosocomial pathogen causing infections on indwelling devices such as catheters and heart valves, in the peritoneal cavity and in the urinary tract among others (1, 2). Important characteristics involved in enterococci virulence include intrinsic resistance to antibiotics, the ability to transfer antibiotic resistance genes via conjugation (a form of cell-cell communication), and the ability to form biofilms on many different surfaces. The interplay between these characteristics is of clinical importance and is the focus of my thesis research.

**Cell-cell communication in bacteria**

The ability of bacterial cells to communicate via secreted molecules is a significant aspect of microbial biology and an important area of research. Inter- and intra-species signaling can coordinate multicellular behaviors such as light production, virulence gene production, biofilm formation, and competence development among others (3-6). Quorum sensing, a type of bacterial communication, involves sensing of bacterial population densities via the use of secreted small molecules. In the case of gram negative bacteria, these small molecules are generally secreted homoserine lactones whereas gram positive bacteria generally employ small peptides.
Recent research has focused on the effects of quorum sensing on altering bacterial transcription patterns and behaviors, but less is known about the role of changing environments on cellular communication. For example, quorum sensing has been shown to enhance biofilm formation in *Pseudomonas aeruginosa* (6), but a larger question that remains is once the biofilm is formed, what are the effects on further cell signaling? For *P. aeruginosa*, growth medium has been found to alter expression of quorum sensing systems greatly (7) demonstrating that although quorum sensing may alter bacterial responses to the environment, the quorum sensing circuits are, in turn, affected by environmental cues.

**Bacterial biofilms**

Bacterial biofilms are communities of bacteria growing attached to a surface. They are encased in a matrix of their own production that is made up of different combinations of proteins, polysaccharides, and DNA. Many serious bacterial infections have been traced to biofilms including infections on indwelling medical devices and human tissues. In the case of *E. faecalis*, biofilms have been shown to form on heart valves and cause endocarditis. As an important nosocomial pathogen, *E. faecalis* biofilms that form on medical devices are of the utmost importance in human health.

In microbial biofilms, cell-cell communication and coordinated multicellular interactions are of paramount importance. For example, quorum sensing and related forms of intercellular signaling play critical roles in biofilm
development and dispersal (4). Small molecule inhibitors of these signaling interactions show promise as compounds for drug development (8). In addition to communication by extracellular signals, the biofilm environment may contribute to microbial evolution by serving as an important niche for horizontal genetic transfer by transformation or conjugation (9). Moreover, expression of the gene products involved in these processes has been linked to enhanced biofilm formation (5). To determine the extent to which biofilm growth plays a role in transcriptome control, I decided to use a technique known as RNAseq. 

**RNAseq as a tool for examining the effects of the environment on bacterial gene expression**

Recently, RNAseq has emerged as a powerful new technique to examine the transcriptome of bacteria under different conditions. RNAseq is a process that involves high-throughput sequencing to analyze cDNA copied from cellular RNA. The technique allows measurement of total cellular RNA and comparison of transcriptome expression between differing conditions. Transcription profiles of eukaryotic cells including yeast and *Arabidopsis* cells, were published in 2008 (10, 11). As the technique became more popular, articles outlining the transcriptome of eukaryotic organisms became more abundant. In 2010, the first paper was published outlining the use of RNAseq to determine the primary transcriptome of a prokaryotic organism, *Helicobacter pylori* (12). Not only did the authors determine the transcriptome profile *H. pylori*, they also mapped
transcription start sites of these RNAs to determine the extent to which
antisense transcription was playing a role in gene expression (12).

As it becomes more accessible, RNAseq will be a powerful tool in the
analysis of gene expression profiles of organisms in the near future. Very
recently, a paper was published demonstrating the use of this technique to
determine gene expression in *P. aeruginosa* biofilms compared to planktonically
growing cells (13). With this in mind, I decided to investigate gene expression in
biofilm cells and compare it to planktonic cells. Examination of gene expression
in *E. faecalis* biofilms has been done on a smaller scale in the last ten years. Data
have been collected from promoter activation assays, transposon mutagenesis
studies and specific knockout strains to determine gene expression in biofilms
but comprehensive genomic analysis on *E. faecalis* biofilms has not been done
(14-17).

*The role of the biofilm environment on regulation of conjugation*

In spite of the plethora of research linking biofilm development to
intercellular signaling and genetic transfer, there is limited information about the
molecular aspects of cell-cell signaling in biofilms. The inherent heterogeneity of
the biofilm environment and extensive gene transcriptome changes may have
important effects on both signal production and response. Furthermore,
relatively little is known about the molecular interactions between signal
molecules and the biofilm matrix and the resulting effects of these interactions
on transmission of signals. With regard to gene transfer, transformation in
biofilms has been examined (3) and there is considerable overlap between
regulation of competence and biofilm development (18). Numerous studies have
described the regulation of conjugation in broth cultures, and it has been shown
that the conjugative F pilus promotes biofilm formation (5). Conjugative transfer
frequencies have also been measured in biofilms (19), but controlled studies of
specific effects of biofilm growth on conjugative gene transfer are limited.

In enterococci, the most efficient conjugation systems are plasmid-borne,
and their induction in donor cells is mediated by peptide sex pheromones
excreted by plasmid-free recipient cells. Our lab studies the E. faecalis pCF10
system, where the signaling molecule is the heptapeptide cCF10 (LVTLVFV)(20).
A detailed image describing the induction of conjugation in the pCF10 system is
shown in Figure 1.

Since production of cCF10 is encoded within the core genome, the
plasmid encodes functions to prevent self-induction of donor cells by
endogenous pheromone. These include a membrane protein PrgY which acts to
reduce pheromone production by donor cells (21) and the peptide iCF10
(AITLIFI), which acts as a competitive inhibitor of cCF10 (22). Both iCF10 and
cCF10 are processed by a protease, Eep, before being exported into the
extracellular environment (23). They are both then re-imported into donor cells
via the Opp system either alone or in conjunction with the peptide-specific
importer PrgZ (24) (Figure 1). The peptides bind to the same site on PrgX, the
master regulator controlling initiation of transcription of the conjugal transfer
genes. Although both peptides bind to the same site, they cause different structural changes in PrgX, ultimately stabilizing (iCF10) or destabilizing (cCF10) a tetramer transcription-repressing complex of PrgX and its target sites on pCF10, the primary and secondary binding sites (25) (Figure 1).

The direct effect of pheromone binding to PrgX is to alleviate repression of the prgQ operon, which encodes iCF10 and positive regulatory RNAs at its 5’end and the molecular machinery for conjugation in downstream regions, by increasing transcription initiation from the PQ promoter (Figure 1). Even in the absence of exogenous pheromone, prgQ transcription occurs, but it terminates upstream of the conjugation genes, producing a 380 nt RNA called QS (26).

Antisense interactions of nascent Q transcripts with a small, complementary RNA called Anti-Q, transcribed from a convergent promoter PX, shifts the secondary structure of Q to a terminator (27, 28). In uninduced cells there is adequate Anti-Q to block formation of any transcripts longer than QS, whereas in the presence of pheromone, the increased production of Q transcripts titrates all of the Anti-Q and leads to longer transcripts (QL). Some of these transcripts extend into the genes for conjugation such as prgB, whose product promotes donor-recipient aggregation (Figure 1).

Interestingly, the transcription unit encoding PrgX and Anti-Q is convergent with prgQ, with 220 bp of overlap. This unique arrangement means that several levels of reciprocal post-transcriptional regulation operate in this system. These mechanisms serve to amplify the primary effect of pheromone,
and recent studies suggested that these multiple layers of regulation confer properties of a bistable genetic switch to the system (29). From the perspective of a donor cell, this peptide signaling system may be viewed as a form of telesensing (30), where donor cells secrete a mixture of the two peptides, with iCF10 present in 50-100 fold excess (22), and monitor their relative concentrations over time. If the ratio changes in favor of cCF10, as is the case when recipients are in close proximity, the conjugation system is activated. It was previously shown that an *in vivo* biofilm, namely an endocarditis vegetation, could serve as niche for high-frequency transfer of pCF10 (31, 32). No previous studies examining the effects of biofilm growth on conjugation have been undertaken in *E. faecalis*.

I became interested in a systematic analysis of peptide signaling and plasmid transfer in biofilms using the pCF10 system as a model. To initiate these studies, I generated a pheromone-inducible fluorescent reporter construct and compared pheromone responses in planktonic versus biofilm cultures. A *gfp* reporter was transcriptionally fused in-frame to *prgB*, the gene encoding Aggregation Substance (Asc10) in the pCF10 conjugative plasmid (Figure 1). As Asc10 production facilitates conjugal transfer of pCF10, the expression of GFP is used as a reporter of conjugation readiness (Figures 1-2). Using this reporter strain, I found a remarkable difference in the population dynamics of the pheromone response under different growth conditions, consistent with the
hypothesis that biofilm growth results in formation of distinct cell types that impact the behavior and regulation of pCF10 transfer.

Following up on these studies and with the help of our mathematical modeling collaborators, I generated a hypothesis regarding the role of plasmid copy number in conjugation. We hypothesized that if the biofilm environment alters the copy number and/or distribution of pCF10 plasmids, this would have a significant impact on induction of conjugation.

In the upcoming chapters, I will describe three separate but related areas of study I undertook during my thesis research. In the first chapter, I will describe the development of a biofilm system in which *E. faecalis* cells are able to grow and form large biofilm communities on glass and polycarbonate surfaces. This system was used to generate biofilm and planktonic cells which were then tested for gene expression using RNAseq. The results of the RNAseq studies are described in the final part of the first chapter.

The second chapter of the thesis describes the induction of the pCF10 conjugation system in *E. faecalis* and the effects of growth in a biofilm on said system. Using an inducible GFP reporter of conjugation readiness, the pattern of conjugation induction is shown to be altered when cells are grown in a biofilm. Using a mathematical model of pCF10 conjugation, we developed a hypothesis that plasmid copy number was altered by growth in a biofilm. The final section of
this chapter describes analysis of plasmid copy number of pCF10 as well as 4 other enterococcal plasmids in biofilm or planktonic populations.

The third and final chapter of the thesis deals with the peptide signals of conjugation and their dual role in directing conjugation and distinguishing between self-/non-self population densities. In the case of cCF10, the peptide can turn on conjugation and allow cells to determine recipient density and, contrarily, iCF10 can both turn off conjugation and signal density of the donor cells. This type of dual role in quorum sensing and conjugation regulation has not previously been identified in other systems.
Figure 1: Induction of pCF10. Conjugation in pCF10 is induced by a peptide pheromone produced by recipient cells. As levels of cCF10 in the cell increase, it displaces an inhibitor peptide, iCF10, in the PrgX binding pocket allowing greatly increased transcription from the \( P_Q \) promoter and transcription of downstream genes involved in conjugation. The position of the \( gfp \) gene fusion used in later studies is shown in the bottom panel.
**Figure 2: Cloning of gfp into pCF10.** Using two-step PCR primers (Table 3) and pCJK41 (33), a plasmid was created containing the 3’ end of prgB and the 5’ end of prgC flanking a gfp gene fused to the PrgB ribosomal binding site. After a double crossover reaction, the pheromone inducible plasmid pCF10-GFP (pCF10-LC1) was created.
MATERIALS AND METHODS

Bacterial strains and media

All bacterial strains are listed in Table 1 and plasmids are listed in Table 2.

*Escherichia coli* strains were grown in Luria Bertani medium (Gibco). Todd Hewitt (Gibco) broth (THB) was used to make electrocompetent cells. When required, the following antibiotics were used to supplement the medium in overnight cultures of *E. coli*: 20 µg/mL chloramphenicol (Cl) and 100 µg/mL erythromycin (Erm). MM9YEG medium (referred to as M9) (34), a semi-defined M9-based medium supplemented with 0.3% yeast extract, 1% casamino acids, 20mM glucose, 1mM MgSO$_4$ and 0.1mM CaCl$_2$ was used for all experiments with *E. faecalis* unless otherwise stated. When required, the following antibiotics were used to supplement the medium in overnight cultures of *E. faecalis*: 10µg/mL tetracycline (Tet), 200 µg/mL rifampicin (Rif), 20 µg/mL chloramphenicol (Cl), 50 µg/mL erythromycin (Erm), 1000 µg/mL kanamycin (Kn), spectinomycin (Sp) and streptomycin (St). X-Gal was added at a concentration of 250 µg/mL for blue/white screening in *E. faecalis*. *E. faecalis* strains were grown at 37°C without shaking unless specified.
<table>
<thead>
<tr>
<th>Name</th>
<th>Relevant Characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Escherichia coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC1000</td>
<td><em>E. coli</em> cloning host, provides RepA in trans; Erm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>(35)</td>
</tr>
<tr>
<td>DH5α</td>
<td><em>lacZΔM15, recA1, endA1</em> mutations for yield and stability</td>
<td>Lab Stock</td>
</tr>
<tr>
<td><strong>E. faecalis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OG1</td>
<td>Reference strain</td>
<td>(36)</td>
</tr>
<tr>
<td>OG1S</td>
<td>OG1 St&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This work</td>
</tr>
<tr>
<td>OG1SSp</td>
<td>OG1 Sp&lt;sup&gt;r&lt;/sup&gt;, St&lt;sup&gt;r&lt;/sup&gt;</td>
<td>(37)</td>
</tr>
<tr>
<td>OG1RF</td>
<td>OG1 spontaneous mutant; Rif&lt;sup&gt;r&lt;/sup&gt;, Fusidic acid&lt;sup&gt;r&lt;/sup&gt;</td>
<td>(37)</td>
</tr>
<tr>
<td>OG1Sp</td>
<td>OG1 spontaneous mutant; Sp&lt;sup&gt;r&lt;/sup&gt;</td>
<td>(33)</td>
</tr>
<tr>
<td>CK111</td>
<td>OG1Sp <em>upp4::P&lt;sub&gt;23&lt;/sub&gt;repA4</em></td>
<td>(33)</td>
</tr>
<tr>
<td>100-5</td>
<td>OG1Sp contains one copy of <em>prgX</em> in the chromosome</td>
<td>(38)</td>
</tr>
<tr>
<td>JRC104</td>
<td>OG1RF ΔccfA</td>
<td>(33)</td>
</tr>
<tr>
<td>JRC106</td>
<td>OG1RF Δeep10</td>
<td>(33)</td>
</tr>
<tr>
<td>OG1S-tdtomato</td>
<td>OG1S contains P&lt;sub&gt;23&lt;/sub&gt;<em>tdtomato</em> inserted</td>
<td>This work</td>
</tr>
</tbody>
</table>
Table 2. Cloning vectors and plasmids

<table>
<thead>
<tr>
<th>Name</th>
<th>Relevant Characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBK1</td>
<td>prgQ::lacZ no lacZ regulation</td>
<td>(39)</td>
</tr>
<tr>
<td>pBK2</td>
<td>prgXQ::lacZ pheromone inducible lacZ reporter</td>
<td>(39)</td>
</tr>
<tr>
<td>pLC1</td>
<td>Cm', pBK1 ΔlacZ::gfpmut3b</td>
<td>(40)</td>
</tr>
<tr>
<td>pLC2</td>
<td>Cm', pBK2 ΔlacZ::gfpmut3b</td>
<td>(40)</td>
</tr>
<tr>
<td>pCJK47</td>
<td>Conjugative donor, Erm', oriT&lt;sub&gt;PCF10&lt;/sub&gt;, P-pheS*, requires RepA in trans for replication</td>
<td>(33)</td>
</tr>
<tr>
<td>pCJK47-LC1</td>
<td>pCJK47-5′ prgB::gfpmut3b::3′ prgC</td>
<td>(40)</td>
</tr>
<tr>
<td>pCJK141</td>
<td>Allelic exchange suicide vector for chromosomal insertions into OG1RF, Conjugative donor, Erm', oriT&lt;sub&gt;PCF10&lt;/sub&gt;, P-pheS*, requires RepA in trans for replication</td>
<td>C. Kristich, unpublished</td>
</tr>
<tr>
<td>pCF10</td>
<td>E. faecalis pheromone inducible conjugative plasmid, Tet'</td>
<td>(41)</td>
</tr>
<tr>
<td>pCF10-101</td>
<td>pCF10 ΔoriT2</td>
<td>(42)</td>
</tr>
<tr>
<td>pCF10-LC1 (pCF10-GFP)</td>
<td>pCF10 prgB RBS::gfpmut3b inserted between prgB and prgC</td>
<td>(40)</td>
</tr>
<tr>
<td>pMSP3535VAX</td>
<td>prgX insertion into pMSP3535VA, Kn'</td>
<td>(43)</td>
</tr>
<tr>
<td>p043lacZ</td>
<td>prgXQ::lacZ</td>
<td>(25)</td>
</tr>
<tr>
<td>pMSP043</td>
<td>prgXQ::lacZ</td>
<td>(44)</td>
</tr>
<tr>
<td>pINY8101</td>
<td>pWM402 prgB::lacZ</td>
<td>(45)</td>
</tr>
<tr>
<td>pJBA110</td>
<td>pUC18Not-P&lt;sub&gt;A1/04/03&lt;/sub&gt;-RBSII-gfp(LAA)-T&lt;sub&gt;0&lt;/sub&gt;-T&lt;sub&gt;1&lt;/sub&gt;</td>
<td>(46)</td>
</tr>
<tr>
<td>pJBA111</td>
<td>pUC18Not-P&lt;sub&gt;A1/04/03&lt;/sub&gt;-RBSII-gfp(LVA)-T&lt;sub&gt;0&lt;/sub&gt;-T&lt;sub&gt;1&lt;/sub&gt;</td>
<td>(46)</td>
</tr>
<tr>
<td>pJBA112</td>
<td>pUC18Not-P&lt;sub&gt;A1/04/03&lt;/sub&gt;-RBSII-gfp(AAV)-T&lt;sub&gt;0&lt;/sub&gt;-T&lt;sub&gt;1&lt;/sub&gt;</td>
<td>(46)</td>
</tr>
<tr>
<td>pJBA113</td>
<td>pUC18Not-P&lt;sub&gt;A1/04/03&lt;/sub&gt;-RBSII-gfp(ASV)-T&lt;sub&gt;0&lt;/sub&gt;-T&lt;sub&gt;1&lt;/sub&gt;</td>
<td>(46)</td>
</tr>
<tr>
<td>pBK2-ASV</td>
<td>pLC2 Δgfpmut3b::gfp(ASV)</td>
<td>This work</td>
</tr>
<tr>
<td>pBK2-LAA</td>
<td>pLC2 Δgfpmut3b::gfp(LAA)</td>
<td>This work</td>
</tr>
<tr>
<td>pBK2-LVA</td>
<td>pLC2 Δgfpmut3b::gfp(LVA)</td>
<td>This work</td>
</tr>
<tr>
<td>pBK2-AAV</td>
<td>pLC2 Δgfpmut3b::gfp(AAV)</td>
<td>This work</td>
</tr>
<tr>
<td>ptdTomato</td>
<td>Cloning vector containing tdTomato</td>
<td>Clontech</td>
</tr>
</tbody>
</table>
**Cloning and DNA manipulation**

Recombinant plasmids were propagated in *E. coli* DH5α cells prior to transformation into *E. faecalis* unless specified. Transformation was done via electroporation and competent cells were made by diluting overnight cultures 1:10 into 90mL THB and growing for ~2 hours until the OD\(_{600}\) was between 0.5-1.0. After 20 minutes of chilling on ice, the cells were pelleted and washed in ice cold 10% glycerol, pelleted again and resuspended in 500 µL of 25 µg/mL lysozyme in 10mM Tris pH 8.0, 20% sucrose, 10mM EDTA and 50mM NaCl. Cells were incubated at 37°C for 20 minutes and then pelleted and washed three times with 1 mL ice cold electroporation buffer (EB) (0.5M sucrose, 10% glycerol). Pellets were resuspended in 200 µL EB, split into 5 aliquots and frozen at -80°C for storage until use.

