This is to certify that the

thesis entitled

ESCHERICHIA COLI RIBOSOMAL RNA OPERON MULTIPLICITY:
GENETIC INTERACTIONS AND FITNESS COSTS

presented by

Joel Ginji Hashimoto

has been accepted towards fulfillment
of the requirements for

M.S. degree in Microbiology and
Molecular Genetics

Major professor

Date 11/27/00

ESCHERICHIA COLI RIBOSOMAL RNA OPERON MULTIPLICITY: GENETIC INTERACTIONS AND FITNESS COSTS

By

Joel Ginji Hashimoto

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Microbiology and Molecular Genetics

2000

ABSTRACT

ESCHERICHIA COLI RIBOSOMAL RNA OPERON MULTIPLICTY: GENETIC INTERACTIONS AND FITNESS COSTS

By

Joel Ginji Hashimoto

Recombination between *rrnB* and the other six rRNA operons of *Escherichia coli* resulted in the removal of a *sacB-neo* insertion in *rrsB* at a rate of $5x10^{-9}$ per cell division. The loss of the insertion was dependent on flanking regions of sequence homology to the other rRNA operons. The relative fitness of the recombinant isolates, measured in head-to-head competition experiments with a wild-type strain, was ca. 1% higher than the strain carrying the insertion. Two of 14 recombinant isolates from apparent gene conversion events resulted in *rrnB* operons with alternate 16S-23S internally transcribed spacer transfer RNAs. These findings are discussed in the context of a conceptual model describing the maintenance of sequence heterogeneity in co-evolving genes.

Previously constructed chromosomal deletions of *rrnA* were complemented by reinserting the *rrnA* operon in its native site using two different approaches. Wild-type or near wild-type fitness was not restored in either of the complemented strains, indicating the large fitness loss seen in strains with *rrnA* deleted is not specifically attributable to the deletion of *rrnA* but rather to differences in adjacent genes transferred by P1 transduction during the strain construction. An alternative approach to generate rRNA operon deletions is proposed that should circumvent the problems with the earlier strain constructions.

ACKNOWLEDGEMENTS

I would first like to thank my parents Merle and Andrew Hashimoto for their assistance and support in the many avenues I have pursued. To Meri, Noelle, and Riki: Thank you for the belief you have shown in my abilities even when I was doubtful. I would like to thank my advisor Dr. Thomas Schmidt for all of the excellent advice, for showing the elegance of scientific inquiry, and the passion with which it should be conducted. Thank you to my committee members, Dr. Patterson and Dr. Jackson, for their guidance. To the Schmidt and Breznak labs: Dr. Brad Stevenson, Dr. Dan Buckley, Joel Klappenbach, Kristen Huizinga, Les Dethlefsen, Joe Graber, and Keri Byzek. Thank you all for the fun and interesting times in and out of the lab. Thank you Kate Delaney and Ann Stafford for the countless number of plates poured, miscellaneous support, and for keeping the project going. And finally to Bryan, Shalette, Ryan, and Sean, for helping me remember what I set out to do and the importance of the time we are given.

TABLE OF CONTENTS

LIST OF TABLES....vi
LIST OF FIGURES....vii

CHAPTER 1....1

CHAPTER 2....6
2.1 Introduction....6
2.2 Materials and Methods....9
Strains and genotype verification....9
Growth conditions and media....10
Recombination rate measurements....12
Relative fitness assay....13
rRNA operon sequence alignment....14
2.3 Results....15
Rate of recombination involving rrnB....15
Fitness effects of recombination between rrn operons....21
Correlation analysis of pair-wise *rrs* and *rrl* sequence difference and distance....23
2.4 Discussion....25
rRNA operon recombination results in the loss of the *sacB-neo* insertion....26
Fitness effects of rRNA recombination and rRNA operon interruption....27
Factors affecting the co-evolution of rRNA operons in *E. coli*....28
Conclusions....33

CHAPTER 3....35
3.1 Introduction....35
3.2 Materials and Methods....37
Cloning, XL-PCR and general methods....37
Version 1 plasmid constructions....38
Version 1 strain construction....43
Version 2 strain construction....47
Chemostat competition experiments....47
Batch culture competition experiments....48
3.3 Results....48
Version 1 competition experiments....48
Version 2 batch culture competition experiments....50
3.4 Discussion....53
Lack of complementation in *rrnA* replacement strains....53
Rational for constructing *rrnA* deletions and replacements from the same transductant....54
Future directions....55
Conclusions....56

CHAPTER 4.....61

LITERATURE CITED.....64

LIST OF TABLES

Table 2.1. Recombination and *sacB* mutation rates

Table 2.2. Values used in equations 1(A) and 2(B) to determine recombination and *sacB* mutation rates

Table 3.1. Plasmids used for strain constructions

Table 3.2. *E. coli* strain constructions

LIST OF FIGURES

Figure 2.1. Location and organization of *E. coli* rRNA operons. A. *E. coli* chromosome with location (mbp) and orientation of rRNA operons. B. General organization of *rrn* operons. P1 and P2 are the tandem promoters, *rrs* encodes 16S rRNA, ITS represents the internally transcribed spacer, *rrl* encodes 23S rRNA, *rrf* encodes 5S rRNA, and *ter* is the rho-independent termination region. C. Table of rRNA operons and their associated ITS and distal tRNA species.

Figure 2.2. BSV1 and BSV2 *rrnB* operon constructions containing the *sacB-neo* cassette. A. BSV1 has the *sacB-neo* cassette inserted into the *Bgl*II site of *rrsB*, upstream sequence similarity to the other *rrn* operons, and a transcriptionally active operon. B. BSV2 was constructed through the removal of sequences upstream of *rrnB*, *rrsB*, the ITS region, and the 5' end of *rrl* for the *sacB-neo* cassette leaving no upstream sequence similarity to the other *rrn* operons and a transcriptionally inactive operon. The *murI* gene is unique to the *rrnB* region and represents the end of the upstream sequence similarity to the other rRNA operons.

Figure 2.3. Southern hybridization of BSV1 and BSV2 isolate genomic DNA digested with *Pvu*II and probed for a conserved region of *rrs*. In BSV1, the band corresponding to *rrnB* is shifted when containing the *sacB-neo* cassette inserted in *rrs* (lane 1), after recombination with a *rrn* operon with a $tRNA^{Ile}$-$tRNA^{Ala}$ ITS

(lane 3), and back to wild-type after recombination with a *rrn* operon with a $tRNA^{Glu}$ ITS (lanes 4-8). BSV1 *sacB* isolates are sucrose resistant and kanamycin resistant and still harbor the *sacB-neo* cassette in *rrsB* (lanes 9-12). BSV2 120 and BSV2 126 are sucrose resistant and kanamycin sensitive (recombinant phenotype) but have not regained the deleted region of *rrsB* (lanes 13-14). BSV2 118 and BSV2 142 are sucrose resistant and kanamycin resistant (*sacB* phenotype) and have not regained the deleted region of *rrsB* (lanes 15-16).

Figure 2.4. PCR amplification of the *rrnB* region of the isolates used in the relative fitness assay. Amplification of the *rrnB* operon with the *sacB-neo* cassette produces a 10.4 kb fragment as in the BSV1 *sacB* isolates. Amplification of the wild-type or a recombinant *rrnB* region produces a 6.6 kb fragment. The BSV1 recombinant isolates (rec) all have PCR amplification products of 6.6 kb, indicating the removal of the *sacB-neo* cassette. The BSV1rec#31 isolate was excluded due to the strange amplification size and unexpected banding pattern in the Southern hybridization (Figure 2.5).

Figure 2.5. Southern hybridization of *Pvu*II digsted genomic DNA from BSV1 isolates used in the relative fitness assay. The banding pattern confirms the recombinant isolates (rec) have lost the *sacB-neo* cassette and *sacB* mutants (sacB) still have the cassette in *rrnB*. Due to the unusual banding pattern of

BSV1rec#31, it was excluded from further analysis. BSV1 rec#5 has a banding pattern indicative of an alternate ITS in *rrnB*.

Figure 2.6. Relative fitness of BSV1, *sacB*, and recombinant isolates. The recombinant isolates are significantly different from both BSV1 and *sacB* groups but BSV1 and *sacB* are not statistically distinguishable. Error bars represent +/- 1 standard deviation for each group.

Figure 2.7. Correlation of pair-wise sequence differences of *rrs* and *rrl* with distance between operons of all available bacterial and archeal genome sequences with multiple rRNA operons. Pair-wise sequence differences were tallied and distances between operons were calculated based on their published locations. *E. coli* comparisons are represented as open diamonds and the other organism comparisons are represented as closed diamonds.

Figure 3.1. *Xba*I digestions of pUCA21 plasmid constructions. *rrnA* was cloned as an *Xba*I fragment into pUC19 giving pUCA21 L1 and L2. pUCA21 constructions were further digested with restriction endonucleases to yield the resulting deletion plasmids. A. *Xba*I digestions of pUCA21 plasmids. B. Restriction sites used in the creation of the pUCA21 deletion plasmids.

Figure 3.2. Diagram of steps involved in the complementation of BS301 and BS300 creating A+KOBA* and A-KOBA*. The creation of JHB and JHAb strains

follow these steps starting at step 3 using plasmids pMAKA14 and pKTS102, respectively.

Figure 3.3. PCR amplification of the *rrnA* region of the complementation and deletion strains. The *rrnA* region amplifies as a 7.3 kb fragment as indicated in the BS4548 lane. The deleted *rrnA* region amplifies as a 2.6 kb fragment. All band mobilities for the deletion and complementation strains show the expected size.

Figure 3.4. Southern hybridization of *Pvu*II digested genomic DNA from *rrnA* deletion and complementation strains probed for a conserved region of *rrs*. All strains have *rrnB* deleted. A+KOBA* and A-KOBA* are the complementation strains starting from BS301 and BS300 with the *rrnA* operon reinserted as outlined in Figure 3.2. A+JHB122, A-JHB122, A+JHAB621, and A-JHAB311 were constructed using the intermediate strains BS311 and BS310 with the Ara+ and Ara- marker, respectively. JHB strains have *rrnA* reinserted into the chromosome. JHAB strains have *rrnA* deleted from the chromosome.

Figure 3.5. Chemostat competitions of complementation strain A+KOBA*. The A+KOBA* vs. A-KOAB competition suggests the presence of *rrnA* provides a small fitness advantage. The A+KOBA* vs. A-KOB competition shows the fitness loss seen in the KOAB strains (BS301, BS300) is not complemented by the reinsertion of *rrnA*. The KOAB vs. KOB competition represents data from

reference (49) and is included to show the similarity to A+KOBA* vs. KOB competition. All chemostat competitions were run in four parallel chemostat vessels with each time point representing the mean ratio of Ara+ to Ara- with error bar representing plus and minus one standard deviation. The chemostats were run at an average dilution rate of 0.11 hr^{-1}.

Figure 3.6. Relative fitness of A-KOBA* and A-KOAB strains. A+Control is the parental strain (BS4548). Error bars represent +/- one standard deviation. The A-KOBA* competition clearly illustrates that the reinsertion of *rrnA* does not improve fitness.

Figure 3.7. Relative fitness in batch culture of JHAB and JHB strains. Error bars represent +/- one standard deviation. The large difference between the control competitions (A+JHAB vs. A-JHAB and A+JHB vs. A-JHB) indicate that a large amount of variability is introduced during the creation of the strains.

Figure 3.8. Ammons-Fox protocol for generating chromosomal deletions. Plasmid and strain information is summarized in Tables 3.1 and 3.2. The first two steps generate the required plasmids with the FRT-kan^r-FRT cassette inserted between the up/down sequence allowing site-specific integration of the pasmid into the *E. coli* chromosome. The plasmid pKO3 has a temperature sensitive origin of replication (ts-ori) and the *sacB* gene from *Bacillus subtilis* conferring sucrose sensitivity. *E. coli* B strains REL4548 and BS4548 are then

transformed with pKO3-delA (or -delB) and cointegrates are isolated by growing transformants at 42°C in the presence of kanamycin and chloramphenicol. Growth of cointegrants at 30°C in the presence of 6% sucrose and kanamycin selects cells with the plasmid excised and lost from the cell due to the lethal effects of the *sacB* gene in the presence of sucrose. Kanamycin resistant cells are then transformed with pCP20 carrying the Flp recombinase resulting in the site-specific excision of the kan^r gene flanked by FRT sites (Flp recognition sequence). After Flp removal of the kan^r gene, a partial FRT sequence remains at the site of the deletion.

Chapter 1

INTRODUCTION

rRNA operons and Bacterial physiology

In most bacterial species, ribosomal RNA (rRNA) is transcribed from an operon with a *rrs-rrl-rrf* (16S-23S-5S) gene organization, in some instances with transfer RNA (tRNA) genes located between *rrs* and *rrl* (Internally Transcribed Spacer; ITS) and distal to *rrf* (6). The organization of all three species of rRNA into an operon allows for their coordinated regulation preventing the production of an excess of one type of rRNA. In *Escherichia coli*, all seven rRNA operons have two promoters (P1, P2) with transcription from the P1 promoter subject to upstream activation, regulation by the stringent response, and growth rate dependent regulation, while transcription from the P2 promoter occurs at a lower, constitutive level (14).

