Transcriptional regulation of incompatibility type A/C plasmids

A Thesis
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Kevin S. Lang

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

Timothy J. Johnson, Randall S. Singer

May 2015

© Kevin S. Lang 2015

Acknowledgements

So many people to thank, so little space in which to do so.

First, I'd like to thank my adviser, Tim Johnson. He chose to take me into his lab when most would have chosen not to and provided unrelenting support along the way.

Many thanks to my other committee members, Randall Singer, Michael Murtaugh and Gary Dunny. Their guidance and conversations were absolutely crucial to my success.

My mother deserves some credit for this work as well. She taught me how to use a microscope when I was six. Thus began a lifetime of curiosity.

I would like to thank Jonathan B. Clayton, the only other graduate student during my tenure in the Johnson Lab. We worked very closely on some projects and it was a pleasure to work with someone who was like-minded in terms of goals and vision. Hopefully we will have many opportunities to work together as our careers move forward.

Many thanks to the rest of the Johnson Lab for their help and support. And many thanks to the many folks in the Veterinary Biomedical Sciences department and the Comparative and Molecular Biosciences program. I've probably received help from every single person at some point, far too many times for my memory to recall.

I would like to thank my better half, Lindsey, who supported me through this entire endeavor. I'm not sure I would say I couldn't do it without her, but it surely wouldn't have been as fun.

Finally, I thank the microbes. This is their world, we just happen to live in it. They hold all of the answers, you just need to ask the right questions[1].

[1] I don't take credit for this line, it is a play on a quote that I cannot recall the author of: "Biology has all the answers, you just need to know what question to ask"

Dedication

This thesis is dedicated to my undergraduate adviser in the College of Biological Sciences, Sarah Corrigan, who told me I would never get in to graduate school.

Abstract

Plasmids are extrachromosomal DNA elements that often carry beneficial phenotypes for the bacterial host. Incompatibility type A/C (IncA/C) plasmids are large (~100 ~ 200 kilobases (kb)), conjugative plasmids that are carried by Gram-negative bacteria. IncA/C plasmids often carry numerous genes that confer resistance to antimicrobials and have been isolated from many types of bacteria that pose significant risks in both human and animal medicine.

Although IncA/C plasmids have been described in the literature for many years, little is known about their basic biology. For the past decade, many fully sequenced IncA/C variants have been described. There has been a lack of work concerning core functions of these plasmids, such as, replication, conjugative transfer and maintenance. This dissertation focuses on how these plasmids are regulated on a transcriptional level.

We used pAR060302, a prototypical IncA/C plasmid, to conduct several experiments investigating exactly what genes are transcribed and how different conditions affect their transcriptional landscape. RNA-Seq was used to understand how antimicrobial exposure can affect the way genes are expressed on IncA/C plasmids. We found that, under the conditions we tested, antimicrobials have little effect on the transcription of genes on pAR060302. However, this initial study was the first to explore genes carried IncA/C plasmids in terms of their expression.

Further RNA-Seq experiments were carried out in several different bacterial genera all carrying the same IncA/C plasmid, pAR060302. These experiments attempted to characterize how a broad host-range plasmid, such as IncA/C, might be differentially

regulated in different hosts. We found that only subtle changes occur in the expression of plasmid genes. Carriage of IncA/C plasmids was found to have diverse effects on chromosomal gene expression. Genes involved in 2-carbon metabolism in *Escherichia coli* are up-regulated due to plasmid carriage. Our results suggest that plasmid encoded factors might serve varying levels of importance depending on the host chromosomal background.

Experiments were carried out on *E. coli* carrying plasmids with mutations in a group of predicted transcriptional regulators to determine their function. We identified the positive regulators of conjugative transfer in IncA/C plasmids, *acaD* and *acaC*. We also found a repressor of transfer, *acr2*, which encodes an H-NS-like protein. We further show that *acr2* might regulate genes beyond those that are involved in conjugative transfer.

This dissertation builds on our understanding on what mechanisms are important for the maintenance of large plasmids in Gram-negative bacteria. Understanding how plasmids might specifically tune host metabolism to improve competitive fitness would impact what evolutionary processes were involved in their emergence. Characterization of regulatory networks that govern core plasmid processes, such as conjugation, might assist in the development of new models of how these plasmids disseminate throughout bacterial populations.

Table of Contents

Acknowledgments i

Dedication iii

Abstract iv

Table of Contents vi

List of Tables vii

List of Figures viii

Chapter 1: Introduction 1

Chapter 2: Transcriptome mapping of pAR060302, a blaCMY-2-positive broad host range IncA/C plasmid 8

Summary 9

Introduction 10

Results 13

Discussion 23

Materials and Methods 31

Chapter 3: Transcriptome modulations due to IncA/C plasmid acquisition 34

Summary 35

Introduction 36

Results 37

Discussion 50

Materials and Methods 53

Chapter 4: A model for the regulation of IncA/C plasmid conjugation 56

Summary 57

Introduction 58

Results 62

Discussion 77

Materials and Methods 82

Discussion and Future Directions 98

Bibliography 105

Appendix A 117

List of Tables

Table 1: RPKM values calculated for transcribed genes of interest. 15

Table 2: Fold change and p-value of genes on pAR060302...41

Table 3: COGs enriched in differentially expressed genes from strains carrying pAR060302 compared to plasmid free strains 45

Table 4: Details of RNA-Seq reads obtained 55

Table 5: Strains and plasmids used in this study 84

Table 6: Oligonucleotides used in this study 85

Table 7: Susceptible colonies of *E. coli* K12 carrying either pAR060302 or pAR060302Δ*acr2* after passing 50 days in LB with no antibiotic selection 118

Table 8: Susceptible colonies of *Salmonella* Newport carrying either pAR060302 or pAR060302Δ*acr2* after passing 50 days in LB with no antibiotic selection 119

Table 9: PCR results of 3 ampicillin susceptible colonies of K-12 (pAR060302Δacr2) 120

List of Figures

Figure 1: Heat map displaying transcriptome map of pAR060302 14

Figure 2: Transcription of regions not previously annotated 16

Figure 3: Transcriptome maps of plasmid backbone genes of interest. 18

Figure 4: Transcriptome map of the *sul2* containing region of pAR060302 20

Figure 5: Transcriptome map of the *bla*$_{\text{CMY-2}}$ containing Tra1 region of pAR060302 21

Figure 6: Transcriptome map of the class I integron containing region of pAR060302...22

Figure 7: Transcriptome map of pAR060302 carried by different hosts 38

Figure 8: Differential expression data from strains DH10 and MR-1 overlaid onto the glyoxylate bypass pathways 48

Figure 9: Transcriptome map of *S. oneidensis* megaplasmid in MR-1 with and without pAR060302 49

Figure 10: Model of SXT regulation of transfer 59

Figure 11: Development of pARK01 63

Figure 12: Regulation of conjugation locus of IncA/C plasmids 65

Figure 13: Conjugation frequency of *acr2* mutants..67

Figure 14: *acr2* regulon in pARK01 69

Figure 15: Expression of *E. coli* chromosomal genes carrying either pAR060302 or pAR060302Δ*acr2*......70

Figure 16: Acr2 binds promoter region of *acr1* 72

Figure 17: Predicted secondary structure of the 5' end of the predicted *acr1* transcript 73

Figure 18: Development of Acr2 variants......74

Figure 19: Activity of Acr2 variants......76

Figure 20: Proposed model for the regulation of IncA/C conjugation 78

Figure 21: Schematic for the development of *acr1-lacZ* fusion reporter......95

Chapter 1

General Introduction[2]

[2] A version of this chapter was published in:
T. J. Johnson and Kevin S. Lang. IncA/C plasmids: An emerging threat? Mobile Genetic Elements. 2012. 2:55-58

Since the discovery of the first multidrug-resistant *Shigella* strains in 1955, plasmids encoding multidrug resistance have been extensively studied and appreciated for their diversity and plasticity (1). Incompatibility type A/C (IncA/C) plasmids are just one of many Inc types. They were first described more than 4 decades ago, but are only now being intensively studied. IncA/C plasmids were first identified among multidrug resistant *Aeromonas hydrophila* and *Vibrio spp.* causing disease in cultured fish in the 1970s, where it was hypothesized that such plasmids emerged in response to the common practice of therapeutic antibiotic treatment in response to disease (2, 3). These are the first reports of IncA/C plasmids being associated with animal agriculture practices.

In the 1990s, the plasmid-encoded β-lactamase gene, $bla_{\text{CMY-2}}$ was identified among *Klebsiella pneumoniae* (4), and several reports in the early 2000s identified $bla_{\text{CMY-2}}$ on a conjugative element in *Escherichia coli* and *Salmonella enterica* from food-producing animals (5–10). Soon following this, $bla_{\text{CMY-2}}$ was found to be harbored by *S. enterica* and *K. pneumoniae* that were causing disease in humans (11, 12). As the genome sequencing era commenced, several fully sequenced plasmids were generated, enabling comparative plasmid genomics. It became increasingly evident that $bla_{\text{CMY-2}}$ was often carried on IncA/C plasmids, and that these plasmids also encoded resistance to numerous other antimicrobials (13–16). Studies by the U.S. Food and Drug Administration examining clinical isolates from 1940-2000 showed that IncA/C plasmids and their associated multidrug resistance phenotypes emerged in *S. enterica* serovar Newport isolates after 1980. Taken together, the literature suggests that IncA/C plasmids conferring resistance

to antimicrobials emerged in aquatic bacteria and subsequently moved into enteric bacteria carried by terrestrial animals. The problem with this conclusion is that strains carrying IncA/C plasmids have historically been isolated by screening for resistance phenotypes first, followed by identification of the IncA/C replicon. Still, Gram-negative pathogens harboring IncA/C plasmids are now common, and these plasmids are notorious for carrying genes that confer resistance to multiple classes of antimicrobials.

How IncA/C plasmids were able to rapidly disseminate throughout genera of enteric bacteria is still an open question. Certainly, their ability to be carried by many different species of Gammaproteobacteria is probably a factor (17). There are some IncA/C variants that have possibly lost the ability to transfer by conjugation (18), however, the origin of transfer has been described and can facilitate mobilization by *trans* acting conjugative machinery (19). There are documented cases of the short-term ability of these plasmids to disseminate in hospital settings. For example, a 2-year outbreak study in a Tunisian hospital involving multidrug-resistant *Providencia stuartii* identified IncA/C plasmid variants that had circulated among three bacterial clones (20).

IncA/C plasmids are unique among other plasmid types in enteric bacteria. They seem to be a chimera of the remnant of an integrative conjugative element similar to that of SXT/R391 (21) and an iteron-containing theta replicon. Comparative sequence analyses of IncA/C plasmids have revealed that their backbone is punctuated by several hotspots for integration of a variety of mobile genetic elements such as IS*26*, IC*Ecp1*, and

class 1 integrons (15, 16, 22–25). These mobile elements carry with them a variety of genes that confer the resistance to multiple classes of antimicrobials. Given these hotspots, the potential for gene shuffling with this plasmid type is high, and this has certainly been observed in comparative genomics studies (16, 23, 26, 27). IncA/C plasmids now have been sequenced from isolates of *E. coli*, *S. enterica*, *Vibrio cholerae*, *Yersinia pestis*, *Photobacterium damselae* subsp. *piscicida*, *Yersinia ruckeri*, and *Klebsiella pneumoniae*. Multiple lineages of IncA/C plasmids now have subsequently been identified, however, recent papers have suggested that there are simply two major lineages, one containing the original pRA plasmid from aquatic pathogens, and a second containing all contemporary plasmids (26). The low variability in the backbone sequences, combined with the multitude of combinations of mobile elements that have been described in the hotspot regions, leads one to believe that the all of the different variants of IncA/C plasmids have rapidly evolved from a common ancestor. For example, IncA/C plasmids acquiring the bla_{NDM-1} gene encoding the New Delhi metallo-β-lactamase have successfully disseminated among clinical *K. pneumoniae* and *E. coli* originating in India (28, 29). Therefore, it seems likely that both horizontal gene transfer and clonal dissemination must play an important role in the success of IncA/C plasmids, although the abundance of these plasmids in the environment and their full range of bacterial hosts is not yet appreciated.

Initial comparative analyses concluded that the backbones of IncA/C plasmids were upwards of 99% conserved on the nucleotide level, with major differences being only in

the resistance-associated insertion elements in their hotspot regions. This remains largely true amongst plasmids originating from enteric bacteria. With the growing number of fully sequenced IncA/C plasmids from other sources, both published and unpublished, that number is down to 72-94% (26, 30). The diversity of IncA/C plasmids is likely to be underappreciated. This is likely due to a primary focus on plasmids conferring resistance to multiple antimicrobials. If this is true, the high conservation among IncA/C plasmids of enteric bacteria indicates a common ancestral plasmid origin and subsequent rapid dissemination.

IncA/C plasmids possess a number of putative transcriptional regulators with sequence similarity to proteins of the H-NS, HU-β, Xre, LysR, GntR and LuxR families (16). The presence of such proteins on large, low copy plasmids is not unexpected. Other large plasmids, such as IncP, IncH, and IncX, also encode H-NS-like proteins and other transcriptional regulators (31–33). Some of these proteins have been shown to exhibit silencing effects on the bacterial chromosome, suggesting that this might be a key activity in the successful dissemination of these plasmids to a broad range of bacterial hosts. However, the high number of putative regulatory genes that IncA/C plasmids possess suggests that strict regulation of both the plasmid and naïve host chromosome is critical for their broad-host dissemination. Interestingly, the closest database matches to these protein sequences are from non-enteric bacteria such as *Vibrio spp.* and *Aeromonas spp.*, indicating that a possible origin of this plasmid type could have been soil or water bacteria (16). One could speculate that these plasmids are well suited to persist in certain

subsets of reservoir bacteria, in which they provide a lower imposed fitness cost to their bacterial host and are better able to persist in a microbial community. It has been demonstrated that IncA/C plasmids in hosts such as *E. coli* and *Salmonella* might require selective pressure for their long-term persistence (34). However, if other reservoir bacteria are better suited to carry IncA/C plasmids regardless of antimicrobial selection, they may persist for long periods of time. This would afford IncA/C plasmids the opportunity to reside in a microbial community in the absence of benefit to an enteric host, to persist, and to subsequently emerge in pathogens.

Many questions remain related to IncA/C plasmid biology. Until very recently, none of these aspects have been functionally characterized. Early studies identified the minimum replicon of IncA/C plasmids, showing that the *repA* gene was functional and was sufficient and necessary for replication (24, 25). Only two other studies, both published during the preparation of this thesis, have examined IncA/C plasmids in molecular detail. One detailed the origin of transfer (19) and the other focused on the regulation of transfer (35). This thesis describes the results of three studies that aim to further characterize IncA/C plasmids in detail. Chapters 2 and 3 use total RNA sequencing (RNA-Seq) to map and characterize global gene expression in cells carrying an IncA/C plasmid. Specifically, Chapter 2 looks at the effect of antimicrobials on plasmid gene transcription. Chapter 3 attempts to better understand the influence that host chromosomal background has on plasmid gene transcription, as well as what effect plasmid carriage has on chromosomal gene transcription. In Chapter 4 we characterize

the function of a set of predicted transcriptional regulators encoded on IncA/C plasmids and their role in the regulation of conjugation.

The results presented in this thesis provide an outline for studying IncA/C plasmid regulatory networks in molecular detail. This thesis merely scratches the surface of such networks, as there are many more predicted DNA binding proteins encoded on IncA/C plasmids whose functions have yet to be explored.

Chapter 2

Transcriptome mapping of pAR060302, a *bla*$_{CMY-2}$-positive broad host range IncA/C plasmid[3]

[3] This chapter has been previously published in:
Kevin S. Lang, J. L. Danzeisen, W. Xu, and T. J. Johnson.Transcriptome mapping of pAR060302, a *bla*$_{CMY-2}$ positive IncA/C plasmid. Applied and Environmental Microbiology. 2012 78:3379-3386.

Summary

The multidrug resistance-encoding plasmids belonging to the IncA/C incompatibility group have recently emerged among *Escherichia coli* and *Salmonella enterica* in the United States. These plasmids have a unique genetic structure, a broad host range, and propensity to acquire large numbers of antimicrobial resistance genes via their accessory regions. Using *E. coli* strain DH5α harboring the prototype IncA/C plasmid pAR060302, we sought to define the baseline transcriptome of IncA/C plasmids under laboratory growth and in the face of selective pressure. The effects of ampicillin, florfenicol or streptomycin exposure were compared to cells left untreated at logarithmic phase using Illumina sequencing (RNA-Seq). Under growth in Luria-Bertani broth lacking antibiotics, much of the backbone of pAR060302 was transcriptionally inactive, including its putative transfer regions. A few plasmid backbone genes of interest were highly transcribed, including genes of a putative toxin-antitoxin system and an H-NS-like transcriptional regulator. In contrast, numerous genes within the accessory regions of pAR060302 were highly transcribed, including the resistance genes *floR*, $bla_{\text{CMY-2}}$, *aadA*, and *aacA*. Antibiotic treatment with ampicillin or streptomycin resulted in no genes being differentially expressed compared to controls lacking antibiotics, suggesting that many of the resistance-associated genes are not differentially expressed due to exposure to these antibiotics. In contrast, florfenicol treatment resulted in the up-regulation of *floR* and numerous chromosomally encoded genes. Overall, the transcriptome mapping of pAR060302 suggests that it mitigates the fitness costs of carrying resistance-associated genes through global regulation with its transcriptional regulators.

Introduction

Antimicrobial resistance is a growing public health concern, as the emergence and rapid dissemination of multidrug resistance (MDR) phenotypes among human and animal pathogens presents limited options for disease treatment. The spread of MDR-encoding genetic determinants has been attributed to indiscriminate use of antimicrobial compounds in animal agriculture and human health, although it is still unclear if there are factors aside from antimicrobial selection pressure that drive the persistence of these resistance genes (36–38). Often these genes reside on mobile genetic elements that have been circulating among bacteria for far longer than antibiotics have been used by humans(39, 40). Horizontal gene transfer of these elements provides a mechanism for rapid spread of MDR among bacteria, thus the factors driving the persistence and dissemination of these mobile elements play a large role in the circulation of resistance genes in the absence of selection pressure.

Horizontal gene transfer plays a significant role in the introduction of new genes into microbial populations (41). Conjugative plasmids, integrons and transposons all contribute as vehicles for this process. Plasmids are extrachromosomal genetic elements that orchestrate their own replication. In some cases they are able to move via conjugative transfer. The sizes of plasmids can vary widely (1.5 kb to >600kb) and are generally considered modular in nature. Core components of plasmids are required for their replication, stability and, in some cases, their transfer. The acquisition of accessory elements by plasmids can result in a wide variety of phenotypes to their host recipient,

such as antimicrobial and heavy metal resistance, metabolism of organic compounds, and virulence (42–45). Plasmids belonging to the A/C incompatibility plasmid group (IncA/C) are a mosaic of a highly conserved core backbone with variable accessory elements. These are large (ca 80 – 160 kb) broad-host-range plasmids and have been isolated from diverse groups of *Proteobacteria* found in the environment, food animals, food products and human pathogens (43, 46–48). IncA/C plasmids are of particular concern because they carry genes that encode resistance to many different antimicrobial compounds, including $bla_{\text{CMY-2}}$, which encodes an AmpC-like β-lactamase that confers resistance to 3rd-generation cephalosporins (49). The physical linkage of this gene to other resistance genes encoded on IncA/C plasmids presumably allows it to persist in the absence of cephalosporin use. The ability of IncA/C plasmids to spread $bla_{\text{CMY-2}}$ and co-transferring resistance genes to a diverse range of bacteria poses a threat to both human and animal health.

Numerous IncA/C plasmids have been fully sequenced (15, 16, 23), though little is known about their basic biology. One such plasmid, pAR060302, is a ~166 kb MDR-encoding plasmid that was isolated from a dairy cow *Escherichia coli* strain AR060302 in 2003 (16). The backbone of this plasmid appears to contain genes required for replication, maintenance, and conjugative transfer. Additionally, three distinct accessory elements are encoded on this plasmid: an IS*CR2*-containing element that carries the *floR, tetA, strBA,* and *sul2* genes; an IS*Ecp1*-containing element that carries the $bla_{\text{CMY-2}}$ gene; and a Tn*21*-like, class I integron-containing element that carries several different gene

cassettes, including *aadA*, which confers aminoglycoside resistance. Although many of the phenotypes have been described for this plasmid, the extent to which these genes are transcribed and regulated remains unknown. Similar to many other plasmids, the transcriptional pattern of the plasmid backbone genes and accessory elements are likely regulated by a combination of plasmid encoded transcriptional regulators and host factors (31). Environmental conditions could also play a role in this regulation. The magnitude of the transcription of these genes can impact phenotypes encoded on these plasmids (50). Due to the presence of multiple antimicrobial resistance genes on IncA/C plasmids, their propensity to acquire new genes, their broad host range, and their rapid dissemination, there is an immediate need to understand the underlying mechanisms of this plasmid's regulation.

The present study aimed to characterize the transcriptome of pAR060302, a prototypical bla_{CMY-2} positive IncA/C plasmid, under laboratory conditions using Illumina based RNA sequencing methods (RNA-Seq). We hypothesized that the level of transcription would vary between plasmid backbone genes and those encoded on accessory elements and that the transcriptional landscape of pAR060302 would differ under antimicrobial selection pressure. Therefore, we mapped the transcriptome of pAR060302 to provide initial insights on the regulation of gene expression of this plasmid.

Results

Transcription of the plasmid core regions. Much of the putative backbone of pAR060302 was transcriptionally inactive under the conditions examined (Figure 1). Only about 18 percent of the coding sequences located in core regions were transcribed. The *repA* replication initiation gene was expressed at a low level (RPKM of ~3700-4700). This gene was used for subsequent comparisons because it is involved with plasmid replication and was assumed to be expressed consistently during exponential growth due to continuous cell division. Just one other locus, containing ORF_051-053, was expressed on the backbone of the plasmid in a similar pattern to *repA* (Table 1). BLASTp analysis of ORF_051-053 identified two of these putative proteins as having functions pertaining to stability and maintenance. Two additional regions were transcribed at high levels, but had not previously been annotated (Figure 2). These two regions have short predicted ORFs (< 65 amino acids) however it is unclear which small peptides, if any, are expressed by these transcripts. Nucleotide BLAST (megablast) analysis of the regions showed similarity only to other IncA/C plasmids, which indicates that it is unlikely that these reads were mapped incorrectly. These regions did not have similarity to the host *E. coli* genomic sequences and therefore did not originate from chromosomal transcripts.

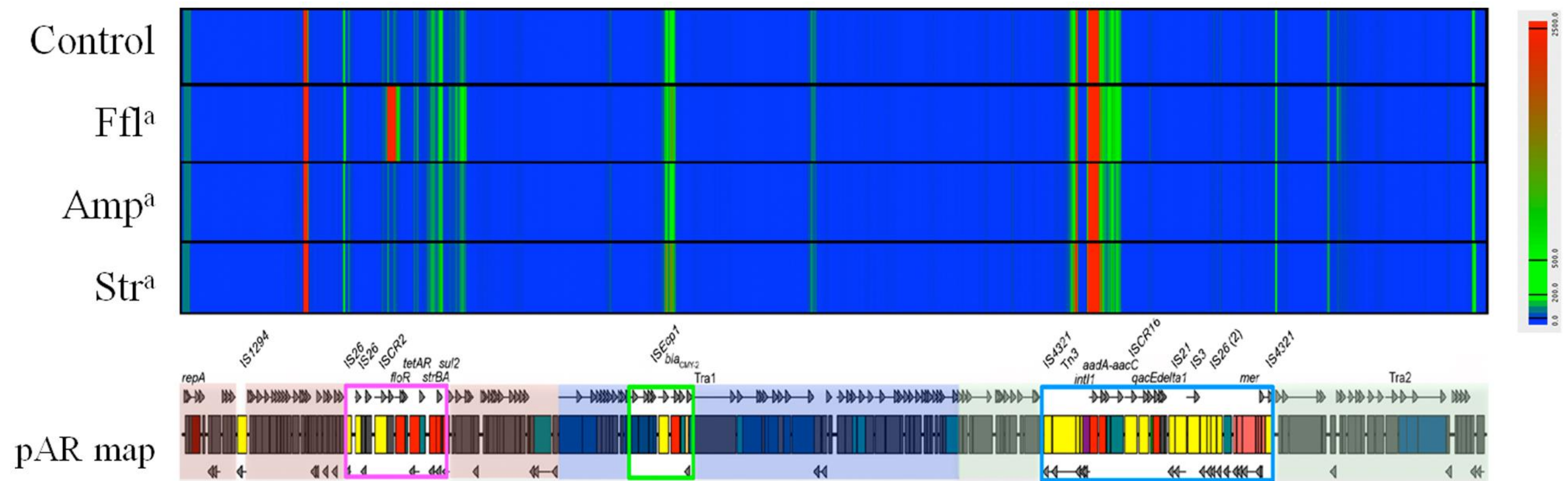

Figure 1: Heat map displaying transcriptome map of pAR060302. Blue indicates no or low expression (0-75 reads aligned per nucleotide position), green indicates moderate expression (250-500) reads aligned per nucleotide position), and red indicates high expression (>2500 reads aligned per nucleotide position). In the pAR060302 (pAR) map, the shaded regions indicate the plasmid backbone as follows: red (IncA/C replicon and hypothetical genes), blue (Tra1 region), and green (hypothetical genes and Tra2 region). Open boxes depict the accessory regions as follows: pink (*flo-tet-sul* containing region; enlarged in figure 4), green (*bla*$_{CMY-2}$ containing region; enlarged in figure 5), and blue (Tn*21*-like region which contains a class 1 integron; enlarged in figure 6). [a]Antibiotics are abbreviated as follows: ampicillin (Amp), florfenicol (Ffl), streptomycin (Str).

Table 1: RPKM values calculated for transcribed genes of interest. RPKM is calculated as reads aligned per kilobase of gene per millions of total reads.

	Genes or ORFs	**Function or predicted function**	RPKM values			
			Control	Ffl[a]	Amp[a]	Str[a]
Putative maintainence and stability genes	*repA*	replication initiation	4376	3780	4772	4753
	pAR060302_0051	putative ParA partition protein	6942	11,306	7534	7066
	pAR060302_0052	putative ParB partition protein	14,265	20,154	13,480	14,939
Putative DNA binding genes	pAR060302_0023	putative IHF transcriptional regulator	5130	2097	5583	5243
	pAR060302_0025	putative antitoxin component protein	206,070	361,900	201,615	203,271
	pAR060302_0026	putative toxin component protein	86,543	148,982	85,362	81,898
	pAR060302_0027	putative DNA binding protein	3176	3074	3431	3769
	pAR060302_0188	putative HNS-like protein	23,351	14,876	26,414	27,179
Resistance genes	*floR*	phenicol resistance	8543	71,346*	8463	8845
	tetA	tetracycline resistance	991	1544	958	993
	bla_{CMY-2}	extended spectrum beta-lactamase	34,314	35,415	32,488	32,308
	aadA	aminoglycoside resistance	323,889	217,434	334,860	330,472
	aacA	aminoglycoside resistance	64,841	61,992	56,338	61,287
Genes within the putative transfer loci	pAR060302_0075	hypothetical protein	8277	3343	8070	8170
	pAR060302_0097	hypothetical protein	13,364	6383	14,597	13,921
	pAR060302_0183	hypothetical protein	3452	5266	3985	4104

[a]Antibiotics are abbreviated as follows: ampicillin (Amp), florfenicol (Ffl), streptomycin (Str).

*Indicates a > 3 fold change in expression that is statistically significant (q-value < 0.05)Its location is similar to the locations of the regulators of the transfer regions of SXT-like integrative conjugative elements (ICEs) that share significant homology with IncA/C plasmids (41). However, no work has been done previously to characterize this gene.