Plasmid DNA was purified using the Qiagen Spin Miniprep Kit (Qiagen Inc, Valencia, CA). Restriction digests and ligations were done according to accompanying protocols and enzymes were purchased from New England Biolabs (Ipswich, MA). Pfu Ultra II polymerase (Stratagene, Santa Clara, CA) was used for polymerase chain reaction (PCR) DNA amplification on a Techne thermocycler (Bibby Scientific US, Burlington, NJ).

**Plasmid and strain construction and induction**

*plLC1 and plLC2*

Plasmids plLC1 and plLC2 were created by PCR amplification of *gfpmut3b* using the *gfp* Forward BamHI and *gfp* Reverse EcoRI (primers listed in Table 3)
followed by restriction digest of the *gfp* gene and the pBK1 or pBK2 plasmid. Digestion of the plasmid with these enzymes cut out the *lacZ* gene. The plasmids were then gel purified using the Qiagen Gel Purification Kit (Qiagen Inc, Valencia, CA). The plasmids and *gfp* gene were ligated together to create pLC1 and pLC2 respectively.

**pCF10-LC1**

To create pCF10-LC1 (Figure 2), the *gfpmut3b* gene was inserted into the pCF10 plasmid using a double crossover method previously described (33). Briefly, approximately 800 base pairs of the downstream region of *prgB*, the ribosomal binding site of *prgB*, the *gfp* gene, and approximately 800 base pairs of the upstream region of *prgC* were joined, in that order, via a two-step PCR method (primers listed in Table 3). This construct was then inserted into the NotI site of the multiple cloning region of the pCJK47 plasmid creating pCJK47-LCC1. The plasmid was electroporated into *E. coli* strain EC1000 for propagation and then into *E. faecalis* strain CK111-pCF10-101. This strain was then conjugated with OG1RF recipient cells to ensure delivery of pCJK47-LCC1 into OG1RF cells. Transconjugants were selected based on antibiotic resistance and were then propagated without selection until colonies with the double crossover phenotype were selected for. These colonies were grown overnight, diluted 1:20 in fresh M9 medium and induced using 10 ng/mL cCF10 pheromone. Correct insertion of GFP gave rise to colonies that fluoresced green following a 60 minute induction as measured by a fluorometer using a blue optical kit.
(excitation/emission 490/510-570) (Turner Biosystems/Promega, Sunnyvale, CA).

The construct was then validated by sequencing.

**pBK2-LAA, LVA, AAV, and ASV**

To create plasmids pBK2-LAA, LVA, AAV and ASV, the *gfp* Forward *BamHI* primer and the relevant *gfp* Reverse primer (listed in Table 3 as *gfp* Reverse LAA, LVA, AAV, and ASV respectively) were used to amplify the *gfp* genes from pJBA110, pJBA111, pJBA112 and pJBA113. Following amplification, the tagged *gfp* genes were restriction digested with *BamHI* and *EcoRI* and ligated into similarly digested pBK2 as above with pLC2. To determine unstable GFP half-lives, overnight cultures grown in M9 + antibiotics were washed and diluted 1:10 into fresh M9 medium. After 60 minutes of growth, 10ng/mL cCF10 was added to the cultures and they were allowed to induce for 60 minutes. Cultures were then washed again and 10ng/mL iCF10 as added to stop induction and shutdown the expression of GFP (downshift). GFP fluorescence was measured using a fluorometer with a blue optical kit (Turner Biosystems/Promega, Sunnyvale, CA) at various times following downshift.

**OG1S-tdtomato**

pCJK141 was created by Dr. Christopher Kristich to make insertions into the OG1 chromosome between EF1117 and EF1116. P_{23}, a constitutive gram positive promoter, was inserted into pCJK141 at the NotI cloning site. Plasmid ptdTomato (Clonetech Laboratories, Mountain View, CA) was propagated in and purified from *E. coli* strain DH5α. Primers were designed with flanking restriction
sites and the *tdtomato* gene was excised from the plasmid and purified using the Qiagen Gel Purification Kit (Qiagen Inc, Valencia, CA). The *tdtomato* gene was fused to the $P_{23}$ promoter in pCJK141 via digestion and ligation. pCJK141- $P_{23}$-tdtomato was propagated in and purified from EC1000. The plasmid was then electroporated into CK111. A strain of OG1 was selected for streptomycin resistance to create a strain OG1S. OG1S was mated with CK111 to transfer the plasmid into OG1S. After homologous recombination, a double crossover was selected for such that $P_{23}$ fused to *tdtomato* was inserted into the OG1S chromosome. Insertion of the $P_{23}$-*tdtomato* fusion into the chromosome was verified by sequencing and fluorescence was measured using a fluorometer (Turner Biosystems/Promega, Sunnyvale, CA) with an optical kit with excitation/emission spectra of 525/580-640. Since the excitation/emission spectra of tdtomato are 554/581, this optical kit was chosen for measuring fluorescence (Clonetech Laboratories, Mountain View, CA).
## Table 3. Oligonucleotides

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>gfp Forward BamHI</td>
<td>GGGGATCCATATGCCTAAAGGAGAAGAATTTC</td>
</tr>
<tr>
<td>gfp Reverse EcoRI</td>
<td>CGGAATTCTTATATGCTAAGCTTTGTCGTTTGT</td>
</tr>
<tr>
<td>gfp Reverse LAA</td>
<td>TCTCGAATTCTATAGCTTATCGTACATGCAGCGTCGTTTGT</td>
</tr>
<tr>
<td>gfp Reverse LVA</td>
<td>TCTCGAATTCTAAGCTTTGTCGTTTGT</td>
</tr>
<tr>
<td>gfp Reverse AAV</td>
<td>TCTCGAATTCTATAGCTTATAGTTCA</td>
</tr>
<tr>
<td>tdtomato Forward</td>
<td>TGGGGCCGCGCGAGGCTTTTGTATAGTTCA</td>
</tr>
<tr>
<td>tdtomato Reverse</td>
<td>ATGGGGGCTATTTGACAGCTT</td>
</tr>
</tbody>
</table>

### Two-step cloning primers

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>prgB Forward</td>
<td>GTGGCCGCGCGCGATATTCAGAATCAT</td>
</tr>
<tr>
<td>prgB Reverse</td>
<td>ACCATTTGATGCTGAGG</td>
</tr>
<tr>
<td>prgB RBS Forward</td>
<td>CGTCAATGGTTGATGTGTCGTT</td>
</tr>
<tr>
<td>prgB RBS Reverse</td>
<td>TCTCAGTTAGTGGTCGTTG</td>
</tr>
<tr>
<td>gfp mut3 Forward</td>
<td>TAAAAATCGAGGAGAATGATACATGAATCAAATGCGT</td>
</tr>
<tr>
<td>gfp mut3 Reverse</td>
<td>CTCTTGAATGTTTTCTTCGGTAATTATTTGTATGTTCA</td>
</tr>
<tr>
<td>prgC Forward</td>
<td>ATTACAAGCCGAGATGGAAGATCAAAACAGCG</td>
</tr>
<tr>
<td>prgC Reverse</td>
<td>GTAGGCAGCCCGCTGATTAGTTG</td>
</tr>
</tbody>
</table>

### qPCR primers

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse primer</th>
<th>Amplicon length</th>
</tr>
</thead>
<tbody>
<tr>
<td>prgX</td>
<td>GCTTTGGCTCAATCCCTTT</td>
<td>CTGACTTCGACCAGATGTT</td>
<td>198 bp</td>
</tr>
<tr>
<td>Qc</td>
<td>CATGTATATGTTCCTCCCGCTTTT</td>
<td>CGGCTCTTACGAGTTGTTCA</td>
<td>156 bp</td>
</tr>
<tr>
<td>gyrB</td>
<td>ACCAACACCCTGCAAGCC</td>
<td>CAAGCCAAACACAGTCGCC</td>
<td>111 bp</td>
</tr>
</tbody>
</table>
**Genomic DNA (gDNA) preparation**

DNA preparations were completed using the Qiagen DNeasy Blood and Tissue Kit (Qiagen Inc, Valencia, CA) following the procedure for extraction of gDNA from gram positive cells. Nucleic acid quantification was done using either a Nanodrop (Thermo Scientific, Wilmington, DE) or, for smaller amounts of DNA, the Quant-iT PicoGreen dsDNA kit (Invitrogen, Carlsbad, CA). Using this kit, samples were mixed in black 96 well plates and read on a fluorometer (Turner Biosystems).

If frozen prior to analysis, the samples were immediately pelleted and cell pellets were frozen no more than 16 hours before gDNA preparation. To determine the copy number of pCF10 in planktonic and biofilm populations, gDNA was obtained from 24 hour planktonic cultures, 24 hour biofilms on polycarbonate coupons, sorted cells, and a reference strain, 100-5 containing one copy of prgX in the chromosome.

To determine the amount of DNA/colony forming unit (CFU), cultures of OG1RF+pCF10 were grown overnight in M9 medium at 37°C. The overnight culture was then diluted into the CBR and grown as a biofilm (see biofilm growth). The overnight culture was also diluted 1:10^5 into a new conical tube with 15 mL fresh M9 medium. This tube was grown overnight at 37°C. The next day, the coupon biofilms were each vortexed in 15 mL conical tubes with 3 mL PBS + 2mM EDTA to remove adherent cells. 15mL of the overnight cultures and 15mL of the planktonic liquid from the CBR were also collected. Liquid from the
overnight culture (labeled planktonic liquid), the liquid portion of the CBR (planktonic CBR), and vortexed cells from the biofilm (biofilm), were all spun down to pellet the cells. The pellets were then resuspended in the original volume of medium. One milliliter of each culture was removed and diluted to determine total CFUs/mL. Another 1 mL of each culture was diluted 1:10 and 1:100. The DNeasy Tissue kit (Qiagen, Valencia, CA) was then used to extract total gDNA from 1 mL of each culture (undiluted) as well as 1 mL of each dilution for each of the three samples. The extracted gDNA was then quantified using the Quant-iT PicoGreen dsDNA kit (Invitrogen, Carlsbad, CA). Three dilutions of each sample were compared with the plate counts to estimate total gDNA extracted per CFU for three different cell concentrations. These numbers were averaged to attain the final concentration of gDNA extracted per CFU.

**RNA extraction and cDNA preparation**

Cells were spun down and resuspended in 500 µL PBS + 2mM EDTA and added to 1 mL RNA Protect (Qiagen). After a 5 minute incubation at room temperature, the cells were spun down and either frozen at -80°C or immediately treated to enzymatic lysis (14). Following lysis, RNA was extracted as described using the Qiagen RNeasy Mini Kit. Two to ten micrograms of total RNA was subjected to Turbo DNase treatment using the rigorous method as described (Ambion). RNA was stored at -80°C. cDNA was prepared using 100-500 ng of total RNA as per the manufacturer protocol using Superscript III first strand
synthesis system kit (Invitrogen) using random hexamers. cDNA was diluted 1:5 in sterile water prior to qPCR.

Quantitative real-time PCR (qRT-PCR)

qRT-PCR was done on gDNA or cDNA using the SYBR green Supermix and an iCycler iQ5 (Bio-Rad) instrument. Each well contained 25µL of total reaction with 200nM primer concentration. Primer efficiencies and C_t measurements were carried out by iQ5 Optical System Software Version 2.0 (2006, Bio-Rad). All experiments met the following criteria for statistical analysis; efficiencies ≥80%, calibration curve r^2 values of ≥0.98, and C_t values for no template control samples ≥35 cycles. Data were analyzed using the Pfaffl method for relative quantification (47). Technical replicates showed no intra-assay variation and were not used in the final statistical analysis.

Biofilm growth and induction

Biofilms were grown using the CDC Biofilm Reactor (CBR) (Bio Surface Technologies Corp., Bozeman, MT). Experiments were done at 37° C except vortexing which was done at 4° C. M9 medium was used as the growth medium for overnight cultures as well as for the batch culture medium. 10% M9 was used in the carboy for the continuous flow portion of growth. The CBR protocol followed the instruction manual guidelines with few minor changes. Briefly, after autoclaving, the reactor and carboy setup were placed in a 37°C room with the reactor on a stir plate rotating at a speed of approximately 125 rpm. Two milliliters of overnight culture at a cell density of approximately 2x10^9 CFU/mL
were added to the reactor aseptically. The reactor was incubated as a batch culture for 4 hours to allow bacterial attachment to the coupons. At this point, the coupons were either sampled or subjected to a continuous flow of medium at a rate of 8 mL/min for 20 hours, enough to wash out growing planktonic cells. To retrieve the coupons, one polypropylene rod was removed from the reactor, and the three coupons were unscrewed and placed in one well of a 6 well dish. Five milliliters of sterile water were then added to each well, and the plate was placed onto a rotating platform for three minutes to remove planktonic cells.

Coupons were moved to a 24 well plate. Pheromone cCF10 was diluted to various concentrations in 50% M9. One milliliter of the pheromone mixture was added to each well to completely submerge the coupon, and the plate was placed on a rocking platform at 37°C for 0-120 minutes (60 minutes unless otherwise stated). Following induction, coupons were removed to 15 mL conical tubes containing 3 mL of PBS + 2mM EDTA and vortexed for 3 minutes at 4°C to remove attached biofilm cells and break up cell clumps. Preliminary data showed that vortexing for 3 minutes at 4°C gave the highest cell yield and lowest variability compared to sonication and vortexing at room temperature. Quantification of cells was done by making dilutions into PBS + 2mM EDTA and plating dilutions onto BHI plates with appropriate antibiotics and incubating overnight at 37°C.
Serial biofilm growth

Biofilms were grown as above, and all cells were collected after 24 hours. One milliliter of the planktonic cells was used to measure CFU/mL. One coupon was vortexed in 3 mL PBS + 2mM EDTA for 3 minutes. After vortexing, the coupon was removed from the liquid and replaced by a new biofilm coupon. This process was repeated twice to have a total of three coupons worth of cells in 3 mL PBS + 2mM EDTA. Cell counts were done on this liquid and divided by three to determine the approximate CFU/coupon of biofilm cells. The remaining coupons were vortexed as above. All collected biofilm cells were spun down and resuspended in M9 to a final OD$_{600}$ of 0.75. One milliliter of this culture was then added to a new reactor. This process was repeated for 4 days for a total of five days of biofilm growth.

RNASeq

To purify RNA from biofilm and planktonic cells for RNASeq, biofilms were grown in the CBR as described above. After 4 hours of batch growth, the biofilm cells were removed from the coupon surface by vortexing (as above), and 10 mL of the surrounding planktonic fluid were also removed from the reactor and used for planktonic cell RNA. The biofilm and planktonic cells were pelleted and washed once in 1 mL PBS + 2mM EDTA. RNA was extracted and treated with DNase as above. Following DNase treatment, the RNA was treated using the MICROBExpress kit (Ambion, Foster City, CA) as per manufacturer’s instructions to purify mRNA and deplete 16S and 23S rRNA. Library creation and RNASeq was
performed by the core facility at the University of Minnesota. RNAseq data analysis was performed by a collaborator (Suzanne Grindle).

RNAseq quality control analysis included trimming of the sequences. This was followed by mapping the sequences to the OG1RF reference genome using bowtie (48). Differential gene expression was determined using the Cufflinks to measure the relative abundance of transcripts in fragments per kilobase of exon per million fragments mapped (FPKM). Cuffdiff was used to take cufflinks expression levels and test them for significant differences. Genes were listed as differentially expressed if the p value ≤ 0.01.