In a growing cell, ribosome biosynthesis requires a large proportion of the cellular energy expenditures. Ribosomal RNA accounts for over half of the total RNA transcribed during rapid exponential growth illustrating the need for lots of ribosomes to meet the translational requirements of the cell (6, 14). Thus, the high cellular requirement for rRNA, especially during transitions to favorable growth conditions, can only be met by the presence of strong promoters and, in the case of many organisms, by having multiple rRNA operons.

rRNA operon copy number and ecological strategy

Ribosome biosynthesis is an energy intensive process that is tightly regulated to match the cellular requirements for translation. However, the cellular concentration of RNA is not proportional to the translational requirements of the cell at slow growth rates. Excess ribosomes were detected in *E. coli*, which has seven rRNA operons, at a growth rate of 0.1 doubling per hour suggesting the increased transcriptional capacity gained by having multiple rRNA operons may be a cost to the cell under slow growth conditions (31).

Recent observations have strengthened the apparent correlation of rRNA operon copy number and general ecological strategies. In *E. coli* strains, an increase in response time, or the interval of time required to enter exponential growth upon the shift to more favorable growth conditions, was seen with increasing numbers of rRNA operons disrupted by antibiotic cassettes suggesting a link between high rRNA operon copy number and rapid response times (12). These results, however, are complicated by the presence of several antibiotic cassettes in the cell and active promoters producing truncated or disrupted rRNA transcripts. The potential metabolic expense incurred by the presence of multiple rRNA operons under slow growth conditions was also studied in strains of *E. coli* that harbored additional plasmid-born copies of *rrnB*. The additional rRNA operons increased doubling times by 33% and was attributed to the diversion of energy to the synthesis of excess ribosomes (50). The general observation of rRNA operon copy number as a predictor for ecological strategy also holds true for environmental microbial isolates. Bacteria

isolated from the soil that respond quickly to the influx of nutrients and form colonies early on agar plates had an average of 5.5 *rrs* genes while late appearing colonies had an average of 1.4 *rrs* gene copies (30).

The pattern of species with few rRNA copies employing an ecological strategy characterized by efficient growth under low, sustained nutrient conditions and species with high numbers of rRNA operons characterized by quick response times to favorable conditions suggests a key predictive role of rRNA operon copy number for different ecological strategies. Both strategies have benefits and trade-offs. For the low copy number species, while they are efficient at growing at slow growth rates under constant nutrient conditions, they are unable to respond quickly to an abundance of nutrients and may be out competed by high copy number organisms when the environment is characterized by periodic fluctuations in favorable growth conditions. For high copy number organisms, while they may be able to respond quickly to an influx of nutrients due to their higher capacity to produce rRNA, the constitutive production of rRNA under slow growth conditions would favor low copy number organisms able to grow more efficiently. While several observations suggest an important association between rRNA operon copy number and ecological strategy, a direct experimental link has yet to be established.

Recombination and rRNA operons

When present in multiple copies, rRNA operons are a conspicuous deviation from the norm in Bacteria where almost all chromosomal genes are present in a single copy (6, 39). While potential ecological explanations for the

presence of multiple rRNA operons are discussed above, an additional implication of multiple rRNA operons is the potential for recombination between the operons. Such recombination events between rRNA operons have been shown to lead to inversions, deletions, duplications, and the exchange of sequence between rRNA operons in *E. coli* (22-27). Thus, multiple rRNA operons may also play an important role in the organization and reorganization of the chromosome of an organism.

Homologous recombination between multiple copies of rRNA operons can also result in the concerted evolution of the sequences and repair of operons that have been disrupted or may be non-functional (2-4, 7, 10, 22). The co-evolution of rRNA sequence is evident in the similarity of the multiple sequences within a species as compared to closely related species. For example, the seven rRNA operons of *E. coli* are more similar to one another (~99% identity) than to the seven rRNA operons of *Salmonella typhimurium* (~97% identity) (3, 4, 7, 10). In addition, several of the rRNA operons of *E. coli* have identical sequences but the number and identity of the operons with identical sequence can vary in different *E. coli* strains (10). Thus, despite decades of inquiry into the interactions between the rRNA operons of *E. coli*, a coherent view of the underlying dynamics which govern such interactions has yet to be addressed beyond the frequency with which the phenomena occur.

Overview of Thesis

In Chapter 2 of this thesis I report the measurement of recombination rates between *rrnB* and the remaining six rRNA operons of *E. coli*. Isolates that had undergone recombination events to remove an insertion in the *rrsB* gene were further examined to determine the fitness effects of recombination in batch culture competitions with a wild-type strain. The results from the competition experiments and recombination rate measurements are discussed in the context of developing a conceptual model for the maintenance of sequence heterogeneity in the rRNA operons of *E. coli*.

Chapter 3 summarizes two attempts to establish a growth rate specific role for the *rrnA* operon by reinserting *rrnA* into its native position in strains of *E. coli* that have *rrnA* alone or *rrnA* and *rrnB* deleted from the chromosome. The resulting strain constructions were tested for rescue of the wild-type fitness levels in batch culture and chemostat competitions. The shortcomings of the complementing strain constructions will be discussed and a potential new method for generating rRNA operon deletions will be proposed.

Chapter 2

RATE AND FITNESS EFFECTS OF RECOMBINATION BETWEEN rRNA OPERONS OF *ESCHERICHIA COLI*

2.1 Introduction

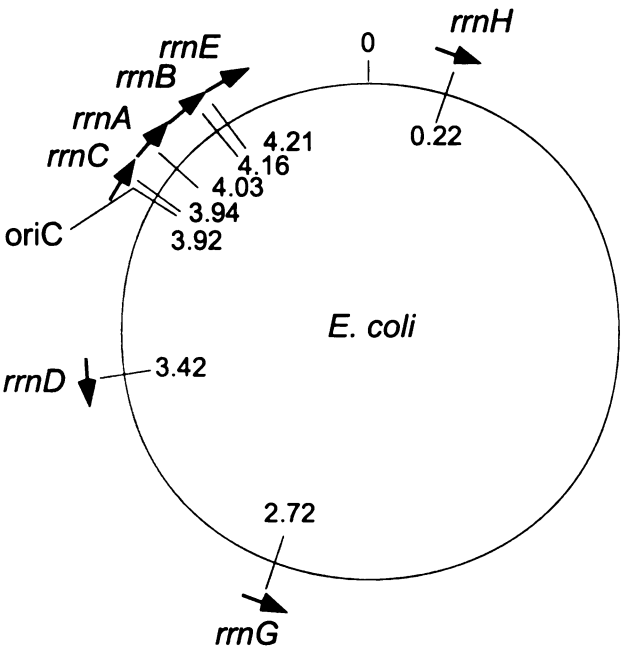
Escherichia coli has seven ribosomal RNA (rRNA) operons located asymmetrically on the chromosome that are highly similar in gene organization and nucleotide sequence (Figure 2.1) (6). The large regions of sequence similarity allow for inversions, duplications, and sequence exchange between the rRNA operons at frequencies of 10^{-3}-10^{-5} in overnight liquid cultures (6, 22, 24-27). High recombination frequencies between rRNA operons suggests that the *E. coli* chromosome is dynamic, yet the overall arrangement of rRNA operons in the chromosome is very stable. For example, the number, chromosomal location, and orientation of rRNA operons has been preserved in *E. coli* and *Salmonella typhimurium* (*S. enterica* serovar Typhimurium) since their divergence (6, 28). The observation of high recombination frequencies and the conservation of chromosome and operon organization suggests strong selection for the maintenance of *rrn* organization and location over evolutionary time scales.

While most genes in Bacteria and Archea are present in single copies, the genes encoding the rRNAs are often present in multiple copies allowing for concerted evolution of the sequences through sequence exchange (6, 34). Recombination between genes present in multiple copies on a chromosome can homogenize sequences by gene conversion (non-reciprocal exchange) or by

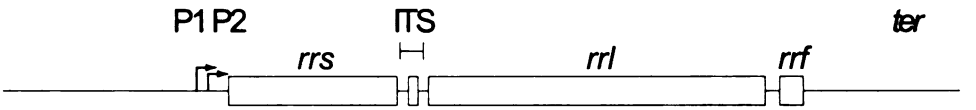
homologous recombination (reciprocal exchange) with sister chromosomes, resulting in the substitution of one gene sequence for another (36). The exchange of *rrn* sequences within the *E. coli* lineage has resulted in the seven *rrn* operons of *E. coli* appearing more similar to one another than to any of the seven *rrn* operons of *S. typhimurium*. If recombination between the rRNA operons that resulted in homogenization and co-evolution of the sequences did not occur, multiple *rrn* operons within a species would be expected to accumulate neutral mutations at a rate comparable to changes observed between species, which is clearly not the case. Apparent exceptions to this phenomenon exist in *Thermomonospora chromogena, Thermobispora bispora,* and *Haloarcula marismortui* where *rrs* or *rrl* sequences differ by up to 10% with the high degree of heterogeneity potentially explained by the lateral transfer of rDNA (17, 42, 53, 54).

This chapter focuses on the factors that affect the concerted evolution of rRNA sequences. The rate of recombination involving *rrnB* and the six other rRNA operons was measured to estimate the homogenizing force of recombination. The resulting fitness effects of inter-operon recombination were then measured to determine the dynamics affecting the eventual fixation or loss of the homogenized operons. The results were then used to develop a conceptual model of rRNA operon evolution, focusing on a mechanism for the maintenance of sequence heterogeneity between rRNA operons.

A.



B.



C.

rrn Operon	ITS tRNA(s)	Distal tRNA(s)
B, E, G	Glu_2	-
C	Glu_2	Asp_1, Trp
A	Ile_1, Ala_{1B}	-
D	Ile_1, Ala_{1B}	Thr_1
H	Ile_1, Ala_{1B}	Asp_1

Figure 2.1. Location and organization of *E. coli* rRNA operons. A. *E. coli* chromosome with location (mbp) and orientation of rRNA operons. B. General organization of *rrn* operons. P1P2 are the tandem promoters, *rrs* encodes16S rRNA, ITS represents the internally transcribed spacer, *rrl* encodes 23S rRNA, *rrf* encodes 5S rRNA, and *ter* is the termination region. C. Table of rRNA operons and their associated ITS and distal tRNA species.

2.2 Materials and Methods

Strains and genotype verification

Insertional inactivations of the *rrnB* operon were produced by linear transformation of a *recD*$^-$ strain of *E. coli* (D308) with a DNA fragment containing a *rrnB* operon and a *sacB-neo* cassette (8) inserted in the *Bgl*II site of *rrsB*. Generalized transduction (P1*vir*) was used to move the *rrnB* operon disrupted by the *sacB-neo* cassette from D308 into a strain of *E. coli* B that had been propogated in the lab for 10,000 generations (33). The resulting strain, BSV1, contains a *rrnB* operon with intact promoters and homology upstream and downstream of the *sacB-neo* cassette to the other rRNA operons (Figure 2.2). A *sacB-neo rrnB* allele with no upstream homology to the other rRNA operons was used to construct strain BSV2 with the promoters (P1 and P2), *rrsB*, and the 5' end of *rrlB* deleted, leaving a transcriptionally inactive *rrnB* region with no homology upstream of the *sacB-neo* cassette to any of the other rRNA operons (Figure 2.2) (49).

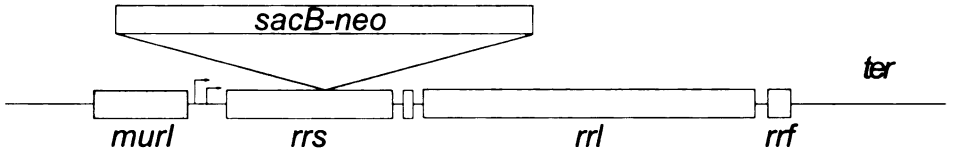
BSV1 and BSV2 genotypes were verified by both PCR and Southern hybridization. PCR amplification of *rrnB* was carried out using the upB forward (5'-CCGAATTACATATGACCGTGCTGGTGTTTGAC-3') and dnB reverse (5'-CTACTCTAGACCTGATGCAAAAACGAGGCTAGTTTA-3') primers and r*Tth* DNA Polymerase XL (PE Applied Biosystems; Foster City, CA) under the following cycle conditions: 94°C 3 min, 94° C 30"; 60° C 1 min; 72° C 10 min; 30 cycles, 72° C 10 min, hold at 4° C. The rRNA operons were visualized using Southern hybridization on *Pvu*II digested genomic DNA probed with a digoxignin-

dUTP-labeled DNA probe for a conserved region of *rrs* (positions 8-536 of *rrsB*) (30).