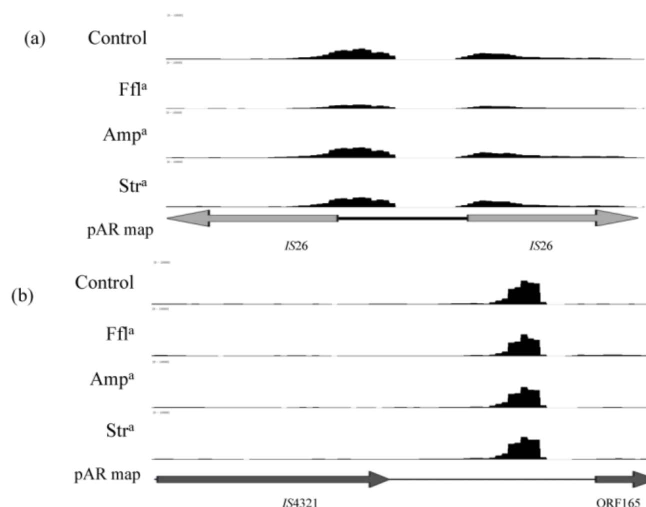

Figure 2: Transcription of regions not previously annotated. (a) Region between two IS*26* elements upstream of the IS*CR*2 elements. (b) Intergenic transcript downstream of the class I integron. Antibiotic treatment conditions are labeled along the x-axis. The x-axis has a scale of 0 – 2500 reads aligned per base. Arrows indicate direction of transcription. [a]Antibiotics are abbreviated as follows: ampicillin (Amp), florfenicol (Ffl), streptomycin (Str).

The majority of the genes located in the Tra1 and Tra2 regions of pAR060302 were not transcribed under the conditions we assayed. Only three genes within these loci, ORFs 75, 97 and 183, were transcribed under the conditions examined here (Table 1). The expression levels of these genes were comparable to the level of *repA* and are all annotated as hypothetical proteins. The gene locus pAR060302_183 is annotated as a hypothetical protein. This ORF is conserved among IncA/C plasmids, but shares no similarity to proteins with known functions.

A few genes on the backbone of pAR060302 were expressed at high levels. One set of genes transcribed at a level 47- and 20-fold higher than *repA* was pAR060302_25 and pAR060302_26, respectively (Table 1, Figure 3a), which display BLAST similarity genes of a toxin-antitoxin system, although this system has not been experimentally characterized as having toxin-antitoxin function. pAR060302_25 is a putative XRE-like transcriptional regulator and the putative antitoxin component, while pAR060302_26 is a putative toxin component of the system. These predicted proteins share amino acid similarity (70%) to a chromosomal locus found in *Salmonella enterica* serovar Paratyphi B strain SPB7. A third gene, pAR060302_188, was transcribed at a level 5-fold higher than *repA* (Table 1, Figure 3b) and is annotated as a putative H-NS-like DNA-binding protein. Analysis of the predicted amino acid sequence of pAR060302_0188 showed 52% sequence similarity to its closest match and did not show any significant similarity to other H-NS proteins encoded on other gram-negative plasmids; however it is conserved among all sequenced IncA/C plasmids (10).

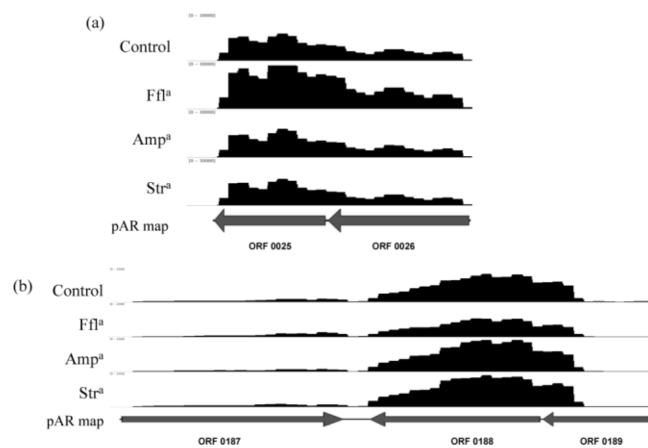

Figure 3: Transcriptome maps of plasmid backbone genes of interest. (a) pAR060302_0025 and pAR060302_0026 are transcribed at extremely high levels. The x-axis has a scale of 0 – 30,000 reads aligned per nucleotide position. Arrows indicate direction of transcription. (b) pAR060302_0188 is transcribed as a single gene transcript. The x-axis has a scale of 0 – 2500 reads aligned per nucleotide position. Arrows indicate direction of transcription.

[a]Antibiotics are abbreviated as follows: ampicillin (Amp), florfenicol (Ffl), streptomycin (Str).

Transcription pattern of the pAR060302 accessory regions. There are at least three hotspots for insertion of mobile genetic elements on IncA/C plasmids (6, 10, 12). In pAR060302, the first hotspot carries an ISCR*2*-containing mobile genetic element that contains the *floR*, *tetRA*, *strBA*, and *sul2* resistance gene cluster; as well as several putative genes with predicted regulator capabilities. This inserted module is highly conserved among bla_{CMY-2} positive IncA/C plasmids (6). The *floR* gene was transcribed on a transcriptional unit that includes three ORFs and this transcript was apparently promoted by a dual promoter system (Figure 4). A transcriptional start site (TSS) was located upstream of ORF_0040, a conserved hypothetical protein, and an inducible promoter was located just upstream of the *floR* ORF. The transcription of *floR* was increased ~8 fold (q-value < 0.05) from this promoter site after exposure to florfenicol for 30 min. Immediately downstream of *floR*, pAR060302_0042 was also up-regulated. This gene likely resides on the same transcript give the pattern of aligned reads that was observed (Figure 4). pAR060302_0042 is annotated as a hypothetical protein. Blastp

analysis of the predicted amino acid sequence of this ORF showed a predicted helix-turn-helix domain and shares homology with the LysR family of transcriptional regulators. No changes in transcription of this region were observed in the presence of ampicillin or streptomycin. Located downstream of the *floR* locus on the opposite strand is the *tetRA* gene cluster, one of the most abundant genetic factors for antimicrobial resistance in Gram-negative bacteria. The transcriptional regulation of *tetRA* has been previously characterized (51) and similar phenomena were observed here. That is, only the repressor gene *tetR* was transcribed while the *tetA* gene was under repression. This gene cluster has been shown previously to be up-regulated in response to tetracycline exposure (51) but was unaffected by the antimicrobial compounds we tested here. The *strBA* and *sul2* genes, encoding streptomycin and sulfonamide resistance respectively, are located on the IS*CR2* element upstream of *tetRA* and were transcribed together on the same transcript originating with a TSS located just upstream of *sul2*. Florfenicol, ampicillin or streptomycin treatment had no effect on the abundance of this transcript.

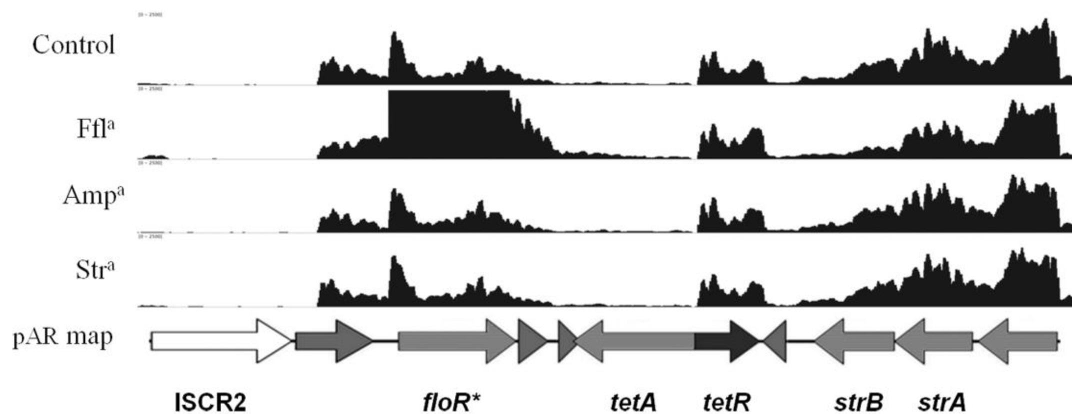

Figure 4: Transcriptome map of the *sul2* containing region of pAR060302. Antibiotic treatment conditions are labeled along the x-axis. The y-axis has a scale of 0 – 2500 reads aligned per base. Arrows indicate direction of transcription.
[a]Antibiotics are abbreviated as follows: ampicillin (Amp), florfenicol (Ffl), streptomycin (Str).
*Indicates a statistically significant greater or less than $\log_2$(fold change) in expression (q-value < 0.05).

The second hotspot in pAR060302 contains the bla_{CMY-2} gene, which is part of an IS*Ecp1*-containing mobile genetic element also containing *blc1* and *sugE1*. It is inserted amongst plasmid backbone genes predicted to be involved with conjugative transfer (Figure 5). There are 22 ORFs in the previously defined Tra1 gene cluster (16). In all bla_{CMY-2} positive IncA/C plasmids that have been sequenced thus far, the IS*Ecp1* mobile element containing bla_{CMY-2} is inserted in this region (15). In pAR060302 this insertion is located downstream of *traA*, but this site of insertion is not conserved amongst all IncA/C plasmids. The IS*Ecp1*-containing element of pAR060302 appeared to carry its own strong promoter just downstream of IS*Ecp1* conferring a high level of transcription of bla_{CMY-2}, which was not changed in the presence of ampicillin, florfenicol or streptomycin. Downstream of bla_{CMY-2}, both *blc1* and *sugE1* were minimally transcribed compared to the level of bla_{CMY-2} transcription, indicating that there may be a transcription stop site between bla_{CMY-2} and these two genes.

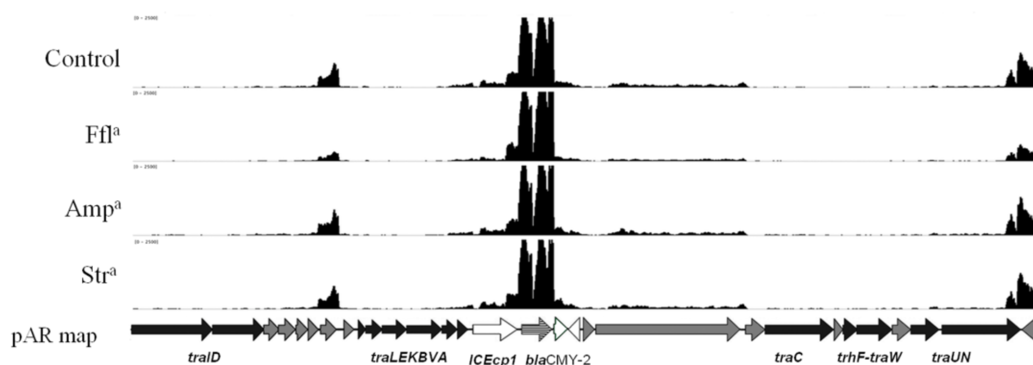

Figure 5: Transcriptome map of the *bla*$_{CMY-2}$ containing Tra1 region of pAR060302. Antibiotic treatment conditions are labeled along the x-axis. The x-axis has a scale of 0 – 2500 reads aligned per base. Arrows indicate direction of transcription.

[a]Antibiotics are abbreviated as follows: ampicillin (Amp), florfenicol (Ffl), streptomycin (Str)

Nearly all sequenced *bla*$_{CMY-2}$ positive IncA/C plasmids carry a Tn*21*-like class I integron region in their third accessory hotspot (6). The class I integron of pAR060302 contains aminoglycoside resistance genes, *aadA* and *aacC*; heat-shock chaperones *groSEL*; the *qacEΔl* and *sul1* genes; and the *merDBAPTR* mercury resistance operon. Our data suggest that a transcript started from an apparently strong promoter upstream of *aadA* and continued until just down downstream of *groEL* where transcription seemingly halted. *aadA* was transcribed at a level 74-fold higher compared to *repA*, while *aacC* and *groSEL* were transcribed at decreasing levels (Figure 6). This position-dependent transcription of integrons has been previously observed (52). The transcription of this region was not affected by treatment with ampicillin, florfenicol, or streptomycin. The

merDBAPTR gene cassette, previously shown to be activated in the presence of mercuric compounds (53), was unaffected by treatment with the three antibiotics used here.

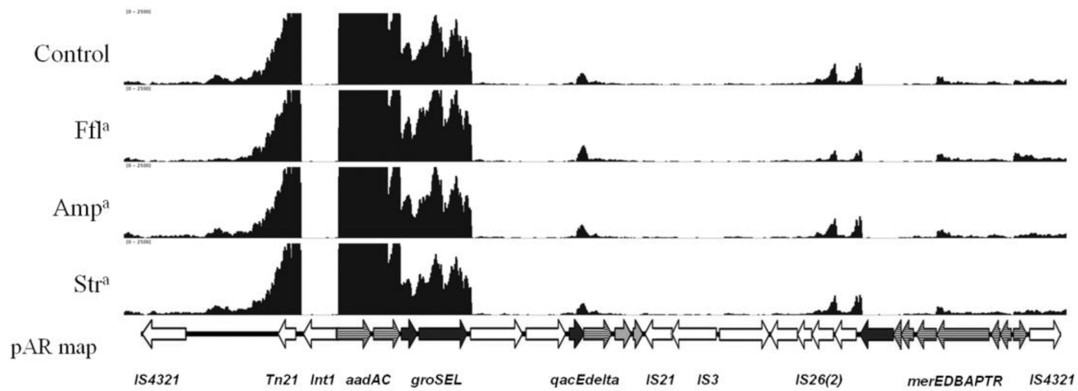

Figure 6: Transcriptome map of the class I integron containing region of pAR060302. Antibiotic treatment conditions are labeled along the x-axis. The x-axis has a scale of 0 – 2500 reads aligned per base. Arrows indicate direction of transcription.

[a]Antibiotics are abbreviated as follows: ampicillin (Amp), florfenicol (Ffl), streptomycin (Str).

***E. coli* chromosomal transcriptional response to exposure to antimicrobials.**

Reads obtained from each treatment were also mapped to the chromosome of the *E. coli* DH5α host. Both the ampicillin and streptomycin treatments yielded no differentially transcribed genes on the chromosome that were statistically significant. However, the florfenicol treatment resulted in 286 chromosomally encoded genes which were transcribed statistically significantly different levels compared to the control cells. Most of these genes (~250) were down regulated, whilst the remaining genes were up regulated. Among the genes that had increased expression were *rbfA*, *deaD, cspA* and *cspG*, which are annotated as being induced upon cold shock. Two transmembrane

tranporters, *proP and mdtK,* were also up-regulated. Several genes annotated as part of the σ^s-directed stress response were down-regulated. These include the *dps* gene, which encodes a histone-like protein that forms clusters with DNA, the acid response genes *gadC* and *hdeAB*, and the *cfu* gene, which is involved with alterations of the cell membrane (54).

Discussion

IncA/C plasmids have been increasingly associated with the dissemination of antimicrobial resistance genes in populations of enteric bacteria from humans and animals. To gain insights on the biology of these plasmids we conducted RNA sequencing experiments, resulting in the first report of the transcriptional landscape of these large plasmids. IncA/C plasmids are distributed globally, yet remain highly similar aside from their accessory elements (15, 16, 23). Data generated in this study, therefore, may be applied broadly across the many sequenced IncA/C derivatives. Illumina based RNA-Seq provided a robust approach to mapping the transcriptional levels down to the single base pair. This provided an opportunity to amend the current annotation of pAR060302 and gain insights into the genetic determinants of plasmid maintenance and stability.

When studying plasmid biology the host carrying any given plasmid is inherently important. Plasmids rely on their host cell for aspects of DNA replication and gene expression. In the present study *E. coli* strain DH5α was used as the host cell. This

common lab strain is well adapted to growth in lab media. It is probable that the extent to which DH5α is able to respond to external stimuli is severely muted compared to that of wild-type bacteria. In the present study, we observed that only florfenicol treatment altered chromosomal transcription, which was unexpected. No clear pattern emerged from the analysis of the differentially expressed chromosomal genes. This may be evidence that *E. coli* DH5α is ill-equipped to respond to environmental stresses. Another possibility could be that our strain most likely represents a naïve host-plasmid relationship and this interaction may be confounding the effects of antimicrobial exposure. Indeed, other hosts would likely have a larger impact on plasmid transcriptional profiles due to the overlap of their spectrum of regulatory networks. The ability of chromosomal regulatory networks to interact with plasmid encoded regulatory networks ultimately governs plasmid gene expression. Therefore, while the present study sought to define IncA/C transcription in the absence of complex interactions between the host chromosome and plasmid, it will be necessary in future work to study the effects of IncA/C transcription in additional bacterial hosts and growth conditions to fully appreciate the plasmid-host relationship.

Two of the three accessory regions of pAR060302 were transcribed at high levels compared to the backbone of the plasmid, while the *floR-tetRA-strAB-sul2* region was not. Evidence of transcription of these genes supports phenotypic observations of their ability to confer resistance against high concentrations of antimicrobial compounds. However, there could also be consequences associated with their expression at high

levels in the absence of antibiotic selection pressure. A recent study by Call *et al.* (34) showed that the bla_{CMY-2} region of pAR060302 is prone to stochastic excision and, in the absence of selection pressure, populations of *E. coli* harboring bla_{CMY-2} were out-competed by bla_{CMY-2}-lacking plasmids over many generations. Other studies have shown a significant cost of carrying other variants of AmpC-like beta-lactamase genes (55, 56). It was hypothesized that this cost is a result of unregulated expression and that this cost of carriage results in reduced dissemination. It is also possible that IncA/C plasmids mitigate the fitness cost conferred by bla_{CMY-2}, to some extent, to allow its persistence and spread.

Call *et al.* also showed that the *flo-tet* region was prone to apparent excision over time, but required significantly more generations to be lost from pAR060302-containing bacterial populations. We show here that *floR* gene transcription is transcribed at relatively low levels compared to the other resistance genes of pAR060302, but is seemingly induced by the presence of florfenicol. Florfenicol has been previously shown to increase expression of other genes and this was found to be a result of increased half-life of mRNA due to the presence of florfenicol (57). Still, the transcript containing *floR* was the only plasmid encoded mRNA that was significantly up-regulated in response to antibiotic treatment. Other mechanisms resulting in the induction of expression of resistance genes by exposure to florfenicol have been described before (58). The putative transcriptional regulator immediately downstream of *floR* may represent a control mechanism for the regulated transcription of *floR*. No previous studies have addressed

this. However, since florfenicol was the only treatment that resulted in altered chromosomal expression this may instead be a result of chromosome-plasmid "cross-talk". Given that many genes on the chromosome were differentially transcribed, why only a specific mRNA on the plasmid was affected remains in question. It is unclear at this time what the molecular basis for the increased *floR* transcripts we observed might be. Others have shown that the *tetRA* region is induced by tetracycline (51), which is likely the case in pAR060302 given our results. While *strAB* was transcribed and was not induced by the addition of streptomycin to the media in our experiments, it was transcribed at a much lower level than that of bla_{CMY-2}, *aadA*, or *aacC*. Thus, the increased stability of the *flo-tet* region observed by others could be correlated to a lower fitness cost for the carriage due to its apparent regulation. This is in contrast to the apparently uncontrolled transcription of bla_{CMY-2} and genes within the Tn*21*-like region, which were highly transcribed under all of our study conditions. The stability of the Tn*21*-like region over long periods of time in the absence of antimicrobial selection pressure has yet to be measured, although we expect that portions of this region would also be deleted over time due to the relative cost of high transcription. This could, in part, explain why some sequenced IncA/C plasmids, such as pAM04528 (16), only possess the *mer* operon and lack other Tn*21*-associated genes. Further analysis is required to verify the impact of gene expression on the evolution of IncA/C plasmids.

Several genes have been identified here that might be key to this plasmid's success, including several putative plasmid-encoded transcriptional regulators that were

transcribed at high levels compared to other backbone components. In other plasmid systems, similar genes have been shown to have a wide array of affects. For example, the IncP-7 plasmid group encodes an H-NS like protein Pmr, and recent studies have shown that Pmr alters the transcriptional landscape of the host chromosome (31). R27, an IncH plasmid isolated from *Salmonella enterica* servar Typhimurium, encodes Sfh, another H-NS-like protein. Deletion of this gene resulted in a nearly 4-fold increase in the number of chromosomally encoded genes that were differentially expressed due to plasmid acquisition (32). These H-NS-like proteins have been termed "stealth-like" and have been hypothesized to silence the effects of plasmid acquisition on the host chromosome. A recent survey of sequenced plasmids found that similar genes reside on many known broad-host-range plasmids (59). All sequenced IncA/C variants encode a putative H-NS like protein that could represent a genetic factor enabling transfer and/or stability among diverse host backgrounds. H-NS orthologs might provide alleviation of the cost of plasmid carriage incurred by the host cell. However, the putative H-NS-like regulator of pAR060302 shares low amino acid similarity to its closest known match (16), so it is still necessary to elucidate the role this gene plays, if any, in the success of IncA/C plasmids in their diverse hosts.

The putative toxin-antitoxin system of pAR060302 was also highly transcribed under the conditions studied. Such systems are well known for their involvement in post-segregational killing in plasmid-free daughter cells (60). The transcription of this system in pAR060302 suggests that IncA/C plasmids ensure their stability in bacterial

populations by potentially killing plasmid-free progeny that may arise. However no previous work has been done to describe the function of these genes found in IncA/C plasmids.

Within the putative conjugal transfer regions of pAR060302, only a few genes of unknown function were transcribed in our study. pAR060302 has been shown to be transferrable by our group and by others (15). The apparent conjugal transfer machinery of IncA/C plasmids is unlike that of any other known plasmids, but is highly syntenic to the transfer loci of SXT/R391 ICE elements (21). These ICEs have a highly regulated transfer mechanism, suggesting that the same is true for IncA/C plasmids. Yet, the key regulators of R391-like ICEs are missing from IncA/C plasmids, pointing to potentially novel mechanisms of regulation (61–63). Several scenarios could explain the regulation of the transfer region in IncA/C plasmids. One possibility is that the region may be simply unregulated, however, this is unlikely due the inability to observe transcription of the *tra* genes. Another possibility is that the transfer loci may be under host cell control as is the case for plasmids that undergo zygotic induction; this too seems unlikely for a plasmid which such a broad-host range since relying on host-specific factors would result in an evolutionary dead-end if the plasmid were to transfer to a host unfit for such regulation. A third possibility is that IncA/C plasmids may encode their own genes for regulation of the transfer loci. This seems to be the most likely scenario, since these plasmids harbor an array of putative regulators, some of which might regulate the expression of the transfer loci. Other plasmid systems rely on a signal from suitable

recipients to initiate plasmid transfer, and although this has not been described for plasmids in Gram-negative bacteria it remains to be tested for IncA/C plasmids (9). Given the role that IncA/C plasmids play in the dissemination of resistance genes among *Enterobaceriaceae*, the mechanisms by which they are transferred deserve further study.

The data generated in this study afforded us the opportunity to examine the annotation of pAR060302 based on experimental evidence. The two short transcripts in areas not previously annotated could represent small peptides encoded on pAR060302. They could, alternatively, represent small non-coding RNAs (ncRNA). In *E. coli,* ncRNAs have previously been shown to be up to 300 bp in length (64, 65). Although the experimental design of our study was not specifically created to isolate ncRNAs, these transcripts are of appropriate length to be observed under our conditions. Further work is needed to identify the nature of these transcripts.

A limitation to this work is that the data only represent a fraction of pAR060302 transcription potential, since we only analyzed a single growth point of *E. coli* harboring pAR060302 in a single growth media. This provides insight to the minimum genetic components required for IncA/C plasmid maintenance under laboratory growth conditions. Our first hypothesis about transcriptional differences between plasmid accessory and backbone elements held true; it is remarkable that much of the backbone of pAR060302, including regions predicted to be involved in plasmid maintenance and stability, was transcriptionally inactive. This could imply the possibility of extensive

regulation and crosstalk between plasmid and host or self-regulation of the plasmid itself. We also hypothesized that the transcriptional profile of the plasmid would respond to antibiotic treatment, which was only true in the case of florfenicol treatment. It is probable that changes could occur if different classes of antibiotics were used, but it is interesting that many of the resistance genes were transcribed at high-levels in the absence of antibiotics.

Overall, RNA-Seq offers a robust approach to assess global scale transcription. In our data we can observe a polarity to the transcripts, which can offer clues to the location of the transcriptional start site. However, a different RNA-Seq experimental design than that used here would be needed to accurately map the start of nascent transcripts (66). Further studies such as these will move toward a global understanding of the expression of the genes on IncA/C plasmids and the implication of such expression on their overall biology. Host strain-dependent transcriptional profiles of other plasmids, such as IncP, have been previously reported (31, 32). Although we expect differences in transcription of pAR060302 in its wild-type host, *E. coli* strain DH5α was used in this study to obtain a baseline transcriptional profile of pAR060302. Our observations of pAR060302 transcription indicates that it is likely highly regulated, allowing for its dissemination to a broad range of bacterial hosts and its stability within these hosts, despite the likely costs of encoding highly transcribed accessory regions encoding antibiotic resistance. Future efforts will use these data to better understand the impacts of host background and plasmid-encoded genes on its transcription and regulation.

Materials and Methods

Bacterial strains and growth conditions. *E. coli* strain DH5α harboring pAR060302 was grown in 10 mL Difco™ Luria-Bertani (LB) broth aliquots at 37º C with shaking until an OD_{600} of 0.5. A total of 8 cultures were independently grown representing two biological replicates per condition tested. Six of the cultures were amended, 2 cultures per antibiotic, with ampicillin (50 µg/mL final concentration), florfenicol (30 µg/mL final concentration), or streptomycin (50 µg/mL final concentration) and allowed to incubate at 37º C with shaking for an additional 30 min. Two cultures were not amended with any antibiotic. Cells were pelleted and RNA was purified using a commercially available RNA extraction kit (Qiagen). RNA preparations were then subjected to a DNase treatment to eliminate DNA contamination from the sample (Qiagen). A treatment was also included to deplete ribosomal RNA using a commercially available kit (MicrobExpress, Ambion). The two biological replicates for each growth condition were pooled for sequencing.

Ilumina sequencing for transcriptome mapping. cDNA libraries were generated with an insert size of 100 bp and sequenced with 76-base cycles of single-end reads using a Genome Analyzer II (Illumina) platform according to manufacturer's protocols at the Biomedical Genomics Center (University of Minnesota, Minneapolis, Minnesota, USA). Approximately 160,000 plasmid-mapped reads each were obtained for the ampicillin and streptomycin treated samples, and 260,000 plasmid-mapped reads each for the control

and florfenicol treatment samples. Genome-mapped read counts were as follows: control, ~6.4 million reads, florfenicol treatment, ~5.7 million reads, ampicillin treatment, ~1.7 million reads, and streptomycin treatment, ~5.2 million reads. We only used those reads uniquely mapped on plasmid or chromosomal DNA for global normalization and further analysis.

RNAseq data analysis. cDNA reads were trimmed so that the quality at each base position was above 30 (~15-20 bp) and then mapped either to the *E. coli* K-12 MG1655 published genome sequence (Genbank accession no. NC_000913) or to the pAR060302 published sequence (Genbank accession no. NC_092692) using BOWTIE (67). The *E. coli* strain DH5α has an incomplete annotation and for this reason the *E. coli* K-12 annotation was used, representing an estimation of differentially expressed genes due to exposure of antimicrobials. For each condition, graphs representing the number of mapped reads per nucleotide were generated and visualized using the Integrated Genome Viewer (IGV) (68). Images were created using XplasMap (http://www.iayork.com/XPlasMap/) and IGV. The reads mapped per kilobase of gene per million (RPKM) reads was calculated using either the *E. coli* chromosome or the pAR060302 annotation and was used for global normalization (69). The per kilobase cDNA length normalized the effect of different length of cDNAs such that the sequence reads have a equal chance to map on the long cDNA regions and the short cDNA regions. After RPKM normalization, each sample is comparable to each other. An R package, DEGseq (70), was used to identify differentially expressed genes between the control and

each antibiotic treatment condition. A cutoff of q-value < 0.05 and a fold change of > 3 were used to measure statistical significance (70).