**Conjugation experiments**

Creation of certain strains required conjugative mating between strains of *E. faecalis*. Matings were done on Brain Heart Infusion (BHI) agar plates (solid mating). Donor and recipient strains were grown to late exponential phase (approximately 2x10⁹ CFU/mL) in BHI and then diluted 1:10 into fresh BHI medium. After 90 minutes of growth at 37°C, the populations were mixed at a donor:recipient ratio of 1:9. Eight hundred microliters of this mixture was spun down and plated onto BHI agar plates without antibiotic selection. Plates were incubated at 37°C for 16 hours. Cells were removed from the plate by addition of 1 mL of PBS + 2mM EDTA followed by scraping. EDTA is used in these and other experiments to break clumps of cells that arise after mating pair formation into single cells. Serial dilutions were plated onto agar medium selective for
donors, recipients, or tranconjugants and incubated for 24 hours at 37°C before counting of colonies.

For measurement of mating transfer efficiencies of biofilm and planktonic cells, overnight cultures of donors and recipients were diluted 1:10 into fresh BHI medium and grown for 90 minutes at 37°C. The cells were then mixed at a donor:recipient ratio of 1:9, and 1 mL of the mixture was added to each well of a 6 well dish. Autoclaved circular aclar membranes (Honeywell Inc., Morristown, NJ) of approximately 1 cm in diameter were added to each well, providing a surface for biofilm formation. The 6 well dish was placed on a shaker at 37°C for 4 hours. Cells remaining in the liquid phase were left in the plate. The aclar membranes were then removed, rinsed twice in sterile water, and place into a new 6 well plate with 1 mL of fresh sterile BHI medium. Both 6 well dishes were then placed back on the shaker at 37°C for 20 hours. Following 24 hours of mating, the liquid cells were collected from the original 6 well dish. Aclar membranes were removed from the remaining plate and rinsed twice in sterile water. The membranes were then place in 2 mL microcentrifuge tubes with 1 mL of PBS + 2mM EDTA and vortexed for 3-5 minutes at 4°C to remove biofilm cells. Biofilm and planktonic cells were then spun down and washed twice in 1 mL PBS + 2mM EDTA. Donors, recipients, and tranconjugants were enumerated as described above.

For measurement of mating efficiencies in donor density experiments, cultures of donor (OG1RF+pCF10) and recipient (OG1SSp) cells were grown
overnight at 37°C in 10mL of M9 medium with antibiotics. Overnight cultures were spun down and washed once in 1mL PBS + 2mM EDTA. The cells were then suspended in 10 mL fresh M9 medium. Donor cells were diluted 1:1, 1:10, or 1:100 in 4.5 mL M9 and 500 µL of recipients were added to each donor dilution. The mating tubes were placed at 37°C. After 30 minutes of incubation, 100 µL was removed from each tube and placed at 4°C. The remaining cultures were placed back at 37°C to continue incubating. One hundred microliter samples were removed at each of the designated time points. After all of the samples were collected, dilutions of the 100 µL samples were plated on donor, recipient, or transconjugant specific agar plates and incubated overnight at 37°C before enumeration. The data shown are averages based on ≥4 biological replicates.

**Flow cytometry and fluorescent activated cell sorting (FACS)**

Cell suspensions were analyzed via flow cytometry using a FACSCalibur flow cytometer (BD Biosciences, Rockville, MD). Cells were analyzed by size and granularity, and dead cells and debris were gated out of the population. Cells did not require filtering, and the vast majority of cells fell into the same region on the forward and side scatter plot indicating that the majority of cell clumps were broken up by EDTA treatment. Expression of GFP was measured for 100,000-200,000 live cells using the 488 nm laser excitation line. Data from flow cytometry experiments are shown as histograms and contour plots. Data were analyzed using FlowJo (Tree Star, Ashland, OR). FACS was performed using a
FACSAria II and cells were gated, analyzed, and sorted using FACSDiva software (BD Biosciences, Rockville, MD).

**Statistical analysis**

qPCR data are based on n≥3 biological replicates. In each experiment, ≥2 dilutions were used for each sample to compare to ≥4 dilutions of the control 100-5 DNA. All samples were run in triplicate, and any wells which had standard deviation of ≥0.3 from the other two samples were not used. If removal of one well did not decrease the standard deviation below 0.3, the dilution was not used. Statistical significance was determined by a paired two-sample t test (Excel 2010, Microsoft, Redmond, WA). P values were considered statistically significant if they were ≤0.01. Exact P values are shown in figure legends.

**Microscopy**

Coupons with GFP-expressing *E. faecalis* cells were counterstained with a red fluorescent Alexa Fluor 594: Wheat Germ Agglutinin conjugate (Invitrogen) labeling the cell envelope and mounted in VECTASHIELD (Vector Laboratories, Burlingame, CA) immediately prior to image acquisition. Images were acquired with a Cascade 1K EMCCD camera (Photometrics, Tucson, AZ) as widefield z-stacks with a 60× 1.4 NA objective (Nikon Instruments, Melville, NY). Z-stacks were taken at 0.2 µm intervals, and were deconvolved using Huygens Pro (Scientific Volume Imaging, The Netherlands). Processed images were aligned and projected as a best-focus composites using MetaMorph (Molecular Devices, Sunnyvale, CA).
**Pulse field gel electrophoresis**

Pulse field gel electrophoresis (PFGE) was done as previously described with minor alterations (49). Biofilm and planktonic cells from an overnight CBR experiment were grown and harvested as above. Cells were then washed twice and resuspended in 1 mL PBS + 2mM EDTA and chilled on ice. The cells were diluted to final OD$_{600}$ values of 0.86 for biofilm cells and 0.43 for planktonic cells using sterile TE Buffer (10 mM Tris-Cl, pH 8.0, 1 mM EDTA). Plate counts showed that the CFU/mL values at these empirically derived ODs were similar. The gel standard used in this experiment was *Salmonella enterica* serotype Braenderup H9812 (ATCC BAA-664) (provided by the Minnesota Department of Health); *XhoI* was the primary restriction enzyme. The gel was run on a CHEF-DR III System (Bio-Rad, Hercules, CA) at an initial switch time of 2.0 seconds and a final switch time of 10.0 seconds for 13 hours followed by an initial switch time of 20.0 seconds and final switch time of 25.0 seconds for 6 hours. A Gel Doc XR System with Quantity One software (Bio-Rad, Hercules, CA) was used to capture and convert the gel images.

**iCF10 donor density and de-induction cell preparation**

OG1RF cells containing pCF10 were grown overnight in M9 medium with antibiotics at 37°C. Overnight cultures were spun down and washed once in 1 mL PBS + 2mM EDTA. Cells were resuspended in the original volume of medium and diluted 1:10 into tubes with 4.5 mL M9 medium. The samples were treated as described below prior to RNA extraction.
For donor density experiments, cells were grown for 1 hour at 37°C, and cCF10 was added to the tubes. Four tubes each received no cCF10, 0.1ng/mL, 1ng/mL or 10ng/mL cCF10. The tubes were placed by at 37°C on a shaking platform. After 15 minutes, one tube containing each cCF10 concentration was removed from the incubator, spun down, and washed once in 1mL PBS + 2mM EDTA. This process was repeated at 30, 60, and 120 minutes.

For de-induction experiments, a T=0 sample was taken before 10ng/mL cCF10 was added to the remaining tubes which were then placed at 37°C on a shaking platform for 60 minutes. The five tubes were then removed from the incubator and one tube was labeled as T=60 0i (for 0ng/mL iCF10). These cells were spun down and treated with RNAProtect prior to RNA extraction. The remaining tubes were labeled 120m 0i, 1i, 10i, and 100i. iCF10 was added to the corresponding tubes at 0 ng/mL, 1 ng/mL, 10 ng/mL, or 100 ng/mL. The four remaining tubes were then placed back on the shaking platform at 37°C on a shaking platform for an additional 60 minutes prior to RNA extraction.

**Mathematical Modeling**

A set of ordinary differential equations based on pCF10 genetic network were developed (Figure 1). Steady state and dynamic simulations were performed in MATLAB (version 2008a MathWorks, Natick, MA). Steady state solutions of non-linear algebraic equations were obtained using MATLAB solve function. ODEs were integrated using function *ode23s* in MATLAB to obtain
dynamic behavior. Mathematical modeling parameters and figure were
generated by Drs. Wei-Shou Hu and Anushree Chatterjee.
RESULTS

I. Development of a reproducible system to examine differential gene expression in *E. faecalis* biofilms

*E. faecalis* biofilms were actively growing and no rapid mutagenic changes selected for biofilm growth

I carried out initial experiments to assess the growth of planktonic and biofilm populations grown in the CDC biofilm reactor (CBR) system. In the 4-8 hour time frame, the population densities of both the biofilm and planktonic cells increased to similar extents and then leveled off (Figure 3). Determination of precise generation times for the two populations in this experiment was difficult because planktonic cells were continually diluted after the initial 4 hours of static incubation, and because both growth and adherence could contribute to the population of biofilm cells. Nonetheless, the parallel increases in both populations suggested that they were both actively growing at similar rates, especially for the first 8-10 hours of the experiments. In this and future experiments, planktonic cells refer to those cells growing in the liquid phase of the CBR system (unless otherwise stated) although similar results were obtained with overnight planktonic cultures.
Figure 3: Growth of biofilm and planktonic cells in the CDC Biofilm Reactor.

After 4 hours of batch culture growth, the reactor was operated as a continuous flow system at a rate of 8 mL/min (arrow). The viable populations of biofilm and planktonic bacteria were enumerated by plate counts as described in Materials and Methods. Error bars show standard deviation using n≥3 biological replicates.
To determine whether selection could increase biofilm yield, I performed serial biofilm experiments. In these experiments, cells were grown in a biofilm for 24 hours, removed, washed, and then used to re-inoculate a fresh biofilm. This process was repeated up to 4 additional days. After each day of growth, cells were enumerated on a per coupon basis. Figure 4 shows biofilm and planktonic cell yield from each of the five days. Biofilm cell yield hovered between $10^7$ and $10^8$ CFU/coupon with no significant increase over the course of five days. This suggests that there was no significant selection for variants with increased biofilm growth during the 24 hour time course of the experiment. Changes in gene expression that occur during early stages of biofilm formation and growth are of great interest, and we used the system described above to obtain early biofilm cells for transcriptional analysis.
Figure 4. Serial transfer and regrowth of biofilm cells for five days does not significantly increase biofilm cell yield. The light gray line shows planktonic cell counts/mL and the darker line represents average CFU/biofilm coupon. Bars represent standard deviation for n≥3 biological replicates.
Whole transcriptome RNA sequencing showed differential gene expression between biofilm and planktonic cells

RNAseq is emerging as a powerful tool to study the transcriptome of microbial cells under different conditions. We used this technique to analyze the *E. faecalis* transcriptome during the early stages of biofilm growth. We determined the transcriptome using cells growing in a biofilm for 4 hours and compared them to cells growing planktonically in the same CBR. I identified 24 genes down-regulated and 127 genes up-regulated during biofilm growth compared to expression in planktonic growth. Down-regulated genes are listed in Table 4, and up-regulated genes are found in Table 5. The groups of genes that were most often found to be up-regulated in biofilms were ABC transporters, DNA replication and repair enzymes, hypothetical genes, and genes involved in metabolism (Figure 5).

Fifty-one incidences occurred in which genes identified as up-regulated in biofilms according to our RNAseq data had also been previously identified in published and unpublished genetic screens including mutagenesis, transposon studies and promoter trap assays (Table 5). This served as a verification of our data as well as a way to conclude which genes might be strongly associated with biofilm formation. Genes that were very highly up-regulated in RNAseq and also previously found in other studies include those encoding the enterococcal pilus *ebpA-C*, transcriptional repressor *copY*, *recA*, *oppA*, the quorum sensing system *fsrBC*, and dihydroorotase *pyrC*, among others.
Besides validating previous studies, analysis of the RNAseq data provides insight into the activity of many genes not previously identified to be involved in biofilm regulation. For example, multiple genes in the oligopeptide permease (Opp) system, involved in the transport of peptides, were found to be highly up-regulated in biofilms. As Opp proteins are involved in the transport of both iCF10 and cCF10, the upregulation of Opps could have a significant impact on the pheromone response system (Figure 1). In *Bacillus subtilis*, Opp proteins transport peptides involved in competence development and sporulation (50). The up-regulation of Opp proteins could have important implications in the signaling of cells in biofilms and should be considered an important area of study.

Genes involved in the metabolism and transport of metals such as copper, manganese, and lead were also found to be highly up-regulated in biofilms. It is possible that these genes are up-regulated by cells located in portions of the biofilms that are starved for nutrients. More research is required to determine whether up-regulation of metal metabolism genes is a universal phenomenon or relegated to certain biofilm subpopulations.

Taking a closer look at the metabolism genes, many different pathways were identified to be either up or down-regulated by growth in a biofilm. The pyrimidine metabolism pathway appears to be highly up-regulated with nine genes identified as being differentially expressed (EF1712-1719, EF2396, and EF2738 (Table 5)). Pyrimidine metabolism has been previously identified as being
up-regulated in *Staphylococcus aureus* and *P. aeruginosa* biofilms as well, but the reason for this has not been determined (51, 52). While the EF1712-1719 genes are all up-regulated in response to growth in a biofilm, EF2396 and EF2738 are down-regulated. It is possible that this apparent down-regulation of some RNA transcripts is the result of anti-sense RNA. The presence of anti-sense RNA in the RNAseq data will be addressed in future studies to determine strand specificity of the transcriptome.

Another pathway found to be up-regulated in biofilm growth is involved in pyruvate metabolism. Again this pathway has been previously identified as being up-regulated in *S. aureus* and *P. aeruginosa* biofilms (53, 54). In the case of *E. faecalis*, EF1140 (lactoylglutathione lyase gloA), EF1355 (pyruvate dehydrogenase aceF), and EF1356 (dihydrolipoamide dehydrogenase lpdA) were highly up-regulated in biofilms. These genes are responsible for the metabolism of pyruvate to acetyl-CoA. EF1613, a formate acetyltransferase, *pflB*, was shown to be down-regulated in biofilms. *PflB* plays a direct role in the conversion of pyruvate to formate. EF1612, the pyruvate formate-lyase activating enzyme *pflA*, responsible for activating *pflB* is also down-regulated in biofilms. This is surprising as the *pflB* homologue was found to be up-regulated in *S. aureus* biofilms and important for glucose consumption and formate production in these cells (53, 55). EF2158, a pyruvate-flavodoxin oxidoreductase was also found to be down-regulated. Further studies will be necessary to characterize whether these genes are truly down-regulated or if these are antisense
transcripts. It is obvious from these data that the metabolism of both pyrimidines and pyruvate play an important role the biofilm environment.