Growth conditions and media

Cultures were typically grown in Luria-Bertani liquid medium (LB), or with 1.5% Bacto Agar (Difco, Laboratories; Detroit, MI) added for solid medium. When selecting for sucrose resistance, sucrose was added at a final concentration of 6% to LB with NaCl omitted (LBsuc) (8). Where noted, kanamycin was added to LBsuc at a final concentration of 50μg/ml (LBsuc-kan). Overnight cultures consisted of 2 ml of LB in 16 x 125 mm culture tubes, inoculated from freezer stocks. Competition experiments, conditioning cultures, and cultures for rate measurements were grown in either 10 ml of media in a 50 ml Erlenmeyer flask or 50 ml of media in a 250 ml Erlenmeyer flask shaking at 225 rpm at 37° C. The conditioning cultures and competition experiments used to measure relative fitness were grown in Davis-Mingoli broth (DM) supplemented with 2.0 mg/L thiamine hydrochloride and 25 mg/l of glucose (DM+25) (33).

A. BSV1



B. BSV2

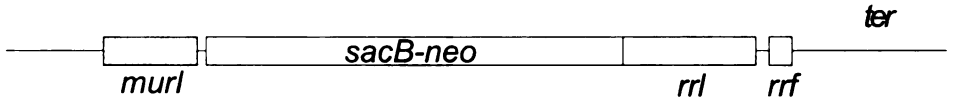


Figure 2.2. BSV1 and BSV2 *rrnB* operon constructions containing the *sacB-neo* cassette. A. BSV1 has the *sacB-neo* cassette inserted into the *Bgl*II site of *rrsB,* upstream sequence similarity to the other *rrn* operons, and a transcriptionally active operon. B. BSV2 was constructed through the removal of sequences upstream of *rrnB*, *rrsB*, the ITS region, and the 5' end of *rrl* for the *sacB-neo* cassette leaving no upstream sequence similarity to the other *rrn* operons and a transcriptionally inactive operon. The gene *murI* is unique to the *rrnB* region and represents the end of upstream sequence similarity to the other rRNA operons.

Recombination rate measurements

The rate measurements for recombination and *sacB* mutation were carried out as outlined by Crane *et. al* (15), and is a modification of the fluctuation analysis of Luria and Delbruck (38). Eleven flasks with 10 ml of LB were inoculated with ~100 cells from an overnight LB culture. The cultures were incubated at 37 °C shaking at 225 rpm for 15 hours. The stationary phase cultures were sampled by spreading 75µl of the culture on 4 LBsuc plates, 4 LBsuc-kan plates, and diluted and spread on 4 LB plates. After overnight incubation at 30° C, the number of colonies on each plate was counted. For each replicate culture, the size of the recombinant population was calculated as the number of colonies on LBsuc minus the number of colonies on LBsuc-kan. The *sacB* mutant population was represented as the number of colonies on LBsuc-kan.

To determine the rate of recombination and *sacB* mutation, the median value for each population was determined from the 11 replicate cultures. The following median estimator of mutation rates was used to determine rates of mutation and recombination (15, 29):

$$\lambda_{med} = ((r_m/s) - \ln 2)/(\ln(r_m/s) - \ln(\ln 2)) \qquad \text{(equation 1)}$$

where r_m is the median number of recombinants or mutants per culture, s is the dilution factor, and λ_{med} is the median number of recombination events or mutation events in the culture. The recombination or mutation rate was then determined by dividing λ_{med} by the median of the total number of cells per culture.

For the rate of double mutation in *sacB* and *neo* in BSV2 (recombinant phenotype; sucrose resistant, kanamycin sensitive), a different equation was used because many of the replicate cultures produced no colonies with the recombinant phenotype. The following equation is more suited for results with many 'mutant free' cultures (15, 29):

$$m_g = ((1-p) \ln (z/C))/(p \ln (p)) \qquad \text{(equation 2)}$$

where p is the dilution factor, z is the number of cultures devoid of mutants, C is the total number of cultures, and m_g is the number of mutants per culture.

Relative fitness assay

Overnight LB cultures of each competing population were inoculated from freezer stocks. The following day, conditioning cultures, consisting of 10 ml of DM+25 in 50 ml Erlenmeyer flasks were inoculated with a final dilution of 1:10,000 from the overnight cultures. Competition cultures were initiated 24 hours after inoculating the conditioning cultures by adding 100µl of a 1:1 mix of the two competitors to 5, 50ml flasks containing 10 ml of DM+25. BSV1, BSV2, and the resulting mutants and recombinants carry a neutral arabinose utilization marker (Ara^+) which enables the enumeration of the competitor versus the control strain (Ara^-), because on TA plates Ara^+ colonies appear white and Ara^- colonies appear red (33). Cultures were transferred at 24 hr intervals for 4 days and monitored daily by plating a $5x10^{-6}$ dilution of the each replicate culture on 3 TA plates.

Relative fitness measurements were calculated using the following equations:

$$m_1 = \ln [N_1(1)/N_1(0)]/(1\ d) \qquad \text{(equation 3)}$$

$$W_{1:2} = m_1/m_2 \qquad \text{(equation 4)}$$

where $N_1(0)$ and $N_1(1)$ are the population size of Ara^+ strains at 0 hours and 24 hours, respectively, and m_1 is the Malthusian parameter for the Ara^+ strain. The relative fitness of the Ara^+ strain to the control Ara^- strain was calculated using equation 4 where m_1 is the Malthusian parameter of the Ara^+ strain and m_2 is the Malthusian parameter of the Ara^- control strain (33, 49). Statistical analysis of the relative fitness of BSV1, the BSV1 recombinant isolates, and the BSV1 *sacB* isolates included ANOVA and pair-wise comparisons using the Bonferroni multiple comparisons procedure.

rRNA operon sequence alignments

Sequences of *rrs* (16S) and *rrl* (23S) rRNA genes were obtained from GenBank, The Institute for Genomic Research Microbial Genome Database (www.tigir.org/tdb/mdb/mdbcomplete.html), and the National Center for Biotechnology Information Microbial Genome Database (www.ncbi.nlm.nih.gov/PMGifs/Genomes/bact.html) and analyzed with the Wisconsin Package Version 10.0, GCG software package (Madison, WI). The following microbial genomes containing multiple rRNA operons were used in the analysis and represent all available sequenced genomes with multiple rRNA operons at the time of preparation: *Aquifex aeolicus* VF5, *Bacillus subtilis* ATCC 23857, *Campylobacter jejuni* ATCC 700819, *E. coli* ATCC 10798, *Haemophilus*

influenzae ATCC 51907, *Helicobacter pylori* 26695, *Methanococcus jannaschii* DSMZ 2661, *Methanobacterium thermoautotrophicum* ATCC 29096, *Neisseria menigitidis* MC58, *Treponema pallidum* ATCC 25870, *Ureaplasma urealyticum* serovar 3, *Vibrio cholerae* ATCC 39315, and *Xyella fastidiosa* 9a5c. Pair-wise differences were calculated for the *rrs* and *rrl* genes of each organism. Pair-wise sequence differences were scored and distances between rRNA operons were calculated based on the published locations of the operons (9, 16, 19, 20, 32, 43, 44, 46, 48, 51, 52).

2.3 Results

Rate of recombination involving *rrnB*

Fluctuation analysis of BSV1 showed the average rate of recombination removing the *sacB-neo* cassette in *rrnB* of BSV1 to be 4.75×10^{-9} per cell division (Table 2.1) in two independent measurements using equation 1 and the values in Table 2.2. Of 14 independent BSV1 recombinant isolates screened by Southern hybridization, two (14%) had recombined with a rRNA operon containing the $tRNA^{Ile1}$-$tRNA^{Ala1B}$ ITS spacer (*rrnA*, *rrnD*, or *rrnH*), resulting in a hybrid *rrnB* operon with an alternate ITS region. In BSV2, the recombinant phenotype occurred at a rate of 1.4×10^{-10} per cell division (equation 2, Table 2.2), and is within the range of values measured for the spontaneous inactivation of both phenotypes of the *sacB-neo* cassette (8). All BSV2 isolates with the recombinant phenotype (sucrose resistant, kanamycin sensitive) that were screened by PCR and Southern hybridization maintained the 3.8 kb insertion in *rrsB*, indicating the

loss of both phenotypes of the *sacB-neo* cassette was a result of mutations rather than the removal of the cassette by recombination. The average *sacB* mutation rate was 5.7×10^{-8} per cell division in BSV1 and BSV2 (Table 2.1). In all *sacB* mutants (sucrose resistant, kanamycin resistant) examined by PCR and Southern hybridization, the *sacB-neo* cassette was present in the *rrnB* operon (Figures 2.3, 2.4, 2.5).

Table 2.1. Recombination and sacB mutation rates.

Strain	Recombination Rate (per cell division)	sacB mutation rate (per cell division)
BSV1 rep. 1	4.8×10^{-9}	4.6×10^{-8}
BSV1 rep. 2	4.7×10^{-9}	7.4×10^{-8}
BSV2	0*	5.1×10^{-8}

*The rate of the recombinant phenotype (sucrose resistant, kanamycin sensitive) in BSV2 cultures was 1.4×10^{-10} per cell division. However, all BSV2 isolates with the recombinant phenotype were found to have maintained the insertion in the *rrnB* gene.

Table 2.2. Values used in equations 1(A) and 2(B) to determine recombination and sacB mutation rates

A. Strain	r_m	s (dilution factor)	λ_{med}	Population
BSV1rec (rep#1)	41	0.03	180	3.76×10^{10}
BSV1rec (rep#2)	20	0.004	562	1.20×10^{11}
BSV1sacB (rep#1)	523	0.03	1720	3.76×10^{10}
BSV1sacB (rep#2)	424	0.004	8879	1.20×10^{11}
BSV2 sacB	447	0.03	1493	2.91×10^{10}

B. Strain	p	z	C	m_g	Population
BSV2rec	0.03	7	11	4.17	2.91×10^{10}

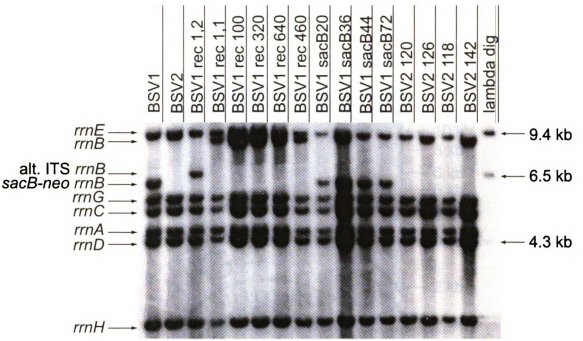


Figure 2.3. Southern hybridization of BSV1 and BSV2 isolate genomic DNA digested with *Pvu*II and probed for a conserved region of *rrs*. In BSV1, the band corresponding to *rrnB* is shifted when containing the *sacB-neo* cassette inserted in *rrs* (lane 1), after recombination with a *rrn* operon with a $tRNA^{Ile}$-$tRNA^{Ala}$ ITS (lane 3), and back to wild-type after recombination with a *rrn* operon with a $tRNA^{Glu}$ ITS (lanes 4-8). BSV1 *sacB* isolates are sucrose resistant and kanamycin resistant and still harbor the *sacB-neo* cassette in *rrsB* (lanes 9-12). BSV2 120 and BSV2 126 are sucrose resistant and kanamycin sensitive (recombinant phenotype) but have not regained the deleted region of *rrsB* (lanes 13-14). BSV2 118 and BSV 142 are sucrose resistant and kanamycin resistant (*sacB* phenotype) and have not regained the deleted region of *rrsB* (lanes 15-16).

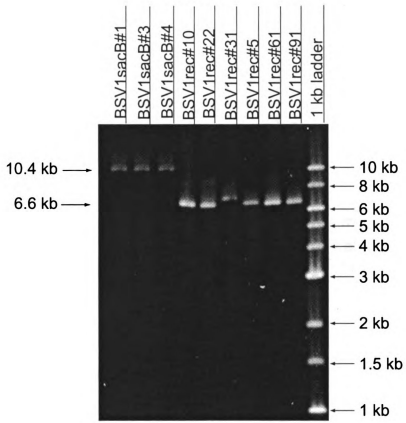


Figure 2.4. PCR amplification of the *rrnB* region of the isolates used in the relative fitness assay. Amplification of the *rrnB* operon with the *sacB-neo* cassette produces a 10.4 kb fragment as in the BSV1*sacB* isolates. Amplification of the wild-type or recombinant *rrnB* region produces a 6.6 kb fragment. The BSV1recombinant isolates (rec) all have PCR amplification products of 6.6 kb indicating the removal of the *sacB-neo* cassette. The BSV1rec#31 isolate was excluded due to the strange amplification size and unexpected banding pattern in the Southern hybridization (Figure 2.5).

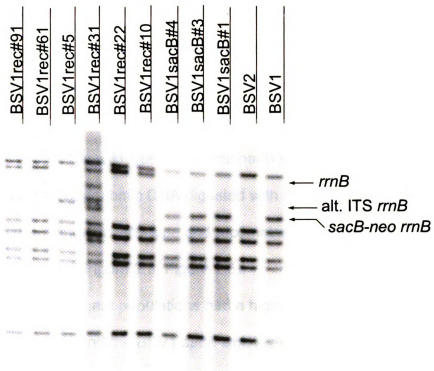


Figure 2.5. Southern hybridization of *Pvu*II digested genomic DNA from BSV1 isolates used in the relative fitness assay. The banding pattern confirms the recombinant isolates (rec) have lost the *sacB-neo* cassette and *sacB* mutants (sacB) still have the cassette in *rrnB*. Due to the unusual banding pattern of BSV1rec#31, it was excluded from futher analysis. BSV1rec#5 has a banding pattern indicative of an alternate ITS in rrnB.