Chapter 3

Transcriptome modulations due to IncA/C plasmid acquisition[4]

[4] This chapter is currently in review at the journal Plasmid
Kevin S. Lang and Timothy J. Johnson. Transcriptome modulations due to IncA/C plasmid acquisition.

Summary

Plasmids play an important role in driving the genetic diversity of bacteria. Horizontal gene transfer via plasmids is crucial for the dissemination of antimicrobial resistance genes. Many factors contribute to the persistence of plasmids within bacterial populations, and it has been suggested that epistatic interactions between the host chromosome and plasmid contribute to the fitness of a particular plasmid-host combination. However, such interactions have been shown to differ between bacterial hosts. In this study, RNA-Seq was performed in six different strains to characterize the influence of host background on the IncA/C plasmid transcriptome. In five of these strains, chromosomal transcriptomes were compared in the presence and absence of IncA/C plasmid pAR060302. Host-specific effects on plasmid gene expression were identified, and acquisition of pAR060302 resulted in changes in the expression of chromosomal genes involved in metabolism and energy production. These results suggest that IncA/C plasmid fitness could, in part, be dependent on host chromosome content, as well as environmental factors.

Introduction

Antimicrobial resistance among Gram-negative bacterial pathogens is a public health crisis (71). The relative lack of drug discovery and the rapid emergence of multidrug-resistant (MDR) bacterial pathogens is a truly alarming combination. Horizontal gene transfer (HGT) is a major contributor to the spread of genes conferring multidrug-resistance. Understanding the factors driving the persistence of the elements facilitating HGT is crucial for the ability to better understand how to curtail the spread of multidrug resistance.

Incompatibility group A/C (IncA/C) plasmids are a group of large (~100 ~ 200 kb), broad-host-range plasmids that often carry genes conferring resistance to multiple classes of antimicrobials (13, 15, 72, 73). They are commonly isolated from Gram-negative Proteobacteria associated with aquatic and terrestrial animals, food products and human infections (24, 43, 48, 74, 75). Because IncA/C plasmids commonly confer multidrug resistance, it is easy to understand how this plasmid type persists in environments under selective pressures such as antibiotic therapy. It has been proposed that this plasmid type may even require antimicrobial selection pressure to persist in microbial populations (34), however these experiments focused on only one host strain. The interactions between plasmids and host chromosome have been reported in other plasmid types (76–80). These studies demonstrated that chromosomal background is a significant factor in the maintenance of plasmids. Crosstalk between the broad host-range IncA/C plasmid type and its host chromosome has yet to be explored.

The present study attempts to better understand the IncA/C plasmid-host relationship through the use of total RNA sequencing (RNA-Seq). To determine if there were species-specific changes to the transcriptome of broad-host range plasmid pAR060302, or if there were common changes to the chromosomal transcriptional program during plasmid carriage, we examined the transcriptome of several different strains, spanning 3 genera of Gammaproteobacteria carrying pAR060302. We compared the transcriptome of pAR060302 in each host to that of the strain it was originally isolated in, AR060302. Finally, we compared the transcriptomes of the chromosomes of five strains with and without plasmid pAR060302.

Results

Analysis of the transcriptome of pAR060302 carried by different bacterial hosts. To understand how chromosomal background influences the pAR060302 transcriptome, a total of six strains carrying the plasmid were subjected to RNA-Seq. The sequencing reads were mapped to the pAR060302 reference sequence. Figure 7 shows the transcriptome map of plasmid pAR060302 while being carried by different bacterial hosts. Similar to what we have previously reported in *E. coli* strain DH5α (81), the genes most highly transcribed amongst all host species were resistance genes $bla_{\text{CMY-2}}$, *floR*, *strA, strB*, *aadA2/aadA1* and *aac*(3)-VIa, as well as the putative toxin-antitoxin module (locus tags pAR060302_0025 and 0026).

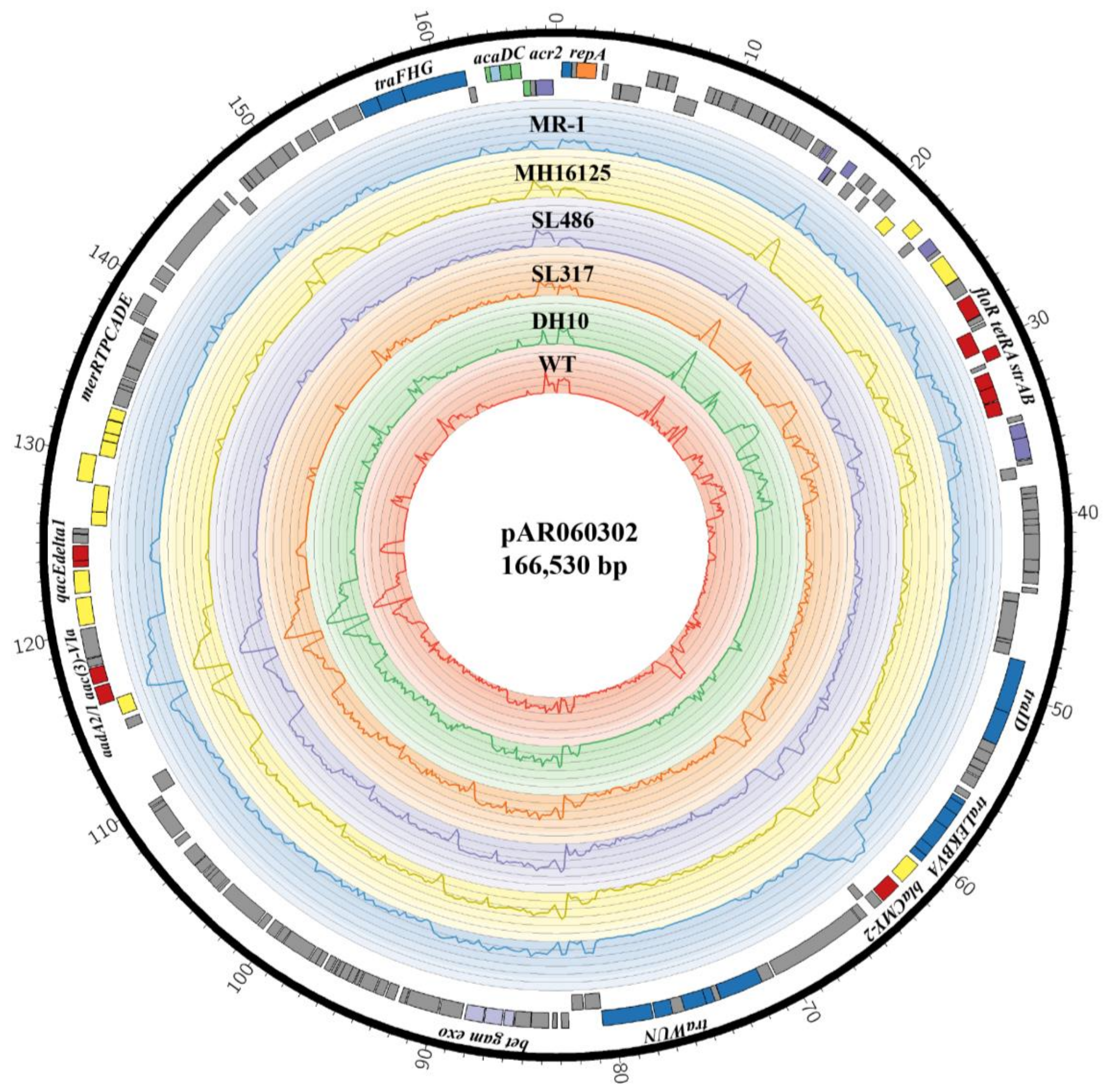

Figure 7. Transcriptome map of pAR060302 carried by different hosts. The rings from outside-in: 1) plasmid base pair coordinates; 2) gene map of pAR060302; 3-8) transcriptome maps of pAR060302 carried by *S. oneidensis* strain MR-1, *S. enterica* serovar Enteritidis strain MH16125, *S. enterica* serovar Heidelberg strain SL486, *S. enterica* serovar Newport strain SL317, *E. coli* strain DH10B, and *E. coli* strain AR060302. Genes on the map are colored by function: gray, hypothetical protein; red, antibiotic resistance; yellow, horizontal gene transfer; green, transcriptional regulation;

light purple, recombination; purple, DNA binding; blue, conjugal transfer; light blue, transglycosylase; orange, replication. Line plots represent average reads aligned to 250 bp windows along pAR060302, normalized for sequencing library size and log-transformed. Each line on the transcriptome maps represents 0.5 log.

Table 2 displays differentially expressed genes on plasmid pAR060302 when comparing their transcription in the wild-type strain it was originally isolated in (AR060302) versus their expression when carried by a new host. Interestingly only 2 genes, *qacEΔ1* and *sul1*, were differentially expressed across all host bacteria. Both genes were significantly down-regulated in all strains compared to their expression in AR060302. Given their organization on pAR060302, these two genes are probably transcribed from the same promoter. The genome sequence of strain AR060302 is not known and it cannot be ruled out that these genes might be expressed from somewhere on the chromosome, as well as pAR060302. Interestingly, the expression of many of the copies of IS*26* are significantly down-regulated in all strains except DH10B. However, given how similar each copy of IS*26* is, this result may be an artifact of the read aligner not being able to assign the read to the true copy its cognate transcript was transcribed from. Still, the overall number of reads originating from IS*26* transcripts are lower in abundance in these strains compared to AR060302. It is possible that strain AR060302 has copies of IS*26* somewhere on its chromosome that may contribute to the total number of reads aligning to copies of IS*26* on the plasmid. In either scenario, IS*26* may be an active player in gene transfer events in the original host AR060302. Conversely, in *S. oneidensis,* the putative transposase genes

of IS*CR2*, IS*CR16*, and IS*Ecp1* are all significantly up-regulated, suggesting that genetic rearrangements are possible in this strain via mechanisms apart from IS*26*. Also of note, an entire region of genes from locus tag pAR060302_0098-0133 is significantly up-regulated in strain SL486. Many of these genes have unknown functions and are dissimilar to any coding sequences in Genbank with known functions. A few of the coding sequences in this region do have predicted functions, most notably a phage-like site specific recombination operon, *ssb-bet-exo*.

Summary of chromosomal transcriptome changes due to carriage of pAR060302. IncA/C plasmids are large, broad host-range plasmids (16, 23, 26). To better understand how acquisition of such a plasmid would influence chromosomal gene expression we compared chromosomal transcriptomes in five strains with and without pAR060302. In order to characterize broad shifts in chromosomal gene expression, lists of differentially expressed genes were subjected to COG ontology analysis using DAVID (82, 83). Table 3 shows the top COG ontologies enriched in both the down- and up-regulated genes for each chromosomal background. Across all hosts, carriage of pAR060302 resulted in changes in expression of COGs related to oxidation/reduction reactions, cellular metabolism and/or metal cofactor binding. This suggests that plasmid carriage alters the metabolic demands of cells, which has been reported as a general phenomenon of plasmid carrying bacteria (84). Aside from central metabolism, the carriage of pAR060302 appears to change the expression of genes involved in localization, motility and adhesion in *Salmonella* strains, which may indicate a conserved interaction in that genus.

Table 2. Fold change and p-value of genes on pAR060302.

Locus Tag	Product	SL317 log2FC	SL486 log2FC	MH16125 log2FC	MR-1 log2FC	DH10 log2FC
pAR060302_0005	hypothetical protein	-2.76	0.80	0.26	2.57	1.70
pAR060302_0025	hypothetical protein	-0.56	-1.21	-0.18	-1.02	1.83
pAR060302_0034	IS26	-1.87	-2.08	-2.25	-3.01	-0.70
pAR060302_0035	IS26	-2.64	-2.68	-3.12	-3.02	-1.69
pAR060302_0039	ISCR21 TnpA	0.85	0.63	1.09	2.65	1.31
pAR060302_0040	hypothetical protein	-0.55	-1.30	-0.50	-0.06	1.85
pAR060302_0041	FloR	2.11	0.36	0.68	0.94	3.67
pAR060302_0042	hypothetical protein	2.10	0.21	0.59	1.71	3.28
pAR060302_0044	TetA(A)	1.30	0.04	0.14	2.27	1.49
pAR060302_0045	TetR(A)	-0.82	-1.98	-1.41	-0.53	0.40
pAR060302_0069	TraI	0.59	1.05	0.06	-0.95	-2.29
pAR060302_0072	hypothetical protein	1.00	1.31	0.57	-0.94	-2.57
pAR060302_0075	hypothetical protein	0.31	0.40	-0.19	-2.08	0.20
pAR060302_0076	hypothetical protein	1.49	2.59	1.68	-0.69	-2.74
pAR060302_0078	TraE	0.98	1.68	0.99	-1.73	-2.56
pAR060302_0079	TraK family protein	0.80	1.44	0.68	-0.72	-2.24
pAR060302_0081	TraV	1.53	1.89	0.99	-0.47	-1.41
pAR060302_0082	TraA	1.84	2.33	1.28	0.82	-0.79
pAR060302_0083	ISEcp1 TnpA	-0.38	-0.03	-0.68	3.03	-1.23
pAR060302_0084	CMY-2	-0.93	-1.21	-1.29	1.60	-3.93
pAR060302_0085	outer membrane lipoprotein Blc	0.05	-1.33	-0.96	3.21	-0.56
pAR060302_0086	SugE1	0.20	-1.49	-0.85	3.65	-0.99
pAR060302_0087	hypothetical protein	-0.20	-0.40	-1.98	1.27	-9.56
pAR060302_0089	hypothetical protein	0.21	0.94	0.22	-0.54	-2.28

pAR060302_0090	TraC	0.54	1.10	0.45	-0.89	-2.42
pAR060302_0091	hypothetical protein	0.09	1.31	0.47	-0.90	-4.43
pAR060302_0092	TrhF	0.31	0.89	0.39	-0.96	-2.60
pAR060302_0093	TraW	0.88	1.20	0.66	-0.77	-2.69
pAR060302_0098	hypothetical protein	0.37	1.77	0.31	-0.17	-0.80
pAR060302_0099	hypothetical protein	1.62	2.43	1.76	-0.25	1.41
pAR060302_0100	hypothetical protein	1.36	2.24	1.12	0.14	1.41
pAR060302_0101	ATPase family protein	1.71	2.51	1.65	0.18	0.86
pAR060302_0102	hypothetical protein	1.64	2.45	1.60	0.02	1.03
pAR060302_0103	Ssb	1.82	2.26	1.34	0.13	0.94
pAR060302_0104	Bet	1.68	2.48	1.63	0.33	1.04
pAR060302_0105	exo	1.60	2.24	1.36	0.17	1.04
pAR060302_0106	hypothetical protein	1.66	2.31	1.59	0.18	0.70
pAR060302_0107	hypothetical protein	1.65	2.22	1.47	0.24	0.51
pAR060302_0109	hypothetical protein	1.81	2.78	1.81	0.49	-1.09
pAR060302_0110	hypothetical protein	1.92	2.86	1.96	0.39	-0.42
pAR060302_0111	hypothetical protein	1.87	2.65	1.70	0.50	-0.54
pAR060302_0112	ferredoxin family member protein	1.89	2.77	1.84	1.08	0.37
pAR060302_0113	hypothetical protein	1.82	2.60	1.64	0.70	-0.58
pAR060302_0114	hypothetical protein	1.97	2.71	1.69	0.96	0.05
pAR060302_0115	hypothetical protein	1.86	2.87	2.08	1.07	0.52
pAR060302_0116	hypothetical protein	1.61	2.61	1.83	0.58	0.65
pAR060302_0117	hypothetical protein	1.77	2.64	1.73	1.03	0.79
pAR060302_0118	hypothetical protein	1.82	2.38	1.94	0.46	1.92
pAR060302_0119	hypothetical protein	2.14	2.66	2.10	0.46	2.07
pAR060302_0120	Dcm	1.97	2.72	1.95	0.61	1.81
pAR060302_0121	hypothetical protein	1.25	2.50	1.92	0.69	1.11
pAR060302_0122	hypothetical protein	1.85	2.53	1.60	0.82	1.20

pAR060302_0123	hypothetical protein	1.47	2.04	1.58	-0.22	0.18
pAR060302_0124	phosphoadenosine phosphosulfate reductase	1.34	1.91	1.08	0.28	0.66
pAR060302_0125	hypothetical protein	2.06	2.32	1.60	-0.21	-0.40
pAR060302_0126	HD phosphohydrolase family protein	1.91	2.57	1.72	-0.12	0.44
pAR060302_0127	hypothetical protein	1.55	2.21	1.31	0.07	0.32
pAR060302_0128	hypothetical protein	2.01	2.55	1.42	-0.29	0.01
pAR060302_0129	hypothetical protein	2.08	2.84	1.93	-0.04	0.57
pAR060302_0130	hypothetical protein	1.59	2.09	1.20	-0.08	0.26
pAR060302_0131	hypothetical protein	2.11	2.64	1.95	-0.14	1.85
pAR060302_0132	primase-helicase family member protein	1.72	2.65	1.73	-0.41	1.45
pAR060302_0133	hypothetical protein	1.86	2.67	1.93	0.05	1.72
pAR060302_0138	IntI1	-0.50	-1.60	-1.91	-0.94	0.67
pAR060302_0142	chaperonin GroEL	0.86	-1.23	-0.05	1.98	2.09
pAR060302_0143	ISCR16 InsE	0.58	-0.40	1.59	2.65	2.31
pAR060302_0145	QacEΔ1	-2.74	-3.77	-4.14	-2.69	-3.15
pAR060302_0146	Sul1	-3.03	-4.21	-4.61	-2.47	-4.29
pAR060302_0147	hypothetical protein	-2.23	-3.87	-3.70	-1.74	-4.35
pAR060302_0149	IS1326 IstB	-3.71	-2.17	-2.27	0.20	-1.33
pAR060302_0150	IS1326 IstA	-1.61	-1.87	-1.55	0.73	-0.10
pAR060302_0154	IS26	-2.41	-3.04	-3.04	-2.77	-1.76
pAR060302_0155	IS26	-1.69	-2.09	-2.30	-2.09	-0.73
pAR060302_0158	transcriptional regulator MerD	-0.71	-0.25	0.78	1.62	2.46
pAR060302_0160	putative mercuric reductase	-0.30	-1.01	0.36	0.88	1.88
pAR060302_0163	putative transcriptional regulator MerR	0.71	0.10	1.56	1.12	2.39
pAR060302_0165	hypothetical protein	-1.89	-0.20	-0.32	-0.19	-2.24
pAR060302_0171	hypothetical protein	-0.84	-1.39	2.48	-1.48	1.23
pAR060302_0172	hypothetical protein	-1.70	-1.47	2.45	-1.39	0.22
pAR060302_0173	hypothetical protein	-1.15	-1.49	2.13	-0.82	0.28

pAR060302_0174	site-specific recombinase	-0.98	-1.54	1.96	-1.14	0.25
pAR060302_0175	exonuclease family protein	-0.91	-1.18	1.98	-0.91	-0.10
pAR060302_0188	Acr2	-1.09	-0.70	-0.99	-2.56	-1.04
pAR060302_0190	StbA	0.50	0.81	0.07	-2.13	-2.12

logFC indicates the log2(fold-change). Boxes in gray indicate genes that were statistically significantly different compared the genes expression in AR060302 (FDR < 0.05).

Table 3. COGs enriched in differentially expressed genes from strains carrying pAR060302 compared to plasmid free strains.

Downregulated COGs			**Upregulated COGs**		
DH10B (n = 445)			**DH10B (n=370)**		
	# of genes	P-value		# of genes	P-value
Biological Process:			***Biological Process:***		
Protein binding	75	3.80E-23	Reponse to stimulus	56	3.70E-15
Lyase activity	40	4.10E-09	Catabolic process	47	2.70E-13
Active transmembrane transport	42	2.40E-06	Response to stress	40	8.70E-13
Cellular Compartment:			***Cellular Compartment:***		
Plasma membrane	118	4.70E-10	Cytoplasm	94	1.50E-12
Molecular Function:			***Molecular Function:***		
Metal ion binding	66	7.30E-05	Metal ion binding	56	3.70E-06
Cell surface activity	7	3.50E-04	Glycolate dehydrogenase activity	3	3.50E-05
rRNA binding	6	8.90E-04	Unfolded protein binding	4	2.90E-02
MH16125 (n = 24)			**MH16125 (n =72)**		
	# of genes	P-value		# of genes	P-value
Biological Process:			***Biological Process:***		
Nitrate metabolic process	3	3.40E-04	Localization	17	5.50E-03
Oxidation reduction	8	6.00E-04	Ion transport	7	8.20E-03
Cellular respiration	3	1.80E-02	Pathogenesis	4	9.40E-03
Cellular Compartment:			***Cellular Compartment:***		
Nitrate reductase complex	4	1.50E-05	Plasma membrane	14	6.20E-04
Molecular Function:			***Molecular Function:***		
Oxidoreductase activity	5	2.70E-07	Transporter activity	14	7.10E-03
Formate dehydrogenase activity	3	4.00E-04			4.00E-04
Metal binding	4	4.10E-02			4.10E-02
SL317 (n = 154)			**SL317 (n = 160)**		
	# of genes	P-value		# of genes	P-value
Biological Process:			***Biological Process:***		
Locomotion	28	3.90E-28	Histidine biosynthesis	8	2.90E-09
Pathogenesis	26	1.90E-23	Ketone metabolism	21	1.10E-05
Chemotaxis	19	2.60E-17	Cellular biosynthetic process	35	1.70E-02
Cellular Compartment:			***Cellular Compartment:***		
Flagellum	24	4.50E-24	Cytoplasm	25	3.40E-03
Molecular Function:			***Molecular Function:***		
Motor activity	8	3.50E-10	Metal ion binding	22	8.70E-05
Structural molecule activity	9	3.90E-05	Lyase activity	12	9.80E-03
Unfolded protein binding	4	1.10E-03	Petosyl transferase activity	4	3.40E-03

Downregulated COGs			Upregulated COGs		
SL486 (n = 245)			**SL486 (n = 73)**		
	# of genes	P-value		# of genes	P-value
Biological Process:			***Biological Process:***		
Pathogenesis	13	2.50E-06	Localization	20	1.90E-03
Cellular catabolic process	16	8.30E-06	Transporter activity	20	1.90E-03
Carbohydrate transport	21	4.10E-03	Siderophore transport	3	3.80E-03
Cellular Compartment:			***Cellular Compartment:***		
Cytoplasm	34	9.90E-03	Plasma membrane	13	6.70E-03
Molecular Function:			***Molecular Function:***		
Oxidoreductase activity	38	1.40E-04	Iron ion transport	5	5.50E-05
Transporter activity	46	4.60E-03			
Unfolded protein binding	4	5.00E-03			
MR-1 (n = 269)			**MR-1 (n = 169)**		
	# of genes	P-value		# of genes	P-value
Biological Process:			***Biological Process:***		
Oxidation reduction	40	4.50E-12	Generation of precursor metabolites	15	4.70E-04
Anaerobic respiration	15	5.10E-10	Iron ion homeostasis	4	1.30E-03
Cellular respiration	22	1.00E-08			
Cellular Compartment:			***Cellular Compartment:***		
Formate dehydrogenase complex	8	2.10E-05	N/S		
Molecular Function:			***Molecular Function:***		
Iron binding	36	1.80E-15	Metal binding	21	5.40E-03
Electron carrier activity	37	9.70E-14	Lyase activity	10	1.10E-02
Oxidoreductase activity	55	6.50E-13			

Number of genes entered in each analysis are indicated. P-value is based on significance testing of enrichment based on total number of genes involved in each COG. N/S – Not significant.

Specific pathways up-regulated in cells bearing pAR060302. The two main functional classes of genes upregulated in DH10 in response to carriage of pAR060302 were those involved in two-carbon and fatty acid metabolism and amino acid degradation (Figure 8). All genes involved in glyocolate metabolism (*glcDEFGBA*) and glyoxylate cycle (*aceBAK*) were up-regulated in DH10. These two operons can be induced by either acetate or glycolate and are used to replenish pools of TCA cycle intermediates lost due

to synthesis of amino acids (85). Similarly, in the case of MR-1, which lacks the *glc* operon, genes in the glyoxylate cycle are up-regulated (*aceBAK*). Other genes involved with acetate production were also up-regulated (*pta* and *ackA*). MH16125 exhibited increased expression of genes involved in transportation across the cell membrane and those involved with pathogenicity. Some of the transport genes are involved with movement of carbon sources (*wza*, *glpF*, *glpT* and *proV*). The other genes are all located on pathogenicity islands and are involved in type three secretion (*ssaJ*, *ssaK*, *ssaL*, *invH* and *sicP*), inter-macrophage survival (*mgtB* and *mgtC*), and iron uptake (*iroN*). SL486 and MR-1 also up-regulated iron transport genes, although these are not associated with pathogenicity. SL317 exhibited increased expression of several amino acid biosynthesis genes as well as genes involved in metabolism of alternative carbon sources such as propionate (*prpR* and *prpE*).

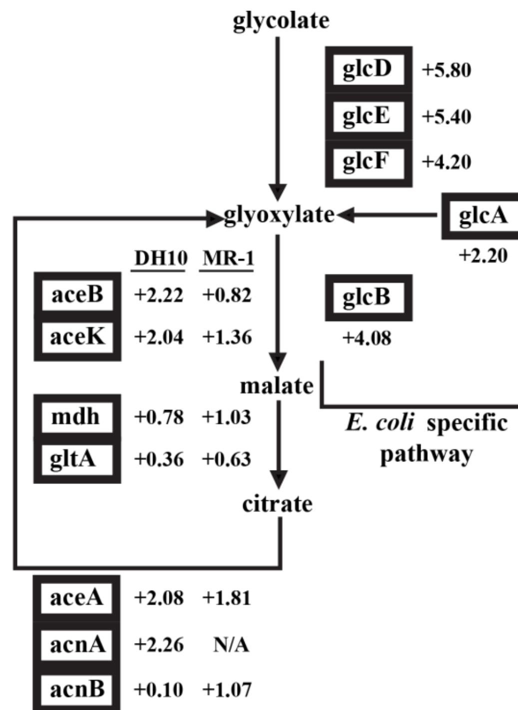

Figure 8. Differential expression data from strains DH10 and MR-1 overlaid onto the glyoxylate bypass pathways. Data were analyzed as described in the Materials and Methods. Values given as $\log_2$(fold-change). All values had an adjusted p-value of < 0.05.

Specific pathways down-regulated in cells bearing pAR060302. Several genes involved in the biosynthesis of various amino acids were down-regulated in DH10. Also, the *citCDEFXG* operon, which converts malate into acetate and oxaloacetate, was down-regulated. In MH16125, two operons involved in anaerobic respiration were down-regulated in response to plasmid carriage (*fdnI, fdnH, fdnG, narH, narJ, narG*). Several homologs of these anaerobic respiration genes were also down-regulated in MR-1. Whereas in strain MH16125 a few genes involved in pathogenesis were up-regulated, in SL317 and SL486 several genes spanning multiple pathogenicity islands were down-regulated. The functions of these genes include both invasion and type III secretion. In strain SL317 many genes involved in chemotaxis and motility were also down-regulated.