One gene category found to be highly up-regulated in biofilms was that of ABC transporters with over 13 transporters up-regulated in biofilms and two transporters down-regulated. Table 6 lists the types of ABC transporters identified in the RNAseq analysis along with their putative transport substrates. Among these transporters, three had been previously identified in our lab in both in vitro and in vivo recombinase based in vivo technology (RIVET) screens (14). The identified ABC transporters point to an increased role in nutrient sensing and acquisition in biofilm cells. Because nutrients may be limiting in biofilms, it would make sense for the bacteria to increase their ability to import substrates such as amino acids, metal ions, and sugars (Table 6).
<table>
<thead>
<tr>
<th>EF number</th>
<th>Description</th>
<th>Planktonic</th>
<th>Biofilm</th>
<th>p value</th>
<th>Previously Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF0104</td>
<td>Arginine deiminase, arcA</td>
<td>99.976</td>
<td>32.607</td>
<td>0.00192905</td>
<td></td>
</tr>
<tr>
<td>EF0108</td>
<td>Arginine/ornithine antiporter ArcD</td>
<td>42.702</td>
<td>13.824</td>
<td>0.00424342</td>
<td></td>
</tr>
<tr>
<td>EF0262</td>
<td>S1 RNA binding domain protein</td>
<td>91.504</td>
<td>278.26</td>
<td>0.00136608</td>
<td></td>
</tr>
<tr>
<td>EF0452</td>
<td>AMP-binding family protein</td>
<td>29.775</td>
<td>8.606</td>
<td>0.00560481</td>
<td>KF RIVET, (14)</td>
</tr>
<tr>
<td>EF0633</td>
<td>Tyrosyl-tRNA synthetase , tryS-1</td>
<td>105.62</td>
<td>36.469</td>
<td>0.00298658</td>
<td></td>
</tr>
<tr>
<td>EF0798</td>
<td>hypothetical protein</td>
<td>94.839</td>
<td>28.242</td>
<td>0.00321544</td>
<td>(14, 56)</td>
</tr>
<tr>
<td>EF0978</td>
<td>Methenyl-tetrahydrofolate cyclohydrolase, folD</td>
<td>16.273</td>
<td>70.516</td>
<td>0.00113086</td>
<td></td>
</tr>
<tr>
<td>EF1495</td>
<td>V-type ATPase, subunit E</td>
<td>171.41</td>
<td>51.422</td>
<td>0.00141929</td>
<td></td>
</tr>
<tr>
<td>EF1496</td>
<td>V-type ATP synthase subunit C</td>
<td>162.08</td>
<td>51.77</td>
<td>0.00160776</td>
<td></td>
</tr>
<tr>
<td>EF1498</td>
<td>V-type ATP synthase subunit A</td>
<td>146.76</td>
<td>54.211</td>
<td>0.0099222</td>
<td></td>
</tr>
<tr>
<td>EF1510</td>
<td>Protein of unknown function DUF59</td>
<td>155.47</td>
<td>471.35</td>
<td>0.00140751</td>
<td></td>
</tr>
<tr>
<td>EF1533</td>
<td>hypothetical protein</td>
<td>351.71</td>
<td>111.12</td>
<td>0.0017833</td>
<td></td>
</tr>
<tr>
<td>EF1612</td>
<td>Pyruvate formate-lyase activating enzyme, pflA</td>
<td>266.18</td>
<td>81.106</td>
<td>0.00108234</td>
<td></td>
</tr>
<tr>
<td>EF1613</td>
<td>Pyruvate formate-lyase, pflB</td>
<td>490.52</td>
<td>147.4</td>
<td>0.00511915</td>
<td></td>
</tr>
<tr>
<td>EF1752</td>
<td>hypothetical protein</td>
<td>898.72</td>
<td>310.05</td>
<td>0.00337257</td>
<td></td>
</tr>
<tr>
<td>EF2050</td>
<td>ABC transporter, ATP-binding protein</td>
<td>200.88</td>
<td>75.832</td>
<td>0.00634159</td>
<td>(14)</td>
</tr>
<tr>
<td>EF2158</td>
<td>Pyruvate-flavodoxin oxidoreductase</td>
<td>31.531</td>
<td>11.282</td>
<td>0.00568888</td>
<td></td>
</tr>
<tr>
<td>EF2396</td>
<td>Uridylate kinase, pyrH</td>
<td>73.088</td>
<td>227.34</td>
<td>0.00121922</td>
<td></td>
</tr>
<tr>
<td>EF2547</td>
<td>hypothetical protein</td>
<td>128.07</td>
<td>837.29</td>
<td>0.00111879</td>
<td></td>
</tr>
<tr>
<td>EF2642</td>
<td>Glycine betaine ABC transport system, OpuAC</td>
<td>85.474</td>
<td>27.989</td>
<td>0.00198982</td>
<td></td>
</tr>
<tr>
<td>EF2738</td>
<td>Alkyl hydroperoxide reductase protein F, thioredoxin reductase</td>
<td>142.45</td>
<td>52.487</td>
<td>0.00898846</td>
<td></td>
</tr>
<tr>
<td>EF3062</td>
<td>Rod shape-determining protein mreC</td>
<td>55.502</td>
<td>170.09</td>
<td>0.00134989</td>
<td></td>
</tr>
<tr>
<td>NEW7</td>
<td>tRNA pseudouridine synthase A</td>
<td>24.257</td>
<td>8.9404</td>
<td>0.00734664</td>
<td></td>
</tr>
<tr>
<td>EF Number</td>
<td>Description</td>
<td>Planktonic</td>
<td>Biofilm</td>
<td>p value</td>
<td>Previously identified</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>EF0004</td>
<td>DNA recombination and repair protein RecF</td>
<td>54.553</td>
<td>139.75</td>
<td>0.008472 02</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0005</td>
<td>DNA gyrase subunit B</td>
<td>65.608</td>
<td>202.4</td>
<td>0.005397 72</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0063</td>
<td>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein oppA</td>
<td>9.7143</td>
<td>60.345</td>
<td>1.15E-05</td>
<td></td>
</tr>
<tr>
<td>EF0066</td>
<td>Holliday junction DNA helicase RuvA</td>
<td>18.243</td>
<td>80.546</td>
<td>0.002674 88</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0067</td>
<td>Holliday junction DNA helicase RuvB</td>
<td>39.806</td>
<td>100.13</td>
<td>0.006865 59</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0297</td>
<td>transcriptional repressor CopY</td>
<td>0</td>
<td>160.94</td>
<td>7.08E-06</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0298</td>
<td>Copper-translocating P-type ATPase</td>
<td>16.391</td>
<td>275.16</td>
<td>3.39E-13</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0368</td>
<td>Aspartokinase</td>
<td>18.813</td>
<td>48.096</td>
<td>0.007019 53</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0675</td>
<td>L-proline glycine betaine binding ABC transporter protein proX/ Osmotic adaptation</td>
<td>21.06</td>
<td>61.284</td>
<td>0.002135 91</td>
<td>KF RIVET (14)</td>
</tr>
<tr>
<td>EF0682</td>
<td>DNA repair exonuclease family protein YhaO</td>
<td>22.982</td>
<td>64.61</td>
<td>0.002948 21</td>
<td></td>
</tr>
<tr>
<td>EF0701</td>
<td>Peptide chain release factor 3, prfC</td>
<td>16.706</td>
<td>49.115</td>
<td>0.002080 7</td>
<td></td>
</tr>
<tr>
<td>EF0704</td>
<td>lipoprotein, putative</td>
<td>8.6872</td>
<td>42.38</td>
<td>0.002835 62</td>
<td></td>
</tr>
<tr>
<td>EF0758</td>
<td>Lead, cadmium, zinc and mercury transporting ATPase</td>
<td>40.674</td>
<td>239.21</td>
<td>6.52E-06</td>
<td></td>
</tr>
<tr>
<td>EF0759</td>
<td>sapB protein, putative</td>
<td>55.175</td>
<td>440.5</td>
<td>2.87E-08</td>
<td></td>
</tr>
<tr>
<td>EF0807</td>
<td>Oligopeptide ABC transporter, oppA</td>
<td>32.715</td>
<td>151.14</td>
<td>2.71E-05</td>
<td></td>
</tr>
<tr>
<td>EF0809</td>
<td>hypothetical protein</td>
<td>200.58</td>
<td>602.62</td>
<td>0.002706 46</td>
<td></td>
</tr>
<tr>
<td>EF0822</td>
<td>hydrolase, haloacid dehalogenase-like family</td>
<td>29.388</td>
<td>79.276</td>
<td>0.009643 8</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF0858</td>
<td>phage infection protein, pip</td>
<td>28.378</td>
<td>68.776</td>
<td>0.002494 74</td>
<td></td>
</tr>
<tr>
<td>EF0940</td>
<td>Conserved hypothetical protein</td>
<td>39.623</td>
<td>124.01</td>
<td>0.002740 91</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF1030</td>
<td>endonuclease/exonuclease/phosphatase family protein</td>
<td>15.909</td>
<td>47.763</td>
<td>0.005273 27</td>
<td></td>
</tr>
<tr>
<td>EF1033</td>
<td>6-aminohexanoate-cyclic-dimer hydrolase, putative</td>
<td>10.626</td>
<td>50.122</td>
<td>2.16E-05</td>
<td></td>
</tr>
<tr>
<td>EF1072</td>
<td>galR, galactose operon repressor, LacI family transcriptional regulator</td>
<td>6.5923</td>
<td>45.02</td>
<td>0.004501 27</td>
<td></td>
</tr>
<tr>
<td>EF1084</td>
<td>Universal stress protein family</td>
<td>18.135</td>
<td>136.7</td>
<td>0.002752 66</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>E-value</td>
<td>Blast score</td>
<td>Probability</td>
<td>RIVET Status</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>EF1085</td>
<td>hypothetical protein</td>
<td>0</td>
<td>235.82</td>
<td>0.00097289</td>
<td></td>
</tr>
<tr>
<td>EF1091</td>
<td>ebpA endocarditis and biofilm-associated pili A</td>
<td>10.131</td>
<td>104.74</td>
<td>7.51E-10</td>
<td>(14-16)</td>
</tr>
<tr>
<td>EF1092</td>
<td>ebpB endocarditis and biofilm-associated pili B</td>
<td>9.17</td>
<td>175.99</td>
<td>2.89E-11</td>
<td></td>
</tr>
<tr>
<td>EF1093</td>
<td>ebpC endocarditis and biofilm-associated pili C</td>
<td>15.326</td>
<td>288.1</td>
<td>1.46E-13</td>
<td></td>
</tr>
<tr>
<td>EF1097</td>
<td>hypothetical protein</td>
<td>126.68</td>
<td>612.88</td>
<td>1.62E-05</td>
<td></td>
</tr>
<tr>
<td>EF1117</td>
<td>Polar amino acid ABC transporter, permease protein</td>
<td>77.891</td>
<td>337.36</td>
<td>4.88E-05</td>
<td></td>
</tr>
<tr>
<td>EF1118</td>
<td>amino acid ABC transporter, permease protein</td>
<td>53.562</td>
<td>324.01</td>
<td>1.04E-06</td>
<td></td>
</tr>
<tr>
<td>EF1119</td>
<td>amino acid ABC transporter, amino acid-binding protein</td>
<td>68.763</td>
<td>399.15</td>
<td>1.97E-06</td>
<td></td>
</tr>
<tr>
<td>EF1120</td>
<td>amino acid ABC transporter, ATP-binding protein</td>
<td>75.022</td>
<td>363.41</td>
<td>1.49E-05</td>
<td></td>
</tr>
<tr>
<td>EF1133</td>
<td>2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase , dapD</td>
<td>68.775</td>
<td>184.84</td>
<td>0.00402721</td>
<td></td>
</tr>
<tr>
<td>EF1140</td>
<td>Lactoylglutathione lyase, gloA</td>
<td>169.4</td>
<td>423.02</td>
<td>0.00744091</td>
<td></td>
</tr>
<tr>
<td>EF1141</td>
<td>MutT/nudix family protein</td>
<td>62.541</td>
<td>171.02</td>
<td>0.00623522</td>
<td></td>
</tr>
<tr>
<td>EF1154</td>
<td>DNA replication protein dnaD</td>
<td>30.332</td>
<td>80.854</td>
<td>0.00764152</td>
<td></td>
</tr>
<tr>
<td>EF1155</td>
<td>Endonuclease III, nth</td>
<td>36.02</td>
<td>95.971</td>
<td>0.00767048</td>
<td></td>
</tr>
<tr>
<td>EF1189</td>
<td>Integral membrane protein</td>
<td>21.504</td>
<td>73.349</td>
<td>0.00459217</td>
<td></td>
</tr>
<tr>
<td>EF1202</td>
<td>Hypothetical protein possible functionally linked with Alanyl-tRNA synthetase</td>
<td>449.08</td>
<td>1205.8</td>
<td>0.00585493</td>
<td></td>
</tr>
<tr>
<td>EF1203</td>
<td>Putative Holliday junction resolvase</td>
<td>241.21</td>
<td>648.78</td>
<td>0.00688697</td>
<td></td>
</tr>
<tr>
<td>EF1214</td>
<td>Alpha-acetolactate decarboxylase, budA</td>
<td>19.848</td>
<td>99.152</td>
<td>0.0001694</td>
<td></td>
</tr>
<tr>
<td>EF1270</td>
<td>COG0779: clustered with transcription termination protein NusA, orthologous to ribosome maturation factor RimP</td>
<td>62.101</td>
<td>236.64</td>
<td>0.00018588</td>
<td></td>
</tr>
<tr>
<td>EF1355</td>
<td>dihydrolipoamide acetyltransferase, aceF</td>
<td>168.07</td>
<td>522.54</td>
<td>0.00821873</td>
<td></td>
</tr>
<tr>
<td>EF1356</td>
<td>Dihydrolipoamide dehydrogenase, lpdA</td>
<td>196.45</td>
<td>698.39</td>
<td>0.00311097</td>
<td></td>
</tr>
<tr>
<td>EF1372</td>
<td>thioesterase family protein, contains CBS domain</td>
<td>28.397</td>
<td>72.651</td>
<td>0.00596077</td>
<td></td>
</tr>
<tr>
<td>EF1400</td>
<td>cadmium-translocating P-type ATPase</td>
<td>43.514</td>
<td>183.26</td>
<td>0.00027343</td>
<td></td>
</tr>
<tr>
<td>EF1406</td>
<td>Exonuclease ABC subunit C, uvrC</td>
<td>11.686</td>
<td>31.697</td>
<td>0.005036</td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>E Value</td>
<td>p Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1505</td>
<td>conserved hypothetical protein</td>
<td>31.887</td>
<td>0.000210 32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1579</td>
<td>SOS-response repressor and protease LexA</td>
<td>125.42</td>
<td>0.000120 65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1587</td>
<td>MutT/nudix family protein</td>
<td>39.886</td>
<td>0.000750 17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1596</td>
<td>lipoprotein, putative</td>
<td>28.371</td>
<td>0.000744 83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1597</td>
<td>Catalase/peroxidase, katA</td>
<td>17.424</td>
<td>0.000222 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1599</td>
<td>TPR domain transcriptional regulator, Cro/Cl family</td>
<td>60.617</td>
<td>0.003766 82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1616</td>
<td>Succinyl-CoA synthetase, alpha subunit-related enzymes</td>
<td>97.003</td>
<td>0.000937 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1710</td>
<td>transcriptional regulator, LysR family</td>
<td>68.576</td>
<td>0.003205 73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1711</td>
<td>carbonic anhydrase, putative</td>
<td>42.16</td>
<td>0.000605 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1712</td>
<td>Orotate phosphoribosyltransferase, pyrE</td>
<td>55.9</td>
<td>3.34E-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1713</td>
<td>Orotidine 5'-phosphate decarboxylase, pyrF</td>
<td>81.975</td>
<td>3.11E-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1714</td>
<td>Dihydroorotate dehydrogenase, catalytic subunit, pyrD-2</td>
<td>35.43</td>
<td>3.46E-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1715</td>
<td>Dihydroorotate dehydrogenase electron transfer subunit, pyrDII</td>
<td>43.349</td>
<td>6.49E-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1716</td>
<td>Carbamoyl-phosphate synthase large chain, carB</td>
<td>65.653</td>
<td>2.72E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1717</td>
<td>Carbamoyl-phosphate synthase small chain, pyrA</td>
<td>80.383</td>
<td>0.000549 07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1718</td>
<td>Dihydroorotase, pyrC</td>
<td>55.639</td>
<td>0.000159 03 (14-16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1719</td>
<td>Aspartate carbamoyltransferase, pyrB</td>
<td>91.591</td>
<td>0.008181 84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1755</td>
<td>Phosphate transport ATP-binding protein pstB</td>
<td>27.122</td>
<td>0.009121 37 (KF RIVET 14-16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1758</td>
<td>Phosphate transport system permease protein pstC</td>
<td>12.556</td>
<td>0.007842 55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1794</td>
<td>hypothetical protein</td>
<td>33.804</td>
<td>0.008001 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1817</td>
<td>Staphylococcal serine proteinase homologue</td>
<td>81.816</td>
<td>3.76E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1818</td>
<td>Coccolysin, gelE</td>
<td>66.45</td>
<td>9.67E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1820</td>
<td>fsrC, histidine kinase</td>
<td>45.055</td>
<td>1.32E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1821</td>
<td>fsrB,</td>
<td>24.339</td>
<td>1.58E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1949</td>
<td>hypothetical protein</td>
<td>49.71</td>
<td>0.006072 05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF1979</td>
<td>ATPase, recombination factor rarA</td>
<td>46.541</td>
<td>0.002143 77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>P Value</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
<td>-----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>EF2072</td>
<td>Cysteine desulfurase, class V aminotransferase</td>
<td>20.99</td>
<td>55.253</td>
<td>0.006183</td>
<td></td>
</tr>
<tr>
<td>EF2074</td>
<td>Manganese ABC transporter, ATP-binding protein SitB</td>
<td>165.61</td>
<td>842.06</td>
<td>6.16E-05</td>
<td></td>
</tr>
<tr>
<td>EF2075</td>
<td>Manganese ABC transporter, inner membrane permease protein SitD</td>
<td>154.73</td>
<td>536.62</td>
<td>0.002086</td>
<td></td>
</tr>
<tr>
<td>EF2159</td>
<td>Glutamine synthetase type I, glnA</td>
<td>122.07</td>
<td>500.83</td>
<td>0.000737</td>
<td></td>
</tr>
<tr>
<td>EF2160</td>
<td>regulatory protein glnR</td>
<td>129.59</td>
<td>438.32</td>
<td>0.000506</td>
<td></td>
</tr>
<tr>
<td>EF2203</td>
<td>transcriptional regulator, TetR family</td>
<td>218.88</td>
<td>611.33</td>
<td>0.008166</td>
<td></td>
</tr>
<tr>
<td>EF2394</td>
<td>Iron-sulfur cluster assembly ATPase protein SufC</td>
<td>16.762</td>
<td>59.251</td>
<td>0.003507</td>
<td></td>
</tr>
<tr>
<td>EF2395</td>
<td>Ribosome recycling factor, frr</td>
<td>111.97</td>
<td>292.09</td>
<td>0.005918</td>
<td></td>
</tr>
<tr>
<td>EF2400</td>
<td>RNA methyltransferase, TrmH family</td>
<td>16.702</td>
<td>61.062</td>
<td>0.004394</td>
<td></td>
</tr>
<tr>
<td>EF2419</td>
<td>Hypothetical protein</td>
<td>33.818</td>
<td>84.332</td>
<td>0.008195</td>
<td></td>
</tr>
<tr>
<td>EF2427</td>
<td>hypothetical protein</td>
<td>19.603</td>
<td>109.08</td>
<td>7.06E-06</td>
<td></td>
</tr>
<tr>
<td>EF2438</td>
<td>PTS system, IIA component</td>
<td>35.837</td>
<td>126.81</td>
<td>0.002202</td>
<td></td>
</tr>
<tr>
<td>EF2457</td>
<td>Cell division protein ftsW</td>
<td>21.732</td>
<td>54.345</td>
<td>0.008946</td>
<td></td>
</tr>
<tr>
<td>EF2673</td>
<td>GTP pyrophosphokinase</td>
<td>38.866</td>
<td>97.32</td>
<td>0.009908</td>
<td></td>
</tr>
<tr>
<td>EF2687</td>
<td>hypothetical protein</td>
<td>20.912</td>
<td>65.561</td>
<td>0.002699</td>
<td></td>
</tr>
<tr>
<td>EF2702</td>
<td>hypothetical protein</td>
<td>111.47</td>
<td>377.01</td>
<td>0.000515</td>
<td></td>
</tr>
<tr>
<td>EF2704</td>
<td>A/G-specific adenine glycosylase, mutY</td>
<td>13.744</td>
<td>46.43</td>
<td>0.001910</td>
<td></td>
</tr>
<tr>
<td>EF2705</td>
<td>Regulatory protein recX</td>
<td>23.579</td>
<td>84.169</td>
<td>0.000825</td>
<td></td>
</tr>
<tr>
<td>EF2720</td>
<td>ABC transporter, ATP-binding protein, las</td>
<td>38.157</td>
<td>112.94</td>
<td>0.002497</td>
<td></td>
</tr>
<tr>
<td>EF2756</td>
<td>DNA-damage inducible protein P, dinP</td>
<td>18.209</td>
<td>73.576</td>
<td>0.000176</td>
<td></td>
</tr>
<tr>
<td>EF2868</td>
<td>conserved hypothetical protein</td>
<td>59.49</td>
<td>142.73</td>
<td>0.009993</td>
<td></td>
</tr>
<tr>
<td>EF2870</td>
<td>Hydrolase (HAD superfamily), yqeK</td>
<td>39.103</td>
<td>101.18</td>
<td>0.009845</td>
<td></td>
</tr>
<tr>
<td>EF2872</td>
<td>RNA binding protein</td>
<td>0</td>
<td>99.509</td>
<td>0.001641</td>
<td></td>
</tr>
<tr>
<td>EF2885</td>
<td>3-oxoacyl-[acyl-carrier-protein] synthase III, fabH</td>
<td>141.35</td>
<td>550.27</td>
<td>0.000838</td>
<td></td>
</tr>
<tr>
<td>EF2886</td>
<td>transcriptional regulator, MarR family</td>
<td>338.19</td>
<td>1007.5</td>
<td>0.005544</td>
<td></td>
</tr>
<tr>
<td>EF2908</td>
<td>glycosyl transferase, group 2</td>
<td>45.483</td>
<td>135.18</td>
<td>0.001671</td>
<td></td>
</tr>
</tbody>
</table>