Fitness effects of recombination between rRNA operons

Independent isolates of BSV1 cultures with the recombinant phenotype (sucrose resistant, kanamycin sensitive) were confirmed by PCR and Southern hybridization to have lost the *sacB-neo* cassette from *rrsB* (Figure 2.4, Figure2.5). The amplification product from the *rrnB* operon with the *sacB-neo* cassette in BSV1 is 10.4 kb, while the amplification product from the *rrnB* operon without the cassette (either wild type or recombinant) is 6.6 kb (Figure 2.4). Southern hybridization of genomic DNA digested with *Pvu*II and probed for *rrs* produces a 8.4 kb fragment in wild-type and recombinant isolates of *E. coli* and a 6.1 kb fragment in BSV1 (Figure 2.3, Figure 2.5).

The BSV1 recombinant populations had a higher relative fitness than both the parental strain BSV1 and the BSV1 *sacB* isolates ($p = 0.0089$, $p = 0.0034$; respectively). The relative fitness difference between the BSV1 recombinant isolates and either BSV1 or the BSV1 *sacB* isolates was about 1% (Figure 2.6). There was no significant difference between BSV1 and theBSV1 *sacB* isolates ($p = 0.2823$). One of the BSV1 recombinant isolates had an abnormal Southern hybridization pattern and a PCR amplification product at an unexpected size. Due to this discrepancy, the data from BSV1rec#31 was excluded from our analysis. One of the recombinant isolates used in the relative fitness measurement (BSV1rec#5) was found to have recombined with a rRNA operon with a $tRNA^{Ile}$-$tRNA^{Ala}$ ITS by Southern hybridization. The relative fitness of the alternate ITS recombinant was indistinguishable from the other recombinants and is included in the BSV1 recombinant group.

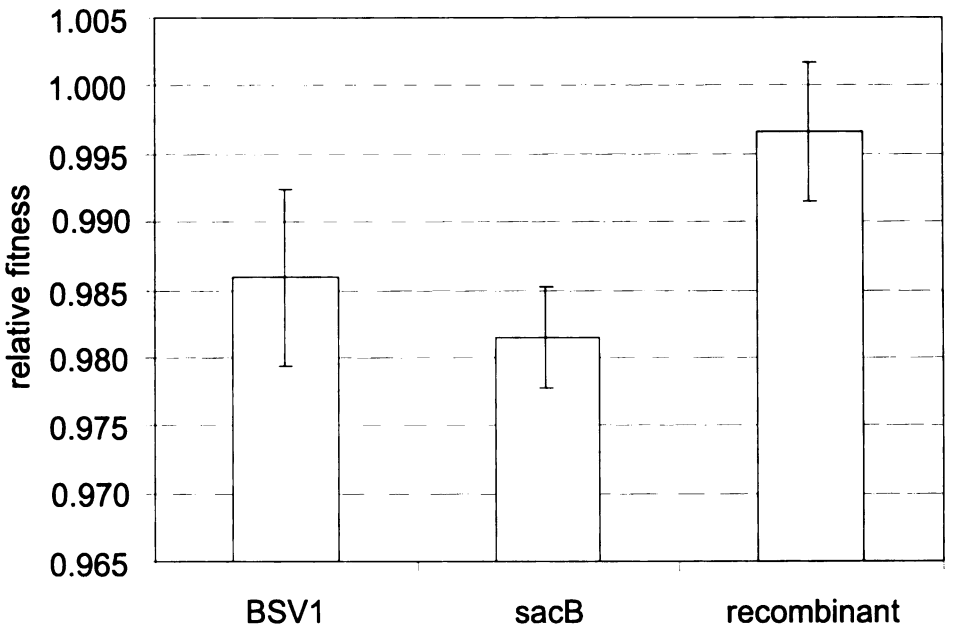


Figure 2.6. Relative fitness of BSV1, sacB, and recombinant isolates. The recombinant isolates are significantly different from both BSV1 and sacB groups but BSV1 and sacB are not statistically distinguishable. Error bars represent +/- 1 standard deviation for each group.

Correlation analysis of pair-wise *rrs* and *rrl* sequence difference and distance

The results of the pair-wise sequence comparisons of *rrs* and *rrl* from sequenced microbial genomes with multiple rRNA operons are summarized in Figure 2.7. There is a significant positive correlation between the total pair-wise differences present in *rrs* and *rrl* and the distance between the operons compared, suggesting a potential link between chromosomal distance between operons and recombination rate.

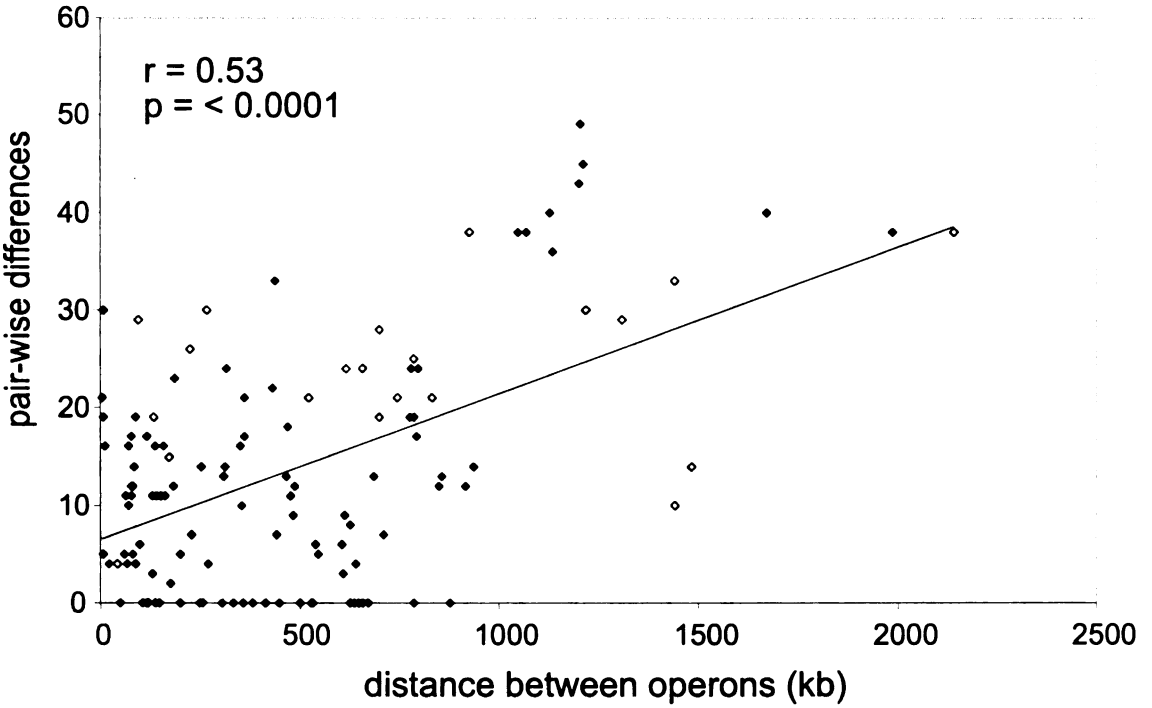


Figure 2.7. Correlation of pair-wise sequence differences of *rrs* and *rrl* with distance between operons of all available bacterial and archeal genome sequences with multiple rRNA operons. Pair-wise sequence differences were tallied and distances between operons were calculated based on their published locations. *E. coli* comparisons are represented as open diamonds and the other organism comparisons are represented as closed diamonds.

2.4 Discussion

The objective of this investigation was to address and quantitatively measure some of the factors that affect the co-evolution of rRNA sequences in *E. coli*. The presence of reiterated sequences in a genome allow for various types of recombination with a variety of outcomes. One result of recombination between repeated sequences is the potential for sequence exchange resulting in the co-evolution of the sequences. Ribosomal RNA operons are often present in multiple copies in bacterial and archeal genomes setting them apart from most other genes which are normally present in single copies. Prokaryotic species with multiple copies of the rRNA operon have been shown to have rDNA sequences that are more similar within a species than between species which clearly indicates that the rRNA genes are co-evolving (34).

Gene conversion is typically invoked to explain the concerted evolution of gene families. The clearest cases for gene conversion are from Ascomycetes fungi where all the products of meiosis are easily examined (40, 41). Our system, however, does not allow the easy isolation and identification of all the products of a given recombination event and therefore any claim as to the nature of the specific recombination mechanism would be tentative, at best. The type of recombination events described here, however, clearly would lead to the concerted evolution of rRNA sequences as a heterogeneity (*sacB-neo* cassette) has been removed from the *rrnB* operon through recombination with one of the other rRNA operons.

rRNA operon recombination results in the loss of the *sacB-neo* insertion

The involvement of recombination in genetic alterations is typically shown by the requirement of one or a combination of recombination genes (i.e. *recA*, *recB*, *recF*). Here we have not shown a requirement for any of the recombination genes but two lines of evidence clearly show the involvement of recombination in our system. First, the loss of the 3.8kb *sacB-neo* insertion was observed only when flanking homology to the other rRNA operons was present. Second, the loss of the *sacB-neo* cassette from *rrnB* was observed to coincide with the exchange of internally transcribed spacer (ITS) regions resulting in an altered mobility of the *rrnB* fragment in Southern hybridizations (Figure 4). In 14% of the BSV1 recombinant isolates (2 of 14), Southern hybridization indicated that the *rrnB* operon had recombined with an operon containing the $tRNA^{Ile1}$-$tRNA^{Ala1B}$ ITS (*rrnA*, *rrnD*, or *rrnH*) resulting in an *rrnB* operon with the $tRNA^{Ile1}$-$tRNA^{Ala1B}$ ITS. The presence of the $tRNA^{Ile1}$-$tRNA^{Ala1B}$ ITS region in *rrnB* could only occur through recombination between *rrnB* and a rRNA operon with the $tRNA^{Ile1}$-$tRNA^{Ala1B}$ ITS and, therefore, clearly implicates recombination in the loss of the *sacB-neo* cassette.

The recombination rate measured for BSV1 is strikingly similar to other rates reported for the exchange of sequence via recombination. The *sacB-neo* cassette was lost from BSV1 at a rate of $5x10^{-9}$ per cell division which is within the range of recombination rates ($3x10^{-9}$ to $2x10^{-8}$) observed between the *tuf* genes of *S. typhimurium* (1). Mutation and recombination events are commonly reported as a frequency of a population, which is a measure of the rate and the

subsequent growth of the recombinant or mutant sub-populations. The reporting of recombination and mutation events as frequencies and not rates makes comparisons to other studies difficult. In a strain of *E. coli* with only one $tRNA^{Glu2}$ spacer, the exchange of ITS regions resulting in an increase in the number of $tRNA^{Glu2}$ spacers occurred at a frequency of $5x10^{-5}$ (22). The apparent discrepancy between our measured rate of recombination and previously measured frequencies can be attributed in part to the large differences in the growth rate of the recombinant and parental populations in the experiments by Harvey and Hill (22) caused by the large disparity of ITS tRNA species and their lack of control for "jack-pot" cultures.

Fitness effects of rRNA recombination and rRNA operon interruption

The loss of the *sacB-neo* cassette from *rrnB* through homologous recombination with one of the other six intact rRNA operons resulted in a 1-2% increase in relative fitness (Figure 2.6). Thus, the cumulative effects of *rrnB* interruption, carrying the *sacB-neo* cassette, and recombination involving *rrnB* is approximately 1%. It has been previously reported that antibiotic cassette interruptions of either *rrnA*, *rrnB*, or *rrnD* resulted in no significant change in growth rate in various types of media and temperatures (11, 12). Our results do not contradict these earlier studies because the comparisons of relative fitness made here were measured in head-to-head competition experiments which detects small differences in lag, exponential growth, and stationary phase, enabling us to resolve the differences summarized in Figure 2.6.

The BSV1 *sacB* isolates had a relative fitness indistinguishable from BSV1 and significantly lower than the recombinant isolates indicating that a functional *sacB* gene has little impact on relative fitness. The relative fitness of the recombinant isolate with an alternate ITS region was found to be within the range of the other recombinant isolates, suggesting the exchange of spacer ITS has little effect on relative fitness. Because of the paucity of alternate ITS recombinants, statistically supported comparisons could not be made for this group.

Factors affecting the co-evolution of rRNA operons in *E. coli*

The similarity of the seven rRNA operons of *E. coli* and preservation of chromosome and rRNA operon organization since the divergence from *S. typhimurium* (6, 28) suggests that selection against recombination between rRNA operons is negligible when resulting in the exchange of short stretches of sequence with no alteration of the organization of the chromosome or the rRNA operons. Furthermore, apart from short insertions and deletions from the ITS regions of the seven rRNA operons of *E. coli* K-12, the degree of heterogeneity in the ITS is similar to the degree of heterogeneity seen in *rrs* and *rrl* (3, 4), indicating that sequence exchange and homogenization affects the entire rRNA operon and is not limited to the structural genes.