Summary of changes to S. oneidensis megaplasmid due to carriage of pAR060302. *S. oneidensis* strain MR-1 stably maintains a 161,613 bp plasmid termed a "megaplasmid" (86). We mapped the transcriptome of MR-1's megaplasmid in strains with and without plasmid pAR060302 (Figure 9). Surprisingly, the MR-1 strain used suffered a ~12 kb deletion which encodes several megaplasmid genes including that of *repA*. This could mean either 1) the plasmid has integrated into the genome of MR-1 or

2) the megaplasmid still exists as an extrachromosomal element that replicates via some alternative mechanism. Regardless, much of the gene content in our MR-1 strain remained. Table 4 indicates megaplasmid genes that were significantly differentially expressed when pAR060302 was present. Remarkably, few changes were observed, but of those few changes was an up-regulation in genes involved in toxin secretion.

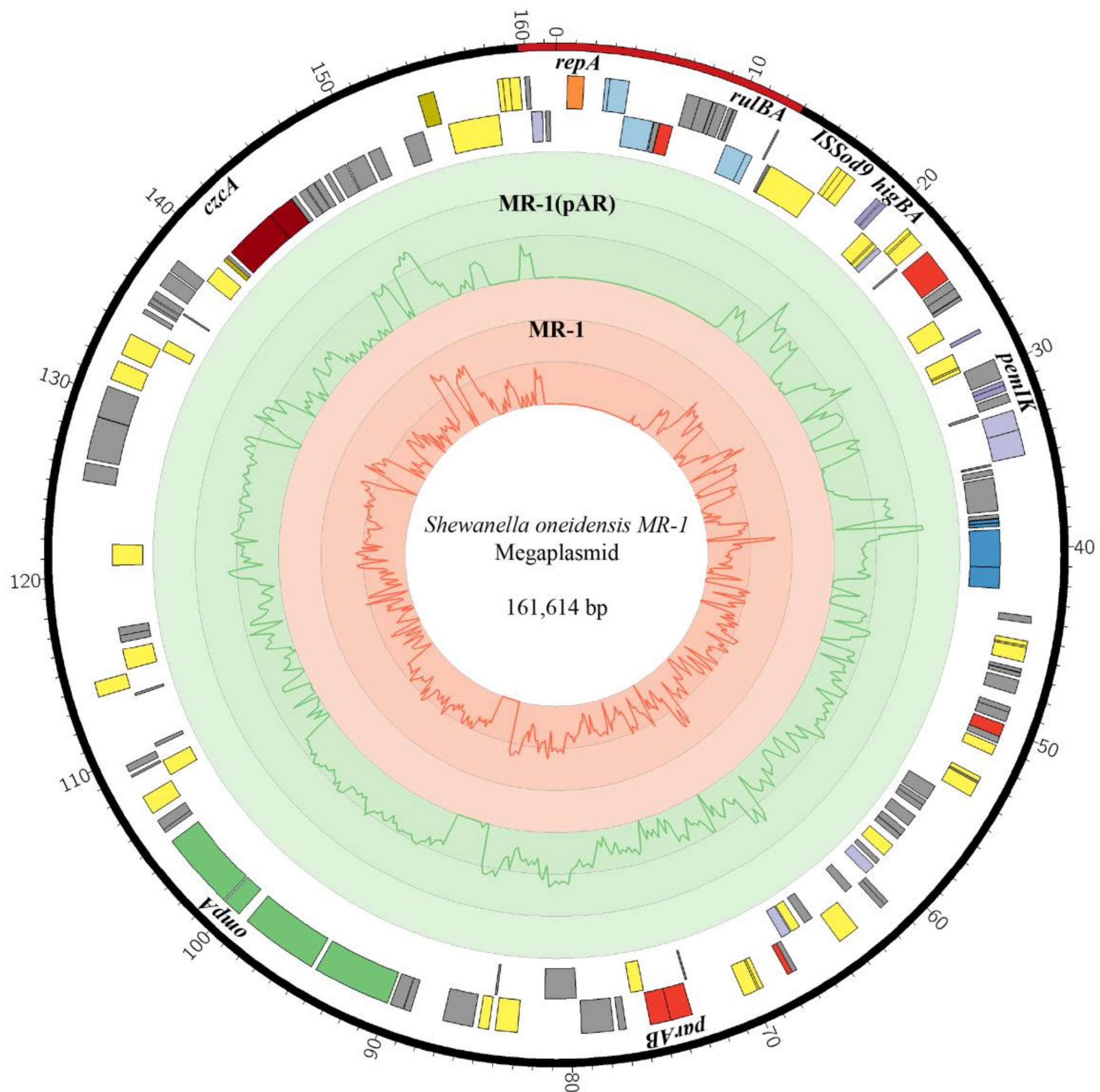

Figure 9. Transcriptome map of *S. oneidensis* megaplasmid in MR-1 with and without pAR060302. The rings from outside-in: 1) plasmid base pair coordinates (red indicating

possible deletion); 2) gene map of the megaplasmid; 3-4) transcriptome maps of the megaplasmid in strain MR-1(pAR) (green) and MR-1 (red). Genes on the map are colored by function: gray, hypothetical protein; yellow, mobile genetic elements; light purple, recombination; orange, replication; red, plasmid partitioning; light blue, DNA restriction/modification; purple, toxin/antitoxin; blue, bacterocin related; green, lipoprotein; dark red, metal efflux; dark yellow, transcriptional regulator. Line plots represent average reads aligned to 250 bp windows along the megaplasmid, normalized for sequencing library size and log-transformed. Each line on the transcriptome maps represents 1.0 log.

Discussion

The purpose of this work was to characterize how the acquisition of broad-host-range IncA/C plasmid pAR060302 might change the transcriptional network within the host cell. In addition, we were interested in how the host chromosome might influence the transcription of genes on the plasmid. Analysis of the transcriptome map of pAR060302 revealed that there are chromosome-plasmid interactions that are specific for individual genomes. It is interesting that in all but one strain, the IS*26* transposase gene is down-regulated compared to that of its expression in wild-type strain AR060302. This suggests that IS*26*-mediated gene shuffling might play a significant role in genetic rearrangements in the host strain. Unfortunately there were not more *E. coli* strains used as hosts in our experiments as this may be a phenomenon specific to IncA/C plasmid carriage by *E. coli*. This, in part, could explain the vast amount of genetic rearrangements that define the IncA/C plasmid type (16, 23). The differential expression in genes

ORF0099-0133 on pAR060302 observed in strain SL486 and, in part, MR-1 and DH10, is curious as the genes in this region have received relatively little attention concerning their biological roles. Once thought to be completely conserved, as more IncA/C plasmids have been sequenced this has proven not to be true (87). More work focused on the molecular function of genes in this region is needed to further elucidate the role they play in the biology of this plasmid type.

In order to analyze changes in chromosomal gene expression conferred by acquisition of plasmid pAR060302 we compared the transcriptomes of five strains with and without the plasmid. We used gene ontology analysis to distill meaning from large lists of differentially expressed genes. It is clear that plasmid acquisition is altering the energy flux of the cell. This has been reported elsewhere, primarily in the biotechnology literature (84, 88). These studies suggest that plasmid containing cells require more oxygen and increase the rate of glucose uptake in order to generate the additional ATP needed to replicate the plasmid, and to express any genes that may be present on the plasmid. In the present study, all strains were grown in LB broth, which is nutritionally rich, but does not contain glucose. The vast array of carbon-containing compounds available to cells may explain why expression of genes involved in metabolism was altered, but with no clear pattern when comparing all of the strains. Strains DH10 and MR-1 shared up-regulation of genes involved in glycolate metabolism and the glyoxylate bypass. These pathways are used to generate precursors for the TCA cycle (85). Conversely, the majority of the genes is these pathways were unchanged in the

Salmonella strains tested in this study. Both MR-1 and MH16125 showed down-regulation of anaerobic metabolism genes, which suggests that IncA/C plasmid acquisition might also influence oxygen flux, but that the rate might be dependent on the host chromosome. Strain-specific metabolic effects has been reported for plasmid containing strains (84). These previous studies examining the effects of plasmid carriage on host metabolism have primarily been conducted using small cloning vectors. Many of the differentially regulated metabolic pathways conflict with the results that we observed, indicated that there are likely plasmid-specific influences on the host chromosome. There is a clear need for understanding how large plasmids, such as IncA/C, with many genes of unknown function, might be affecting host metabolism. Future experiments should focus on strictly defined media to better understand utilization of specific carbon sources. Work in this area might help elucidate new barriers to host range or explain fitness landscapes in natural environments.

In our experiments, many genes located on pathogenicity islands in *Salmonella* were differentially expressed upon acquisition of pAR060302. These effects were not consistent across all strains of *Salmonella*. Two studies have demonstrated the ability of IncA/C plasmids to mobilize genomic islands located on the *Salmonella* genome (35, 89). Carraro *et al*. confirmed that this is due to the positive regulators of IncA/C conjugative transfer, AcaDC. It remains to seen whether or not IncA/C plasmids might also encode regulators that could alter pathogenic phenotypes, such as invasion or secretion. The activation or repression of virulence genes could be tied to other plasmid functions, such

as replication. The scenario where a plasmid that confers resistance to multiple antibiotics, and also alters expression of virulence genes, is truly an alarming one. Future work is needed to further characterize the myriad of putative transcriptional regulators encoded on IncA/C plasmids.

Changes to host metabolic pathways observed in the present study beg the question: are these changes merely due to new energy requirements associated with plasmid acquisition or are they due to specific interactions of genes encoded on the plasmid directed at altering host gene expression? Plasmid pAR060302 carries several interesting hypothetical proteins that could directly mediate such interactions with chromosomal metabolic pathways: ORF0053 shares similarity to MhpR, an activator of phenylpropionate degradation; ORF0108 shares similarity with BglG (BglC), an activator of the sugar-transporting phosphotransferase *bgl* operon; and ORF0112 shares similarity to iron-sulfur cluster proteins involved in electron transport. Other plasmid types have been shown to encode novel catabolic proteins (90). It is tantalizing to consider that IncA/C plasmids may carry genes that might finely tune host metabolic pathways to reduce their fitness burden. More molecular genetics based experiments will be needed to clearly define the function of these putative protein and substantiate these broad ideas.

Materials and Methods

Bacterial strains and growth conditions. The bacterial strains used were as

follows: *Escherichia coli* strain DH10B (DH10), *E. coli* strain AR060302 (AR060302) (15), *Salmonella. enterica* serovar Newport strain SL317 (SL317) (14), *Salmonella. enterica* serovar Heidelberg strain SL486 (SL486) (14), *Salmonella enterica* serovar Enteritidis strain MH16125 (MH16125) (14), and *Shewanella oneidensis* strain MR-1 (MR-1) (86). A nalidixic acid resistant mutant of each strain (except for DH10 (rifampicin resistant) and AR060302) was selected and used as a plasmid host. Plasmid pAR060302 was moved by conjugation into each strain, except AR060302, which already harbors the plasmid. Successful transconjugants were selected for using Difco^{TM} Luria-Bertani (LB) agar amended with either naldixic acid (30 ug/mL) or rifampicin (100 ug/mL) and florfenicol (16 ug/mL). All strains were grown in 10 mL LB broth aliquots at 37° C with shaking (200 RPMS), except MR-1, which was grown at 30° C, until an OD_{600} of 0.5 was achieved. Cells were pelleted and RNA was purified using a commercially available RNA extraction kit (Qiagen). Treatments were included to remove DNA contamination (Qiagen) and ribosomal RNA (MicrobExpress, Ambion). Two biological replicates for each strain were pooled for paired-end library sequencing (either 50 or 100 bp reads) via Illumina Genome Analyzer II at the Biomedical Genomics Center at the University of Minnesota. The details of each sequencing run and subsequent read mapping are reported in Table 4. All of the sequencing data are publically available under the NCBI BioProject ID PRJNA273283.

RNA-Seq Analysis. All Perl scripts and other computational biology resources used in this study can be found at https://github.com/kevinslang. cDNA reads were first trimmed so that the quality at each base position was above 30 and then mapped to the

appropriate genome or plasmid sequence (for pAR060302, GenBank accession no. NC_092692, for DH10, the *E. coli* K-12 MG1655 published sequence was used GenBank accession no. NC_000913, for all *S. enterica* strains the LT2 published sequence was used, GenBank accession no. NC_003197, for *S. oneidensis*, GenBank accession no. NC_004347 and its megaplasmid, GenBank accession no. NC_004349). Read mapping was done using BOWTIE (67). For each host, transcriptome maps of pAR060302 were constructed using Circos (91). To achieve this, a table was generated containing the average number of reads mapped per 250 bp of plasmid sequence. Each average was then normalized per 1 million total reads in the cognate sequence library. These averages were then log transformed and plotted as a line plot.

Table 4. Details of RNA-Seq reads obtained.

Strain	Number of Reads	Read size	Accession Number	Mapped to Genome	Mapped to pAR
DH10	1.94E+07	100 bp	SRR1768060	1.60E+07	N/A
DH10(pAR)	1.49E+07	100 bp	SRR1768074	1.40E+07	9.90E+04
SL317	1.44E+07	100 bp	SRR1768368	1.30E+07	N/A
SL317(pAR)	1.32E+07	100 bp	SRR1768369	1.20E+07	1.30E+05
SL486	2.58E+07	50 bp	SRR1768404	2.30E+07	N/A
SL486(pAR)	3.32E+07	50 bp	SRR1768405	3.00E+07	1.80E+05
MH16125	2.20E+07	50 bp	SRR1769820	2.00E+07	N/A
MH16125(pAR)	3.00E+07	50 bp	SRR1769821	2.80E+07	2.10E+05
MR-1	4.40E+07	100 bp	SRR1772150	4.30E+07	N/A
MR-1(pAR)	4.20E+07	100 bp	SRR1772169	4.00E+07	1.10E+05
AR060302	4.00E+07	100	SRR1772148	N/A	3.30E+05

For statistical testing of differentially expressed genes, the total number of reads mapped to each coding sequence (CDS) was calculated using Perl. These values were then analyzed using the R package EdgeR (92, 93). We conservatively estimated the dispersion at 0.001. A fold-change cutoff of > 2 or < -2 and an adjusted p-value of < 0.05 were used to define significantly differentially expressed genes. Clusters of orthologous groups (COGs) analysis was done using the online tool DAVID (82, 83).

Chapter 4

A model for the regulation of IncA/C plasmid conjugation

Summary

Conjugation plays a large role in the horizontal movement of DNA between bacteria species and even genera. Large conjugative plasmids in Gram-negative bacteria are associated with multi-drug resistance and have been implicated in the spread of these phenotypes to pathogenic organisms. IncA/C plasmids often carry genes that confer resistance to multiple classes of antibiotics. Although fully sequenced IncA/C plasmids have been described for almost a decade, many questions still remain concerning their basic biology. In this work, we investigated the function of genes that were involved in the regulation of IncA/C conjugation. We developed an IncA/C variant, pARK01, by removing resistance genes carried by the plasmid in order to make a suitable model for genetic based experiments. Using pARK01, we constructed additional mutations that show that *acaDC* encode positive regulators of conjugations. Furthermore, we conducted RNA-Seq and ChIP-Seq experiments to elucidate the regulon of *acr2*, an H-NS-like gene that is a negative regulator of conjugative transfer. This study begins to describe the functional roles of transcriptional activators encode by IncA/C plasmids and shows that these regulators may extend their function to genes encoded on the host chromosome.

Introduction

Several decades ago, IncA/C plasmids were observed in florfenicol-resistant bacterial isolates related to aquaculture farms in Asia receiving antibiotics (2, 94). Subsequently, the IncA/C plasmid type has spread globally and is now a major contributor to multidrug resistance in enteric pathogens (16, 22, 23, 26). Most studies, until recently, have focused on comparative genomics of these plasmids. These reports clearly showed that IncA/C plasmids carry a core set of backbone genes that are highly conserved (16, 19, 26). The functions of some IncA/C genes are easily assigned due to their similarity with genes of known function, particularly those involved with the type IV pilus and its assembly. However, there are numerous genes with similarity only to hypothetical proteins stored in GenBank databases. There have been no attempts to study these genes at a functional level and because of this very little is known about the mechanisms that are important for plasmid carriage and spread.

Regulation of SXT conjugation in Vibrio spp. SXT is an Integrative Conjugative Element (ICE) that is commonly found in *Vibrio*. SXT is closely related to IncA/C plasmids, in that they encode a similar type IV secretion apparatus used for conjugation as well as some antimicrobial resistance genes (21, 95). The regulation of SXT transfer has been investigated in detail (61, 62). SXT transfer is controlled primarily by three genes: *setR, setC* and *setD* (Figure 10). Under steady-state conditions, SetR negatively regulates a promoter upstream of *setC* and *setD*, P_L (Figure 10a). SetR is susceptible to degradation by RecA, thus, during the SOS response when RecA becomes active within

the cell, SetR mediated repression of P_L is relieved (Figure 10b).

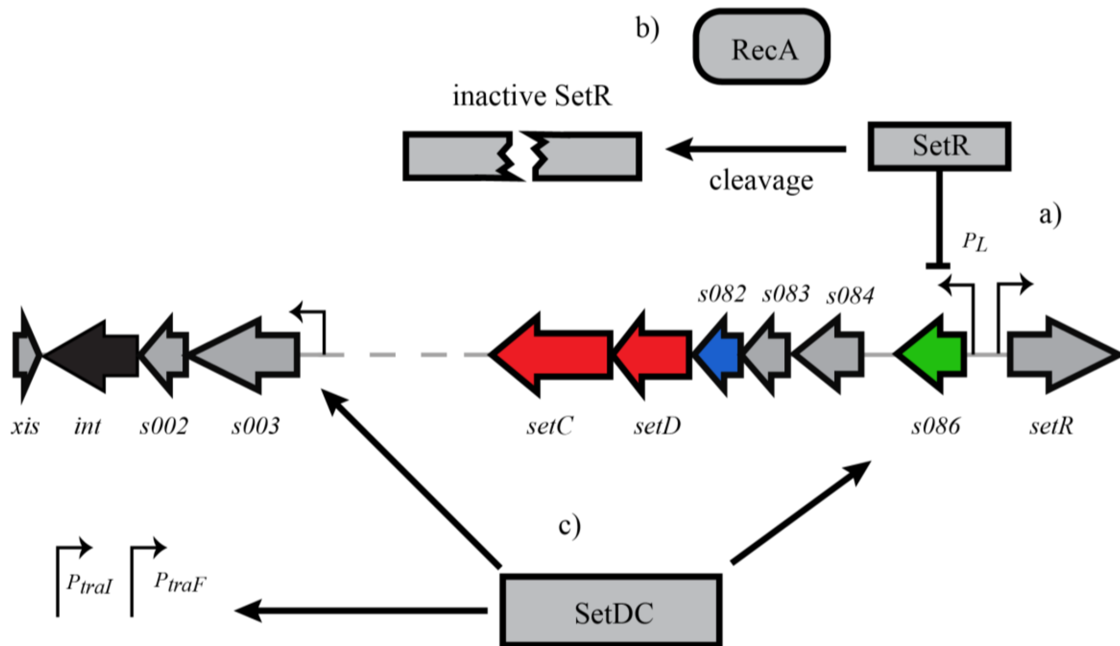

Figure 10. Model of SXT regulation of transfer. Adapted from (62). Red color indicates homology with *flhDC*. Blue color indicates predicted transglycosylase. Green color indicates homolog with *cro*. a) SetR represses at P_L. b) During SOS response activated RecA degrades SetR. c) SetDC activates transcription of various promoters.

SetC and SetD are homologous to FlhC and FlhD, respectively. FlhDC forms a complex that positively regulates motility in *E. coli* (96). Similarly, SetDC positively regulates several promoters on SXT, including those upstream of the genes that encode the conjugative transfer machinery (62) (Figure 10c). DNA damage is one of the stimuli that activate the SOS response in bacteria (62). Intuitively, during times of DNA damage it would be beneficial for the survival of a mobile element to transfer to neighboring cells.

IncA/C plasmids don't encode a homolog of SetR, so their conjugative mechanisms may be regulated by a separate set of environmental stimuli. They do, however, encode homologs of SetD and SetC, AcrD and AcrC. The function of these proteins has only been inferred by comparative genomics and has not been tested experimentally.

H-NS-like nucleoid associated proteins. There are several nucleoid-associated proteins encoded by *E. coli*. H-NS is among the most well studied (97). H-NS binds preferentially to promoter sites upstream of genes and DNA with a low G + C content (98). These sequences serve as sites of initial binding. After these initial interactions, other copies of H-NS can bind to each other via protein-protein interactions, as well as the DNA, forming large complexes (98, 99). This is achieved through two distinct domains. The N-terminus of H-NS contains a oligomerization domain and the C-terminus contains a DNA binding domain (100). Primarily, H-NS binding results in silencing of adjacent gene transcription (101). Overcoming H-NS mediated silencing has been shown to be accomplished by binding of other proteins to the DNA, altering the topography (102).

Chromosomal copies of H-NS play an important role in regulation of horizontally acquired DNA, such as pathogenicity islands (103–105). In some cases, mobile elements encode their own copies of H-NS homologs (31, 33, 106). These horizontally encoded homologs of H-NS have been shown to antagonize the binding of the chromosomal copies of H-NS to horizontally acquired DNA (102, 107). These antagonistic H-NS

homologs have been found only on genomic islands of pathogenic *E. coli*, which makes sense, given their specific function. There have been a few studies focusing on plasmid encoded H-NS homologs (31–33, 108). H-NS homologs from plasmid pSfR27 and pCAR (Sfh and Pmr, respectively) seem to play roles in regulating a diverse set of genes, some of which are regulated by the chromosomally encoded H-NS copies (31–33). It has been proposed that uncontrolled expression of these genes, caused by plasmid acquisition, could lead to a reduction in fitness and subsequent loss of the plasmid from the population. The H-NS homolog encoded on the R27 plasmid of *E. coli*, H-NS_{R27}, has been shown to directly interact with the plasmid's origin of replication, *oriT* and other transfer associated genes to regulate conjugation (108). H-NS_{R27} was shown to be involved in an intricate interplay of chromosomally encoded H-NS homologs to thermally regulate the expression of the conjugative transfer apparatus of R27.

Aside from comparative analyses, no work has been done to elucidate the function of the transcriptional regulators encoded on IncA/C plasmids. In this study, we described the function of a cluster of transcriptional regulators. We found that three of these genes regulate conjugative transfer of IncA/C plasmids. One of these genes, *acr2*, encodes an H-NS homolog. RNA-Seq and ChIP-Seq were used to characterize Acr2 binding sites. Using site directed mutations, we found that Acr2 shares a DNA binding motif with that of other H-NS homologs and that this collection of amino acids is critical for its function as a repressor of conjugation.

Results

Development of pARK01. A plasmid such as pAR060302 with its many antimicrobial resistance genes makes it difficult to study systematically using modern genetic techniques. The most common method for making deletions in *E. coli* is via site directed recombination using a dsDNA substrate (recombineering) (109). To prime cells for recombineering, a vector is commonly used, which also requires antimicrobial selection for its maintenance. We first set out to develop a derivative of pAR060302, lacking tetracycline and ampicillin resistance in order to obtain a prototypical IncA/C plasmid that could be used for more in depth genetic analysis. Figure 11 shows the schematic for deletion of *tet(A)*, which confers tetracycline resistance. The details of this are described in detail in the materials and methods. Briefly, we used strain DY331, which carries the genes needed for recombineering on a stable prophage, circumventing our need for maintenance by antimicrobial selection. Subsequently we inserted a module conferring kanamycin resistance and a synthetically lethal copy of *ccdB*, a potent toxin, which is under control of a rhamnose-inducible promoter. This module allows for selection of the initial insertion and subsequent counter-selection of the removal of *ccdB*, allowing for repair of the insertion in any manner we choose using either a ssDNA or dsDNA substrate. In this case we chose to remove the module, leaving only the start and stop codon of *tet(A)* (Figure 11c). A similar strategy was used to eliminate the $bla_{\text{CMY-2}}$ gene as well. Figure 11d demonstrates that the conjugation frequency of pARK01 is no different than that of the intact plasmid. Furthermore the conjugation frequency of pARK01 is similar to that of the WT strain AR060302 under laboratory conditions.

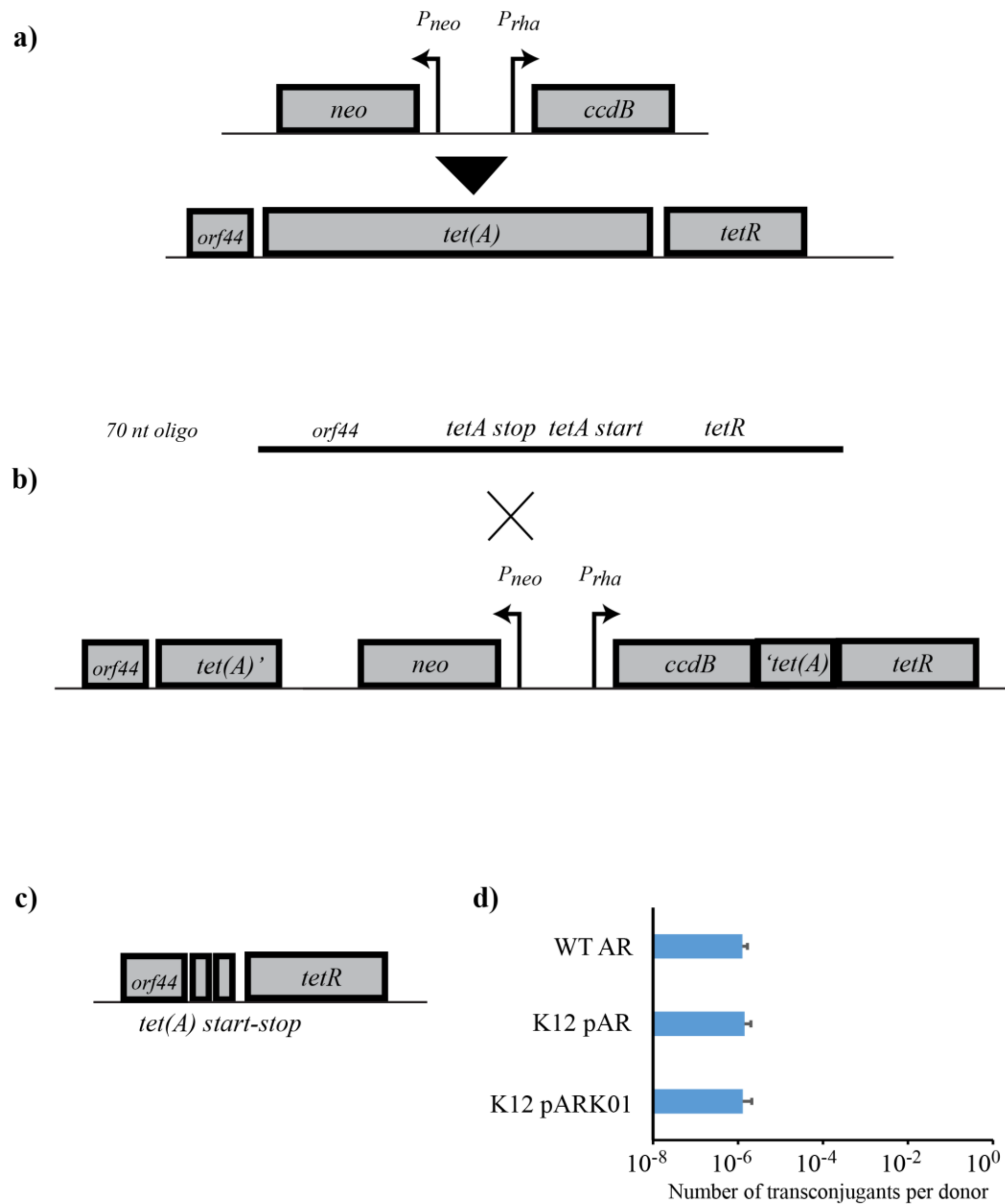

Figure 11. Development of pARK01. a) The *neo-ccdB* cassette was inserted into the *tet(A)* gene. The *neo-ccdB* cassette confers kanamycin resistance as well as lethality on minimal media plates with rhamnose used as the carbon source. b) A 70 nt ssDNA oligo was used to repair the *neo-ccdB* insertion in cells primed for recombineering. c) The repair makes a scarless mutation leaving behind only the start and stop codon of *tet(A)*. d)

Conjugation experiment showing conjugation frequency of donor strains WT strain AR060302 (WT AR), K-12 strain MG1655 carrying pAR060302 (K12 pAR) and K-12 strains MG1655 carrying pARK01 (K12 pARK01). The results represent the means and standard deviations of 3 experiments.