KF RIVET
<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Score</th>
<th>Value</th>
<th>P-Value</th>
<th>KF RIVET</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF2909</td>
<td>hypothetical protein</td>
<td>251.74</td>
<td>708.46</td>
<td>0.002841</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF2932</td>
<td>Thiol peroxidase, Tpx-type, TSA family protein, AhpC</td>
<td>153.54</td>
<td>420.25</td>
<td>0.004871</td>
<td></td>
</tr>
<tr>
<td>EF2982</td>
<td>phosphoglycerate mutase family protein, pgm</td>
<td>20.933</td>
<td>87.333</td>
<td>0.007668</td>
<td></td>
</tr>
<tr>
<td>EF3022</td>
<td>sodium:dicarboxylate symporter family protein</td>
<td>20.967</td>
<td>51.168</td>
<td>0.009155</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF3106</td>
<td>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein oppA</td>
<td>209.15</td>
<td>676.73</td>
<td>0.007067</td>
<td></td>
</tr>
<tr>
<td>EF3107</td>
<td>peptide ABC transporter, permease protein</td>
<td>131.48</td>
<td>445.17</td>
<td>0.002205</td>
<td></td>
</tr>
<tr>
<td>EF3127</td>
<td>Guanylate kinase, gmk</td>
<td>44.978</td>
<td>123.97</td>
<td>0.004140</td>
<td></td>
</tr>
<tr>
<td>EF3131</td>
<td>Hypothetical Protein yicC-like</td>
<td>27.704</td>
<td>81.448</td>
<td>0.003054</td>
<td></td>
</tr>
<tr>
<td>EF3148</td>
<td>CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidylintransferase</td>
<td>35.744</td>
<td>120.47</td>
<td>0.001947</td>
<td></td>
</tr>
<tr>
<td>EF3171</td>
<td>RecA protein</td>
<td>126.07</td>
<td>880.54</td>
<td>2.48E-06</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF3173</td>
<td>hypothetical protein</td>
<td>17.586</td>
<td>54.265</td>
<td>0.001971</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF3177</td>
<td>Hypothetical protein</td>
<td>28.954</td>
<td>91.344</td>
<td>0.0016</td>
<td>(56)</td>
</tr>
<tr>
<td>EF3196</td>
<td>Autolysis response regulator LytR</td>
<td>10.332</td>
<td>78.039</td>
<td>0.002725</td>
<td></td>
</tr>
<tr>
<td>EF3198</td>
<td>lipoprotein, YaeC family</td>
<td>87.178</td>
<td>543.37</td>
<td>1.73E-06</td>
<td></td>
</tr>
<tr>
<td>EF3199</td>
<td>Methionine ABC transporter permease protein</td>
<td>69.845</td>
<td>342.88</td>
<td>1.17E-05</td>
<td></td>
</tr>
<tr>
<td>EF3200</td>
<td>ABC transporter, ATP-binding protein</td>
<td>69.15</td>
<td>330.24</td>
<td>1.63E-05</td>
<td></td>
</tr>
<tr>
<td>EF3209</td>
<td>ABC transporter, ATP-binding protein</td>
<td>9.8698</td>
<td>42.981</td>
<td>0.006016</td>
<td></td>
</tr>
<tr>
<td>EF3270</td>
<td>Glutathione reductase/metabolism, gor</td>
<td>34.133</td>
<td>90.041</td>
<td>0.004837</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>EF3275</td>
<td>N-methylhydantoinase (ATP-hydrolyzing), hydantoinase/oxoprolinase</td>
<td>0</td>
<td>23.761</td>
<td>0.000119</td>
<td></td>
</tr>
<tr>
<td>EF3276</td>
<td>Conservative hypothetical protein probably involved in hydantoin, pyrimidine utilization</td>
<td>0</td>
<td>41.344</td>
<td>0.000196</td>
<td></td>
</tr>
<tr>
<td>EF3283</td>
<td>DNA-binding transcriptional regulator, ctsR</td>
<td>47.978</td>
<td>220.89</td>
<td>7.40E-05</td>
<td></td>
</tr>
<tr>
<td>EF3289</td>
<td>DNA-binding response regulator</td>
<td>133.76</td>
<td>397.55</td>
<td>0.003868</td>
<td></td>
</tr>
<tr>
<td>NEW?</td>
<td>tRNA pseudouridine synthase A</td>
<td>0</td>
<td>48.037</td>
<td>0.002772</td>
<td></td>
</tr>
<tr>
<td>OG1RF00</td>
<td>hypothetical protein</td>
<td>34.998</td>
<td>150.27</td>
<td>5.46E-05</td>
<td></td>
</tr>
<tr>
<td>OG1RF01</td>
<td>DNA-cytosine methyltransferase</td>
<td>127.5</td>
<td>361.99</td>
<td>0.009862</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>--------</td>
<td>--------------------------------------</td>
<td>------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>47</td>
<td>family protein</td>
<td></td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OG1RF0151</td>
<td>Glycosyl transferase, group 2 family protein</td>
<td>19.336</td>
<td>54.404</td>
<td>0.00884778</td>
<td></td>
</tr>
<tr>
<td>OG1RF0191</td>
<td>D,D-carboxypeptidase, vanYG like</td>
<td>5.9544</td>
<td>258.76</td>
<td>0.00027139</td>
<td>KF RIVET</td>
</tr>
<tr>
<td>OG1RF0217</td>
<td>hypothetical protein</td>
<td>29.178</td>
<td>90.416</td>
<td>0.00655996</td>
<td>KF RIVET</td>
</tr>
</tbody>
</table>
Figure 5. Gene categories found in RNAseq analysis. A pie chart showing categories of genes found in RNAseq to be either upregulated (A) or downregulated (B) when cells are grown in a biofilm relative to expression in planktonically growing cells.
<table>
<thead>
<tr>
<th>EF gene ID</th>
<th>Putative substrate transported</th>
<th>Previously identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF0063</td>
<td>Oligopeptides</td>
<td></td>
</tr>
<tr>
<td>EF0675</td>
<td>Osmoprotectants</td>
<td>(14)</td>
</tr>
<tr>
<td>EF0807</td>
<td>Peptides/nickel</td>
<td></td>
</tr>
<tr>
<td>EF1117</td>
<td>Polar amino acids</td>
<td></td>
</tr>
<tr>
<td>EF1118</td>
<td>Glutamine</td>
<td></td>
</tr>
<tr>
<td>EF1119</td>
<td>Glutamine</td>
<td></td>
</tr>
<tr>
<td>EF1120</td>
<td>Polar amino acids</td>
<td></td>
</tr>
<tr>
<td>EF1755</td>
<td>Phosphate</td>
<td>Unpublished RIVET data – Dunny lab, (14)</td>
</tr>
<tr>
<td>EF1758</td>
<td>Phosphate</td>
<td></td>
</tr>
<tr>
<td>EF2074</td>
<td>Manganese/iron</td>
<td></td>
</tr>
<tr>
<td>EF2075</td>
<td>Manganese/iron</td>
<td></td>
</tr>
<tr>
<td>EF2394</td>
<td>Iron/sulfur cluster</td>
<td></td>
</tr>
<tr>
<td>EF2438</td>
<td>Glucose/phosphate</td>
<td></td>
</tr>
<tr>
<td>EF2720</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>EF3022</td>
<td>Glutamate</td>
<td>Unpublished RIVET data – Dunny lab</td>
</tr>
<tr>
<td>EF3106</td>
<td>Peptide/nickel</td>
<td></td>
</tr>
<tr>
<td>EF3107</td>
<td>Peptide/nickel</td>
<td></td>
</tr>
<tr>
<td>EF3198</td>
<td>D-methionine</td>
<td></td>
</tr>
<tr>
<td>EF3199</td>
<td>D-methionine</td>
<td></td>
</tr>
<tr>
<td>EF3200</td>
<td>D-methionine</td>
<td></td>
</tr>
<tr>
<td>EF3209</td>
<td>Antibiotics</td>
<td></td>
</tr>
</tbody>
</table>
Genes involved in homologous recombination were among those highly up-regulated in biofilm formation. EF0004 (recF) and EF3171 (recA) were up-regulated in biofilms (three and seven fold respectively). Involved in DNA recombination, these genes act during the process of RecA filament formation. RecF binds to single or double-stranded DNA breaks and requires the action of RecA to complete strand exchange and form a Holliday junction. EF0066 (ruvA) and EF0067 (ruvB) were also up-regulated in biofilms approximately four and three fold respectively. These genes are involved in the translocation and resolution of Holliday junctions. All four of these genes are considered members of the SOS-controlled response. Deletion of these genes can lead to decreased resistance to stresses such as H₂O₂ as can others identified in this screen such as katA (57, 58).

Genes encoding UvrC, required for the removal of UV induced damage, and DinP (damage-inducible protein P), an error-prone DNA polymerase, were also found to be up-regulated in biofilms. A multitude of other putative DNA repair genes were identified in the screen including EF0682 (DNA repair exonuclease yhaO), EF1154 (DNA replication protein dnaD), EF1155 (Endonuclease III nth), EF1203 (a putative Holliday junction resolvase), and EF2705 (regulator protein recX). Interestingly, no genes involved in DNA replication or repair were measured to be down-regulated in biofilms. It is apparent that whether or not these cells consider biofilms to be stressful environments, they are highly up-regulating genes involved in their SOS response.
and ensuring appropriate mechanisms for DNA replication and repair are present. Another SOS-controlled gene, lexA, was also up-regulated in biofilms. Interestingly, LexA acts as a repressor during the SOS response repressing its own expression as well as that of recA. Implications of the up-regulation of these genes are expanded on in the discussion section of this work.

Together these data provide us with interesting insights into the transcriptome of cells growing in a biofilm and also act as a complementary approach to the transposon and RIVET studies previously completed in our laboratory (14, 15, 56). The increase in genes involved in metabolism, the SOS response, and environmental sensing will be essential in understanding how bacteria in biofilm communities respond to environmental cues. Further work is underway to determine the strand specificity of the biofilm transcriptome to identify possible biofilm-specific antisense regulators of gene expression.

It is apparent from the RNAseq data that gene expression profiles can vary dramatically when cells are grown in a biofilm. This raises questions as to how regulation of other bacterial functions may be affected by growth conditions. Because of the importance of conjugation systems in E. faecalis biology and pathogenesis, I next decided to examine the role of biofilm formation in the regulation of conjugation.
II. Biofilm growth alters regulation of conjugation by a bacterial pheromone

The pattern of conjugation induction differs between cells growing in liquid culture and those grown in a biofilm

Development of an inducibly fluorescent reporter of conjugation allowed us to examine conjugation induction in cells grown both planktonically and in a biofilm to determine the effects of growth conditions on conjugation. Cells containing the pCF10-GFP fusion expressed GFP in response to addition of exogenous cCF10 in liquid culture as well as in biofilms grown in the CBR.

The induction pattern of cells was quantified using either a range of cCF10 concentrations with an induction time of 60 minutes or a set inducer concentration with varying induction times. Flow cytometric analyses of planktonic cell induction patterns are shown in Figures 6A and 7A. As cells were induced with higher concentrations of cCF10, the entire planktonic population shifted to higher levels of GFP expression, apparently homogeneously (i.e. the detection system of the flow cytometer was not able to resolve two distinct populations) (Figure 6A). The histograms are representative of results obtained with pheromone titrations ranging from 0.1-10 ng/mL cCF10. A time dependent unimodal increase in GFP expression was also observed when identical levels of cCF10 were used to induce planktonic cells for time periods ranging from 0-120 minutes (Figure 7A).
When cells growing in the biofilm state were similarly induced, the induction pattern was markedly different (Figures 6B and 7B). At low concentrations of pheromone, a portion of the cells turned on GFP expression and formed a small subpopulation distinct from the larger population. As inducer concentrations increased, the subpopulation expressing GFP increased and the proportion of cells not expressing GFP concurrently decreased (Figure 6B). This suggests that the cells that were non-responsive at low concentrations of pheromone exposure retain the potential to respond to higher pheromone concentrations. If either inducer concentrations or induction times were increased, >95% of cells expressed GFP. A bimodal pattern of pheromone response of biofilm cells was also observed when the time course of the response was examined (Figure 7B). I subjected pheromone-treated biofilms to propidium iodide staining to assess the viability of the GFP-positive and -negative populations, and found very low numbers of potential non-viable cells in either population (<2% in several different experiments, as illustrated in Figure 6D). These data rule out cell death as a reason for the lack of pheromone response in the GFP-negative cells.

The biofilms used for the induction experiments shown in Figure 6 were grown for 24 hours, which produced sufficient numbers of bacterial cells for analysis from a relatively small number of coupons. By substantially increasing the number of coupons, I was able to do similar induction experiments with 4 hour biofilms, and obtained essentially identical results (Figure 8). This suggests
that differentiation of the biofilm cells into distinct sub-populations occurs early in development, while the adherent bacteria are actively growing (Figure 3). Furthermore, I carried out numerous experiments involving induction of planktonic cells (including the planktonic cells from the same reactors used to harvest the biofilms) where the nutrient content of the medium during pre-growth and induction was varied by diluting the M9 growth medium to various concentrations ranging from 10-100%, or by using tryptic soy broth. In all of these experiments a unimodal induction pattern similar to that depicted in Figure 6A was observed, suggesting that biofilm growth was a more important determinant of the bimodal response than nutrient content or growth rate.

The results described above suggest, that at limiting concentrations of pheromone typically produced by recipient cells, the overall frequency of plasmid transfer might be lower in biofilms than in planktonic cultures. I examined this possibility by comparing transfer frequencies in the planktonic and biofilm subpopulations of CDC reactors containing mixed cultures of donors and recipients, and found that the overall efficiency of transfer was significantly lower in the biofilm phase (Table 7).

*The biofilm matrix is not responsible for differences between biofilm and planktonic conjugation regulation*

A structural component of the biofilm that could cause the biofilm cells to undergo different response patterns from planktonic cells is the biofilm matrix. Most simply, the matrix could inhibit pheromone induction of some cells
by interfering with signal diffusion. The matrix could also serve to concentrate peptides in certain areas to stimulate cell induction in the immediate vicinity. To test for these possibilities, coupons containing biofilm cells were vortexed in PBS + 2mM EDTA to release them from the matrix and suspended in a 50% concentration of M9 prior to pheromone induction. The overall induction pattern of dispersed biofilm cells was the same as that of attached biofilm cells (Figure 6C). This demonstrates that the effects of the biofilm matrix on cCF10 diffusion are not a major factor in the difference responses to pheromone observed between biofilm and planktonic cells. If biofilm cells were resuspended in full strength medium, the population dynamics of pheromone induction reverted to the unimodal mode of normal planktonic cells within 60 minutes.
Figure 6. Growth in a biofilm alters the induction pattern of pCF10 conjugation.

A. Cells grown in a liquid culture were induced for 60 minutes with various concentrations of cCF10. Induced populations shifted to higher GFP expression in a unimodal pattern. B. Coupons containing biofilm cells were induced for 60 minutes with various concentrations of cCF10. Induced cells expressed GFP in a bimodal population distribution. C. Cells grown in a biofilm for 24 hours and then dispersed prior to a 60 minute induction with various concentrations of cCF10 behave as attached biofilm cells showing bimodal response to induction.