If the heterogeneity in the rRNA operons is not selectively maintained, the presence of heterogeneity is surprising, due to the high rate and frequency of recombination between rRNA operons observed in this and other studies. In light of the observations presented here, the maintenance of sequence heterogeneity

in the rRNA operons of *E. coli* can be explained by a model in which the balance of the antagonistic phenomenon of recombination and mutation results in similar yet heterogeneous rRNA operons.

If we assume the 1% fitness increase seen in the recombinant isolates is specifically attributable to the repair and reestablishment of a functional rRNA operon, the fixation of recombination events between rRNA operons would be expected to be a process that occurs very rarely. In an environment similar to the conditions of the relative fitness assay, the probability of fixation of a recombinant with a fitness increase of 1% over the rest of the population is 4.4×10^{-3} and would occur only after ~2500 generations (33). However, the 1% increase in the relative fitness of the recombinant isolates may not be solely attributable to the reestablishment of the *rrnB* operon, but may also be due to the loss of the cassette carrying and the kanamycin resistance phenotype. Thus, homogenization may be occurring at a significant rate but the fixation of the homogenized operons would rarely occur or would occur on the rare occasion that the homogenized operons hitch-hike to fixation with a beneficial mutation located elsewhere in the chromosome.

Comparisons of rRNA operon sequences of available Bacteria and Archea complete genome sequences adds an additional factor to this model. When the data from all of the prokaryotic organisms with sequenced genomes and two or more rRNA operons are analyzed, there is a positive correlation between the number of pair-wise sequence differences (of 16S and 23S genes within an individual organism) and the distance between the operons on the chromosome

(Figure 2.7). In other words, the further apart two rRNA operons are, the more likely they are to have sequence heterogeneity, suggesting that the proximity to other operons may be an important factor in the rate with which operons recombine. Thus, the outcome of the opposing forces of the recombination rate resulting in homogenization and the de novo mutation rate resulting in heterogeneity becomes more complex. Organisms with multiple rRNA operons would have different rates of recombination for each operon depending on the proximity to the other operons. Thus, by being localized distantly from other rRNA operons, an operon may be able to escape some of the homogenization by having a lower rate of recombination resulting in a divergent operon.

The dependence of recombination rate on the distance between the regions participating in the exchange has been shown for plasmid and chromosomal based systems in various types of organisms (18, 37, 45, 47). In *E. coli*, recombination rates were reduced up to two orders of magnitude when homologous regions were separated by 7kb (37). In *S. typhimurium*, recombination between 5 kb regions is reduced by over one order of magnitude when the distance separating the homologies is increased from 2 to 17 minutes (45). In Eukarya, the pattern of decreasing recombination rate with increasing distance between regions has been observed in the nontrascribed spacers (NTS) of ribosomal RNA genes (18). The pattern of heterogeneity in NTS regions of wheat, rye, maize, and mice was better fit by a model that assumes an increased probability of sequence exchange for closely located regions (18). In the bacteriophage T4, the frequency of recombination decreased over one order of

magnitude when the distance between recombining regions was increased from 79 to 1698 bp (47). Thus, the observed positive correlation of sequence difference and distance has empirical support based on recombination rates between homologous sequences separated by varying distances, and may be a widespread phenomenon impacting the concerted evolution of gene families in all domains of life.

The positive correlation of sequence heterogeneity between rRNA operons and distance between the operons raises the question of how proximity would affect recombination rates. The correlation may be a result of local unwinding of the DNA during replication or transcription resulting in easier access to close regions for strand invasion and sequence exchange. Additionally, the proximity of closely located operons may result in an increased local concentration of operons during replication. As a replication fork passes through a region, the operons in that region would be at a higher relative concentration than other operons resulting in an increased probability of sequence exchange.

A subtle yet interesting implication of the observed correlation of increasing sequence heterogeneity with distance is the implied mechanistic nature of the interactions leading to homogenization. While not a direct inference from the data, it seems more likely that the observed correlation would arise primarily through intrachromosomal interactions where replication is not required in order to explain an increased frequency of interaction between closely located operons. Only under near maximal growth rates, growth conditions assumed to be rare and transitory, does gene copy number increase appreciably as a result

of the off-set timing of chromosomal replication and cell division in *E. coli*. Thus, the potential for increased interaction based on the proximity of rRNA operons on the chromosome seems to favor intrachromosomal interactions but in no way excludes interchromosomal interactions from contributing to the observed correlation.

The presence of sequence heterogeneity between the rRNA operons of *E. coli* is probably a result of several factors. First, the lack of a large fitness benefit associated with recombination events between rRNA operons provides a steep barrier to the fixation of the recombinant operons in a population. Thus, very few homogenizing events would be expected to rise to fixation in the population, but the few that do become fixed result in the overall similarity seen in the rRNA operons. In the case of *E. coli*, the rate of fixation of recombinants would have to be lower than the rate of fixation of point mutations within the sequences resulting in highly similar sequences but with some differences. Second, the increased chance of recombination between operons located close to one another would result in some very similar or identical sequences and some operons that contain more sequence differences.

An additional mechanism for the appearance of heterogeneity between rRNA operons other than point mutation is the inter-species lateral transfer of rDNA (54). While not specifically considered by the model presented here, lateral transfer has been suggested as the most likely mechanism for the presence of highly heterogeneous (~10% pair-wise difference) rRNA genes in several species and remains a potential source of introducing sequence

heterogeneity (17, 42, 53, 54). However, in our model, the origin of heterogeneity within the rRNA operons is for the most part irrelevant because once heterogeneity is present, it will be governed by the homogenizing force of recombination in the same way as heterogeneity arising via point mutations.

Conclusions

We have provided direct measurements of recombination between the rRNA operons of *E. coli*. A rate of ~$5x10^{-9}$ per cell division for recombination events involving *rrnB* was measured in two independent fluctuation experiments and is in agreement with similar recombination events in related systems. The fitness effects of recombination between rRNA operons resulting in the reestablishment of an intact rRNA operon was found to be ~1%. These results may be interpreted in an analogous situation where a mutation within a structural rRNA gene arises resulting in a non-functional transcript. In this case, the recombinational repair of such a lesion resulting in the reestablishment of a functional rRNA transcript would be associated with an increase in fitness of ~1%.

The presence of highly similar, but not identical, rDNA sequences can be explained by the rare chance of fixation of genotypes with a small fitness advantage allowing for some homogenization of sequences through recombination and heterogeneity arising by de novo mutation. In the case of *E. coli*, the rate of mutation and fixation of such mutations giving rise to heterogeneity must be higher than the rate of recombination and fixation of the recombinants resulting in the heterogeneity that is present. These observations

in addition to the observation that some organisms have completely identical rRNA operons but others have significant heterogeneity, raises the question of why sequence heterogeneity exist in some organisms and not in others. The degree of heterogeneity may simply be a result of the rates of recombination and mutation and proximity to the other operons. There may also be ecological explanations for the heterogeneity depending on the type of selection and the frequency and severity of population bottlenecks resulting in the fixation of more homogenization events or the increased divergence of certain operons resulting in more heterogeneity.

Chapter 3

ALTERNATE METHODOLOGIES FOR THE REMOVAL OF RIBOSOMAL RNA OPERONS FROM THE CHROMOSOME OF *ESCHERICHIA COLI*

3.1 Introduction

Microbial species with multiple copies of ribosomal RNA (rRNA) operons have been proposed to employ an ecological strategy which differs from organisms with one or a few copies of the rRNA operon (30). Specifically, organisms with a high number of rRNA operons are expected to have the transcriptional potential to quickly produce rRNA and ribosomes allowing a rapid shift-up of growth upon the influx of nutrients resulting in shorter lag times and higher growth rates. Organisms with one or a few copies of the rRNA operon are thought to be more efficient at growing and surviving during periods of low sustained nutrient levels but respond slower to an influx of nutrients. While these observations seem to hold true in many different organisms (30), a causal link between the copy number of rRNA operons and ecological strategy has yet to be made.

Previous work sought to establish a direct link between rRNA operon copy number and ecological strategy by sequentially deleting two of the seven rRNA operons of *E. coli* (*rrnB* and *rrnA*) (49). In batch culture competition experiments of the deletion strains versus the wild-type parental strain of *E. coli* B, there was a decrease in relative fitness with each sequential deletion of a rRNA operon (i.e. wt > Δ*rrnB* = Δ*rrnA* > Δ*rrnAB*) supporting the hypothesis that multiple rRNA operons is beneficial during fluctuating nutrient conditions. In chemostat

competitions, however, where the dilution rate of the culture vessels was 0.11 per hour (dilution rate = growth rate) simulating a constant low nutrient environment, Δ*rrnA* and Δ*rrnAB* strains unexpectedly showed a pronounced decrease in relative fitness (50). In parallel chemostat competitions, Δ*rrnB* strains showed no loss in fitness relative to the wild-type ancestor suggesting a specific role for *rrnA* under slow growth conditions. Previous comparisons of rRNA operon promoter strengths (P1, P2) suggested that *rrnH*, *rrnE*, and *rrnG* may differ in their regulation from the other operons under some conditions (13). Despite the evidence of differential regulation for some of the rRNA operons, it was still surprising to see such a pronounced growth rate specific role for *rrnA*.

A potential reason for the operon specific nature of the decreased fitness lay in the sequence differences between *rrnB* and *rrnA*, the most obvious being the presence of $tRNA^{Glu2}$ in *rrnB* and $tRNA^{Ile1}$ and $tRNA^{Ala1B}$ in *rrnA* (6). By deleting the *rrnA* operon, an imbalance may arise in the cellular tRNA pool under slow growth conditions, where the $tRNA^{Ile1}$ and $tRNA^{Ala1B}$ may be in higher demand relative to $tRNA^{Glu2}$ resulting in a growth rate specific effect of deleting *rrnA*. There are also sequence differences in the structural genes and in the spacer regions that could result in transcripts that are more easily processed into mature rRNA under certain conditions, or even ribosomes produced from the different operons with slightly different translational functions.

An alternate explanation for the results explained above is that a mutation or other disruption occurred during the construction of the Δ*rrnA* and Δ*rrnAB* strains and that the effects are not attributable specifically to the *rrnA* operon. In

order to establish a direct link between the deletion of the *rrnA* operon and the drop in relative fitness in slow dilution chemostat competitions, complementation was necessary. Reestablishing a wild-type *rrnA* operon in the Δ*rrnA* and Δ*rrnAB* strains and observing a rescue of the wild-type or Δ*rrnB* fitness under the given conditions would establish the causal link between the *rrnA* operon and the observed fitness loss in chemostat competitions.

In this chapter, I outline the construction of several strains using two different approaches to complement the Δ*rrnAB* strains and the subsequent measurements of relative fitness. The shortcomings of the strains will be discussed and a new approach to constructing rRNA operon deletions will be introduced that may help establish the link between rRNA operon copy number and fitness.

3.2 Materials and Methods

Cloning, XL-PCR, and general methods

Genomic DNA and plasmids were harvested from overnight cultures using the QIAamp Tissue Kit from Qiagen (Santa Clarita, CA), and the Wizard Plus SV Minipreps from Promega (Madison, WI), respectively. Restriction digestions were carried out using restriction endonucleases from New England Biolabs (Beverly, MA) according to the manufacturers recommended conditions. Gel electrophoresis was carried out using 0.6-1.5% SeaKem LE Agarose (FMC; Rockland, ME) in 0.5xTBE. When necessary, DNA fragments were extracted

from agarose gels using the Qiaex II Gel Extraction Kit from Qiagen (Santa Clarita, CA).

Amplification of genomic or plasmid DNA was carried out with r*Tth* XL-DNA polymerase from PE Applied Biosystems (Foster City, CA) using purified genomic or plasmid DNA. The following parameters were used in a typical 50 or 100µl amplification reaction: 94°C 3 min, 94° C 30"; 60° C 1 min; 72° C 10 min: 30 cycles, 72° C 10 min, hold at 4° C.

Cultures were typically grown in Luria-Bertani (LB) with 1.5% Bacto Agar (Difco Laboratories; Detroit, MI) added for solid media. When necessary, the following antibiotics were added at the specified concentrations: ampicillin (amp; 60µg/ml), chloramphenicol (cam; 30µg/ml), kanamycin (kan, 50µg/ml). For LB sucrose media (LBsuc), a final concentration of 6% sucrose was added to LB with NaCl omitted (8). All competition experiments and conditioning cultures were carried out in Davis-Mingoli broth (DM) (33) supplemented with 2.0 mg/L thiamine hydrochloride and glucose at a final concentration of 25 mg/l (DM+25).

Southern hybridization was used to visualize the rRNA operons using a digoxigenin-dUTP-labeled DNA probe (Roche; Indianapolis, IN) for a conserved region of *rrs* (positions 8-536 of *rrsB*) (30).