The genes* acrD *and* acrC *encode the positive regulators of IncA/C transfer. A cluster of predicted transcriptional regulators has been demonstrated to be conserved amongst all sequenced IncA/C plasmids (16). Given the similarity of the organization of these genes on IncA/C plasmids and SXT elements we investigated whether or not these genes were responsible for the control of IncA/C plasmid conjugation (Figure 12a). The putative *orf183* of IncA/C plasmids shares homology only with hypothetical proteins. Some of these proteins are listed as having permease activity, but this is likely due to inadequate annotation. Although *orf183* shares no amino acid similarity with SetR but because of its orientation and proximity to the other predicted transcriptional regulators, it may be negatively regulating conjugative transfer. Deletion of this gene does result in a slight increase in transconjugants, however not to the extent that would be consistent with it serving a role analogous to *setR* of SXT (Figure 12b). SetR is such a strong repressor of conjugation that deletion of the gene which encodes it is lethal (62). We found that this is not the case with *orf183*.

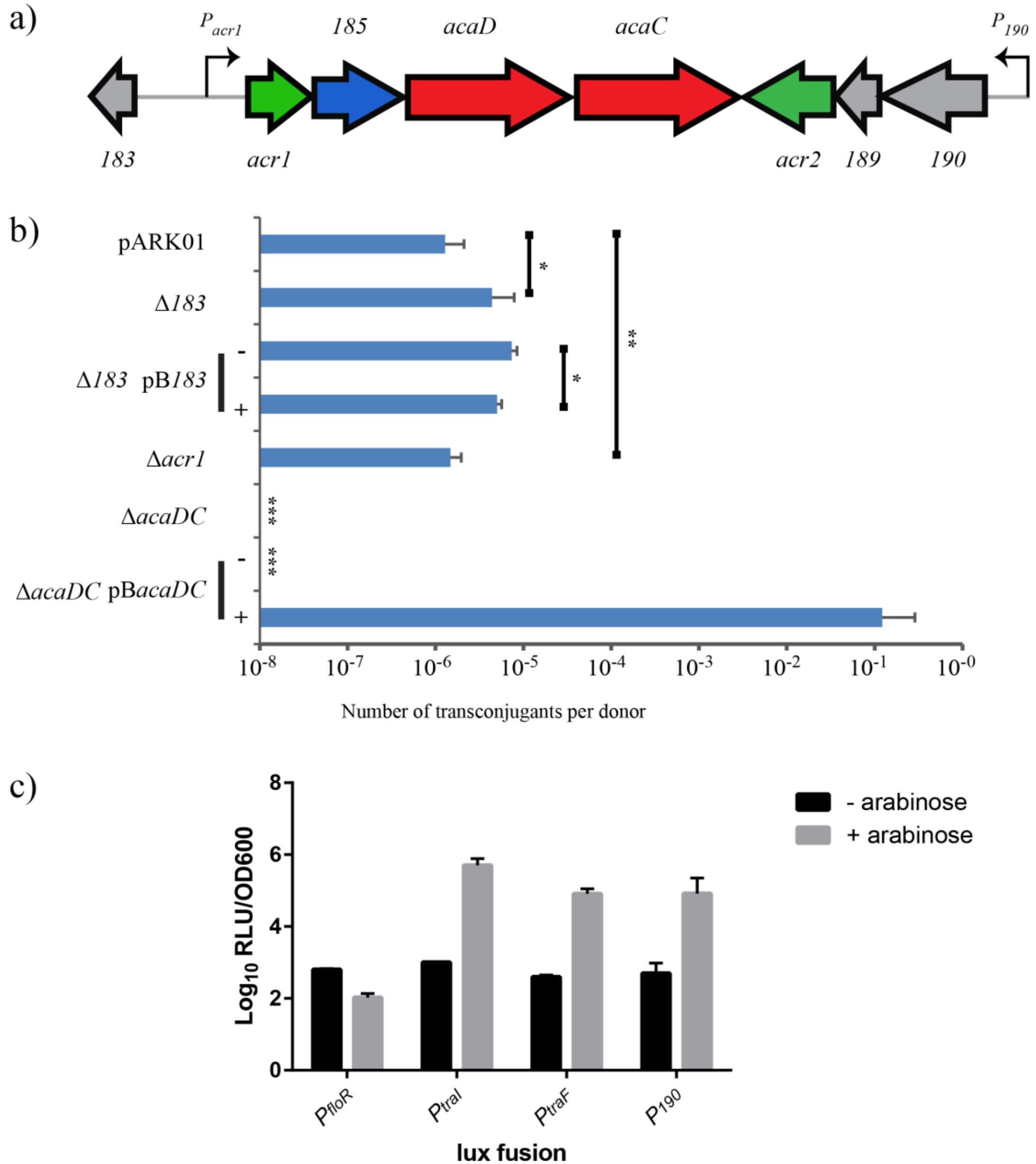

Figure 12. Regulation of conjugation locus of IncA/C plasmids. a) Gene map of the cluster of transcriptional regulators that regulate conjugation. Red indicates homology to *flhDC*. Green indicates predicted DNA binding domains. Blue indicates predicted transglycosylase activity. b) Conjugation frequency of mutants. Results represent mean

and standard deviation of 3 experiments. Means were compared using a 2-tailed Student's T-test. * indicates p-value > 0.05. ** p-value < 0.05. c) Activity of promoter fusions to *lux* operon. Results represent mean and standard deviation of 3 experiments.

Given its similarity to the phage gene *cro* we predicted that *acr1* would be involved in a negative feedback loop, limiting expression of genes involved in conjugation once they have been upregulated. This prediction turned out to be false, as a deletion mutant conferred no change in conjugative frequency (Figure 12b). Due to their similarity with SetDC encoded on SXT, we investigated whether AcrDC served a similar role as positive regulators for IncA/C plasmids. Indeed, deletion of these genes results in a plasmid devoid of transfer and over expression of these genes from an inducible promoter results in increased transfer frequencies (Figure 12b). To further investigate this phenomenon, we cloned promoters upstream of transfer-associated genes into a *lux* reporter vector. Our results demonstrate that AcaDC positively regulates these promoters (Figure 12c). The promoter upstream of *orf190* is analogous to the promoter controlling the expression of the *int* gene in SXT. Curiously, in IncA/C plasmids, the *int* gene has been replaced by a gene encoding a putative H-NS-like NAP, *acr2*. Our results suggest that the same genes responsible for positive regulation of conjugation also positively regulate *acr2*.

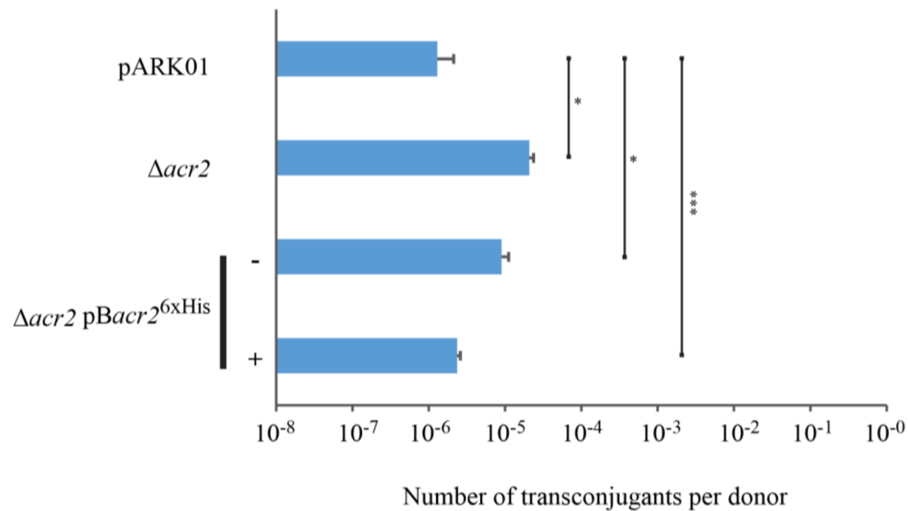

Figure 13. Conjugation frequency of *acr2* mutants. Results represent the mean and standard deviation of 3 experiments. Means were compared using a 2-tailed Student's T-test. * indicates p-value > 0.05; *** indicates p-value < 0.05.

Acr2 is an H-NS-like protein that represses conjugative transfer of IncA/C plasmids. To further investigate the potential role of *acr2* in the conjugative transfer of IncA/C plasmids, we generated a deletion mutant in pARK01 and conducted transfer experiments. Plasmids lacking *acr2* exhibit a 10-fold increase in conjugation frequency compared to plasmids with no *acr2* mutation (Figure 13). We were able to complement this phenotype by expressing *acr2* cloned upstream of an arabinose inducible promoter in the vector pBAD22. The observation that *acr2* negatively regulates conjugation corroborates our RNA-Seq results that show that nearly all of the genes that are predicted to be involved in conjugative transfer are up-regulated in the *Δacr2* strain (Figure 14-4th ring). In addition to genes involved with transfer being up-regulated, the region

containing many hypothetical proteins and putative phage-like proteins (bp positions ~82k-100k) were also up-regulated in our experiment.

Characterization of Acr2 binding sites. The vector used for arabinose inducible expression of *acr2* was constructed in such a way that a polyhistidine epitope was affixed to Acr2. It was crucial that this epitope-tagged Acr2 was able to complement the mutant phenotype to WT levels. We made the assumption that because the complementation strain transfer frequency was not statistically different than that of the WT, $Acr2^{6xHis}$ was likely binding to sites similar to that of the WT protein. This allowed us to perform a ChIP-Seq experiment *in vivo*. To better understand how *acr2* was regulating conjugative transfer we used nickel affinity chromatography to pull down $Acr2^{6xHis}$ cross-linked to DNA, which was subsequently sequenced. Analysis of the ChIP-Seq data revealed that $Acr2^{6xHis}$ binds to several loci on the plasmid (Figure 14, rings 1 and 2). Some of the ChIP peaks overlapped with genes of unknown function, such as near base pair coordinates 19k, 45k, 140k, and 145k. The regions that showed greatest binding were the entire IS*Ecp1* region, the regions within the *traFHG* and the region upstream of *acr1*. Binding upstream of *acr1* would suggest a direct repression of the operon encoding the positive regulators *acrDC*.

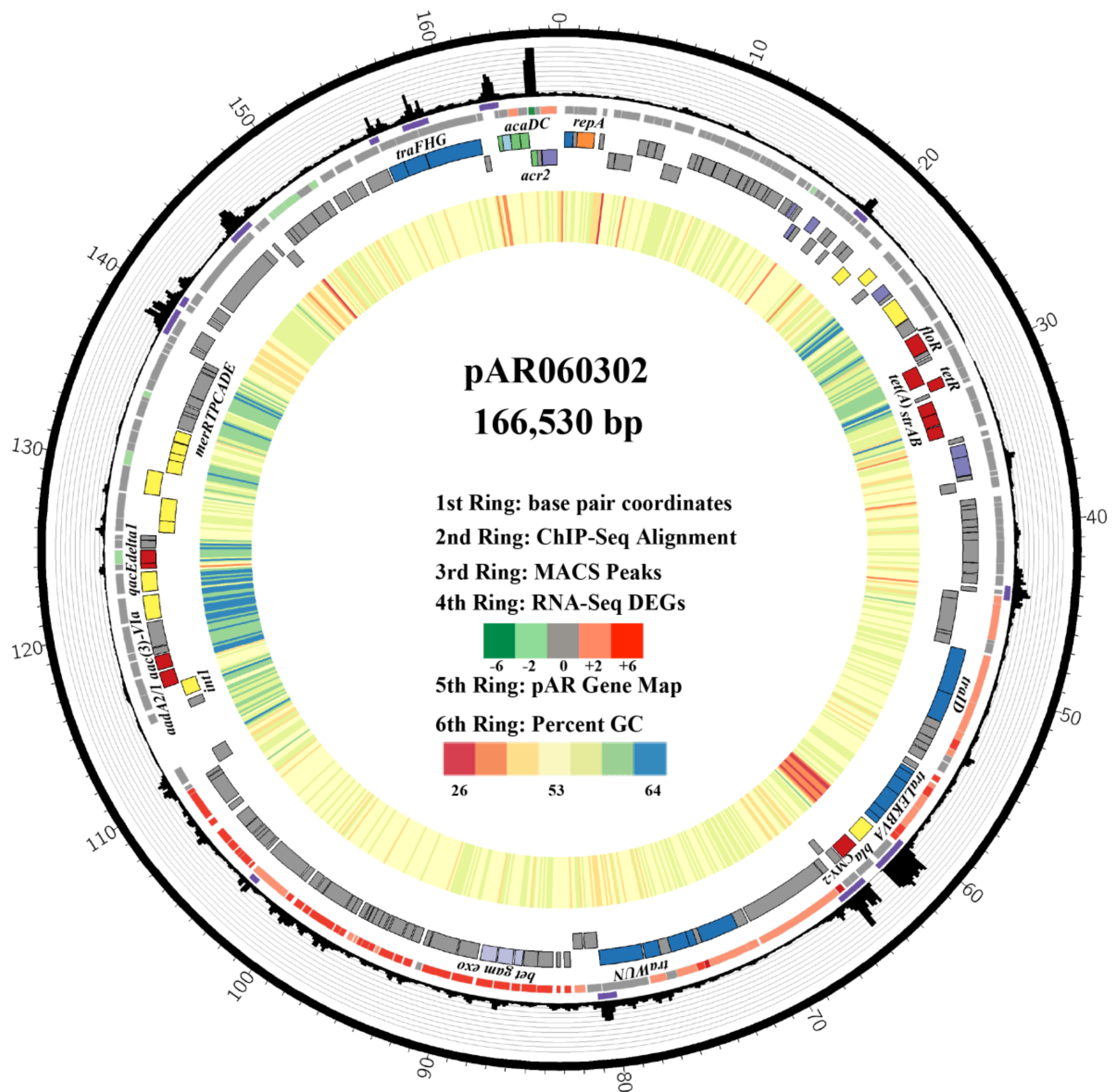

Figure 14. *acr2* regulon in pARK01. Outer ring (Ring 1) Base pair coordinates of pAR060302. (Ring2) Results of ChIP-Seq experiment. Scale (toward the outer circle) is from 0 reads to 1500 on a linear scale. (Ring 3) Enriched peaks (purple bars) as found by MACS (110). (Ring 4) Differentially expressed genes in pAR060302Δ*acr2* compared to pAR060302. Color indicates fold-change and FDR < 0.05. Grey indicates unchanged. (Ring 5) pAR060302 gene map. Genes on the map are colored by function: gray,

hypothetical protein; red, antibiotic resistance; yellow, mobile genetic elements; green, transcriptional regulation; light purple, recombination; purple, DNA binding; blue, conjugal transfer; light blue, transglycosylase; orange, replication. (Ring 6) G+C content along 150 bp windows of pAR060302.

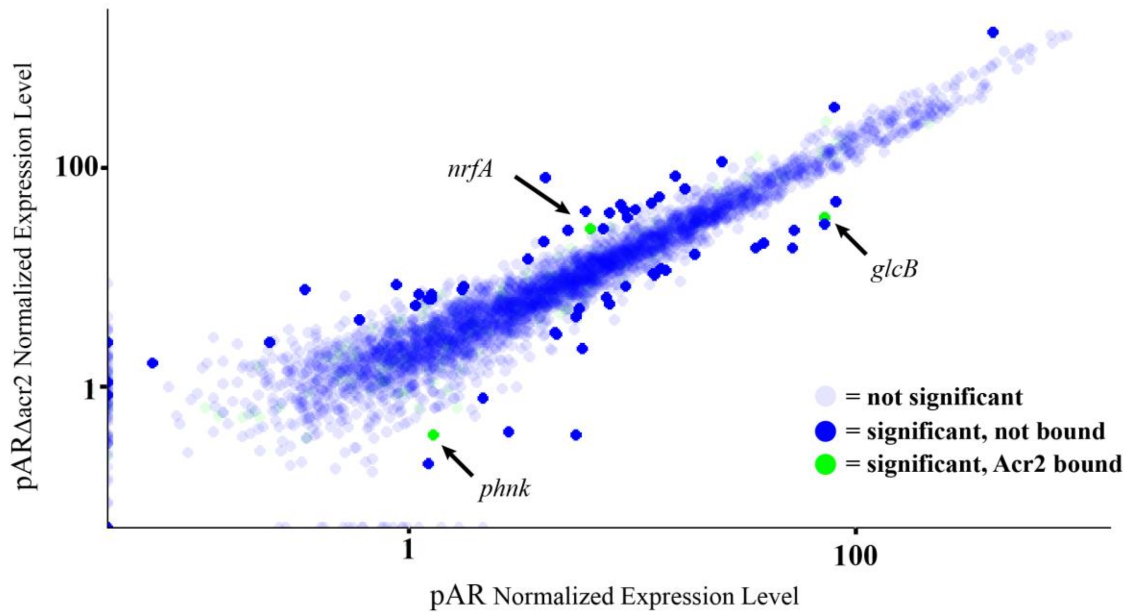

Figure 15. Expression of *E. coli* chromosomal genes carrying either pAR060302 or pAR060302Δ*acr2*. Normalized expression values were calculated using Perl and are displayed as Reads Per Kilobase of gene length per Millions of total reads (RPKM) on a log10 scale. Acr2 enriched peaks were found with MACS and annotated with BedTools.

Our ChIP data does suggest that $Acr2^{6xHis}$ binds to chromosomal DNA sequences. However, like other H-NS-like proteins, the interactions are not easily interpretable. For example, we only found 3 instances of binding of genes that were subsequently found to

be significantly differentially expressed in our RNA-Seq data (Figure 15). Interestingly, all three of these genes have to do with metabolism, most notably the *glcB* gene which is in the an operon that was discussed in Chapter 3 as being up regulated in *E. coli* carrying pAR060302, compared to cells lacking the plasmid.

We used MACS (Model-based Analysis of ChIP-Seq) (110) to determine significantly enriched peaks in DNA pulled down with Acr2^{6xHis}. The Sequence surrounding the summits of each of the significant ChIP peaks found by MACS was submitted to MEME (Multiple Em for Motif Elicitation) to determine if there were common motifs within the significant peaks. The motif discovered in a majority of the peaks is A+T rich, similar to that of other H-NS-like proteins (Figure 16b). We used EMSA to confirm that Acr2^{6xHis} binds to the DNA fragment upstream of *acr1* containing the motif, but not a different A+T rich promoter sequence (Figure 16a). Analysis using the tool Bprom (111) to computationally predict bacterial promoters shows a predicted transcriptional start site 287 bp upstream of the start codon of *acr1* (Figure 16c). If this prediction is correct, 5' end of the *acr1* transcript could conform to a highly structured, stable secondary structure as predicted by our Mfold (112) analysis (Figure 17). These results provide evidence that Acr2 directly represses transcription of the operon containing *acrDC* by binding to the sequence upstream of *acr1*.

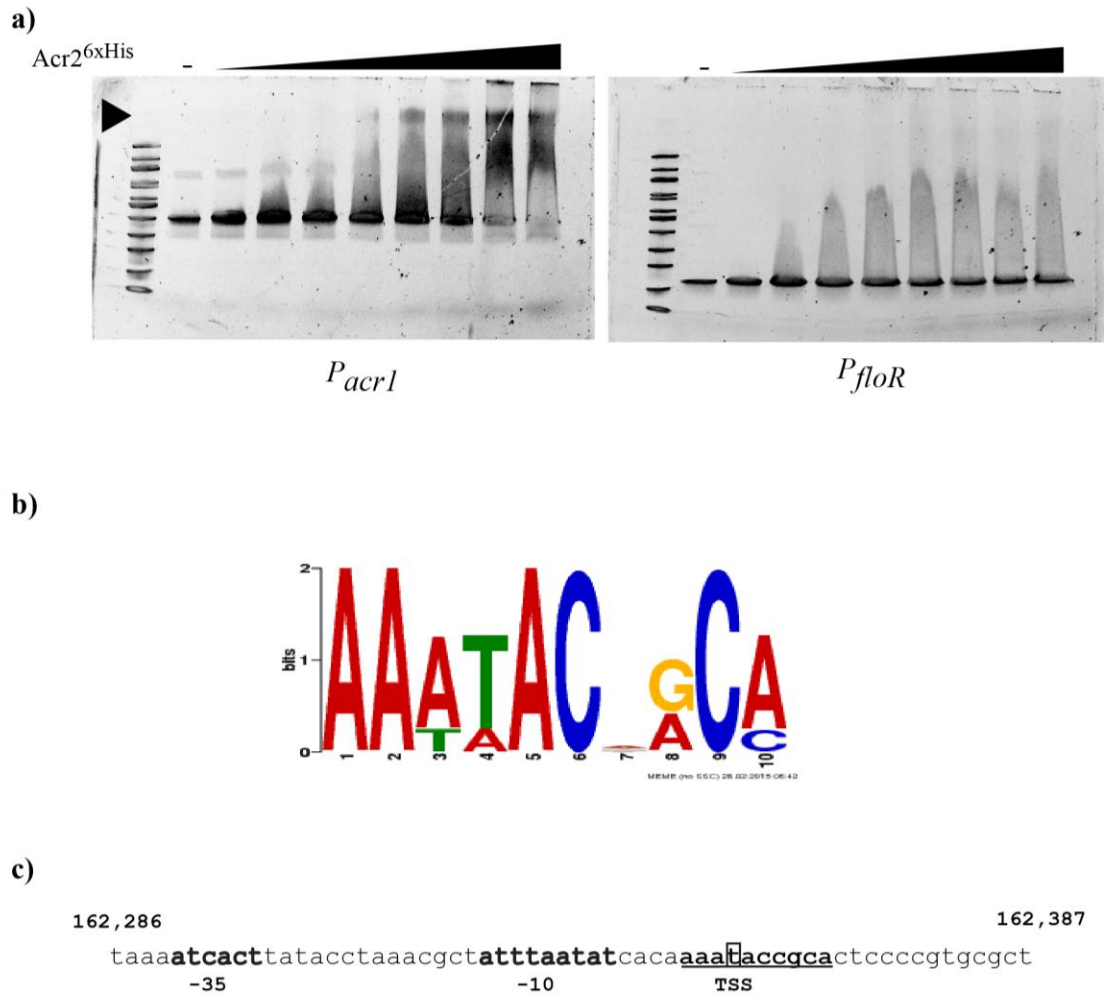

Figure 16. Acr2 binds promoter region of *acr1*. a) Electromobility Shift Assay using either a fragment of DNA upstream of *acr1* (left) or *floR* (right) combined with either no protein, or increasing protein concentrations of 150 nM, 200 nM, 300 nM, 400 nM, 500 nM, 600 nM, 700 nM, or 850 nM. DNA-protein complexes were separated on a 7.5 % native acrylamide gel. The black triangle indicates a bound DNA complex. b) Consensus DNA binding sequence of Acr2 as found by MEME (113). 200 bp of sequence surrounding each peak summit found by MACS and compared by MEME. c) Promoter region of *acr1* 323 bp upstream of *acr1* start codon. -35, -10 and TSS sites indicated were

predicted by bprom (114). The underlined sequence is the predicted binding site of Acr2.

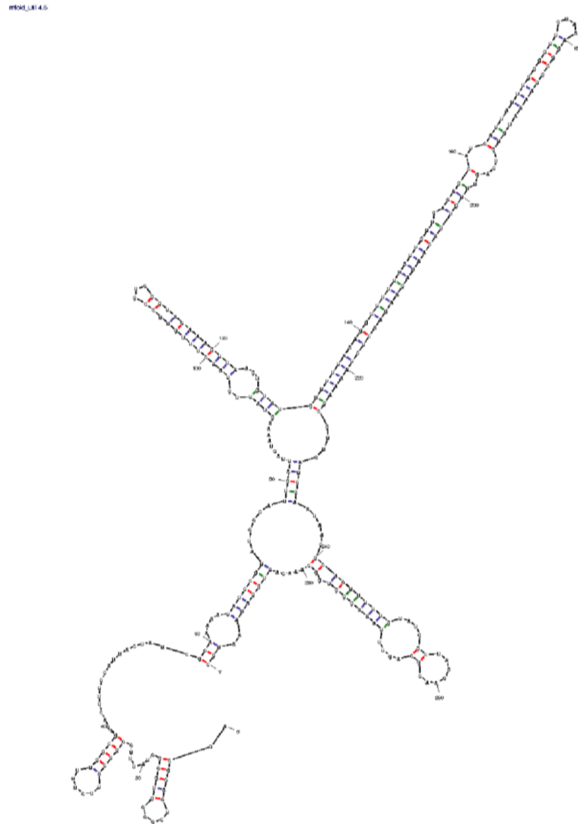

Figure 17. Predicted secondary structure of the 5' end of the predicted *acr1* transcript. 323 bp of sequence upstream of the *acr1* start codon was submitted to Mfold online prediction tool (112). The resulting structures all had a ΔG ~ 73 kcal/mol at 37° C.

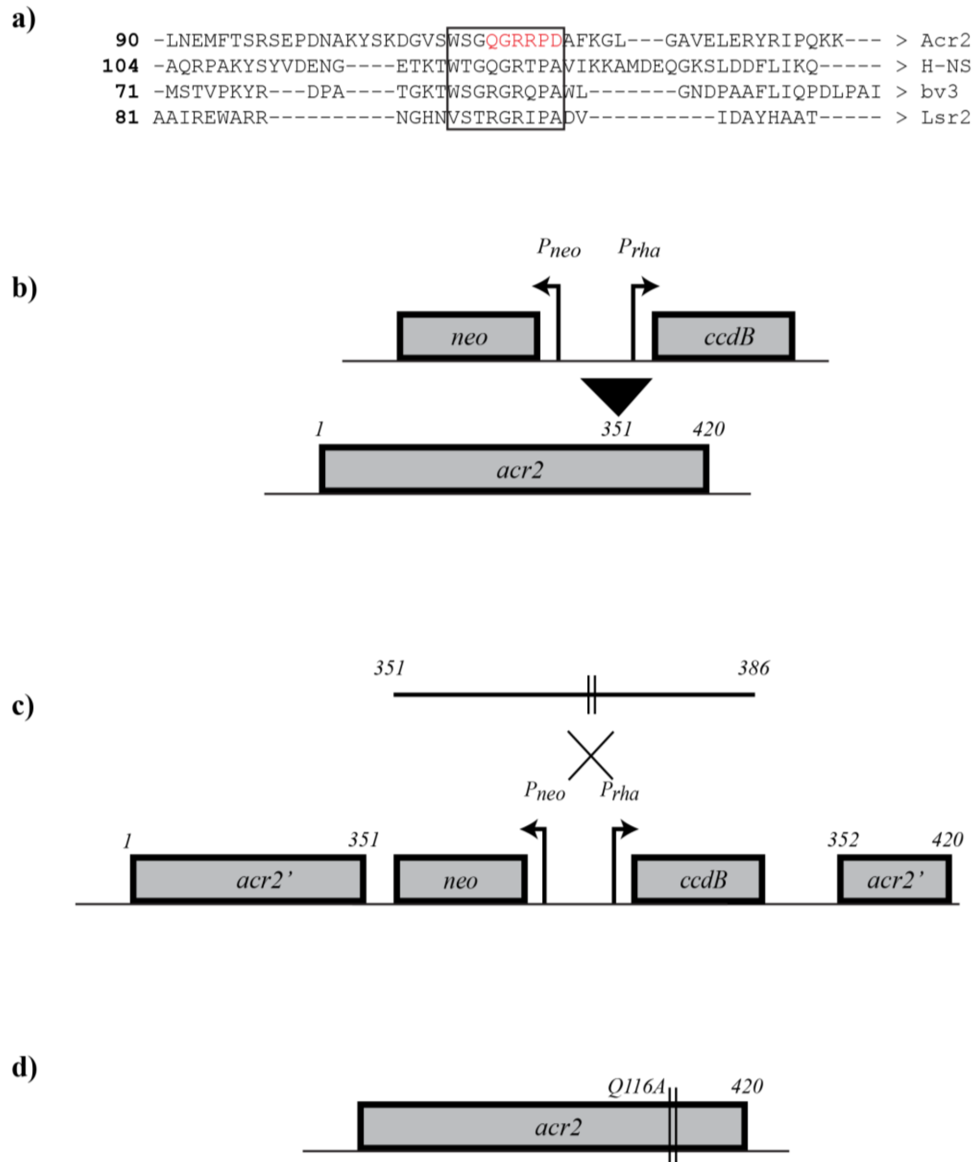

Figure 18. Development of Acr2 variants. a) Alignment of the C-terminal domain of Acr2 and other H-NS-like proteins. Box indicates region of conserved amino acid positions. Red letters indicate the AT hook motif in Acr2. b) The *neo-ccdB* cassette was inserted into the the 3' end of *acr2* via recombineering. c) a 70 nt ssDNA oligo was electroporated into cells primed for recombineering to excise the *neo-ccdB* cassette and replace it with the sequence of *acr2* containing nucleotide changes needed to achieve the

proper (d) amino acid substitutions.