Horizontal axis = GFP (FL1) expression, Vertical axis = % of maximum cell number. D. Flow cytometry analysis of 24 hour biofilms following induction with
1 ng/mL of cCF10 for 60 minutes. Left panel demonstrates the populations gated to remove debris following biofilm dispersal based on size (FSC) and granularity (SSC). Right panel indicating propidium iodide (PI) staining (FL2) on the y-axis and GFP expression (FL1) on the x-axis. Less than 2% of the sorted cells stained with PI; similar numbers were seen with uninduced cells.
Figure 7. Growth in a biofilm alters the time course induction pattern of pCF10 conjugation. Cells were induced using 0.1 ng/mL of cCF10 and the population distribution of GFP fluorescence was measured over time. A. Cells grown in a liquid culture exhibit a unimodal shift pattern in induction. B. A time course of biofilm cell induction shows bimodal induction kinetics. Horizontal axis = FL1 (GFP) expression, Vertical axis= % of maximum cell number.
Figure 8. Growth pattern changes are seen as early as 4 hours post biofilm inoculation. A. Cells grown in a liquid culture (A) or a biofilm (B) for 4 hours were induced for 60 minutes with various concentrations of cCF10. Induced planktonic populations shifted to higher GFP expression in a unimodal pattern. Induced biofilm cells expressed GFP in a bimodal population distribution. Horizontal axis = GFP (FL1) expression, Vertical axis = % of maximum cell number.
We also tested the induction profile of a GFP fusion construct derived from pCF10 where transcription was driven by the same promoter, but the gene encoding pheromone receptor/conjugation repressor protein, PrgX, was deleted. In this case, GFP expression was constitutive, unimodal, and unresponsive to pheromone induction (Figure 9A). Adding prgX in trans rescued the bimodal response (Figure 9B). From this I conclude that the bimodal distribution in GFP expression observed with the pheromone-inducible construct arose from biofilm effects on the pheromone response machinery and was not due to random inhibition of GFP expression in a subpopulation of the biofilm cells.
Figure 9. Bimodal induction is not an artifact of GFP expression from the prgQ promoter. 

A. With a constitutively expressed reporter of the promoter P_Q (due to lack of prgX repressor gene) virtually all of the biofilm cells are in the “on” state and unresponsive to cCF10. 

B. The same constitutive reporter plasmid demonstrates bimodal induction when the prgX is expressed in trans from the chromosome.
Table 7. Biofilm and planktonic mating rates

<table>
<thead>
<tr>
<th></th>
<th>Donor CFU/mL</th>
<th>Recipient CFU/mL</th>
<th>Transconjugant CFU/mL</th>
<th>Transconjugant: Donor Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biofilm</td>
<td>2.46E9</td>
<td>1.28E9</td>
<td>5.6E4</td>
<td>1:43938</td>
</tr>
<tr>
<td>Planktonic</td>
<td>1E9</td>
<td>5.9E8</td>
<td>1E7</td>
<td>1:100</td>
</tr>
</tbody>
</table>
RNaseIII and endogenous pheromone production do not promote bistability in planktonic cells

The role of Eep protease and RNaseIII were also examined to see if a deletion in these genes could promote bimodal induction in planktonic cells. Eep is a protease that is responsible for cleavage of pheromone precursors to form mature peptides. Δeep strains do not produce mature iCF10 or cCF10 (23). RNaseIII cleaves double-stranded RNA and can be an important mediator of RNA decay (59). The pCF10-GFP plasmid was mated into strains in which eep and RNaseIII had been knocked out. These strains were then tested as above for induction patterns in planktonic cells. Deletion of either eep or RNaseIII did not allow bimodal induction in cells grown planktonically (Figure 10A-B).

Using a strain deleted for the chromosomal ccfA gene, encoding for cCF10 production (JRC104) I again tested whether endogenous pheromone production alters induction of biofilm and planktonic cells. As above, pCF10-GFP was mated into JRC104 and the cells were examined for induction patterns under biofilm and planktonic growth conditions. Deletion of the gene responsible for cCF10 production did not modify induction patterns in biofilm or planktonic cells (Figure 10C-D). Together, these data suggest that the rate of peptide production and RNA turnover are not the major determinants in the observed differential regulation of conjugation.
Figure 10. Knockouts of RNaseIII, *eep*, and *ccfA* do not affect the planktonic induction pattern of pLC-2. Planktonic cells deleted for RNaseIII (A), *eep* (B), or *ccfA* (C) or Δ*ccfA* biofilm cells (D) were induced for 60 minutes with various concentrations of cCF10. Induced planktonic populations shifted to higher GFP expression in a unimodal pattern as seen with wildtype cells whereas Δ*ccfA* biofilm cells were induced in a bimodal fashion as seen in wildtype cells.
**pCF10 copy number and heterogeneity are increased in biofilm cells**

A series of mathematical equations were derived by our collaborators Drs. Wei-Shou Hu and Anushree Chatterjee to model the pCF10 genetic network (40). The details of this model are explained in Cook, et al. 2011, and a summary of the modeling parameters and equations are shown in Appendix IV of this thesis. A characteristic S-shaped bistable response of PrgB to extracellular cCF10 was predicted by Dr. Anushree Chatterjee and is shown in Figure 11. The bistable curve is comprised of three sections: the lower and upper parts of the curve correspond to “off” (low level of PrgB) and “on” conjugation states (high level of PrgB) respectively. The middle section is characterized by multiple steady states, two stable states corresponding to “on” and “off” and an unstable steady state, which is not observed experimentally. This predicted bistable behavior is intrinsic to the gene regulatory network and is predicted to be present both in planktonic cells as well as biofilm cells.

Mathematical modeling further suggested that an increase in the copy number of pCF10 could greatly alter induction responses to cCF10 (Figure 11C). The effect of increased plasmid copy number manifests itself through increased number of binding sites (XBS 1 and 2) available for iCF10 bound PrgX tetramers and cCF10 bound PrgX dimers to bind (Equation S10) and an increase in the number of copies of Q₅ and Q₆ transcripts per donor cell, which consequently results in increased production of inhibitor iCF10. Donor cells with fewer copies
of pCF10 are predicted to require lower amounts of cCF10 to turn “on” but respond with a lower PrgB expression level than cells with a higher copy number (Figure 11C). Interestingly, increasing plasmid copy number widens the bistable region, suggesting that the “on” and “off” populations are better separated and more easily distinguishable. The model predicts that cells with high plasmid copy number respond slowest to induction and require a higher concentration of cCF10 whereas low copy number cells respond faster and at a lower concentration of cCF10 (Figure 11D-E). A broader plasmid copy number distribution is predicted to give rise to a bimodal population response to low levels of inducer at longer exposure times, such that cells with high plasmid copy number continue to exist in the “off” state even after long exposure time to inducer, whereas cells with lower copy number switch to “on” state (Figure 11D-E). Our model predicts that a higher pCF10 copy number and copy number heterogeneity would enhance a bimodal response to induction with pheromone.
Figure 11. Modeling the effect of pCF10 plasmid copy number (N) on the cCF10 pheromone response. The data illustrated in this figure were produced by simulations based on the mathematical model described in the text and in Appendix IV by Drs. Wei-Shou Hu and Anushree Chatterjee. A. A simple schematic of the regulatory circuit of pCF10-based conjugation. The prgQ operon encodes the pCF10 conjugation machinery, including an aggregation protein Asc10 encoded by prgB. The prgQ open reading frame encodes the inhibitor peptide iCF10. The promoters for these operons are located within the same region, but on the opposite strands of pCF10, and produce transcripts that are complementary for 220 nt at the 5’ ends. This arrangement leads to reciprocal negative regulation by antisense interactions and transcription interference,
resulting in the double negative feedback loop illustrated by the dashed lines. B. Mathematical modeling predicts that the steady state response of PrgB levels to induction with cCF10 shows bistable switch behavior. C. Bistable switch response of pCF10 to induction with cCF10 for different plasmid copy numbers. The bistable region increases with plasmid copy number along with an increased threshold cCF10 concentration required to turn “on” the genetic switch. D-E. Dynamic response of donor cells with different plasmid copy numbers to low (0.1 ng/mL, (D)) and high (10ng/mL, (E)) inducer concentrations of cCF10. Cells with high plasmid copy number respond slowest to induction and at a higher concentration of cCF10 whereas low copy number cells respond faster and at a lower concentration of cCF10. PrgB levels shown in B-E are normalized to initial state at time t=0 corresponding to “off” steady state levels at 0 ng/mL of cCF10.
The mathematical model predicted that pCF10 copy number changes could alter the pheromone response. Based on this model, I hypothesized that the copy number of pCF10 in biofilms was increased in both average copy number and heterogeneity when compared to planktonic population. To examine this hypothesis experimentally, I performed qPCR on gDNA obtained from pCF10-containing planktonic and biofilm cells to compare the pCF10 copy number of the respective populations. A statistically significant increase in pCF10 copy number of 1.5-2 times that of planktonic cells was observed using qPCR (Figure 12A). A complementary approach using PFGE showed an increase in pCF10 copy number of approximately 1.23 times as determined by band density analysis (Figure 13). The slightly lower estimate from PFGE is understandable based on the band saturation of the pCF10 band in the biofilm cells.

To further examine the relationship between biofilm growth, plasmid copy number and pheromone response, biofilm cells were induced using conditions that generated two approximately equal subpopulations (e.g. Figure 6B). These cells were then sorted based on GFP expression and qPCR was used to examine pCF10 copy number in the “on” and “off” subpopulations. Even though the entire population of biofilm cells was exposed to the same concentration of inducer, the copy number of the pCF10 plasmid differed significantly between cells that expressed GFP and those that did not. Cells not expressing GFP had a statistically higher copy number of pCF10 than cells that expressed GFP during the same induction course (Figure 12B, biofilm). planktonic cells were also
induced with the same level of cCF10 and cells from the single induced peak were sorted. Cells labeled “on” represent cells from 25% of the peak with the highest GFP expression levels and those labeled “off” are cells representative of the 25% of the peak with the lowest expression of GFP. When cells from these populations were examined for copy number heterogeneity, no statistical difference was observed (Figure 12B, planktonic). This demonstrates that biofilm cells have, on average, higher copy numbers of the conjugative plasmid pCF10 than their planktonic counterparts and also possess a greater heterogeneity in copy number compared to planktonic populations consistent with a mathematical model proposed by our collaborators A. Chatterjee and W.S. Hu. A substantial difference in plasmid copy number was also observed between induced biofilm cells and induced planktonic cells. Planktonic cells contained pCF10 at 3-5 copies per chromosome (Figure 12B, planktonic), while induced biofilm cells possessed plasmid copy numbers as high as 8-15 copies per chromosome (Figure 12B, biofilm).
Figure 12. The copy number of pCF10 as well as the population heterogeneity of copy number is increased in biofilm cells compared to planktonic cells. A. Analysis of pCF10 copy number in uninduced biofilm and planktonic cells by qPCR. The average copy number of biofilm cells is 1.81±0.49 times greater than that of planktonic cells. n=5 independent experiments. B. Comparison of pCF10 copy number in biofilm and planktonic cells exposed to cCF10. Biofilm and planktonic cultures were induced with cCF10 for 60 minutes and sorted into “on” and “off” populations. The sorted subpopulations were then analyzed for pCF10 copy number by qPCR. Induced biofilm cells not expressing GFP had a pCF10 copy number 2.77±0.12 times higher than induced cells expressing GFP. Sorted planktonic populations were not statistically different. All planktonic cells contained 3-5 copies of pCF10/chromosome, whereas biofilm cells contain up to 15 copies of pCF10/chromosome with statistically significant heterogeneity between the “on” versus “off” subpopulations. n=4 independent experiments ±
and error bars represent standard deviation ** P value < 0.02 *** P value < 0.003

P values were calculated using a two-tailed t-test assuming equal variance.
Figure 13. Pulsed field gel electrophoresis (PFGE) to examine cell copy number.

PFGE of biofilm and planktonic cells shows that biofilm cells have a higher pCF10 DNA/chromosomal DNA ratio than planktonic cells. Arrow indicates pCF10 band. Using three different chromosomal bands for reference, the ratio of biofilm pCF10/planktonic pCF10 = 1.23±0.01
The copy number of various enterococcal plasmids is altered by growth in a biofilm

We hypothesized that the increased plasmid copy number phenomenon we observed in biofilm cells was not restricted to pCF10. To examine this hypothesis, we chose four plasmids with four different replicons representing both methods of plasmid replication (rolling circle and theta replication) (Table 8). These plasmids all contain the prgX gene, which we used to compare the number of copies of the plasmid to the 100-5 strain which contains one copy of prgX on the chromosome. The biofilm and planktonic growth and copy number analysis were done as previously described for pCF10.

We found that all four plasmids were statistically increased in copy number in biofilm cells when compared to their planktonic counterparts. The plasmid copy number in biofilm cells was between 1.6-2x higher than the copy numbers in planktonic cells (Figure 14A). This phenomenon was present regardless of the native plasmid copy number of the four replicons (Figure 14B).

One explanation for differing copy number measurements is that the amount of gDNA could be higher in planktonic cells. This would cause the ratio of prgX to gyrB to be skewed and thus make it appear that biofilm cells had more copies of the plasmid when in fact they only had fewer copies of the gyrB gene. To test for this possibility, total gDNA/CFU was measured. We found that the total amount of gDNA/CFU was the same between biofilm and planktonic cells.
whether the planktonic cells were harvested from the reactor or from an overnight culture (Figure 15).

Although gDNA content does not appear to play a role in the copy number phenomenon, I examined early biofilms to ensure that growth rate was not affecting our measurements. One plasmid, pMSP3535VAX was chosen and grown in the CBR for 4 hours. Planktonic cells were removed from the liquid portion of the reactor and biofilm cells were collected from the coupons. Copy number analysis was done to determine whether young biofilms, which should retain high growth rates (Figure 3), still have differential copy number expression. Results from four biological replicates show that the copy number of biofilm cells is higher than planktonic cells even as early as 4 hours post-inoculation of the biofilms (Figure 16).
Table 8. Enterococcal plasmids for copy number measure

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Replicon</th>
<th>Replication</th>
<th>Source</th>
<th>Antibiotic resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBK2</td>
<td>pCI305/pBR322</td>
<td>Rolling circle</td>
<td>(60)</td>
<td>Chloramphenicol</td>
</tr>
<tr>
<td>pMSP3535VAX</td>
<td>pVA380-1</td>
<td>Rolling circle</td>
<td>(43)</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>p043lacZ</td>
<td>pAT18/pAMβ1</td>
<td>Theta</td>
<td>(61)</td>
<td>Erythromycin</td>
</tr>
<tr>
<td>pINY8101</td>
<td>pWM402/pIP501</td>
<td>Theta</td>
<td>(45)</td>
<td>Chloramphenicol</td>
</tr>
</tbody>
</table>
Figure 14. The copy number of four different plasmids is increased during biofilm growth versus growth in liquid culture. A. Plasmid copy number is shown as a ratio to planktonic copies. Copy numbers are measured per chromosome. B. The average copy numbers of the four plasmids range widely from ~1-150 copies/chromosome. **=p value ≤ 0.05 ***=p value ≤ 0.001 (original pCF10 data reported in (40) and shown in Figure 12A).
The amount of total gDNA/CFU extracted is the same for biofilm and planktonic cells. Approximately 0.05-0.5 fg of total DNA was extracted per CFU from cells grown in a biofilm, in the liquid phase of the biofilm reactor, or in a liquid overnight culture. No statistical difference was seen between total DNA extracted from the three populations indicating that the number of chromosomes per cell is not significantly altered in these three environments after 24 hours of growth.
Figure 16. Copy number of pMSP3535VAX is increased as early as 4 hours post-inoculation of the biofilm. Four hour samples of biofilm and planktonic cells were analyzed for pMSP3535VAX plasmid copy number. Planktonic cell copy number is set to 1 and biofilm copy number is shown relative to planktonic copies. **=p value ≤ 0.05 n=3 biological replicates.
In cells containing pCF10, I demonstrated that not only do biofilm cells have a higher average copy number; they also have increased heterogeneity of copy number compared with planktonic cells. To determine whether the same behavior is exhibited by other plasmid replicons, I repeated this experiment with pLC2. pLC2 was chosen because it contains the prgX-prgQ operon so cells could be sorted based on induction state, as was done with pCF10. Figure 17A shows a map of the pLC-2 plasmid containing the prgX-prgQ operon and a gfp gene fused downstream of the QL elongation site (see Figure 1 for a more detailed map).

Using this plasmid, I induced biofilm and planktonic cells and then sorted based on GFP expression. As seen earlier, induction of planktonic cells produced a unimodal pattern and induction of biofilm cells resulted in a bimodal distribution (Figure 17B). Following sorting, copy numbers in the populations were examined. As seen in pCF10, the pLC2 plasmid also demonstrated a significant heterogeneity in copy number in biofilm cells that was not present in planktonic cells.
Figure 17. The prgX-Q_L portion of the pCF10 plasmid is sufficient to produce bimodal induction of biofilm cells, and sorting of induced cells demonstrates copy number heterogeneity in biofilms that is not present in planktonic populations. A. A plasmid map of the pLC2 plasmid showing the portion of
pCF10 present in the plasmid. The \textit{gfpmut3b} gene was inserted downstream of Q\textsubscript{i} RNA. B. Following induction with cCF10, only one peak exists in planktonic populations and (left panel) this peak increases in GFP expression. Cells grown in a biofilm (right panel) are induced bimodally following cCF10 induction. C. Following sorting of the extreme ends of the peak(s), there is no statistical significance between the edges of induced planktonic populations with all cells containing \(~2\text{-}6\) copies of the pBK2 plasmid per chromosome. In biofilm populations, the two distinct peaks show increased heterogeneity in their copy number from \(~5\text{-}16\) copies per chromosome. **=p value ≤0.05
III. The dual role of iCF10 as inhibitor and quorum sensing molecule

Donor density affects $Q_L$ expression

Quorum sensing via the use of peptide pheromones has been demonstrated in numerous gram-positive bacteria and has been shown to control coordinated activities such as biofilm formation and competence regulation (3, 6). We examined the role of quorum sensing in the conjugative transfer of the pCF10 plasmid of *E. faecalis*. It has been previously suggested that iCF10 is required to prevent self-induction of the donor cells. Because the iCF10 is produced at a basal level in the absence of pheromone induction by cells containing pCF10 and excreted extracellularly. It is possible that iCF10 functions as a quorum signal to sense the population of surrounding donors. We hypothesize that when there is a large population density of donor cells, only a small number would benefit from expending the energy to undergo conjugation when induced. Thus high iCF10 levels would signal high concentration of surrounding donors and prevent induction of the entire population.