Version 1 plasmid constructions

The *rrnA* operon of BS4548 was amplified using the upA forward (5'-GCGAATTACATATGCCCTCACGCCATCCTCTTTTAT-3') and dnA reverse (5'-CTACTCTAGACGTTCGCATCCGGCATTTTTTT-3') primers. The PCR reaction mixture was subjected to a post amplification incubation at 72° C for 10 minutes

with *Taq* polymerase to add 3' adenine overhangs and cloned into the Invitrogen pCR2.1 cloning vector (Carlsbad, CA). The cloned *rrnA* operon in pCR2.1 (p2.1A2) was moved into pUC19 and the temperature sensitive vector pMAK705 on a *Xba*I fragment (7.3 kb) producing pUCA21 and pMAKA14, respectively.

A series of plasmids were generated starting with pUCA21 with the promoter region, the region spanning the *rrsA*-ITS-*rrlA*, the *rrsA* gene, and the *rrlA* gene deleted from the *rrnA* operon. These constructions are depicted and verified by restriction digestion with *Xba*I in Figure 3.1 and summarized in Table 3.1.

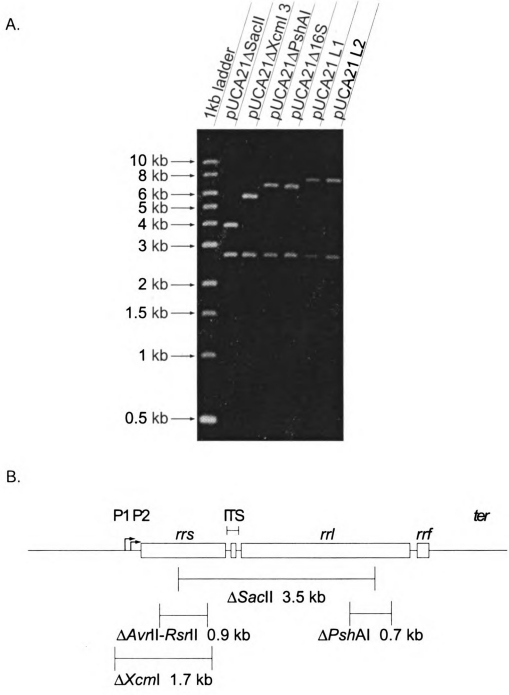


Figure 3.1. *Xba*I digestions of pUCA21 plasmid constructions. *rrnA* was cloned as an *Xba*I fragment into pUC19 giving pUCA21 L1 and L2. pUCA21 constructions were further digested with restrction endonucleases to yield the resulting deletion strains. A. *Xba*I digestions of pUCA21 plasmids. B. Restriction sites used in the creation of the pUCA21 deletion plasmids.

Table 3.1 Plasmids used for strain constructionss

Plasmid	Relevant genotype or phenotype	Source
pUC19	Amp^r, *lacZ'*	Pharmacia
pCR2.1-TOPO	Kan^r, Amp^r, *lacZ'*	Invitrogen
pMAK705	Cam^r, *lacZ'*, rep^{ts}	C.M. Hamilton (21)
pKO3	Cam^r, *sacB*, rep^{ts}	G.M. Church (35)
pCP15	Amp^r, FRT-Kan^r-FRT cassette in pUC19	D. Ammons (2)
pCP20	Flp^+, λ $cI857^+$, λ p_R rep^{ts}, Amp^r, Cam^r	D. Ammons (2)
pKTS102	Cam^r, rep^{ts}, UpA:DnA	UpA:DnA fragment (XbaI, 2,609 bp) cloned in pMAK705
p2.1A2	Kan^r, Amp^r, *rrnA*	*rrnA* cloned in pCR2.1-TOPO
pUCA21	Amp^r, *rrnA*	*rrnA* cloned into pUC19 from p2.1A2 on a *Xba*I fragment
pMAKA14	Cam^r, rep^{ts}, *rrnA*	*rrnA* cloned into pMAK705 from p2.1A2 on a *Xba*I fragment
pUCA21ΔSacII6	Amp^r, *rrnA*ΔSacII	pUCA21 with the SacII fragment removed from *rrnA* (Δ*rrsA*, ITS, *rrlA*)
pUCA21ΔXcmI3	Amp^r, *rrnA*ΔXcmI	pUCA21 with the XcmI fragment removed from *rrnA* (promoters and upstream FIS sites removed)
pUCA21ΔPshAI1	Amp^r, *rrnA*ΔPshAI	pUCA21 with the PshAI fragment removed from *rrnA* (Δ*rrlA*)
pUCA21ΔAvrII-RsrII	Amp^r, *rrnA*ΔAvrII-RsrII	pUCA21 with the AvrII-RsrII fragment removed from *rrnA* (Δ*rrsA*)

Table 3.2 *E. coli* Strain Constructions

Strain	Relevant genotype and phenotype	Source
TOP10	lac$^-$, *recA*, *endA1*	Invitrogen
REL4548	Ara$^-$	R.E. Lenski (33)
BS4548	Ara$^+$	B.S. Stevenson (49)
BS301 (Δ*rrnAB*)	Ara$^+$, Δ*rrnA*, Δ*rrnB*	B.S. Stevenson (49)
BS300 (Δ*rrnAB*)	Ara$^-$, Δ*rrnA*, Δ*rrnB*	B.S. Stevenson (49)
BS301p11	Ara$^+$, Δ*rrnA*, Δ*rrnB*, Suc^s, Kan^r	P1 transductant of BS301 moving the *sacB-neo* cassette into the *rrnA* deleted region
BS300p11	Ara$^-$, Δ*rrnA*, Δ*rrnB*, *sacB*, Kan^r	P1 transductant of BS301 moving the *sacB-neo* cassette into the *rrnA* deleted region
BS311	Ara$^+$, *rrnA*$^-$, Δ*rrnB*, Suc^s, Kan^r	B.S. Stevenson; intermediate strain for BS301 construction, upA:*sacB-neo*:dnA transduced into *rrnA*
BS310	Ara$^-$, *rrnA*$^-$, Δ*rrnB*, Suc^s, Kan^r	B.S. Stevenson; intermediate strain for BS301 construction, upA:*sacB-neo*:dnA transduced into *rrnA*
BS200 (Δ*rrnB*)	Ara$^-$, Δ*rrnB*	B.S. Stevenson
A+KOBA* (HSBA1)	Ara$^+$, Δ*rrnB*, *rrnA* reinserted	allelic exchange of upA:*sacB-neo*:dnA of BS301p11 with *rrnA* of pMAKA14
A-KOBA* (HSBA2)	Ara$^-$, Δ*rrnB*, *rrnA* reinserted	allelic exchange of upA:*sacB-neo*:dnA of BS300p11 with *rrnA* of pMAKA14
A+JHB122	Ara$^+$, Δ*rrnB*, *rrnA* reinserted	allelic exchange of upA:*sacB-neo*:dnA of BS311 with *rrnA* of pMAKA14
A-JHB122	Ara$^-$, Δ*rrnB*, *rrnA* reinserted	allelic exchange of upA:*sacB-neo*:dnA of BS310 with *rrnA* of pMAKA14
A+JHAB621	Ara$^+$, Δ*rrnB*, Δ*rrnA*	allelic exchange of upA:*sacB-neo*:dnA of BS311 with upA:dnA of pKTS102
A-JHAB311	Ara$^-$, Δ*rrnB*, Δ*rrnA*	allelic exchange of upA:*sacB-neo*:dnA of BS310 with upA:dnA of pKTS102

Version 1 strain construction

The construction of the *rrnA* complementation strains A+KOBA* and A-KOBA* was carried out as previously reported (49) and depicted in Figure 3.2. The upA:*sacB-neo*:dnA region of DAS113, which is comprised of upstream sequence (upA) with homology only to the *rrnA* operon, was transduced by the phage P1*vir* into the strains BS300 (Ara^-; Δ*rrnAB*) and BS301 (Ara^+;Δ*rrnAB*). Transductants were selected by kanamycin resistance, screened for sucrose sensitivity, and verified by PCR amplification of the *rrnA* region producing a specific band indicative of *rrnA* containing the *sacB-neo* cassette. The resulting intermediate strains (BS300p11, BS301p11) were transformed with pMAKA14. Growing the transformants at the non-permissive temperature of 42° C in the presence of chloramphenicol forces cointegrate formation at the *rrnA* region of the chromosome. Growth of cointegrates at the permissive temperature of 30° C favors the excision of the plasmid due to the detrimental effects of maintaining multiple active origins of replication within the chromosome (2, 21). Growing cointegrates at 30° C in the presence of 6% sucrose favors plasmid excision with the removal of the *sacB-neo* cassette leaving the wild-type *rrnA* operon in its normal chromosomal location (8, 21, 49).

Isolates with the proper phenotype ($sucrose^r$, cam^s, kan^s) were verified by PCR amplification of the *rrnA* operon region (Figure 3.3) and Southern hybridization (Figure 3.4). The strains that had undergone the proper exchange are designated Ara^+KOBA* and Ara^-KOBA* (Table 3.2).

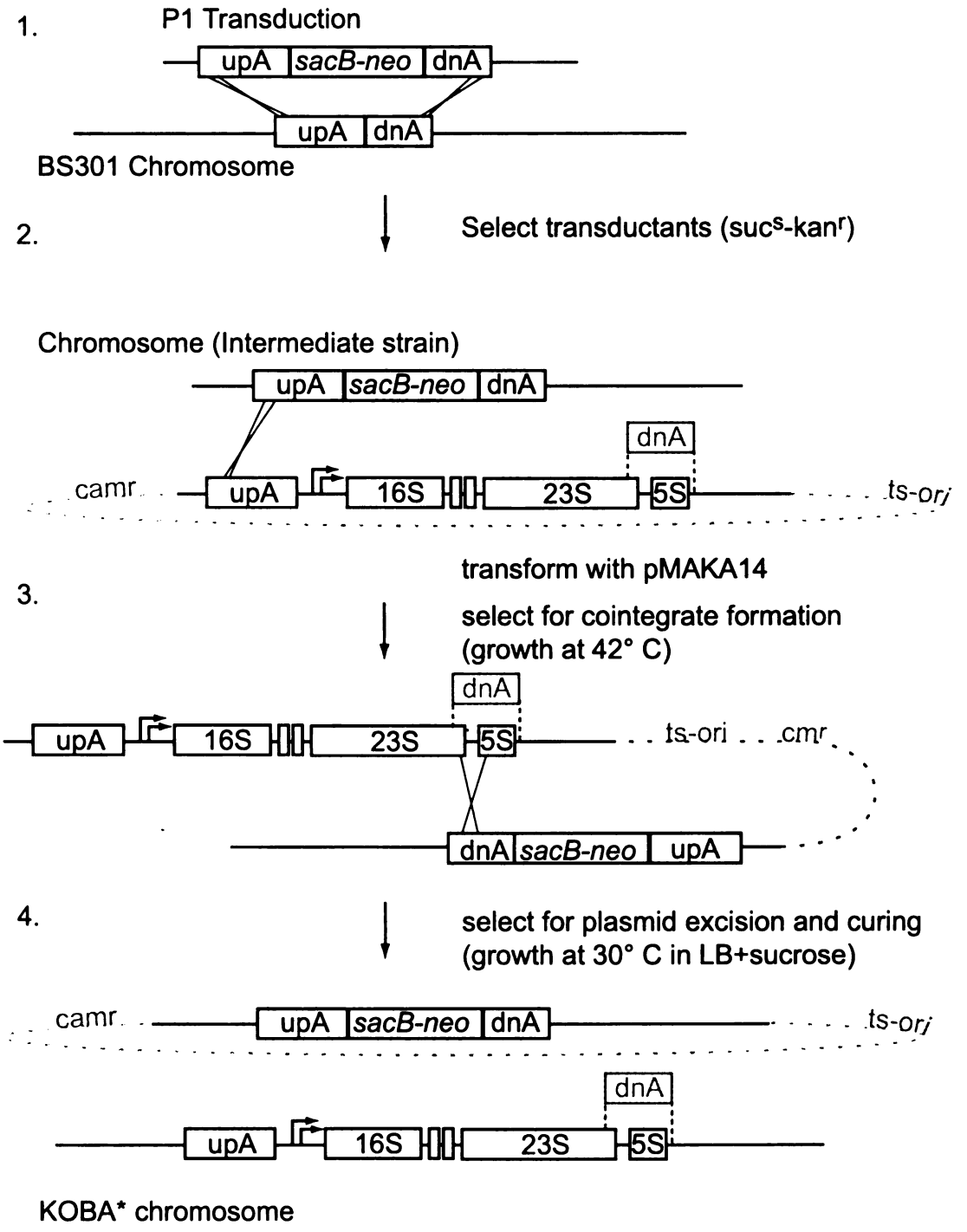


Figure 3.2. Diagram of steps involved in the complementation of BS301 and BS300 creating A+KOBA* and A-KOBA*. The creation of JHB and JHAB strains follow these steps starting at step 3 using plasmids pMAKA14 and pKTS102, respectively.

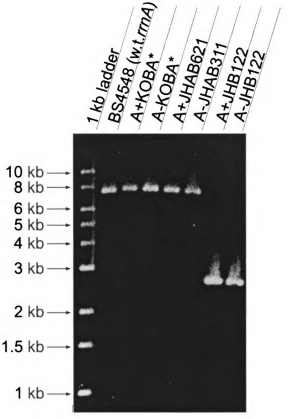


Figure 3.3. PCR amplification of the *rrnA* region of the complementation and deletion strains. The *rrnA* region amplifies as a 7.3 kb fragment as indicated in the BS4548 lane. The deleted *rrnA* region amplifies as a 2.6 kb fragment. All band mobilities for the deletion and complementation strains show the expected size.