The C-terminal domain of Acr2 is crucial for its activity. The C-terminal domain of Acr2 is similar to that of other H-NS-like proteins (Figure 18a). It contains the Q/RGR motif that has been shown in structure studies to be critical to contact with DNA (115). We examined the possibility that this motif was important for Acr2 activity by making mutations in this region using recombination with a ssDNA substrate (Figure 18b, c, and d, detailed in Materials and Methods). An in-frame deletion of the codons for QGRRPD resulted in abolishment of the ability of Acr2 to repress conjugative transfer (Figure 19a). Although substitution of alanine for glutamine at position 116 resulted in increased transfer frequency, it was not as a dramatic as the increase observed for alanine substitutions for the arginine residues at positions 118 and 119 (Figure 19a). Given that these substitutions could have led to proteins targeted for degradation, which could also explain the *acr2*$^{-}$ phenotypes observed, we cloned each of these Acr2 variants in an arabinose inducible vector. We used these constructs to test the ability to repress the LacZ activity of an *acr1-lacZ* fusion construct (Figure 19b). Only the WT Acr2 and the Q116A variant were able to repress LacZ activity. The Q116A result contrasts with that of the conjugation experiment. In the conjugation experiment the Acr2 variants were under native promotion. This is probably much lower than that achieved by arabinose induction of the pBAD vector, however, we cannot rule out the possibility of degradation. Taken together, these results suggest that the QGRR motif is critical for the ability of Acr2 to repress the promoter upstream of *acr1*.

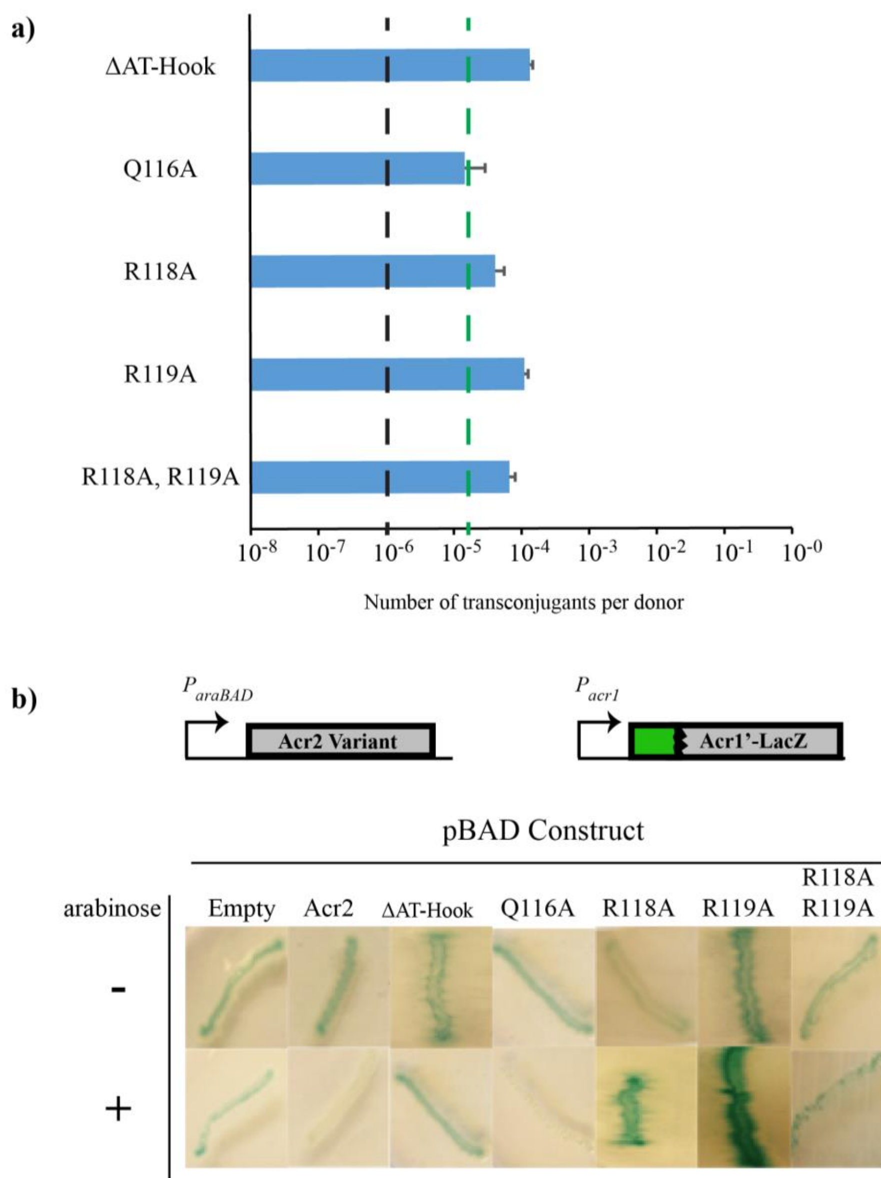

Figure 19. Activity of Acr2 variants. a) pARK01 encoding variants of Acr2 were used in conjugation experiments. Results represent the mean and standard deviation of 3 experiments. Green hashed line indicates the conjugation frequency of pARK01Δ*acr2*. Black hashed line indicates the conjugation frequency of pARK01. b) Alleles encoding Acr2 variants were cloned into pBAD and assayed for their ability to repress a *acr1-lacZ*

fusion construct. Blue color indicates LacZ activity. White color indicates absence of LacZ activity.

Discussion

IncA/C conjugation is regulated by several different transcriptional regulators. Figure 20 displays our proposed model for the regulation of transfer built from the results of our work, as well as others (19, 35). The results of this work demonstrate that Acr2 represses conjugative transfer. We showed that this is, at least in part, due to a direct interaction between Acr2 and the sequence upstream of *acr1*. During the preparation of this manuscript, Carraro *et al.* also showed this interaction using a different IncA/C plasmid isolated from *Vibrio cholera* (35). Our ChIP-Seq data suggest that Acr2 binds upstream of other genes that are predicted to be important for transfer, namely *traF* and *traG*. There are a few hypothetical proteins that are annotated on IncA/C plasmids that are located near or within the transfer-associated loci. For example, *orf97* and *orf98* are located just downstream of *traN,* both of which have no known function nor are similar to any protein sequence with known functions in GenBank. We found a ChIP-Seq peak on the 3' end of *traN* near *orf97*. It is difficult to understand the function of such an interaction due to the ambiguity of purely hypothetical proteins such as *orf97*.

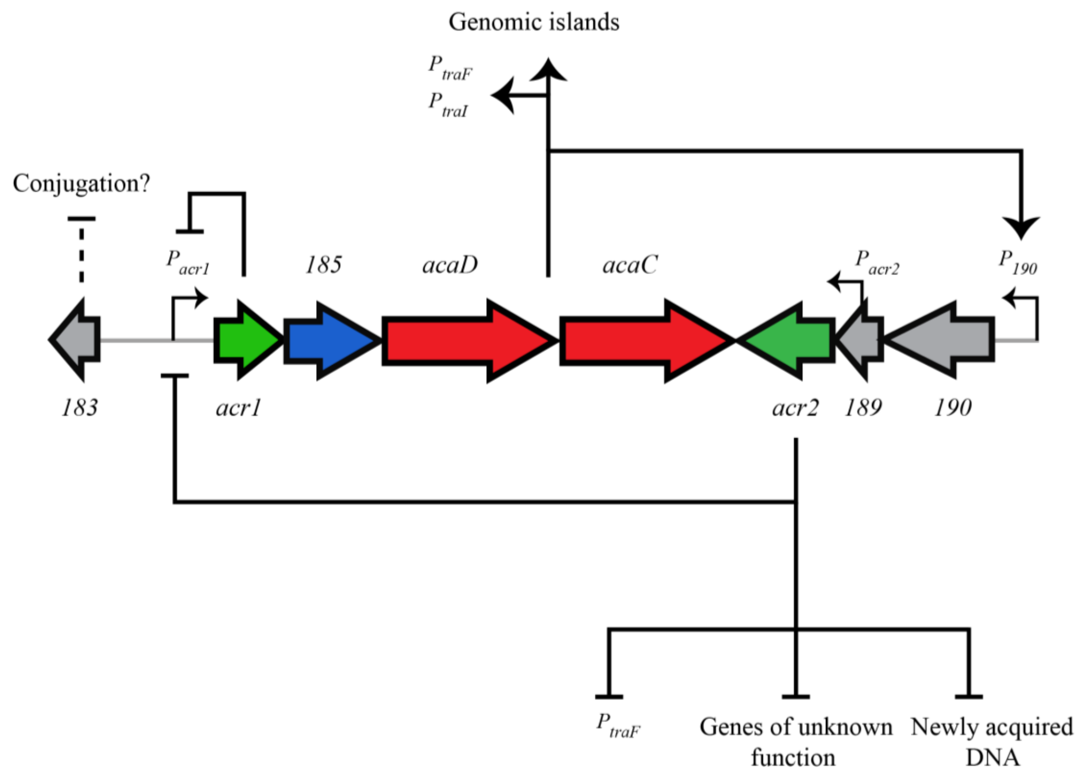

Figure 20. Proposed model for the regulation of IncA/C conjugation. Red indicates homology to *flhDC*. Green indicates predicted DNA binding domains. Blue indicates predicted transglycosylase activity. Acr2 and Acr1 have been shown to repress P_{acr1} and orf183 represses conjugation by some unknown mechanism. Expression of *acaDC* results in the up-regulation of transfer genes as well as mobilization of other genomic islands. This also may express *acr2,* which, in turn, shuts down conjugation. Acr2 acts on other loci on the plasmid including on genes of unknown function and newly acquired DNA.

There are other mechanisms that contribute to the repression of conjugation. In this work we found that *orf183* represses conjugation by a mechanism that is unclear. The predicted protein sequence encoded by *orf183* shares no similarity to proteins with

known functions in protein databases. Certainly, more work needs to be done to understand the gene product of *orf183* to understand if its role is truly to act as a repressor of conjugation. In our experiments, *acr1* deletion did not exhibit an effect on conjugative transfer. This contrasts data recently published showing that Acr1 represses conjugation and directly binds the sequence upstream of *acr1* (35). We still wonder if there might be another factor involved in repressing conjugation of IncA/C plasmids. The *setR* gene of SXT is a strong repressor of conjugation, so much so that its deletion is lethal (62). None of the known repressors of IncA/C transfer, described in this work or elsewhere, exhibit this phenotype. Furthermore, SetR acts as a proxy for detecting DNA damage. Intuitively, if the cost of conjugation is so high, one would conclude that it must be turned on only when needed.

We have shown that conjugation requires the activators AcaDC. This has been described elsewhere and, similar to homologous FlhDC, AcaDC has been shown to form a multimeric complex (35). Our experiments also demonstrate that conjugation frequencies can reach high levels when *acrDC* is over expressed, indicating that the strength of transcription of the operon containing them is still a limiting factor in *acr2* mutants. It is also possible that post-transcriptional regulation plays a role. Our computational approaches to predict the TSS suggest that it is several hundred base pairs upstream of the *acr1* start codon and that the resulting mRNA could be highly structured. The secondary structure of the untranslated portion of this mRNA could be important for regulation of its translation (65). Surely overexpression of this activity might have

negative consequences for the host bacterium. Both our results and those of others (19, 35) begin to paint a picture of how IncA/C plasmids fit in existing regulatory networks encoded by the host chromosome.

Acr2 is an H-NS-like protein that binds several sites on IncA/C plasmids. Our results demonstrate that, like other H-NS homologs, Acr2 binds A+T rich DNA, but that it also has a propensity to bind a specific signature (98, 99). It has been proposed that the amino acid composition in the AT hook motif is the mechanism by which different H-NS homologs distinguish between chromosomal and horizontally acquired DNA (115). It is thought that the number or arrangement of positively charged amino acids, such as arginine, determine the DNA recognition by H-NS-like proteins. Acr2 shares a DNA binding motif with that of other H-NS homologs and our results show that this set of amino acids is critical for its function. The AT hook motif in Acr2 contains an addition arginine residue (R119) that is not observed in other H-NS homologs. We have shown that this arginine is important for the function of Acr2. This difference might result in Acr2 having a distinct propensity to bind specific regions of IncA/C plasmids to accomplish specific tasks, such as repressing conjugation.

Other roles for Acr2. We have generated, at the very least, circumstantial evidence that Acr2 might have other roles in the biology of IncA/C plasmids. Our ChIP-Seq data suggest that Acr2 binds nearly the entire IS*Ecp1* element that carries the *bla*$_{\text{CMY-2}}$ gene. Comparative genomics studies have demonstrated that the acquisition of this transposon

is a recent event in the evolution of IncA/C plasmids (16, 23, 26). Is it possible that Acr2 binds newly acquired mobile elements on IncA/C plasmid backbones? Given the propensity for H-NS to bind genomic islands, it is seems plausible (101). Furthermore, Acr2 binds to the 3' end of the *rhs* gene located downstream of the class 1 integron (bp coordinate ~146,000). The *rhs* genes have repeat sequence elements that implicates them in genome shuffling (116). A recent work from Harmer *et al.* suggests that the *rhs* homolog carried on IncA/C plasmids plays an important role in the diversity of newly integrated mobile elements (26). Could binding by Acr2 inhibit the ability for recombination to happen at the *rhs* locus? H-NS has recently been shown to play an important role in the evolution of *Salmonella,* providing stability for its genomic islands (104). Given broad spatial distribution of highly similar variants of IncA/C plasmids, there must be several mechanisms driving their evolution.

Our ChIP-Seq experiment also yielded evidence that Acr2 binds chromosomal DNA. It is unclear what the true role of this binding is as nearly all genes bound showed no differential expression in RNA-Seq experiments. It will take more sophisticated experiments to tease out the meaning behind these results. There were three bound genes, however, that were differentially expressed in our RNA-Seq data. All three have to do with different metabolic pathways. Interestingly, one of the genes is *glcB*, which is involved in the glyoxylate bypass pathway discussed in Chapter 3. This lends some credence to the ideas posed in Chapter 3 detailing the possibility that IncA/C plasmids encode the ability to specifically alter host metabolic pathways to improve fitness of

plasmid carrying cells. This presents an almost phage-like scenario where the plasmid co-ops the host bacterium to become adept at carrying and disseminating the plasmid. It is an exceedingly interesting area of study that future work must explore.

Materials and Methods

Bacterial strains, plasmids and growth conditions. The bacterial strains and plasmids are listed in Table 5. All strains were routinely grown in Difco™ Luria-Bertani (LB) broth or LB agar at 37° C unless otherwise noted. Broth cultures were grown in a shaking incubator at 37° C with shaking (200 RPM) unless otherwise noted. Supplementation of ampicillin (Amp, 100 µg/mL), chloramphenicol (Cm, 20 µg/mL), nalidixic acid (Nal, 30 µg/mL), kanamycin (Kan, 100 µg/mL), rifampicin (Rif, 100 µg/mL) and tetracycline (12.5 µg/mL) were used as needed. Counter-selections were done using M9 minimal agar was supplemented with 0.2% rhamnose. Arabinose induction of pBAD22 vector constructs was achieved with a final concentration of 0.02 %. X-gal was added to agar plates at a concentration of 40 µg/mL. Diaminopimelic acid (DAP) was add to a final concentration of 300 µM to facilitate growth of WM3064.

Table 5. Strains and plasmids used in this study.

Strain or Plasmid	Characteristics	Source
E. coli DH10B	Rif^R strain used as a conjugation recipient	Lab stock
E. coli MC4100	lac$^-$ strain used for observing LacZ fusion activity	Lab stock
E. coli K12 MG1655	Model *E. coli* strain	Lab stock
E. coli DY331	*E. coli* with stable prophage that expresses *beta-gam-exo*	(117)
E. coli WM3064	*dapA* auxotroph	Lab stock
E. coli WM3064nalR	nal^R mutant of WM3064	This work
E. coli AR060302	Wild type host of pAR060302	(15)
pAR060302	IncA/C plasmid, FloR, CmR,	(15)
pARK01	pAR060302 derivative carrying deletion of *tetA* and *bla*$_{CMY-2}$	This work
pARK01-01	pARK01 carrying deletion of *orf183*	This work
pARK01-02	pARK01 carrying deletion of *acr1*	This work
pARK01-03	pARK01 carrying deletion of *acaDC*	This work
pARK01-04	pARK01 carrying deletion of *acr2*	This work
pARK01-05	pARK01 carrying lacZ fusion of the first 4 codons of *acr1* and *lacZ* and a deletion of *acr2*	This work
pARK01-06	pARK01 carrying the Acr2ΔAThook mutation	This work
pARK01-07	pARK01 carrying the Acr2(Q116A) mutation	This work
pARK01-08	pARK01 carrying the Acr2(R118A) mutation	This work
pARK01-09	pARK01 carrying the Acr2(R119A) mutation	This work
pARK01-10	pARK01 carrying the Acr2(R118A and R119A) mutation	This work
pBBRLux	Plasmid carrying a promoterless *lux* operon; Cm^R	(118)
pLux	pBBRLux with *tetRA* in place of *cat*	This work
pLuxFloR	pLux with *floR* promoter cloned in between SpeI and BamHI sites	This work
pLuxTraI	pLux with *traI* promoter cloned in between SpeI and BamHI sites	This work
pLuxTraF	pLux with *traF* promoter cloned in between SpeI and BamHI sites	This work

pLux190	pLux with *orf190* promoter cloned in between SpeI and BamHI sites	This work
pBAD22	arabinose inducible vector, Amp^R	(119)
pB183	pBAD22 with orf183 cloned	This work
pBacr2	pBAD22 with *acr2* cloned	This work
pBacaDC	pBAD22 with *acaDC* cloned	This work
pBacr2ΔAThook	pBAD22 with *acr2ΔAThook* cloned	This work
pBacr2Q116A	pBAD22 with *acr2Q116A* cloned	This work
pBacr2Q116A	pBAD22 with *acr2Q116A* cloned	This work
pBacr2R119A	pBAD22 with *acr2R119A* cloned	This work
pBacr2R118A,R119A	pBAD22 with *acr2R118A,R119A* cloned	This work
pCP20	plasmid expressing FLP recombinase, Amp^R, *ts rep*	(109)
pSIM5-tet	plasmid expressing *beta-gam-exo* used for recombineering, tet^R	(117)
pKD4	plasmid carring FRT-*neo*-FRT cassette, Kan^R, Amp^R	(109)
pKD45	plasmid carrying *neo-ccdB* cassette, R6K *ori*	(117)

Table 6. Oligonucleotides used in this study.

Name	**Sequence**
pARdeltaTet-fw	CCGGCATTCGCTGCGCTTATGGCAGAGCAGGGAAAGGAATTGCCGGGCTAGAGCCTGACATTTATATTCC
pARdeltaTet-rv	GGCTGCAACTTTGTCATGCTTGACACTTTATCACTGATAAACATAATATGGTCCCGCTCAGAAGAACTC
deltaTetRepairOligo	CTTATGGCAGAGCAGGGAAAGGAATTGCCGGGCTACATATTATGTTTATCAGTGATAAAGTGTCAAGCAT
pARdeltaBlaCMY2-fw	CTGACGGGCCCGGACACCTTTTTGCTTTTAATTACGGAACTGATTTCATGGAGCCTGACATTTATATTCC
pARdeltaBlaCMY2-rv	AAAAAAGAGAAAGAAAGGAGGCCCAATATCCTGGGCCTCATCGTCAGTTAGTCCCGCTCAGAAGAACTC
deltaCMYrprOligo	ACCTTTTTGCTTTTAATTACGGAACTGATTTCATGTAACTGACGATGAGGCCCAGGATATTGGGCCTCCT
183_KOchk-F	TCGACGGAAGCTGGAGACCGT

183_KOchk-R	TGATGCCGTGGTAAACTGGAGGT
183_koCHK-F	CCGCGCAACGCGGATAAAGC
183_koCHK-R	GCTGGCTCCCTTCGGCGTTT
pBAD22_orf183-rv	ATTCCCAGAAGCACCATTCCCAGCAGAAGAATATTGACCACATGTTTCATGGTGAATTCCTCCTGCTAG
pBAD22_orf183-fw	ATTCCTATGATTTTGACCAGCGCTTCCATATTGGAGCTGGAGACGAATAAGTACCCGGGGATCCTCTAG
pbadseq-2	GGCTGAAAATCTTCTCT
pbadseq-1	CTGTTTCTCCATACCCGTT
acr1_KO-F	TGTATGGCTAACAACGCAGTTTAAAGGGTGGCAAACAATGGTGTAGGCTGGAGCTGCTTC
acr1_KO-R	GCCGCAGACAGAACTAGCGCCTTGGCTTTGTTCTTCATCAATGGGAATTAGCCATGGTCC
acr1_KOchk-F	ACAACGCAGTTTAAAGGGTGGCA
acr1_KOchk-R	GCCGCAGACAGAACTAGCGCC
acaDC_KO-F	ATCTATCGCAACCTTCGTGATTTGTGAGGGGGGCGGAATGGTGTAGGCTGGAGCTGCTTC
acaDC_KO-R	GGGCGCTCATCTTCTGGTCCGAAATGTCATAGTCTACTCAATGGGAATTAGCCATGGTCC
acaDC_KOchk-F	GCCGCTATCACGCCTGGGAA
acaDC_KOchk-R	GGGCGCTCATCTTCTGGTCCG
pBAD22_acaDC-rv	TGTCGAGCTGTAACCCAGCGACCCAATGTACCAGAGTTGCCAATGTTCATGGTGAATTCCTCCTGCTAG
pBAD22_acaDC-fw	TATTGAAAAAACTGCAACAGTCTCTACAGGGACGTGGGAGGTGGAGATGAGGCTGTTTTGGCGGATGAGA
pBBR-F	CGCGCGTAATACGACTCACT
pBBR-R	TCGGGAAAGATTTCAACCTG
cat_frt-tet-frt-fw	GACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAGAATTCGAGCTCGGTACCCG
cat_frt-tet-frt-rv	ATTGGCTGAGACGAAAAACATATTCTCAATAAACCCTTTAGCTATGACCATGATTACGCC
BamHI-190p-rv	GGATCCATAACATGAAATCCTCCAAAGATT
SpeI-190p-fw	GATACTAGTTGCTGGAGGGCAAGCGGTTC
BamHI-floRp-rv	GGATCCATAACATGACGAATCAACCCTCG
SpeI-floRp-fw	GATACTAGTAAAGATAATCGGATAAAATG
BamHI-traFp-rv	GGATCCATAACACGTATTTTCCCATGCTCG

SpeI-traFp-fw	GATACTAGTGGACACGACCAAATAGTTTG
BamHI-traIp-rv	GGATCCATAACATGTAGGTTGACCATATTC
SpeI-traIp-fw	GATACTAGTTCGTTCCTCATAAAAAATTG
acr2_ko-F	AGAAGATGAGCGCCCAAACTTTGCGGCGCTCATTTTTTTAGTGTAGGCTGGAGCTGCTTC
acr2_ko-R	GATTCACTTAAAGGAAAACTGATGGGCTCCGTATGACATGATGGGAATTAGCCATGGTCC
acr2_chk-F	AGGAAGGCATGGTGGTGGGGA
acr2_chk-R	ATTCGCCAGTGCAGACTGACGC
pBAD22_acr2_6xhis-fw	AGCTGTCGAGCTGGAGCGCTACCGTATCCCACAGAAGAAGCATCACCATCA CCATCACTAAGGCTGTTTTGGCGGATGAGA
pBAD22_acr2-rv	CAGCCGCTGATAGCGTAGAGAACTCCTCGTGGTCTTTCATGGTGAATTCCTCCTGCTAG
acr1-EMSA-F	CCCTAAAATCACTTATACC
acr1-EMSA-R	CAATTGGCATTTATCCATTG
floRp-rv	ACATGACGAATCAACCCTCG
floRp-fw	AAAGATAATCGGATAAAATG
delATHOOKAcr2-kanccdB-fw	CTCCAGCTCGACAGCTCCCAACCCCTTGAAAGCATCCGGGGTCCCGCTCAGAAGAACTC
delATHOOKAcr2-kanccdB-rv	ATTCAAAAGACGGAGTTTCGTGGAGTGGGCAAGGACGTCGGAGCCTGACATTTATATTCC
delATHOOK-repair	CGCTCCAGCTCGACAGCTCCCAACCCCTTGAAAGCCCCACTCCACGAAACTC CGTCTTTTGAATATTTAG
Acr2Q116A	CTAAATATTCAAAAGACGGAGTTTCGTGGAGTGGGGCGGGACGTC GCCCGGATGCTTTCAAGGGGTTGGGAGC
Acr2R118A	AATATTCAAAAGACGGAGTTTCGTGGAGTGGGCAAGGAGCGCGCC CGGATGCTTTCAAGGGGTTGGGAGCTGT
Acr2R118A-R119A	AATATTCAAAAGACGGAGTTTCGTGGAGTGGGCAAGGA GCGGCGCCGGATGCTTTCAAGGGGTTGGGAGCTGTCGA
Acr2R119A	ATTCAAAAGACGGAGTTTCGTGGAGTGGGCAAGGACGTG CGCCGGATGCTTTCAAGGGGTTGGGAGCTGTCGA

deltalacAY-fw	GCCTGATAAGCGCAGCGTATCAGGCAATTTTTATAATTTAGTGTAGGCTGGAGCTGCTTC
deltalacAY-rv	CAGGCCATGTCTGCCCGTATTTCGCGTAAGGAAATCCATTATGGGAATTAGCCATGGTCC
acr1-lac-rv	TTTGGGCGCTCATCTTCTGGTCCGAAATGTCATAGTCTACCGTGTAGGCTGGAGCTGCTT
acr1-lac1-fw	AACAACGCAGTTTAAAGGGTGGCAAACAATGGATAAATGCCCCGTCGTTTTACAACGTCG

Strain construction via recombineering. All deletions and mutations were done via λ-Red mediated recombination (recombineering) with some variations. Strain WM3064 was grown at 37° C with shaking (200 RPM) in LB broth supplemented with DAP overnight. This culture was used to inoculate a new 10 mL LB+Dap broth culture, which was incubated for 4 hours at 37° C with shaking (200 RPM). The cells were pelleted by centrifugation at 8000 RPM for 10 minutes. The pellet was resuspended with 100 µL of LB broth and used to inoculate an LB agar plate supplement with Nal and Dap. This plate was incubated at 37° C overnight. Isolated colonies were streak purified on a new LB+Nal+Dap agar plate. This strain was struck on an LB+Dap plate and the WT strain AR060302 was struck over the top of WM3064nalR. The plate was incubated overnight. The resulting growth was struck onto an LB+Nal+CM+Dap plate to select for WM3064nalR(pAR0603020) transconjugants. This strain was used to transfer pAR060302 into strain DY331 in a similar manner except selection for transconjugants was achieved by growth on LB+Cm plates. PCR was used to amplify the *neo-ccdB* cassette from pDK45 using the primers pARdeltaTet-fw and rv (12.5 µL Phusion 2x master mix (Life Technologies™), 500 nM of each primer, 1 uL of template (boiled cells), PCR conditions: 95° C 5 min, 25 cycles of 95° C 30 sec, 55° C 30 sec, 72° C° for 2 min; then a final incubation at 72° C for 10 min). The resulting amplicon was purified with a Qiagen™ PCR cleanup kit.