Before beginning these studies, we initially tested ratios of iCF10:cCF10 to determine which ratios allowed for complete shutdown of conjugation protein expression (as measured by GFP expression using the pCF10-GFP plasmid). We determined that a 1:1 ratio (in ng/mL) of iCF10:cCF10 did not turn off expression of conjugation genes completely (Figure 18). This result was expected based on previous data indicating that the PrgX protein has a higher binding affinity for cCF10 than iCF10 (25). Using a range of iCF10 and cCF10
concentrations between 0.01-50ng/mL, we observed that a ratio of approximately 10:1 iCF10:cCF10 was needed to establish complete shutdown of the system (Figure 18).

To determine whether cell induction is dependent on the density of donor cells, we examined transcript levels from the $P_Q$ promoter. Cells were diluted 1:10 (high density) or 1:1000 (low density) and induced with various amounts of endogenous cCF10. At the indicated time points, samples were removed, and RNA was collected. RNA was reverse transcribed into cDNA and analyzed via qRT-PCR using primers against a housekeeping gene $gyrB$ and $Q_L$ (primers listed in Table 3).

As predicted, cells induced at a high donor density expressed high levels of $Q_L$ but were quickly shut back off (Figure 19A). After 15 minutes, maximal induction levels showed 30-100x more $Q_L$ transcripts than uninduced cells. In as little as 30 minutes, the expression level of $Q_L$ had already decreased significantly returning to basal levels by 1-2 hours (Figure 19A). In low density cell populations, immediate induction of cells was still seen in the first 15 minutes but levels reached 3-5 times that of cells induced at high density. The shutdown observed in high density populations was less pronounced and expression of $Q_L$ was detectable for up to two hours at levels between 50-500xs higher than uninduced cells (Figure 19B).
Figure 18. The ratio of cCF10:iCF10 is responsible induction independent of donor density. Cells carrying the pCF10-GFP plasmid were diluted 1:10 or 1:20 and exposed to varying ratios of iCF10:cCF10 and fluorescence levels were measured. At a 1:1 (ng/mL) ratio of iCF10:cCF10, cells were still slightly induced. It takes a ratio of approximately 10:1 iCF10:cCF10 to completely shut off expression of conjugation genes. These results are independent of the initial donor dilution in these ranges.
Figure 19. Donor density affects induction and continued expression of Q_L transcripts. **A.** Cells diluted to high initial donor densities (1:10) were induced to express high levels of Q_L but expression decreased to basal levels by 30-60 minutes. **B.** Cells diluted to low donor densities (1:1000) were induced to higher Q_L expression levels and then remained at high levels for up to two hours post induction. Bars represent the standard deviation for n≥3 biological replicates.
Donor density affects mating frequencies

To determine whether donor density affects conjugation rates, donors were grown overnight and diluted 1:1 or 1:10 before being mixed with recipients (D:R=1:10 v/v). Mating occurred in broth culture for the indicated time, and the samples were plated on antibiotic plates specific for donors, recipients, or transconjugant.

The donor cells that had been diluted 1:1 produced ~1 transconjugant for every 27 donor cells after 3 hours (Figure 20 green line). When the donor cells were diluted 100x lower, 1 transconjugant was made for every ~2.6 donor cells (Figure 20 blue line). The slope of the line between 2-3 hours after mating also demonstrates that mating was still occurring in the dilute donor mating population whereas mating in the high density donor population leveled off (Figure 20).
Figure 20. Mating efficiency is affected by donor density. Low density donor populations produce more transconjugants per donor cell than high density donor populations, and conjugation rates continue to increase at later time points.
*iCF10 acts to repress induction of Qₗ by cCF10*

We observed that donor cell density is a determining factor in the expression of conjugation genes. Decreasing the density of donors supports increased expression of conjugation genes and allows this expression to stay relatively constant for a much longer period of time (Figure 19). Cell mating data show a similar phenomenon in which lower donor density allows mating to occur at a higher rate and continue for a longer period of time (Figure 20). These data support a quorum sensing hypothesis in which donor cells are secreting a small molecule capable of signaling donor cell density to surrounding cells to control the induction of conjugation. We hypothesize that donor-produced iCF10 serves this function as well as specifically regulating the expression of Qₗ.

Although it has long been hypothesized that iCF10 acts as a direct negative regulator of conjugation gene expression, data proving this theory have been somewhat lacking. To determine the role of iCF10 in quorum sensing and negative regulation, I used a strain of OG1RF deleted for the *eep* gene but containing the pCF10 plasmid (JRC106). Because *eep* plays an integral role in the processing of mature peptides, JRC106 produces no measurable endogenous iCF10 and cCF10. Although small amounts of mature peptide may still be made, these amounts are negligible.

Using JRC106, I examined the role of exogenous peptides in donor density sensing as well as shutdown of Qₗ expression. To determine whether iCF10 was in fact responsible for the shutdown of Qₗ and, subsequently,
conjugation, we first induced a high density culture of JRC106 with 10 ng/mL exogenous cCF10 for one hour and then added varying concentrations of exogenous iCF10. Q₉ RNA levels were quantified using qRT-PCR before induction, following induction, and after addition of iCF10. When JRC106 cells were induced with high levels of cCF10, the level of Q₉ transcription increased dramatically (Figure 21A). This result differs from the previous data with the wildtype OG1RF where the Q₉ levels had returned to uninduced levels after 60 minutes (Figure 19B). When no iCF10 was added, Q₉ levels did not depreciably drop even two hours after the induction. When iCF10 was added to the cultures in levels exceeding cCF10, a statistical decrease was seen in Q₉ expression which returned to uninduced levels (Figure 21A). These data strongly suggest that iCF10 is the molecule directly counteracting the inducing effects of cCF10.

To determine whether iCF10 plays a role in donor density sensing, I repeated this experiment using JRC106 at low initial donor densities (1:1000 diluted donors). In one preliminary experiment, the density effects on Q₉ expression were eliminated when cells did not produce endogenous iCF10 (Figure 21B). These experiments indicated that not only is iCF10 responsible for the shutdown of Q₉, but endogenous iCF10 also serves as an indicator of donor density. This data will need to be repeated to confirm the findings.
Figure 21. iCF10 is responsible for shutdown of Q_L expression and abolishing endogenous iCF10 abrogates donor density effects on Q_L. Δeep cells (JRC106) containing pCF10 were diluted 1:10 (A) or 1:1000 (B) and 10ng/mL cCF10 was added at T=0. Cells were grown for 60 minutes with cCF10 before the addition of iCF10 (indicated by red arrow) at varying concentrations (as shown in the graph key).
DISCUSSION

The *E. faecalis* biofilm transcriptome

Preliminary RNAseq data showed that many genes were up-regulated during the first four hours of biofilm formation. The main categories found to be up-regulated include ABC transporters, and genes involved in metabolism and DNA recombination and repair. I hypothesize that the up-regulation of ABC transporters serves as a mechanism to sample the environment and subsequently alter gene expression based on the changing conditions during early biofilm formation.

Transcriptome analysis also showed that the metabolism of the cells changes dramatically in early biofilms. Pyrimidine synthesis and pyruvate metabolism pathways were both highly up-regulated. A paper published in 2010 by Prüss et al. used a pattern-mining algorithm to determine metabolic enzymes related to biofilm formation. The algorithm predicted that pyruvate catabolism was highly involved in biofilm formation and that acetyl-coenzyme A, a product of pyruvate metabolism, increased biofilm formation in *E. coli* strain K-12 (62). The authors do not posit many hypotheses for why acetyl-CoA might induce increased biofilm formation although they suggest that acetylation of proteins could be important. Proteomics data from *S. aureus* biofilms also demonstrates an important role for pyruvate metabolism in biofilms (53). Interestingly, these two studies were both looking at older biofilms from 8-48 hours old. Pyruvate metabolism, then, is not only involved in the formation of early biofilms, but
continues to occur at high levels throughout at least the first two days of biofilm growth. Although the exact mechanism is not understood, it is apparent that the metabolism of pyruvate, acetyl-CoA and acetate plays an important role in the development of biofilms made by both gram-positive and gram-negative bacteria. This would be an interesting area for future research into biofilm cell metabolism.

My RNAseq data also indicated that genes involved in the DNA repair, recombination, and the SOS response were highly up-regulated in biofilms. Of note, the data indicate that both the activator (RecA) and the repressor (LexA) of the SOS response are up-regulated. LexA has previously been identified to be important in biofilm formation in *P. aeruginosa*, and deletion of *lexA* led to reduced biofilm formation (63). When cells are growing normally, LexA binds to the SOS box in the promoter of SOS genes, including *recA*, and represses their expression. During periods of stress and increased DNA damage, LexA cleaves itself with the help of RecA and relieves repression of the SOS genes. The up-regulation of both *recA* and *lexA* in biofilms is of interest because of their differing roles in the SOS response. If cells are attempting to up-regulate stress responses, why would they concurrently up-regulate the SOS activator and repressor? It is possible that by up-regulating both genes, cells are attempting to up-regulate DNA damage repair but also have a mechanism to keep RecA in check.
Interestingly, a 2008 paper by Boles and Singh identified many of these same genes (e.g. *lexA*, *recA*, and *katA*) as being involved in double-stranded DNA (dsDNA) breaks in biofilm cells when in the presence of oxidative stress (64). The increase in dsDNA breaks lead to an increase in cell diversity in *P. aeruginosa* biofilms which is likely also occurring in *E. faecalis* biofilms (64).

Another question that arises from this data is that while we assume that the biofilm is a more natural environment for bacteria such as *E. faecalis*, why are cells behaving as if the biofilm is a stressful environment? Does something about growth in a biofilm cause DNA breakage that would necessitate increased RecA production? Could this pathway play a role in the increase of extracellular DNA seen in early biofilm matrices? Although the results of the RNAseq data have given us many fascinating answers as to how cells are responding to growth in a biofilm, these data have also provided us with a multitude of new questions for which further research is required.

Currently, studies are being undertaken to enhance our understanding of these data including strand specific analysis and comparison of biofilm and planktonic RNA. It is possible that some of the genes that appear to be up-regulated are in fact antisense transcripts. In this case, genes that would seem to be turned on during biofilm formation may actually be under negative control. Looking at the current RNAseq data and comparing it to promoter trap assays shows that some genes described as lowered in biofilms according to RNAseq had upregulated promoters according to RIVET studies (e.g. EF798). Determining
the strand specific RNA transcriptome of biofilm cells will expose many previously unidentified critical biofilm pathways and provide possible gene targets for small molecules to prevent biofilm formation.

RNAseq results demonstrated that a large portion of *E. faecalis* genes are differentially expressed in biofilms when compared to cells grown in liquid culture. These data allowed us to hypothesize that other bacterial systems could be altered by growth in a biofilm. Conjugal gene transfer in enterococci is a significant area of research because of the ability to transfer antibiotic resistance genes. We hypothesized that growth conditions may alter the regulation of conjugative gene induction.

*Effects of the environment on conjugation in pCF10*

Control of expression of conjugation functions in the pCF10 system is complex and involves competing antagonistic activities of two secreted signaling peptides and multiple intracellular regulatory circuits acting at the level of transcription initiation, as well post-transcriptionally (Figure 1 and (65)). There is experimental evidence demonstrating that the two-peptide signaling system increases versatility such that a response can be activated either by the presence of potential recipient cells in close proximity or by growth in the mammalian bloodstream (66), where expression of the pheromone-inducible *prgB* gene increases virulence (31, 67). However, until recently the relative importance of the multiple (potentially redundant) intracellular mechanisms of regulation of expression of the pCF10 *prgQ* conjugation operon was not clear. Our
collaborators carried out modeling studies and quantitative analysis of transcript levels from the prgQ and prgX operons in response to cCF10 in planktonic cultures, using genetic constructs to analyze the specific contributions of the individual regulatory mechanisms to the system (68). These studies suggested that the pheromone response system could function as a sensitive bistable genetic switch, and that disruption of any of the individual regulatory circuits of the system abolished switch behavior.

I used a pheromone-inducible GFP reporter construct to allow for expression analysis on a single cell level, and also examined the effects of biofilm growth on the population dynamics of the response. At first glance, the results reported here for planktonic cells appear to be inconsistent with previous studies (68). However, when the surprising effects of biofilm growth on increasing both the average plasmid copy number and copy number heterogeneity (Figure 12) are considered, the cumulative results are consistent with and validate the predictions of mathematical models put forward by our collaborators (40, 68).

While fluorescent reporter proteins facilitate single cell expression analysis, they do have limitations for our system, including the fact that the fermentative metabolism of enterococci may result in a reduced cytoplasmic environment suboptimal for proper GFP folding. Also, the gfp allele used in these studies encodes a very stable protein which may be toxic to cells at very high expression levels. These factors likely reduced the effective signal to noise ratio...
in our experiments. I suspect that the apparent lack of bimodal response in planktonic cultures might reflect the limitations of the detection system. It is likely that the low copy number of pCF10 in planktonic cells created a lower threshold for switch behavior to occur and also caused the cells to respond at a lower level blurring the distinction between “on” and “off” as predicted by our mathematical model (Figure 11). In any case, there is a clear difference in the biology of the pheromone response in the two types of cells, with biofilm growth requiring an increased level of signal, but resulting in a more vigorous response to pheromone. This suggested that the average frequency of conjugative plasmid transfer in biofilms could be lower than that of planktonic cells, which was observed experimentally (Table 6).

*Plasmid copy number is altered by changing environments*

An unexpected result of this study was the effect of biofilm growth on pCF10 copy number (Figure 12). Mathematical modeling suggested that changes in plasmid copy number could account for the observed effects of biofilm growth on the pheromone response (Figure 11 and (40)). While the average copy number of the entire population of biofilm cells was increased, biofilm cells also exhibited a remarkable increase in the heterogeneity of copy number values, which was revealed by analysis of sorted “on” and “off” populations exposed to threshold inducing concentrations of pheromone (Figure 12). As noted above, cells with more copies of pCF10 are more adequately poised to respond to cCF10 pheromone in an appropriate and highly controlled manner than adjacent cells.
with lower plasmid copy number (40). By requiring higher pheromone concentrations to turn on conjugation, these cells avoid the expenditure of large amounts of energy involved in producing conjugative machinery. Biofilms are often seen as a type of multicellular community and allowing for conjugation to occur only in a subset of potential donor cells may be beneficial to the community as a whole.

While the bacteria in biofilms are sometimes described as non-growing, or slowly growing, I observed the bimodal pheromone response in cells grown for as little as 4 hours on surfaces, when adherent cell densities were quite low and populations were increasing rapidly. This observation coupled with the apparent unimodal response of planktonic cells grown in a variety of nutrient conditions suggests that aspects of biofilm development not directly related to changes in generation time may affect plasmid copy number and lead to the observed changes in the pheromone response.

This observation was further confirmed by examining the chromosomal content of planktonic and biofilm cells. No difference was observed in the chromosomal content of planktonic and biofilm cells on a per cell basis (Figure 15). This indicates that the difference in copy number observed is not a reflection of chromosomal copy number, but rather a real copy number change.

The mathematical modeling developed by our collaborators suggested that an increase in plasmid copy number would cause an increase in negative regulators such as PrgX and iCF10. This, in turn, could cause more prominent
bistable behavior. Using pLC2, I demonstrated that the $P_Q/P_X$ promoter region and surrounding area is sufficient to see bimodal behavior in cells. As the negative regulators of the system are present in that portion of the plasmid, this result fits with the mathematical modeling prediction.

In terms of pCF10, one can postulate that the increase copy number in biofilms may arise from the need for tighter control of conjugation gene expression. Increased copy number leads to an increase in the overall negative regulators in the cell (i.e. PrgX and iCF10). This allows cells to sense a higher level of cCF10 without turning on. This would be advantageous for *E. faecalis* as it would require potential recipient cells to be in very close proximity to donor cells in order to up-regulate conjugation machinery. As *E. faecalis* cells are non-motile, the need for close cell-cell contact in the biofilm prior to conjugation prevents expenditure of energy if a recipient cell is physically touching the donor cell.

The reduced efficiency of pCF10 transfer in biofilms runs contrary to the popular notion of biofilms being the optimal niche for conjugation (69). The explanation likely relates to the anatomy of enterococci and to differences in the cell attachment mechanisms employed by the conjugative transfer machines of gram positive versus gram negative bacteria. *E. faecalis* has no known mechanisms for active motility in liquids or on solid surfaces. When planktonic cells colonize a surface and initiate biofilm growth, or attach and become part of a pre-existing biofilm, they likely remain in the same location until they either die.
or re-enter the planktonic phase by detaching from the biofilm. In the pCF10 system, mating pair formation is mediated by the surface adhesin Asc10 (encoded by prgB), which can stably bind the surfaces of cells that randomly collide. There are no extended sex pili that could attach cells that do not come into direct wall-to-wall contact. The evolution of the pheromone response system may have been driven to allow induction of conjugation in biofilms only when donor cells happen to be in extremely close proximity to recipients (perhaps in direct contact), so that the energetically expensive conjugation machinery is only produced when it can be effectively utilized. In planktonic cultures of sufficient population density, random Brownian motion increases the probability of collision between donors and recipients, and induced donors can form stable mating pairs extremely efficiently under these conditions. In this model, the selective pressures for effective spread of the plasmid are balanced by those operating to minimize the metabolic burden of synthesizing the plasmid-encoded proteins. This hypothesis may be tested experimentally by examining conjugation and the pheromone responses of individual donor cells growing in mixed biofilm communities with recipients. We are currently developing the tools such as red fluorescent chromosome reporters to carry out such experiments.