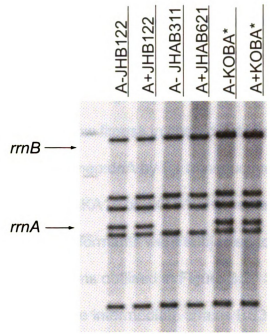


Figure 3.4. Southern hybridization of *Pvu*II digested genomic DNA from *rrnA* deletion and complementation strains probed for a conserved region of *rrs*. All strains have *rrnB* deleted. A+KOBA* and A-KOBA* are the complementation strains starting from BS301 and BS300 with the *rrnA* operon reinserted as outlined in Figure 3.2. A+JHB122, A-JHB122, A+JHAB621, and A-JHAB311 were constructed using the intermediate strains BS311 and BS310 with the Ara+ and Ara- marker, respectively. JHB strains have *rrnA* reinserted into the chromosome. JHAB strains have *rrnA* deleted from the chromosome.

Version 2 strain construction

A second set of strain constructions with *rrnA* deleted and complemented started with strains BS311 and BS310. Strains BS311 and BS310 are the intermediate strains that were originally used to generate the deletion of *rrnA* which had the unexpected low fitness in chemostat competitions (49), and have *rrnA* replaced by upA:*sacB-neo*:dnA by P1 transduction. Both strains were transformed with either pMAKA14 or pKTS102 to replace or delete the *rrnA* operon, respectively. Transformants were subjected to the same regime outlined for the Version 1 constructions outlined in Figure 3.2. Strains with both *rrnA* and *rrnB* deleted starting from the intermediate strains BS311 and BS310 are named A+JHAB621 and A-JHAB31, respectively. Strains with *rrnB* deleted but with a normal intact *rrnA* operon starting from intermediate strains BS311 and BS310 are named A+JHB122 and A-JHB122, respectively (Table 3.2).

Chemostat competitions

Relative fitness measurements under slow growth, continuous culture were carried out in eight parallel chemostat vessels as previously described (49). Relative fitness was measured as the change in the ratio of the two competitors started at a 1:1 ratio. The two competitors are differentiable by a neutral marker (arabinose utilization) where Ara^+ colonies appear white and Ara^- colonies appear red on TA agar (33).

Batch culture competitions

Separate overnight cultures of the two competitors were inoculated into LB from 20% glycerol freezer-stocks. Conditioning cultures, consisting of 10ml of DM+25 in a 50 ml Erlenmeyer flask, were inoculated the following day with a 1:10,000 total dilution from the overnight culture. After 24 hours shaking at 37° C, the two conditioning cultures were mixed in a 1:1 ratio and used to inoculate replicate 50ml flasks containing 10ml of DM+25 at a final dilution of 1:100. Batch culture competitions were carried out in 50ml Erlenmeyer flasks containing 10ml of DM+25 and were allowed to continue for 3-5 days with a sampling and transfer every 24 hours. Competitors were differentiated by the ability or inability to utilize arabinose on TA plates (33, 49).

3.3 Results

Version 1 competitions

The relative fitness of A+KOBA* (HSBA1) versus BS300 (Δ*rrnAB*) in a slow dilution chemostat showed that the complemented strain A+KOBA* has a small fitness advantage over BS300 which has both *rrnA* and *rrnB* deleted as evidenced by the positive slope of the KOBA* vs. KOAB line in Figure 3.5. The chemostat competition between A+KOBA* and BS200 (Δ*rrnB*), however, showed that the presence of an intact *rrnA* operon in A+KOBA* does not rescue the large fitness loss associated with the deletion of *rrnA*. Further analysis shows that the relative fitness of A+KOBA* is nearly identical to that of BS301 (KOAB) versus BS200 (KOB)(Figure 3.5).

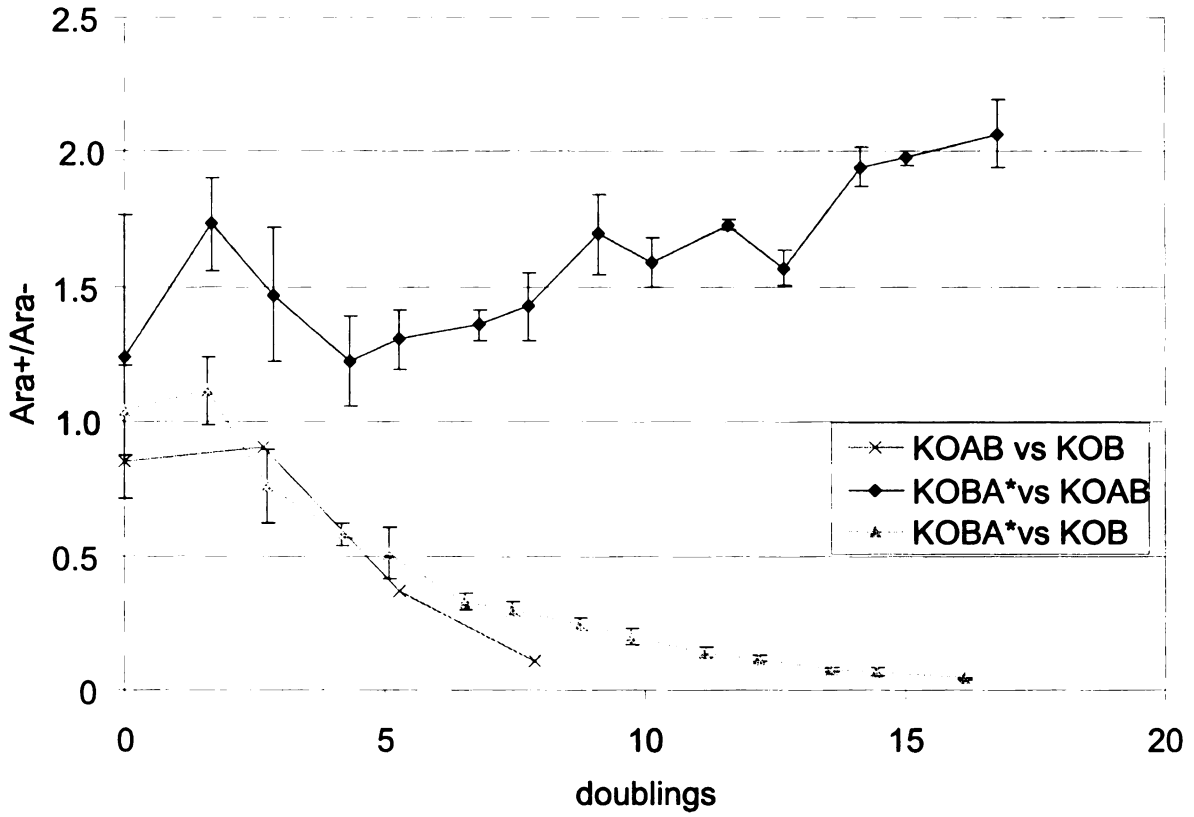


Figure 3.5. Chemostat competitions of complementation strain A+KOBA*. The A+KOBA* vs. A-KOAB competition suggests the presence of *rrnA* provides a small fitness advantage. The A+KOBA* vs. A-KOB competition shows the fitness loss seen in the KOAB strain (BS301, BS300) is not complemented by the reinsertion of *rrnA*. The KOAB vs. KOB competition represents data from reference (49) and is included to show the similarity to A+KOBA* vs. KOB competition. All chemostat competitions were run in four parallel chemostat vessels with each time point representing the mean ratio of Ara+ to Ara- with the error bars representing plus and minus one stadard deviation. The chemostats were run at an average dilution rate of 0.11 hr^{-1}.

Batch culture competition results are summarized in Figure 3.6. The relative fitness of the complemented strain A-KOBA* (HSBA2) versus the wild type REL4548 strain is significantly lower than the strain BS300 which has both *rrnA* and *rrnB* deleted. These results along with the chemostat competition results suggests that the large fitness loss associated with the deletion of the *rrnA* operon in slow growth chemostats is not directly attributable to the presence or absence of the *rrnA* operon.

Version 2 batch culture competitions

The results of the batch culture competitions involving the second version of deletion and complementation strains show no clear pattern associated with the presence or absence of the *rrnA* operon (Figure 3.7). The results show a large difference between the Ara+ and Ara- strains in control competitions with the same number of rRNA operons deleted. Surprisingly, the Ara+ strains were always more than 7% more fit than the Ara- strains regardless of the number of intact rRNA operons present. Due to the large amount of variability between strains with the same number of rRNA operons, chemostat competitions were not conducted for the JHB and JHAB strains.

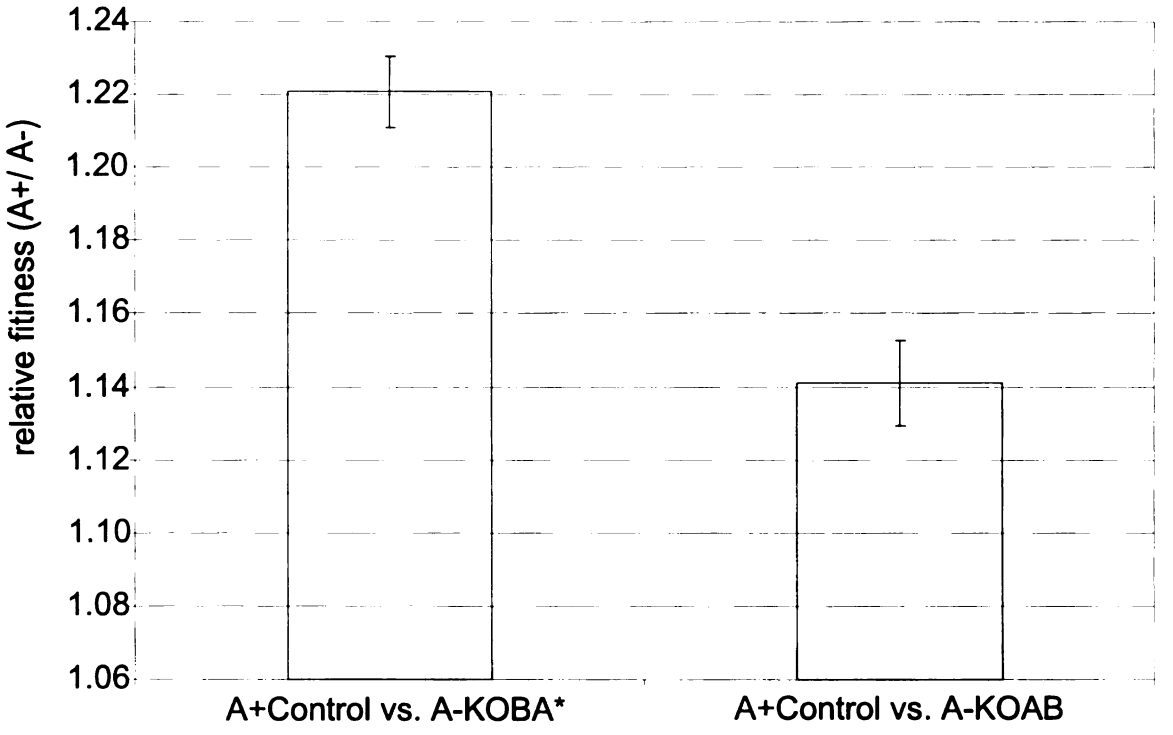


Figure 3.6. Relative fitness of A-KOBA* and A-KOAB strains. A+Control is the parental strain (BS4548). Error bars represent +/- one standard deviation. The A-KOBA* competition clearly indicates that the reinsertion of *rrnA* does not improve fitness.