Strain DY331(pAR) was grown in LB+Cm broth at 32° C with shaking (200 RPM)

for overnight. The overnight culture was used to inoculate a new 50 mL LB+Cm broth culture (1:100 dilution). This was grown until an OD600 ~0.4 at which time it was moved to a 43° C shaking water bath and incubated for 20 min. After the incubation the flask was cooled on ice for 10 min. The cells were pelleted by centrifugation at 8000 RPM for 5 min at 4° C. The cell pellet was resuspended with 25 mL ice cold H_2O. The cells were collected again by centrifugation and washed an additional time. The final cell pellet was resuspended in 200 µL of ice cold H_2O. A 40 µL aliquot of washed cells was electroporated with 0.5 – 1 µg of purified amplicon. Cells were recovered with LB broth and incubated at 32° C for 4 hours. After recovery cells were plated on LB+Kan plates. Isolated colonies were checked for growth on LB+Tet plates to confirm disruption of the *tet(A)* gene. Successful mutants were mated on LB plates with *E. coli* strain K-12 MG1655 carrying pSIM5-Tet at 32° for overnight. Successful transconjugants were selected on LB+Cm+Tet plates. This strain was used to remove the *neo-ccdB* cassette by recombineering using a ssDNA substrate (Figure 11b) (120). Cells were primed for recombineering as done above and the 70 nt oligonucleotide DeltaTetRepairOligo was electroporated. Cells were recovered and plated on M9+Rhamnose+Tet+Cm plates. Isolated colonies were checked for mutations via PCR. The entire insertion-removal process was then done a second time in the K-12 strain to target and remove the $bla_{\text{CMY-2}}$ gene using the appropriate primers.

Deletions of predicted transcriptional regulator on pAR060302 were done recombineering in strain K12(pARK01, pSIM5-tet). The plasmid pKD4 carrying the

FRT-*neo*-FRT cassette was used as a template for PCR (12.5 µL Phusion 2x master mix (Life Technologies™), 500 nM of each primer, 1 uL of template (boiled cells), PCR conditions: 95° C 5 min, 25 cycles of 95° C 30 sec, 55° C 30 sec, 72° C° for 2 min; then a final incubation at 72° C for 10 min). PCR amplicons were subjected to a DpnI digestion to eliminate template plasmid and then electroporated into cells primed for recombineering. Resulting colonies were checked by PCR for mutations. Successful mutants were then grown and made electrocompetent and the plasmid pCP20 was electroporated. These transformants were grown at 32° C for 48 hours to "flip" out the *neo* cassette, leaving a copy of FRT. The plasmid pCP20 was lost via incubation at 37° C for 48 hours. Colonies were screened via PCR and amplicons were verified by sequencing.

Molecular cloning methods. For constructs using pLUXtet, containing a promoterless *lux* operon, cloning was done via double digestion with SpeI and BamHI and ligation with DNA ligase. Oligonucleotides used to generate PCR amplicons of promoter regions with flanking ends containing restriction sites are listed in Table 6. To clone genes from pAR060302 into pBAD22, *in vivo* cloning via recomineering was used (117). In order to generate a template for PCR, pBAD22 was digested overnight using BamHI. The resulting digestion was used as template for PCR (12.5 µL Phusion 2x master mix (Life Technologies™), 500 nM of each primer, 1 uL of template, PCR conditions: 95° C 5 min, 25 cycles of 95° C 30 sec, 55° C 30 sec, 72° C° for 6 min; then a final incubation at 72° C for 10 min). Primers used for *in vivo* cloning were designed to

amplify the vector pBAD22, and add 40 bp regions of homology starting with 2nd codon of the gene of interest to the reverse primer and 40 bp of homology ending with the codon next to the stop codon for the forward primer. PCR amplicons were purified with a Qiagen™ PCR Cleanup Kit and electroporated into K12(pARK01, pSIM5-tet) cells that were primed for recombineering. After 4 hours of recovery in LB broth at 32° C, cells were plated on LB+Amp+Cm plates. Subsequent colonies were check via PCR for insertions in pBAD22. Successfully cloned plasmids were confirmed via DNA sequencing. The pSIM5-tet plasmid was lost by incubation at 37° C for 24 hours.

Conjugation experiments. *E. coli* strain DH10 was used as a recipient in all conjugation assays. Donor and recipient cells were grown overnight at 37° C in LB broth with the appropriate selection. The overnight cultures were used to inoculate new 5 mL cultures in LB broth with no selection. The new cultures were incubated for 4 hours; if arabinose induction was needed it was done at 2 hours of incubation. After 4 hours of incubation, 0.5 mL of donor and recipient cells were added to 1.5 mL centrifuge tube, mixed by pipetting and incubated at 37° C without shaking for 1 hour. The mating reactions were then vortexed and placed immediately on ice. They were subsequently diluted in 1x phosphate buffered saline (PBS) and plated to select for transconjugants and donors.

Lux reporter assays. Overnight cultures were grown with the appropriate selections in LB broth at 37° C with shaking (200 RPM). These cultures were diluted 1:100 in new

broth with selection. These cultures were grown under the same conditions for 2 hours. At that time, the cultures were split in half. One half was treated with arabinose (0.02%) and the other half received no treatment. The cultures were allowed to grow for 2 hours. All cultures were then aliquoted into a 96-well plate, 200 μl per well. Plates were then read on a Bio-Tek plate reader. Cell density and arbitrary light units were measured. Bioluminescence was standardized for cell density by dividing light units by the OD600 absorbance. Each value represents the mean of 3 experiments.

RNA isolation and sequencing for RNA-Seq. Strains DH10B(pAR) and DH10B(pARΔ*acr2*) were grown until an OD_{600} of 0.5 was achieved. Cells were pelleted and RNA was purified using a commercially available RNA extraction kit (Qiagen). Treatments were included to remove DNA contamination (Qiagen) and ribosomal RNA (MicrobExpress, Ambion). Two biological replicates for each strain were pooled for paired-end library sequencing (either 50 or 100 bp reads) via Illumina Genome Analyzer II at the Biomedical Genomics Center at the University of Minnesota. All of the sequencing data are publically available under the NCBI BioProject ID PRJNA273283.

RNA-Seq Analysis. All Perl scripts and other computational biology resources used in this study can be found at https://github.com/kevinslang. cDNA reads were first trimmed so that the quality at each base position was above 30 and then mapped to the appropriate genome or plasmid sequence (for pAR060302, Genbank accession no. NC_012692, for DH10, the *E. coli* K-12 MG1655 published sequence was used Genbank

accession no. NC_000913) Read mapping was done using BOWTIE (67). For each host, transcriptome maps of pAR060302 were constructed using Circos (91). To achieve this, a table was generated containing the average number of reads mapped per 250 bp of plasmid sequence. Each average was then normalized per 1 million total reads in the cognate sequence library. These averages were then log transformed and plotted as a line plot. For statistical testing of differentially expressed genes, the total number of reads mapped to each coding sequence (CDS) was calculated using Perl. These values were then analyzed using the R package EdgeR (92, 93). We conservatively estimated the dispersion at 0.001. A fold-change cutoff of > 2 or < -2 and an adjusted p-value of < 0.05 were used to define significantly differentially expressed genes.

ChIP-Seq experiments. K12(pARK01, pB*acr2*^{6xHis}) was grown overnight in LB+Cm+Amp at 37° C with shaking (200 RPM). The overnight culture was used to inoculate a new 10 mL LB+Cm+Amp culture (1:100 dilution). This was grown for 2 hours at 37° C with shaking (200 RPM). Arabinose was added to a final concentration of 0.02% and the culture was incubated for 2 additional hours. Cells were fixed with the addition of formalin to a final concentration of 1%. Fixed cells were incubated at RT for 20 min. The formalin was quenched by addition of glycine to a final concentration of 0.5 M and incubation for 5 min at RT. Cells were collected by centrifugation at 8000 RPM for 5 min. The cell pellet was washed 1x with 10 mL of cold PBS and collected again. The resulting pellet was resuspended in 1X lysis solution (MagnaHIS Kit™, Promega). The cells were sonicated on ice at 5 watts for 30 second intervals and 1 min rest times.

This was done 10 times. The cell lysates were spun at 12,000 g for 5 min and the supernatant was moved to a 1.5 mL centrifuge tube containing 50 μL of MagneHIS particles. This suspension was taken through the MagneHIS kit protocol. After elution, crosslinked DNA was released via incubation at 65° C for 18 hours. The resulting DNA was purified with a Qiagen™ PCR Clean up Kit and sent for paired-end sequencing using the Illumina MiSeq platform. The library generation and sequencing was done at the UMGC at the University of Minnesota. In total, 2 biological replicates were sequenced separately. As a negative control, two genomic DNA preparations of the same cells were done using a Qiagen™ DNeasy™ kit and and ~400 ng of each was sent for sequencing.

The reads that resulted from the MiSeq run were all trimmed to 50 bp from the 3' end using the tool Trimmomatic (121). The trimmed reads were then mapped to either the pAR060302 sequence or the K12 genome sequence using BWA using the default parameters (122). Reads that were mapped correctly were filtered using SamTools (123). The replicate libraries were then merged using SamTools. The resulting read alignments files were then analyzed using MACS (110). Peak summit coordinates were used to extract 200 bp of sequence surrounding them using BedTools (124). These sequences were combined into a multifasta file and submitted to MEME for motif analysis (113). ChIP-Seq read alignments were also subjected to read counting on 150 bp windows of the pAR060302 sequence for visualization in Circos (91).

Protein purification and EMSA. The pB*acr2*^{6xHis} construct was introduced into strain were transformed into the BL21 (DE3) strain. The resulting strains were cultured in LB+AMP at 37° C for 1 hour with shaking (250 RPM). Arabinose was added to a final concentration of 0.2% prior to growing the cultures for 5 h at 28°C. Cells were spun at 4500×g for 30 min, resuspended in 5 mL cell lysis buffer (20 mM Tris pH 8, 500 mM NaCl, 5 mM imidazole, 5 mM β-mercaptoethanol) and sonicated. The cellular debris was removed by centrifugation at 12,000 g for 15 min. 200 µL of MagnePURE beads were added to the supernatant and these were incubated for 1 h on a rocking platform, washed twice with high salt washing buffer (20 mM Tris pH 8, 1 M NaCl, 5 mM β-mercaptoethanol) and twice with low salt washing buffer (20 mM Tris pH 8, 0.5 M NaCl, 30 mM imidazole, 5 mM β-mercaptoethanol), and then eluted with 100 µL elution buffer (20 mM Tris pH 8, 500 mM NaCl, 500 mM imidazole). Pefabloc™ (Roche) and glycerol were added to final concentrations of 1 mg/mL and 5%, respectively. Purified proteins were analyzed by SDS-PAGE gel electrophoresis and stored at -80° C.

Two DNA fragments were used for EMSA analysis. The region upstream of *acr1* (~480 bp) and the *floR* promotor (~150 bp). These fragments were amplified by PCR using the primers Orf184-EMSA-F and Orf184-EMSA-R (*acr1*) and Flo-F and Flo-R (*floR*). The amplicons were purified using a PCR cleanup kit (Qiagen). Various concentrations of purified Acr2 were incubated with 10 nM DNA in binding buffer (15 mM HEPES pH 7.9, 40 mM KCl, 1 mM EDTA, 1 mM DTT, 5% glycerol) for 30 minutes. The reactions were separated by gel electrophoresis for 2.5 h at 70 V on a 7.5% native polyacrylamide gel at 4°C (buffered with Tris glycine pH 8.0). Gels were stained

with SYBR Green for 20 minutes at room temperature, washed twice with ddH2O, and DNA complexes were visualized with ultraviolet light.

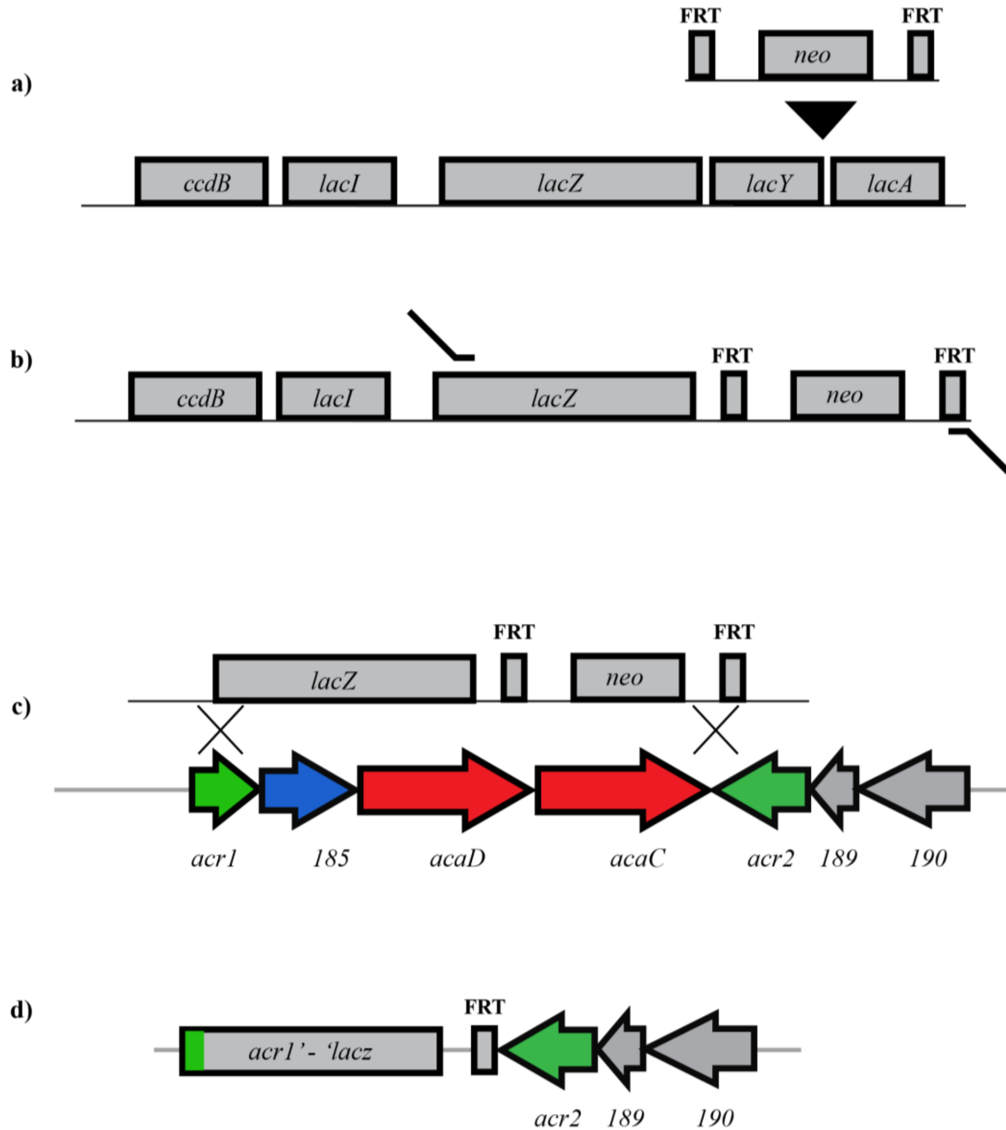

Figure 21. Schematic for the development of *acr1-lacZ* fusion reporter. a) The *lacA* and *lacY* genes were replaced by a FRT-*neo*-FRT cassette. b) The *lacZ*-FRT-*neo*-FRT cassette, starting with the 4th codon of *lacZ,* was amplified with primers adding ends homologous to *acr1* and *acaC* respectively. c) This amplicon was used to replace *aca1*

(at codon 4) through *acaC*. d) The *neo*-FRT cassette was removed by introduction of pCP20 expressing the FLP recombinase.

Development of acr1-lacZ fusion construct and detection of LacZ activity.

Figure 21 shows the strategy for the development of the *acr1-lacZ* fusion. Recombineering was used to replace the genes *lacY* and *lacA* with an FRT-*neo*-FRT cassette (Figure 21a). The *lacZ* gene, starting at the 4th codon was then amplified, along with the FRT-*neo*-FRT cassette with primers that contained 40 bp of homology with *acr1* and the region upstream of *acrC* (Figure 21b). This amplicon was used to replace *acr1-acrC* starting with the 4th codon of *acr1* (Figure 21c) in strain MC4100(pARK01, pSIM5-Tet) via recombineering and the *neo*-FRT cassette was removed by introduction of pCP20 expressing the FLP recombinase (Figure 21d). In the same strain, was disrupted *acr2* via recombineering. These reporter cells were then transformed with constructs over expressing Acr2 and the relevant Acr2 variants. They were struck onto LB+Amp+Xgal plates with and without arabinose.

Chapter 5

Discussion and future directions

The work presented in this dissertation described several aspects of transcriptional regulation of a group of large, conjugative plasmids in Gram-negative bacteria. The purpose of these studies was to move beyond comparative genomics approaches to achieve a more in depth understanding of mechanisms contributing to plasmid persistence and dissemination in bacterial populations. This dissertation represents one of the first comprehensive attempts at better understanding the basic biology of IncA/C plasmids. It provides a framework and tools that can be used to design sophisticated experiments to better understand regulatory mechanisms encoded by IncA/C plasmids.

Studies that focus on comparative analysis of fully or partly sequenced plasmids have their merits. For example, the work presented in this thesis targeted a set of genes that are conserved amongst all IncA/C plasmids known. It is clear from comparative genomics analyses that IncA/C plasmids are quite dynamic and a high level of conservation implies functional importance. The drawback of comparative genomics is that such conclusions assume that genes are expressed and/or are functional. It is possible that this is not the case. Additionally, for IncA/C plasmids in particular, many genes simply cannot be assigned a function during annotation due to the lack of similarity with genes of known function. In this case, comparative genomics sheds little insight into the biological role of such genes. Coupling high throughput genome sequencing, comparative analyses and functional molecular genetics is necessary to provide a deep understanding of how poorly annotated genes might be contributing to plasmid carriage or spread.

Using RNA-Seq, we mapped the transcriptome of pAR060302 under antimicrobial selection and when carried by different host bacteria. We were able to understand what genes are expressed under laboratory conditions and also gained insight into what mechanisms encoded on the host bacterial chromosome might play important roles in the carriage of IncA/C plasmids. We used that information to further inform more specific experiments. After developing mutants of genes of interest we were able to piece together part of the regulatory network responsible for regulating conjugation of IncA/C plasmids.

Modern genetic techniques nearly all rely on the introduction of resistance genes to select for mutations, whether they be generated by PCR or delivered by a phage or transposon. The systematic genetic study of something that confers resistance to nearly every class of antimicrobials was a difficult task to overcome. It is no wonder that, in parallel, another lab developed an IncA/C variant with fewer resistance genes in order to design sophisticated experiments (19). The only difference between their experimental system and the one developed for this work is that pARK01 was made using precise mutations of specific resistance genes. The other IncA/C variant was made by an unknown deletion, which resulted in a plasmid variant nearly identical to other IncA/C plasmids, but missing an unknown portion of DNA. Still, this dissertation, as well as the work of Carraro *et al.*, show the huge potential of coupling high throughput techniques, like RNA-Seq and ChIP-Seq, with more traditional molecular genetics approaches (35). Detailed experiments such as these will be necessary to understand the myriad of hypothetical proteins encoded on IncA/C plasmids.

The future for work focused on characterizing IncA/C plasmids is wide open. Certainly, genome-plasmid crosstalk should be pursued in more detail. Chapter 3 of this dissertation attempted to better understand these processes. A drawback to the design of those experiments was the use of bacterial hosts with unfished and poorly annotated genomes. We attempted to take a comparative transcriptomics approach, which provided a myriad of questions and no answers. Fortuitously, the modulation of 2-carbon metabolism in *E. coli* emerged as a result of Chapter 3 and then came up again in Chapter 4 as a target of Acr2. These results could be used as a direction for future work. A key first step could be simple growth experiments comparing *E. coli* strain K-12 strain with and without plasmid pARK01 in M9 minimal media using glycolate as the sole carbon source. A plasmid variant carrying an *acr2* mutation could also be used to further explore the possibility of direct regulation by Acr2. Further work should include the potential for induction of the *glc* operon by Acr2. The chromosomal factor IHF has previously been shown to positively regulate this operon (125). Precisely how 2-carbon metabolism might provide a fitness advantage in the broader context of IncA/C plasmid acquisition and carriage is a complex question, but specific alteration of host metabolic pathways by plasmid encoded factors suggests that it is important.

Whether or not IncA/C plasmid transfer can be induced by environmental or host cues is another area worth pursuing. This has been studied in enormous detail in the Ti and pCF10 plasmid systems (126, 127). The R27 plasmid has also been shown to sense

temperature (108). As detailed in Chapter 4 of this dissertation, the closely related SXT elements have a phage-like gene *setR* that encodes a repressor that is susceptible to RecA mediated degradation. IncA/C plasmids don't encode a homolog of SetR and their transfer is not dependent on *recA* (19). Given the number of factors contributing to the repression of IncA/C conjugation it would be logical that it be up-regulated by some mechanism. From the knowledge and tools developed in this dissertation, experiments can be carried out to expand the model for the regulation of IncA/C conjugation. For example, it is now known that IncA/C conjugation is dependent on *acaDC* expression. Using the *acr1-lacZ* fusion plasmid described in Chapter 4, a suitable screen could be carried out to look for mutants with increased LacZ activity that would lead to candidate mechanisms. Couple this experiment with a high throughput approach such as Tn-Seq and one could imagine a situation where the model of regulation might expand very rapidly (128).

One area completely unexplored in IncA/C plasmids are the role of small RNAs or small, non-annotated, proteins in the regulation of plasmid mechanisms. Our transcriptome mapping experiments in Chapter 2 first identified the possibility that IncA/C plasmids might express small RNAs or small peptides. This concept is certainly not new to plasmid biology, but given the diversity and complexity of previously described systems, it would be worthwhile to explore (126, 129).

The proposed experiments above are only a few directions future work could take

and only concern data generated in this dissertation. There are many more uncharacterized genes carried by IncA/C plasmids that are waiting to be explored. It would be prudent to study these hypothetical proteins not only for the insight it would provide to IncA/C biology, but also, given their evolutionary distance from proteins with known function, it would potentially produce new functional categories of proteins.

Bibliography

1. **Watanabe T**. 1963. Infective heredity of multiple drug resistance in bacteria. Bacteriol. Rev. **27**:87–115.

2. **Wantanabe T, Aoki T, Ogata Y, Egusa S**. 1971. Anbtibiotics and drug resistance in animals. R factors related to fish culturing. Ann. N. Y. Acad. Sci. **182**:383–410.

3. **Aoki T, Egusa S, Ogata Y, Watanabe T**. 1971. Detection of resistance factors in fish pathogen Aeromonas liquefaciens. J. Gen. Microbiol. **65**:343–349.

4. **Bauernfeind A, Stemplinger I, Jungwirth R, Giamarellou H**. 1996. Characterization of the plasmidic beta-lactamase CMY-2, which is responsible for cephamycin resistance. Antimicrob. Agents Chemother. **40**:221–4.

5. **Allen KJ, Poppe C**. 2002. Occurrence and characterization of resistance to extended-spectrum cephalosporins mediated by beta-lactamase CMY-2 in Salmonella isolated from food-producing animals in Canada. Can. J. Vet. Res. **66**:137–44.

6. **Giles WP, Benson AK, Olson ME, Hutkins RW, Whichard JM, Winokur PL, Fey PD**. 2004. DNA sequence analysis of regions surrounding blaCMY-2 from multiple Salmonella plasmid backbones. Antimicrob. Agents Chemother. **48**:2845–52.

7. **Kang MS, Besser TE, Call DR**. 2006. Variability in the region downstream of the blaCMY-2 beta-lactamase gene in Escherichia coli and Salmonella enterica plasmids. Antimicrob. Agents Chemother. **50**:1590–1593.

8. **Diarrassouba F, Diarra MS, Bach S, Delaquis P, Pritchard J, Topp E, Skura BJ**. 2007. Antibiotic resistance and virulence genes in commensal Escherichia coli and Salmonella isolates from commercial broiler chicken farms. J. Food Prot. **70**:1316–27.

9. **Daniels JB, Call DR, Besser TE**. 2007. Molecular Epidemiology of blaCMY-2 Plasmids Carried by Salmonella enterica and Escherichia coli Isolates from Cattle in the Pacific Northwest. Appl. Environ. Microbiol. **73**:8005–8011.

10. **Zhang Y-N, Peng J, Wang Q, Pei Z-F, Zhang W-J, Niu Z-X**. 2008. Appearance of blaCMY-2)gene-positive Salmonella isolates of pig origin in China. Int. J. Antimicrob. Agents **31**:292–3.

11. **Zaidi MB, Leon V, Canche C, Perez C, Zhao S, Hubert SK, Abbott J, Blickenstaff K, McDermott PF**. 2007. Rapid and widespread dissemination of multidrug-resistant blaCMY-2 Salmonella Typhimurium in Mexico. J. Antimicrob. Chemother. **60**:398–401.

12. **Ding H, Yang Y, Lu Q, Wang Y, Chen Y, Deng L, Wang A, Deng Q, Zhang H, Wang C, Liu L, Xu X, Wang L, Shen X**. 2008. The prevalence of plasmid-mediated AmpC beta-lactamases among clinical isolates of Escherichia coli and Klebsiella pneumoniae from five children's hospitals in China. Eur. J. Clin. Microbiol. Infect. Dis. **27**:915–21.

13. **Welch TJ, Fricke WF, McDermott PF, White DG, Rosso M-L, Rasko DA, Mammel MK, Eppinger M, Rosovitz MJ, Wagner D, Rahalison L, Leclerc JE, Hinshaw JM, Lindler LE, Cebula TA, Carniel E, Ravel J**. 2007. Multiple antimicrobial resistance in plague: an emerging public health risk. PLoS One **2**:e309.

14. **Fricke WF, Mammel MK, McDermott PF, Tartera C, White DG, LeClerc JE, Ravel J, Cebula TA**. 2011. Comparative genomics of 28 Salmonella enterica isolates: Evidence for crispr-mediated adaptive sublineage evolution. J. Bacteriol. **193**:3556–3568.

15. **Call DR, Singer RS, Meng D, Broschat SL, Orfe LH, Anderson JM, Herndon DR, Kappmeyer LS, Daniels JB, Besser TE**. 2010. blaCMY-2-positive IncA/C plasmids from Escherichia coli and Salmonella enterica are a distinct component of a larger lineage of plasmids. Antimicrob. Agents Chemother. **54**:590–6.

16. **Fernández-Alarcón C, Singer RS, Johnson TJ**. 2011. Comparative genomics of multidrug resistance-encoding IncA/C plasmids from commensal and pathogenic Escherichia coli from multiple animal sources. PLoS One **6**:e23415.