These preliminary data suggested to us that the biofilm environment could be affecting the plasmid makeup of the cell. This is potentially a very significant finding, but I first needed to see if this was a pCF10-specific finding or
a more universal phenomenon. To investigate this further, four plasmids with different backbones and representing both theta and rolling circle replication mechanisms were chosen (Table 8). These four plasmids were grown as biofilms in the CBR and biofilm and plasmid copy numbers were measured using the same protocol as pCF10. Surprisingly, all four plasmids showed significantly higher copy numbers per cell in biofilms as in planktonic cells (Figure 14).

Although this may explain the benefit of increased copy number of pCF10, it does not give any indication as to why other plasmids might undergo similar copy number up-regulation in biofilms. After scouring the literature, one other reference to increased plasmid copy number in biofilms was found. Davies et al. found that, in *P. aeruginosa* biofilms, the plasmid copy number was ~ 1.5 times higher in biofilms than in planktonic cells (70). Only one mention of this was made in the paper, published in 1995, and I could find no further indication that research has been done to explain this phenomenon. If Davies et al. was correct in their observation that copy number is also increase in *P. aeruginosa* biofilms, this phenomenon may be much more widespread than we previously imagined.

This may have many implications in the treatment of biofilm related infections as well as the prevention of antibiotic resistance spread. The physiological basis for the heterogeneity of copy number and pheromone response within the biofilm population is of great interest, and it should be feasible to carry out transcriptome analysis of sorted responsive and non-
responsive biofilm subpopulations, as has been reported for biofilm cells of
*Streptococcus mutans* exposed to a competence-inducing peptide pheromone (71).

**Donor density and iCF10: Quorum sensing in E. faecalis conjugation**

In the case of many quorum sensing systems, gene expression is turned on by quorum signals when a certain density is reached. iCF10 signals in *E. faecalis* accumulate at high densities of pCF10-containing donor cells but repress conjugation gene expression rather than activating it. The quorum signal, iCF10, serves not only as an indication of density but also as a direct regulator of conjugation via interaction with PrgX. cCF10, excreted by the plasmid-free recipient cells also acts as a type of quorum sensing molecular, allowing donor cells to sense the presence of recipient cells to initiate conjugation.

Quorum sensing is known to control a wide variety of cellular processes. In our system, data suggest that iCF10 could serve a dual purpose as both a direct negative regulator of conjugation as well as a signal to measure donor density. In this way, *E. faecalis* cells have a unique two-signal system in which cCF10 signals recipient density and iCF10 signals donor density. Both of these signals are required to turn on conjugation and it is their ratio which both signals cell density and activates or represses conjugation (Figure 22).

The data shown here suggest that, in the case of *E. faecalis*, quorum sensing using the iCF10 signal acts to shut down conjugation and thus the spread of antibiotic resistance. Previous studies have suggested that biofilms provide a
niche in which high gene transfer is present (19). The data shown in this thesis refute that assumption for *E. faecalis* biofilms (Table 7). Although the transfer frequencies in biofilm cells are decreased, increased copy number allows them to be uniquely poised to respond very quickly to changes in the cCF10:iCF10 peptide ratio in the environment. We surmise that non-motile cells in a biofilm have adapted a mechanism to precisely monitor the donor and recipient populations in closest proximity to them. This information can lead to increased spread of antibiotic resistance in a population of largely recipient cells. It can also prevent the costly expenditure of individual cell resources on non-productive transfer events or plasmid transfer when the population as a whole is already largely made up of donor cells (Figure 22).

The research presented in this thesis addresses numerous gaps in knowledge about *E. faecalis* biofilms and gene transfer. Data from the RNAseq experiments have provided us with numerous gene targets to investigate biofilm production further in enterococci and related organisms. By linking the biofilm process to conjugation, I have bridged a gap between two seemingly distinct areas of *E. faecalis* biology and pathogenesis. The overlapping systems of quorum sensing, conjugation and plasmid copy number control are now clearer although much is left to be done to fully understand the interplay between these processes.
Figure 22. Model of donor density effect on conjugation. In the left panel, the density of recipient cells greatly exceeds donor cells. In this case the iCF10 levels are decreased so conjugation levels are increased. This allows cells to spread the plasmid containing antibiotic resistance to increase the public good. In the left panel, donor density is higher causing iCF10 levels to increase. This shuts down conjugation allowing cells to decrease the energy output of the donor cells in transferring the plasmid. Thus, the quorum sensing system is a delicate balance for the population as a whole.
REFERENCES


45. Chung JW, Dunny GM. Cis-acting, orientation-dependent, positive control system activates pheromone-inducible conjugation functions at distances greater than 10 kilobases upstream from its target in *Enterococcus faecalis*. Proc Natl Acad Sci. 1992;89:9020-4.


APPENDICES

Appendix I. Unstable GFP variants do not function properly in E. faecalis cells

Four variants of gfp containing peptide sequences on the C-terminal end were tested for fluorescence and half-life in E. faecalis. These four variants were first created and tested in E. coli and P. aeruginosa and shown to have half-lives of approximately 40-60 minutes due to increased degradation of the attached peptide tags and, subsequently, the protein (46). Originally I had hoped to use these variants to study shutdown of the pheromone response.

To determine whether it would be feasible to use these unstable variants as a reporter of gene shutdown, we replaced the gfpmut3b gene in pLC2 with each of the four variants. The GFP variants were named after the last three amino acids on the C terminal peptide tags (AAV, LVA, LAA, and ASV). Cells were induced for one hour and then washed before inhibitor iCF10 was added to the culture to downshift GFP expression. In E. coli, the LVA and LAA mutants lost approximately 90% of their original fluorescence by 120 minutes post downshift (46). The ASV and AAV variants required longer, approximately 220 minutes for AAV and ASV did not lose more than ~60% fluorescence in 220 minutes (46).

In E. faecalis, cells containing the pBK2-LAA fusion, GFP fluorescence did retained wildtype GFP expression levels for 120 minutes. The other three variants, AAV, ASV, and LVA never reached a high enough level of fluorescence to be detectable by fluorometer (Figure 23). Likely the differences between gram-positive and gram-negative bacteria account for the differences in
expression of the fluorescent mutants. Unfortunately, their stability, in the case of LAA, or lack of fluorescence, in the cases of AAV, ASV, and LVA, preclude the use of these GFP variants as gene turnover reporters in our system.
Figure 23. Unstable GFP variants either remain stable or are not expressed properly in *E. faecalis* cells. One unstable GFP variant, LAA, retains high fluorescence and has the same stability as GFPmut3b. Three other GFP variants, AAV, ASV, and LVA, have little to no fluorescence following induction.
Appendix III

_Tdtomato does not currently fluoresce optimally in E. faecalis_

Although _tdtomato_ is constitutively expressed from the chromosome of OG1S, protein expression is patchy with only some of the cells appearing red under the microscope (Figure 24). Cultures were sorted and the 5000 cells with the highest fluorescence were grown up overnight again. When sorted cells were grown overnight, the same patterns of patchy expression of red fluorescence was still seen.

Up to this point, we have been unable to visualize more than ~75% of the population expressing red fluorescence. It is possible that the expression of Tdtomato is costly to the cells and thus they either cause mutations or somehow down-regulate Tdtomato expression. Current experiments examining expression of _tdtomato_ in these cells are ongoing.
Figure 24. TdTomato expression in *E. faecalis* cells is inconsistent. TdTomato is only expressed in a portion of the cells even under control of a constitutive promoter.
Appendix IV

Mathematical modeling parameters

The net rate of production of $Q_S$ and $Q_L$ RNA is given by the first term in Equations 1 and 2 respectively. Expression of PrgB protein is assumed to be proportional to $Q_L$ RNA (Equation 6). The net transcription rate of transcripts $Q_S$, $Q_L$ and $Q_{AR}$ from $P_Q$ and $X$ and $X_{AR}$ from $P_X$ is denoted by the sum of transcription rate in the uninduced and induced states weighted by concentration of looped [O] to unlooped operator sites ([N]-[O]) respectively as shown in Equations 1-9, where N is the plasmid copy number. The concentration of pCF10 DNA in the looped state ([O]) is given by Equation 10 and has been derived elsewhere (29).

The dynamics of extracellular iCF10 (i), intracellular iCF10 (I) and cCF10 (C) are shown in Equations 7-9. The transport of signaling molecules cCF10 and iCF10 across the cell membrane is modeled as a first order reaction (Equations 8-9). Generation of extracellular iCF10 is modeled as first order with respect to both $Q_S$ and $Q_L$ but not from $Q_{AR}$ as it assumed that truncated RNA does not participate in translation (Equation 7). Degradation and dilution due to cell growth is considered (Equations 1-6 and 8-9). Dilution due to growth is not considered for extracellular iCF10, however, extracellular degradation is considered in Equation 7. The ODEs shown in Equations 1-9 were solved for steady state and dynamic solution for fixed extracellular concentration of cCF10.

Steady state response was evaluated for N values from 1 to 25. A characteristic
S-shaped bistable response of PrgB to extracellular cCF10 is predicted (Figure 11B-C).

**All mathematical modeling equations, parameters, and data were derived by our collaborators Drs. Wei-Shou Hu and Anushree Chatterjee.**
### Table 9. Mathematical modeling equations

<table>
<thead>
<tr>
<th>Equation</th>
<th>Number</th>
</tr>
</thead>
</table>
| \[
\frac{d[Q]}{dt} = \left( K_{\gamma_{st}}[O] + K_{\gamma_{sa}}([N] - [O]) \right) \left( \frac{K_{\beta_{st}} - [X_{st}]}{1 + K_{\beta_{st}} [X_{st}]} \right) \cdot [-\gamma_{Q} + \mu][Q] \] | 1      |
| \[
\frac{d[Q_s]}{dt} = \left( K_{\gamma_{st}}[O] + K_{\gamma_{sa}}([N] - [O]) \right) \left( \frac{1}{1 + K_{\beta_{st}} - [X_{st}]} \right) \cdot [-\gamma_{Q} + \mu][Q] \] | 2      |
| \[
\frac{d[Q_{st}]}{dt} = \left( K_{\gamma_{st}}[O] + K_{\gamma_{sa}}([N] - [O]) \right) - K_{\alpha_{st}} [X_{st} + X] \cdot [Q_{st}] \cdot [-\gamma_{Q} + \mu][Q_{st}] \] | 3      |
| \[
\frac{d[X_{st}]}{dt} = K_{\gamma_{st}}[O] + K_{\gamma_{sa}}([N] - [O]) - [K_{\beta_{st}}[O] + K_{\beta_{sa}}([N] - [O]) \left( \frac{K_{\beta_{st}} - X_{st}}{1 + K_{\beta_{st}} - X_{st}} \right) \cdot [-\gamma_{X} + \mu][X_{st}] \] | 4      |
| \[
\frac{d[X]}{dt} = K_{\gamma_{st}}[O] + K_{\gamma_{sa}}([N] - [O]) - K_{\alpha_{st}} [X_{st}] \cdot [Q_{st}] \cdot [-\gamma_{X} + \mu][X] \] | 5      |
| \[
\frac{d[Pr_{GB}]}{dt} = K_{\gamma_{st}}[Q_{st}] \cdot [-\gamma_{Pr} + \mu][Pr_{GB}] \] | 6      |
| \[
\frac{d[I]}{dt} = K_{\gamma_{st}}([Q_{st}] + [Q_{sa}]) \cdot [V_{com}] - K_{\lambda_{st}} \cdot [I] \cdot [-\lambda_{I} + \mu][I] \] | 7      |
| \[
\frac{d[T]}{dt} = K_{\gamma_{st}}[I] \cdot [-\lambda_{I} + \mu][I] \] | 8      |
| \[
\frac{d[C]}{dt} = K_{\gamma_{st}}[C] \cdot [-\lambda_{C} + \mu][C] \] | 9      |
| DNA loop = [O] = \[
\frac{[N][I]}{[I]^2 + K_{\beta_{st}}[C]^2} \] | 10     |
### Table 10: List of variables and parameters used in the mathematical model

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O$</td>
<td>DNA of plasmid in loop form</td>
</tr>
<tr>
<td>$N$</td>
<td>Plasmid copy number, $Q$</td>
</tr>
<tr>
<td>$Q_S$</td>
<td>$Q_S$ mRNA</td>
</tr>
<tr>
<td>$Q_L$</td>
<td>$Q_L$ mRNA</td>
</tr>
<tr>
<td>$X_{AR}$</td>
<td>Truncated $P_X$ RNA interacting with $Q$</td>
</tr>
<tr>
<td>$X$</td>
<td>$X$ mRNA</td>
</tr>
<tr>
<td>$Q_{AR}$</td>
<td>Truncated $P_Q$ RNA interacting with $X$</td>
</tr>
<tr>
<td>$i$</td>
<td>Extracellular iCF10</td>
</tr>
<tr>
<td>$I$</td>
<td>Intracellular iCF10</td>
</tr>
<tr>
<td>$c$</td>
<td>Extracellular cCF10</td>
</tr>
<tr>
<td>$C$</td>
<td>Intracellular cCF10</td>
</tr>
<tr>
<td>$PrgB$</td>
<td>PrgB protein</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value used for bistability</th>
<th>Units</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>Specific growth rate of donor cells</td>
<td>2.58×10^{-4}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{Q,U}$</td>
<td>Transcription rate constant of $Q$ in uninduced state</td>
<td>7.50×10^{-3}</td>
<td>s^{-1}</td>
<td>(29)</td>
</tr>
<tr>
<td>$K_{Q,I}$</td>
<td>Transcription rate constant of $Q$ in induced state</td>
<td>8.87×10^{-2}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{X_{AR},U}$</td>
<td>Transcription rate constant of $X_{AR}$ in uninduced state</td>
<td>1.02×10^{-2}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{X_{AR},I}$</td>
<td>Transcription rate constant of $X_{AR}$ in induced state</td>
<td>1.20×10^{-3}</td>
<td>s^{-1}</td>
<td>(29)</td>
</tr>
<tr>
<td>$K_{X,U}$</td>
<td>Transcription rate constant of $X$ in uninduced state</td>
<td>1.20×10^{-3}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{X,I}$</td>
<td>Transcription rate constant of $X$ in induced state</td>
<td>1.20×10^{-3}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{Q_{AR},U}$</td>
<td>Transcription rate constant of $Q_{AR}$ in uninduced state</td>
<td>1.80×10^{-3}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{Q_{AR},I}$</td>
<td>Transcription rate constant of $Q_{AR}$ in induced state</td>
<td>1.53×10^{-3}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{PrgB}$</td>
<td>Generation rate of PrgB protein</td>
<td>1×10^{-2}</td>
<td>s^{-1}</td>
<td>(72, 73)</td>
</tr>
<tr>
<td>$K_i$</td>
<td>Generation rate of extracellular iCF10</td>
<td>1×10^{-3}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$V_{conv}$</td>
<td>Volume conversion factor</td>
<td>1</td>
<td>-</td>
<td>(29)</td>
</tr>
<tr>
<td>$K_i$</td>
<td>Transport rate constant of iCF10</td>
<td>1×10^{-3}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_c$</td>
<td>Transport rate constant of cCF10</td>
<td>1×10^{-3}</td>
<td>s^{-1}</td>
<td></td>
</tr>
<tr>
<td>$K_{AR}$</td>
<td>Rate constant of interaction between $X$ RNA and $Q_{AR}$ RNA</td>
<td>1×10^{-3}</td>
<td>(nM)^{1/2}, s</td>
<td>(74, 75)</td>
</tr>
<tr>
<td>$K_{eq}$</td>
<td>Equilibrium constant of Q and X$_{AR}$ interaction</td>
<td>4.43</td>
<td>(nM)$^{-1}$</td>
<td>(74, 75)</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------</td>
<td>------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>$K_b$</td>
<td>Equilibrium constant of DNA binding reaction</td>
<td>1×10$^8$</td>
<td>-</td>
<td>(22, 29)</td>
</tr>
<tr>
<td>$\lambda_{Qc}$</td>
<td>Degradation rate of Q$_c$ mRNA</td>
<td>1×10$^{-3}$</td>
<td>s$^{-1}$</td>
<td>(72)</td>
</tr>
<tr>
<td>$\lambda_{Ql}$</td>
<td>Degradation rate of Q$_l$ mRNA</td>
<td>1×10$^{-1}$</td>
<td>s$^{-1}$</td>
<td>(72)</td>
</tr>
<tr>
<td>$\lambda_{Xa}$</td>
<td>Degradation rate of X$_{AR}$ RNA</td>
<td>1.36×10$^{-4}$</td>
<td>s$^{-1}$</td>
<td>(76)</td>
</tr>
<tr>
<td>$\lambda_{Xl}$</td>
<td>Degradation rate of Q$_{AR}$ RNA</td>
<td>1×10$^{-3}$</td>
<td>s$^{-1}$</td>
<td>(72)</td>
</tr>
<tr>
<td>$\lambda_{X}$</td>
<td>Degradation rate of X RNA</td>
<td>1×10$^{-4}$</td>
<td>s$^{-1}$</td>
<td>(72)</td>
</tr>
<tr>
<td>$\lambda_{Pr}$</td>
<td>Degradation rate of PrgB protein</td>
<td>1×10$^{-6}$</td>
<td>s$^{-1}$</td>
<td>(73)</td>
</tr>
<tr>
<td>$\lambda_i$</td>
<td>Degradation rate of intracellular iCF10</td>
<td>1×10$^{-6}$</td>
<td>s$^{-1}$</td>
<td>(73)</td>
</tr>
<tr>
<td>$\lambda_i$</td>
<td>Degradation rate of extracellular iCF10</td>
<td>1×10$^{-6}$</td>
<td>s$^{-1}$</td>
<td>(73)</td>
</tr>
<tr>
<td>$\lambda_c$</td>
<td>Degradation rate of intracellular cCF10</td>
<td>1×10$^{-6}$</td>
<td>s$^{-1}$</td>
<td>(73)</td>
</tr>
</tbody>
</table>