17. **Suzuki H, Yano H, Brown CCJ, Top EEM**. 2010. Predicting plasmid promiscuity based on genomic signature. J. Bacteriol. **192**:6045–55.

18. **Poole TL, Edrington TS, Brichta-Harhay DM, Carattoli A, Anderson RC, Nisbet DJ**. 2009. Conjugative transferability of the A/C plasmids from Salmonella enterica isolates that possess or lack bla(CMY) in the A/C plasmid backbone. Foodborne Pathog. Dis. **6**:1185–1194.

19. **Carraro N, Sauvé M, Matteau D, Lauzon G, Rodrigue S, Burrus V**. 2014. Development of pVCR94ΔX from Vibrio cholerae, a prototype for studying multidrug resistant IncA/C conjugative plasmids. Front. Microbiol. **5**:44.

20. **Arpin C, Thabet L, Yassine H, Messadi A a, Boukadida J, Dubois V, Coulange-Mayonnove L, Andre C, Quentin C**. 2011. Evolution of an Incompatibility Group IncA/C Plasmid Harboring blaCMY-16 and qnrA6 Genes and Its Transfer through Three Clones of Providencia stuartii during a. Antimicrob. Agents Chemother. **56**:1342–9.

21. **Wozniak RAF, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Déry C, Burrus V, Waldor MK**. 2009. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS Genet. **5**:e1000786.

22. **Welch TJ, Fricke WF, McDermott PF, White DG, Rosso M-L, Rasko DA, Mammel MK, Eppinger M, Rosovitz MJ, Wagner D, Rahalison L, Leclerc JE, Hinshaw JM, Lindler LE, Cebula TA, Carniel E, Ravel J**. 2007. Multiple antimicrobial resistance in plague: an emerging public health risk. PLoS One **2**:e309.

23. **Fricke WF, Welch TJ, McDermott PF, Mammel MK, LeClerc JE, White DG, Cebula TA, Ravel J**. 2009. Comparative genomics of the IncA/C multidrug resistance plasmid family. J. Bacteriol. **191**:4750–4757.

24. **Llanes C, Gabant P, Couturier M, Michel-Briand Y**. 1994. Cloning and characterization of the Inc A/C plasmid RA1 replicon. J. Bacteriol. **176**:3403–3407.

25. **Llanes C, Gabant P, Couturier M, Bayer L, Plesiat P**. 1996. Molecular analysis of the replication elements of the broad-host-range RepA/C replicon. Plasmid **36**:26–35.

26. **Harmer CJ, Hall RM**. 2014. pRMH760, a Precursor of A/C2 Plasmids Carrying blaCMY and blaNDM Genes. Microb. Drug Resist. **00**:1–9.

27. **Harmer CJ, Holt KE, Hall RM**. 2014. A type 2 A/C2 plasmid carrying the aacC4 apramycin resistance gene and the erm(42) erythromycin resistance gene recovered from two Salmonella enterica serovars. J. Antimicrob. Chemother. **70**:1021–5.

28. **Carattoli A, Villa L, Poirel L, Bonnin R a, Nordmann P**. 2012. Evolution of IncA/C blaCMY-2-carrying plasmids by acquisition of the blaNDM-1 carbapenemase gene. Antimicrob. Agents Chemother. **56**:783–786.

29. **Chihara S, Okuzumi K, Yamamoto Y, Oikawa S, Hishinuma A**. 2011. First case of New Delhi metallo-beta-lactamase 1-producing Escherichia coli infection in Japan. Clin. Infect. Dis. **52**:153–4.

30. **Chaston JM, Suen G, Tucker SL, Andersen AW, Bhasin A, Bode E, Bode HB, Brachmann AO, Cowles CE, Cowles KN, Darby C, de Léon L, Drace K, Du Z, Givaudan A, Herbert Tran EE, Jewell KA, Knack JJ, Krasomil-Osterfeld KC, Kukor R, Lanois A, Latreille P, Leimgruber NK, Lipke CM, Liu R, Lu X, Martens EC, Marri PR, Médigue C, Menard ML, Miller NM, Morales-Soto N, Norton S, Ogier J-C, Orchard SS, Park D, Park Y, Qurollo BA, Sugar DR, Richards GR, Rouy Z, Slominski B, Slominski K, Snyder H, Tjaden BC, van der Hoeven R, Welch RD, Wheeler C, Xiang B, Barbazuk B, Gaudriault S, Goodner B, Slater SC, Forst S, Goldman BS, Goodrich-Blair H**. 2011. The entomopathogenic bacterial endosymbionts Xenorhabdus and Photorhabdus: convergent lifestyles from divergent genomes. PLoS One **6**:e27909.

31. **Yun C-SS, Suzuki C, Naito K, Takeda T, Takahashi Y, Sai F, Terabayashi T, Miyakoshi M, Shintani M, Nishida H, Yamane H, Nojiri H**. 2010. Pmr, a histone-like protein H1 (H-NS) family protein encoded by the IncP-7 plasmid pCAR1, is a key global regulator that alters host function. J. Bacteriol. **192**:4720–4731.

32. **Doyle M, Fookes M, Ivens A, Mangan MW, Wain J, Dorman CJ**. 2007. An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. Science (80-.). **315**:251–252.

33. **Dillon SC, Cameron ADS, Hokamp K, Lucchini S, Hinton JCD, Dorman CJ**. 2010. Genome-wide analysis of the H-NS and Sfh regulatory networks in Salmonella Typhimurium identifies a plasmid-encoded transcription silencing mechanism. Mol. Microbiol. **76**:1250–65.

34. **Subbiah M, Top EM, Shah DH, Call DR**. 2011. Selection pressure required for long-term persistence of blaCMY-2-positive IncA/C plasmids. Appl. Environ. Microbiol. **77**:4486–4493.

35. **Carraro N, Matteau D, Luo P, Rodrigue S, Burrus V**. 2014. The master activator of IncA/C conjugative plasmids stimulates genomic islands and multidrug resistance dissemination. PLoS Genet. **10**:e1004714.

36. **Alonso A, Sánchez P, Martínez JL**. 2001. Environmental selection of antibiotic resistance genes. Environ. Microbiol. **3**:1–9.

37. **Salyers AA, Amábile-Cuevas CF**. 1997. Why are antibiotic resistance genes so resistant to elimination? Antimicrob. Agents Chemother. **41**:2321–2325.

38. **Singer RS, Ward MP, Maldonado G**. 2006. Can landscape ecology untangle the complexity of antibiotic resistance? Nat. Rev. Microbiol. **4**:943–952.

39. **Hughes VM, Datta N**. 1983. Conjugative plasmids in bacteria of the "pre-antibiotic" era. Nature **302**:725–726.

40. **Koonin E V, Wolf YI**. 2008. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. Nucleic Acids Res. **36**:6688–6719.

41. **Norman A, Hansen LH, Sørensen SJ**. 2009. Conjugative plasmids: vessels of the communal gene pool. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **364**:2275–89.

42. **Johnson TJ, Jordan D, Kariyawasam S, Stell AL, Bell NP, Wannemuehler YM, Alarcón CF, Li G, Tivendale KA, Logue CM, Nolan LK**. 2010. Sequence analysis and characterization of a transferable hybrid plasmid encoding multidrug resistance and enabling zoonotic potential for extraintestinal Escherichia coli. Infect. Immun. **78**:1931–1942.

43. **Lindsey RL, Fedorka-Cray PJ, Frye JG, Meinersmann RJ**. 2009. Inc A/C Plasmids Are Prevalent in Multidrug-Resistant Salmonella enterica Isolates. Appl. Environ. Microbiol. **75**:1908–1915.

44. **Fricke W, McDermott P**. 2009. resistance-conferring plasmids with similarity to virulence plasmids from avian pathogenic Escherichia coli strains in Salmonella enterica serovar Kentucky isolates. Appl. … **75**:5963–5971.

45. **Frost L, Leplae R**. 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. … **3**:722–732.

46. **Galimand M, Guiyoule A, Gerbaud G, Rasoamanana B, Chanteau S, Carniel E, Courvalin P**. 1997. Multidrug resistance in Yersinia pestis mediated by a transferable plasmid. N. Engl. J. Med. **337**:677–680.

47. **Pan J-C, Ye R, Wang H-Q, Xiang H-Q, Zhang W, Yu X-F, Meng D-M, He Z-S**. 2008. Vibrio cholerae O139 multiple-drug resistance mediated by Yersinia pestis pIP1202-like conjugative plasmids. Antimicrob. Agents Chemother. **52**:3829–36.

48. **Welch TJ, Evenhuis J, White DG, McDermott PF, Harbottle H, Miller RA, Griffin M, Wise D**. 2009. IncA/C plasmid-mediated florfenicol resistance in the catfish pathogen Edwardsiella ictaluri. Antimicrob. Agents Chemother. **53**:845–6.

49. **Jacoby GA**. 2009. AmpC {beta}-Lactamases. Clin. Microbiol. Rev. **22**:161–182.

50. **Ramos JL, Marqués S, Timmis KN**. 1997. Transcriptional control of the Pseudomonas TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. Annu. Rev. Microbiol. **51**:341–373.

51. **Hillen W, Berens C**. 1994. Mechanisms underlying expression of Tn10 encoded tetracycline resistance. Annu. Rev. Microbiol. **48**:345–369.

52. **Depardieu F, Podglajen I, Leclercq R, Collatz E, Courvalin P**. 2007. Modes and modulations of antibiotic resistance gene expression. Clin. Microbiol. Rev. **20**:79–114.

53. **Summers AO**. 1986. Organization, expression, and evolution of genes for mercury resistance. Annu. Rev. Microbiol. **40**:607–634.

54. **Storz G, Hengge-Aronis R**. 2000. Bacterial stress responses. ASM Press.

55. **Hossain A**. 2004. Plasmid-encoded functions compensate for the biological cost of AmpC overexpression in a clinical isolate of Salmonella typhimurium. J. Antimicrob. … **53**:964–970.

56. **Marciano DC, Karkouti OY, Palzkill T**. 2007. A Fitness Cost Associated With the Antibiotic Resistance Enzyme SME-1 β-Lactamase. Genetics **176**:2381 –2392.

57. **Blickwede M, Goethe R, Wolz C, Valentin-Weigand P, Schwarz S**. 2005. Molecular basis of florfenicol-induced increase in adherence of Staphylococcus aureus strain Newman. J. Antimicrob. Chemother. **56**:315–323.

58. **Kehrenberg C, Schwarz S**. 2004. fexA, a novel Staphylococcus lentus gene encoding resistance to florfenicol and chloramphenicol. Antimicrob. Agents Chemother. **48**:615–618.

59. **Takeda T, Yun C-S, Shintani M, Yamane H, Nojiri H**. 2011. Distribution of genes encoding nucleoid-associated protein homologs in plasmids. Int. J. Evol. Biol. **2011**:685015.

60. **Van Melderen L**. 2010. Toxin-antitoxin systems: why so many, what for? Curr. Opin. Microbiol. **13**:781–785.

61. **Beaber JW, Hochhut B, Waldor MK**. 2002. Genomic and Functional Analyses of SXT, an Integrating Antibiotic Resistance Gene Transfer Element Derived from Vibrio cholerae. J. Bacteriol. **184**:4259–4269.

62. **Beaber JW, Hochhut B, Waldor MK**. 2004. SOS response promotes horizontal dissemination of antibiotic resistance genes. Nature **427**:72–74.

63. **Beaber JW, Waldor MK**. 2004. Identification of Operators and Promoters That Control SXT Conjugative Transfer. J. Bacteriol. **186**:5945–5949.

64. **Gottesman S**. 2004. The small RNA regulators of Escherichia coli: roles and mechanisms*. Annu. Rev. Microbiol. **58**:303–28.

65. **Waters LS, Storz G**. 2009. Regulatory RNAs in Bacteria. Cell **136**:615–628.

66. **Güell M, Yus E, Lluch-Senar M, Serrano L**. 2011. Bacterial transcriptomics: what is beyond the RNA horiz-ome? Nat Rev Micro **9**:658–669.

67. **Langmead B, Trapnell C, Pop M, Salzberg SL**. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. **10**:R25–R25.

68. **Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP**. 2011. Integrative genomics viewer. Nat Biotech **29**:24–26.

69. **Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B**. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Meth **5**:621–628.

70. **Wang L, Feng Z, Wang X, Wang X, Zhang X**. 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics **26**:136–138.

71. **Levy S, O'Brien T**. 2005. Global antimicrobial resistance alerts and implications. Clin. Infect. Dis. **41 Suppl 4**:S219–220.

72. **Fernández-Alarcón C, Miranda CD, Singer RS, López Y, Rojas R, Bello H, Domínguez M, González-Rocha G**. 2010. Detection of the floR gene in a diversity of florfenicol resistant Gram-negative bacilli from freshwater salmon farms in Chile. Zoonoses Public Health **57**:181–188.

73. **Johnson TJ, Lang KS**. 2012. © 2012 Landes Bioscience . Do not distribute . © 2012 Landes Bioscience . Do not distribute . 55–58.

74. **García P, Guerra B, Bances M, Mendoza MC, Rodicio MR**. 2011. IncA/C plasmids mediate antimicrobial resistance linked to virulence genes in the Spanish clone of the emerging Salmonella enterica serotype 4,[5],12:i:-. J. Antimicrob. Chemother. **66**:543–9.

75. **Carattoli A, Villa L, Poirel L, Bonnin RA, Nordmann P**. 2012. Evolution of IncA/C blaCMY-$_2$-carrying plasmids by acquisition of the blaNDM-$_1$ carbapenemase gene. Antimicrob. Agents Chemother. **56**:783–6.

76. **Bouma JE, Lenski RE**. 1988. Evolution of a bacteria/plasmid association. Nature **335**:351–352.

77. **Harr B, Schlötterer C**. 2006. Gene expression analysis indicates extensive genotype-specific crosstalk between the conjugative F-plasmid and the E. coli chromosome. BMC Microbiol. **6**:80.

78. **Dahlberg C, Chao L**. 2003. Amelioration of the cost of conjugative plasmid carriage in Eschericha coli K12. Genetics **165**:1641–1649.

79. **Sota M, Top EM**. 2008. Host-specific factors determine the persistence of IncP-1 plasmids. World J. Microbiol. Biotechnol. **24**:1951–1954.

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

81. **Lang KS, Danzeisen JL, Xu W, Johnson TJ**. 2012. Transcriptome mapping of pAR060302, a blaCMY-2-positive broad-host-range IncA/C plasmid. Appl. Environ. Microbiol. **78**:3379–86.

82. **Huang DW, Sherman BT, Lempicki RA**. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. **37**:1–13.

83. **Huang DW, Sherman BT, Lempicki RA**. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. **4**:44–57.

84. **Ricci JCO, Hernandez ME, Diaz Ricci JC, Hernández ME**. 2000. Plasmid effects on Escherichia coli metabolism. Crit. Rev. Biotechnol. **20**:79–108.

85. **Clark DP, Jr. JEC, Cronan JE**. 1996. Two-Carbon Compounds and Fatty Acids as Carbon Sources. Escherichia coli Salmonella Cell. Mol. Biol. **1**:343–357.

86. **Heidelberg JF, Paulsen IT, Nelson KE, Gaidos EJ, Nelson WC, Read TD, Eisen JA, Seshadri R, Ward N, Methe B, Clayton RA, Meyer T, Tsapin A, Scott J, Beanan M, Brinkac L, Daugherty S, DeBoy RT, Dodson RJ, Durkin AS, Haft DH, Kolonay JF, Madupu R, Peterson JD, Umayam LA, White O, Wolf AM, Vamathevan J, Weidman J, Impraim M, Lee K, Berry K, Lee C, Mueller J, Khouri H, Gill J, Utterback TR, McDonald LA, Feldblyum T V, Smith HO, Venter JC, Nealson KH, Fraser CM**. 2002. Genome sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis. Nat. Biotechnol. **20**:1118–23.

87. **Meinersmann RJ, Lindsey RL, Bono JL, Smith TP, Oakley BB**. 2013. Proposed model for the high rate of rearrangement and rapid migration observed in some IncA/C plasmid lineages. Appl. Environ. Microbiol. **79**:4806–14.

88. **Ow DS-W, Lee RM-Y, Nissom PM, Philp R, Oh SK-W, Yap MG-S**. 2007. Inactivating FruR global regulator in plasmid-bearing Escherichia coli alters metabolic gene expression and improves growth rate. J. Biotechnol. **131**:261–9.

89. **Douard G, Praud K, Cloeckaert A, Doublet B**. 2010. The Salmonella genomic island 1 is specifically mobilized in trans by the IncA/C multidrug resistance plasmid family. PLoS One **5**:e15302.

90. **Miyakoshi M, Shintani M, Terabayashi T, Kai S, Yamane H, Nojiri H**. 2007. Transcriptome analysis of Pseudomonas putida KT2440 harboring the completely sequenced IncP-7 plasmid pCAR1. J. Bacteriol. **189**:6849–60.

91. **Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA**. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. **19**:1639–45.

92. **Robinson MD, McCarthy DJ, Smyth GK**. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**:139–40.

93. **Robinson M, Mccarthy D, Chen Y, Smyth GK**. 2011. edgeR : differential expression analysis of digital gene expression data User ' s Guide.

94. **Aoki T, Egusa S, Kimura T, Watanabe T**. 1971. Detection of R factors in naturally occurring Aeromonas salmonicida strains. Appl. Microbiol. **22**:716–7.

95. **Toleman MA, Walsh TR**. 2011. Combinatorial events of insertion sequences and ICE in Gram-negative bacteria. FEMS Microbiol. Rev. **35**:912–935.

96. **Fitzgerald DM, Bonocora RP, Wade JT**. 2014. Comprehensive mapping of the Escherichia coli flagellar regulatory network. PLoS Genet. **10**:e1004649.

97. **Dillon SC, Dorman CJ**. 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat. Rev. Microbiol. **8**:185–195.

98. **Fang F, Rimsky S**. 2008. New insights into transcriptional regulation by H-NS. Curr. Opin. Microbiol. **11**:113–120.

99. **Bouffartigues E, Buckle M, Badaut C, Travers A, Rimsky S**. 2007. H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nat. Struct. Mol. Biol. **14**:441–8.

100. **Ali SS, Whitney JC, Stevenson J, Robinson H, Howell PL, Navarre WW**. 2013. Structural insights into the regulation of foreign genes in Salmonella by the Hha/H-NS complex. J. Biol. Chem. **288**:13356–69.

101. **Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC**. 2006. Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. Science (80-.). **313**:236–238.

102. **Stoebel DM, Free A, Dorman CJ**. 2008. Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. Microbiology **154**:2533–2545.

103. **Dorman CJ**. 2007. H-NS, the genome sentinel. Nat. Rev. Microbiol. **5**:157–161.

104. **Ali SS, Soo J, Rao C, Leung AS, Ngai DH-M, Ensminger AW, Navarre WW**. 2014. Silencing by H-NS potentiated the evolution of Salmonella. PLoS Pathog. **10**:e1004500.

105. **Navarre WW, McClelland M, Libby SJ, Fang FC**. 2007. Silencing of xenogeneic DNA by H-NS - Facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. Genes Dev. **21**:1456–1471.

106. **Müller CM, Schneider G, Dobrindt U, Emödy L, Hacker J, Uhlin BE**. 2010. Differential effects and interactions of endogenous and horizontally acquired H-NS-like proteins in pathogenic Escherichia coli. Mol. Microbiol. **75**:280–293.

107. **Bustamante VH, Santana FJ, Calva E, Puente JL**. 2001. Transcriptional regulation of type III secretion genes in enteropathogenic Escherichia coli: Ler antagonizes H-NS-dependent repression. Mol. Microbiol. **39**:664–678.

108. **Forns N, Baños RC, Balsalobre C, Juárez A, Madrid C**. 2005. Temperature-dependent conjugative transfer of R27: role of chromosome- and plasmid-encoded Hha and H-NS proteins. J. Bacteriol. **187**:3950–9.

109. **Datsenko KA, Wanner BL**. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. **97**:6640–6645.

110. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS**. 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol. **9**:R137.

111. **Solovyev V, Salamov A**. 2011. Automatic Annotation of Microbial Genomes and Metagenomic Sequences., p. 61–78. *In* Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies. Nova Science Publishers.

112. **Zuker M**. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31**:3406–15.

113. **Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS**. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. **37**:W202–8.

114. **Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, Jimenez-Jacinto V, Salgado H, Juárez K, Contreras-Moreira B, Huerta AM, Collado-Vides J, Morett E**. 2009. Genome-Wide Identification of Transcription Start Sites, Promoters and Transcription Factor Binding Sites in E. coli. PLoS One **4**:e7526.

115. **Gordon BRG, Li Y, Cote A, Weirauch MT, Ding P, Hughes TR, Navarre WW, Xia B, Liu J**. 2011. Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. Proc. Natl. Acad. Sci. U. S. A. **108**:10690–10695.

116. **Lin RJ, Capage M, Hill CW**. 1984. A repetitive DNA sequence, rhs, responsible for duplications within the Escherichia coli K-12 chromosome. J. Mol. Biol. **177**:1–18.

117. **Lee EC, Yu D, Martinez de Velasco J, Tessarollo L, Swing DA, Court DL, Jenkins NA, Copeland NG**. 2001. A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics **73**:56–65.

118. **Lenz DH, Mok KC, Lilley BN, Kulkarni R V, Wingreen NS, Bassler BL**. 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae. Cell **118**:69–82.

119. **Guzman LM, Belin D, Carson MJ, Beckwith J**. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. **177**:4121–4130.

120. **Sawitzke JA, Costantino N, Li X, Thomason LC, Bubunenko M, Court C, Court DL**. 2011. Probing Cellular Processes with Oligo-Mediated Recombination and Using the Knowledge Gained to Optimize Recombineering. J. Mol. Biol. **407**:45–59.

121. **Bolger AM, Lohse M, Usadel B**. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics btu170–.

122. **Li H, Durbin R**. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**:1754–60.

123. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R**. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–9.

124. **Quinlan AR, Hall IM**. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics **26**:841–2.

125. **Pellicer MT, Fernandez C, Badía J, Aguilar J, Lin EC, Baldom L**. 1999. Cross-induction of glc and ace operons of Escherichia coli attributable to pathway intersection. Characterization of the glc promoter. J. Biol. Chem. **274**:1745–1752.

126. **Dunny GM, Johnson CM**. 2011. Regulatory circuits controlling enterococcal conjugation: lessons for functional genomics. Curr. Opin. Microbiol. **14**:174–180.

127. **Blackwell HE, Fuqua C**. 2011. Introduction to bacterial signals and chemical communication. Chem. Rev. **111**:1–3.

128. **Van Opijnen T, Lazinski DW, Camilli A**. 2015. Genome-Wide Fitness and Genetic Interactions Determined by Tn-seq, a High-Throughput Massively Parallel Sequencing Method for Microorganisms. Curr. Protoc. Microbiol. **36**:1E.3.1–1E.3.24.

129. **Storz G, Vogel J, Wassarman KMM**. 2011. Regulation by small RNAs in bacteria: expanding frontiers. Mol. Cell **43**:880–91.

Appendex A

Summary of experimental evolution experiments using A/C plasmid containing strains.

***IncA/C gene content is more unstable over time when* acr2 .** We evolved *E. coli* K-12 and *Salmonella* Newport carrying pAR060302 and pAR060302Δ*acr2* over 50 days by passing the strains in LB broth cultures. Each strain was used to inoculate a 5 mL LB aliquot and was grown at 37° C overnight with shaking (200 RPM). Every day, 50 µL of the overnight culture was used to inoculate a new 5 mL culture. Every 10 days, 750 µL of the overnight growth was frozen at -80° C with a 20% final concentration of glycerol. The 50 day cultures were allowed to thaw on ice and 100 µL was taken and added to 100 µL of fresh LB. These were grown at 37° C with shaking (200 RPM). After 4 hours, the cultures were serially diluted out to 10^{-8} in PBS and plated on LB plates. Individual colonies were picked and patched onto a 48 spot grid on a new LB plate. After 24 hour incubation, a 48 prong plate replicator was used to replica plate the spots onto LB, LB+TET, LB+CM and LB+AMP plates. For K12, 470 colonies were picked and for *Salmonella* Newport, 235 colonies were picked. Numbers of susceptible colonies were counted. If susceptible result was unclear, the spot was re-struck onto a new LB plate containing the appropriate antibiotic.

Table 7. Susceptible colonies of *E. coli* K12 carrying either pAR060302 or pAR060302Δ*acr2* after passing 50 days in LB with no antibiotic selection.

		pAR060302	pAR060302Δ*acr2*
Chloramphenicol		236	62
Tetracycline		239	64
Ampicillin		1	61

Table 8. Susceptible colonies of *Salmonella* Newport carrying either pAR060302 or pAR060302Δ*acr2* after passing 50 days in LB with no antibiotic selection.

		pAR060302	pAR060302Δ*acr2*
Chloramphenicol		1	235
Tetracycline		11	235
Ampicillin		126	235

Our results indicate that the *floR-tet(A)* cassette is more unstable in the WT plasmid versus the *acr2* mutant when carried by K12 (Table 7). Conversely, this cassette is relatively stable in *S.* Newport (Table 8). In K-12, we found that the bla_{CMY-2} gene is less stable in the *acr2* mutant. None of the ampicillin susceptible colonies were susceptible to any other antibiotic, except in *S.* Newport(pAR060302Δ*acr2)*, where it seems the entire plasmid seems to be unstable. To test this, PCR targeting the *repA* gene on pAR060302 was done on 5 randomly selected ampicillin susceptible colonies of *S.* Newport (pAR060302Δ*acr2)*. All colonies were negative, suggesting complete plasmid loss. In K-12 (pAR060302Δ*acr2)*, 3 of the ampicillin susceptible colonies were selected and PCR targeting 15 genes was conducted. These PCR results suggest that a large mutation occurred that excised the entire Tra1 locus. Out of the 3 colonies selected it appears there are at least 2 separate lineages of deletions because spot 3 contains ORF68 whereas the other spots do not (Table 9).

Table 9. PCR results of 3 ampicillin susceptible colonies of K-12 (pAR060302Δ*acr2*).

Gene	Colony 1	Colony 2	Colony 3	Gene	Colony 1	Colony 2	Colony 3
ORF55	+	+	+	ORF99	-	-	-
ORF68	-	-	+	ORF104	-	-	-
traI	-	-	-	ORF111	-	-	-
traD	-	-	-	ORF111	-	-	-
ORF75	-	-	-	ORF126	-	-	-
traV	-	-	-	*tn3*	-	-	-
$bla_{\text{CMY-2}}$	-	-	-	*int1*	-	-	-
ORF88	-	-	-	*groSEL*	+	+	+

The phenotypic results of the *E. coli* evolution experiments first suggested that Acr2 binding to the IS*Ecp1* locus might have aided its stability over time. After closer examination using PCR, as well as the results of the same experiment in *S.* Newport it seems that this might not necessarily be the case. If *acr2* is deleted, this would result in a subsequent up-regulation of *acaDC*, which would in turn up-regulate the transfer genes. Having the transfer genes up-regulated would likely be costly to those cells giving a predictable advantage to cells that can achieve a large deletion. The same result would be expected in *Salmonella*. However, we did not observe this. We observed total plasmid loss. One possible explanation is that in *Salmonella,* up-regulation of *acaDC* not only up-regulates the transfer genes, but also up-regulates genes on genomic islands (35). This might increase the fitness cost of carrying such a plasmid above some threshold where the barrier of the various stability mechanisms (*psk* and *par*) can be overcome. These

results suggest that there are different fitness outcomes for mutations occurring on the plasmid that are dependent on host chromosomal gene content. If there are more interactions between the plasmid and host chromosome, then that would select for a stable plasmid sequence. This would suggest that A/C plasmids might interact with the host chromosome much more than previously thought, despite their broad host-range. This interplay between the host chromosome and A/C plasmids might account for the relatively low number of mutations observed on the A/C backbone between different fully sequenced plasmids.