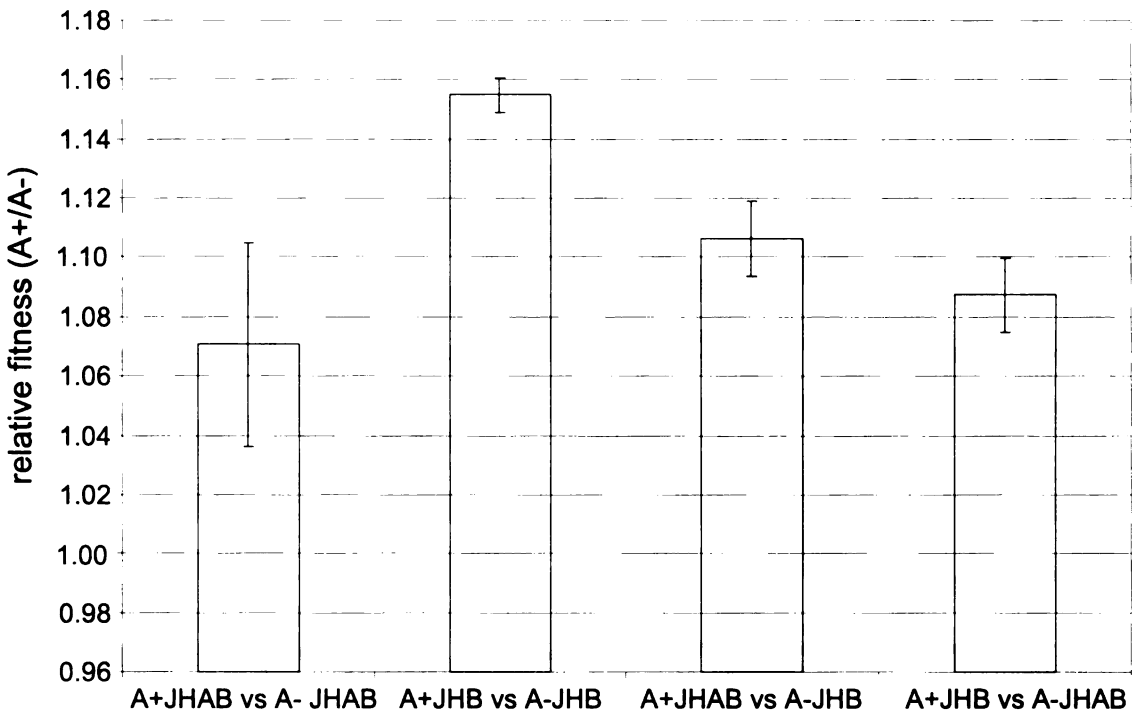


Figure 3.7. Relative fitness in batch culture of JHAB and JHB strains. Error bars represent +/- standard deviation. The large difference between the control competitions (A+JHAB vs. A-JHAB and A+JHB vs. A-JHB) indicate that a large amount of variability is introduced during the creation of the strains.

3.4 Discussion

Lack of complementation in *rrnA* replacement strains

The chemostat competitions with A+KOBA* (HSBA1) clearly show that the large fitness loss in seen for BS301 and BS300 in the slow dilution chemostat is not a direct result of the deletion of the *rrnA* operon (Figure 3.5). The relative fitness of A+KOBA* (Δ*rrnAB* with *rrnA* reinserted) is only slightly better than BS300 (Δ*rrnAB*) which, if the large fitness loss seen in the strains with *rrnA* deleted (BS300, BS301) were a result of the deletion, should have a relative fitness near wild-type levels. The chemostat competition between A+KOBA* (HSBA1) and BS200 (Δ*rrnB*) supports the observation that inserting the *rrnA* operon into BS301 at the original chromosomal location does not rescue the fitness loss.

Batch culture competitions with A-KOBA* (HSBA2) showed that reinserting the *rrnA* operon in BS300 caused a decrease in relative fitness (Figure 3.6). This is contrary to what was expected and indicates that the fitness loss in strains bearing deletions of *rrnB* and *rrnA* are not specifically attributable to the loss of rRNA operons (49). The chemostat and batch culture competition results indicate that something occurred during the construction of these strains that had a large impact on the fitness of the strains and the fitness effects associated with the deletion of rRNA operons is lost in the background variability.

Rational for constructing *rrnA* deletions and replacements from the same transductant

The lack of complementation seen in the KOBA* competitions shows that the large fitness loss in the *rrnA* deletion strains BS301 and BS300 in chemostat competitions is attributable to something other than the presence or absence of the *rrnA* operon. We proposed that the transduction event required to move the *sacB-neo* cassette inserted in *rrnA* from DAS113 (*recD*$^-$) into the well characterized REL4548 and BS4548 (Figure 3.2) strains was the likely cause of the large fitness loss. Generalized transduction using the bacteriophage P1 mobilizes ~80kb from the donor to the recipient (39). For most genetic manipulations this should not be a problem because the resulting strains are not going to be tested for the small fitness changes that are resolved by head-to-head competitions. However, it is apparent that the fitness effects of rRNA operon deletion will be small and the transfer of such large fragments of DNA by P1 transduction may have a large unanticipated effect on fitness. The results can also be interpreted as each P1 transduction decreasing fitness regardless of what is being deleted or complemented.

With this in mind, it was proposed that the variability introduced by the generalized transduction could be overcome by creating both deletion and complementation strains starting from the intermediate strains BS311 and BS310, which contain the *sacB-neo* cassette already transduced into the *rrnA* region, that were used to create the original deletions BS300 and BS301 (Table 3.2). The JHB and JHAB strains were constructed in this fashion creating the

deletion and complementation strains in parallel allelic exchange experiments starting with the same transductant. The effects of deleting *rrnA* was expected to be small, but by starting with the same intermediate strain all of the fitness differences was expected to be attributable specifically to the presence or absence of *rrnA*.

The JHB and JHAB batch culture competition results were rather confounding and not much information regarding the fitness effects of deleting *rrnA* could be gleaned from them (Figure 3.7). The most obvious difference seen in the strains is that the Ara+ strains derived from BS311 have a higher relative fitness regardless of the presence or absence of the *rrnA* operon. Another striking observation is the difference in relative fitness of Ara+ vs. Ara- between the two control competitions (A+JHAB vs. A-JHAB and A+JHB vs. A-JHB). The amount of difference between the control competitions suggests that a change of around 7% is introduced just by taking the strains through the allelic exchange process. These results suggests that the steps involved in the allelic exchange, which includes several population bottlenecks, introduces enough variation in fitness to mask any specific fitness effects of rRNA operon deletion.

Future directions

Clearly, the results from the complementation analysis indicates that the approach to rRNA operon deletion must be rethought. If at all possible, P1 transduction should be avoided and the number of steps required to go from start to finish should be limited. One approach that seems to fit both of the above qualifications is currently being worked out in the lab. The new approach uses a

site specific recombinase (Flp) to remove a kanamycin resistance cassette flanked by FRT sites (Flp recognition sequences) from a deleted region of the chromosome (Figure 3.8).

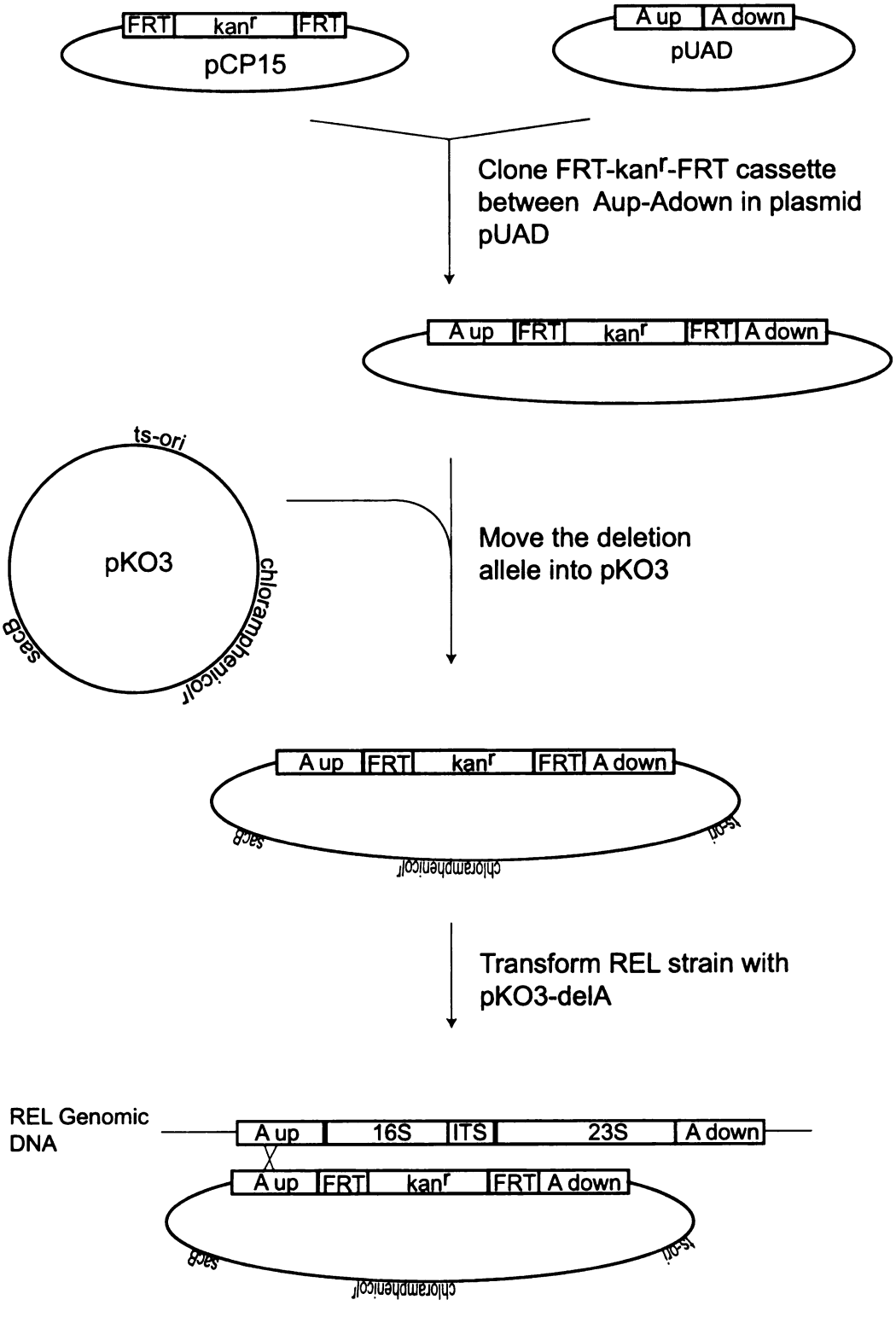
As I stated before, the alternative protocol , developed by Ammons *et. al* (2), does not use generalized transduction to generate an intermediate strain and limits the number of steps required for the allelic exchange (Appendix 1). The main drawback to this approach is the limited ability to complement the *rrn* deletions. The high degree of homology between the seven rRNA operons of *E. coli* would greatly increase the probability of non-specific integration and the presence of the partial FRT sequence after Flp mediated excision of the antibiotic resistance cassette may disrupt the normal operon function.

Conclusions

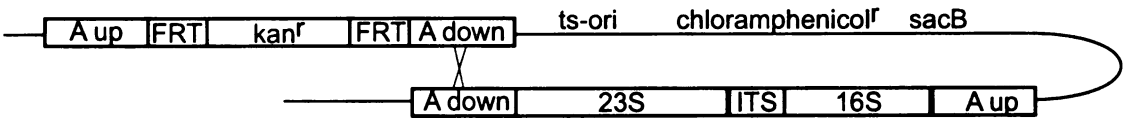
The fitness effects of rRNA operon deletion in *E. coli* are elusive. The glimpse that has been gained through this and other studies have indicated that the fitness effects of deleting one rRNA operon from *E. coli* will be small. Other groups have succeeded in generating strains that have used antibiotic insertions and partial deletions of the rRNA operons but maintained transcriptionally active promoters (5, 11). However, P1 transduction has been used in the creation of these strains which, in addition to the presence of antibiotic resistance genes and active *rrn* promoters, complicates the interpretations. Thus, an accurate picture of the effects of rRNA operon deletion has yet to be shown with respect to the ecological and evolutionary importance of maintaining multiple rRNA operons. The use of the strategy employed by Ammons (2) to generate rRNA operon

deletions should give the cleanest system to elucidate the importance of multiple rRNA operons in *E. coli*. Furthermore, it may be fruitful to explore the fitness effects of rRNA operon deletion in organisms with fewer rRNA operons than the seven which are present in *E. coli.* By deleting an rRNA operon from an organism with fewer rRNA operons, the effect on fitness would presumably be larger due to the smaller range of compensation that would be afforded with a smaller complement of rRNA operons.

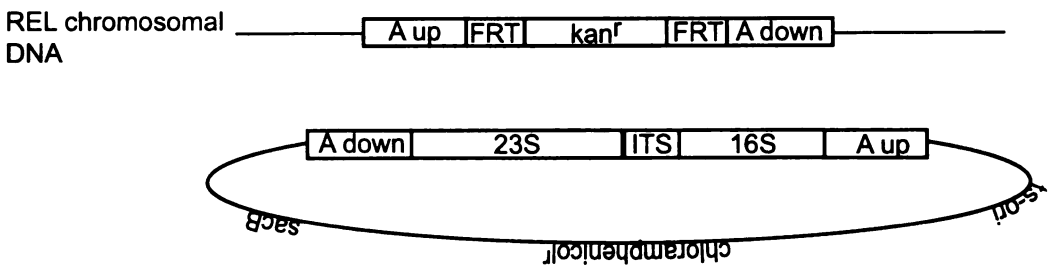
Figure 3.8. Ammons-Fox protocol for generating chromosomal deletions. Plasmid and strain information is summarized in Tables 3.1 and 3.2. The first two steps generate the required plasmids with the FRT-kan^r-FRT cassette inserted between the up/down sequence allowing site-specific integration of the pasmid into the *E. coli* chromosome. The plasmid pKO3 has a temperature sensitive origin of replication (ts-ori) and the *sacB* gene from *Bacillus subtilis* conferring sucrose sensitivity. *E. coli* B strains REL4548 and BS4548 are then transformed with pKO3-delA (or -delB) and cointegrates are isolated by growing transformants at 42°C in the presence of kanamycin and chloramphenicol. Growth of cointegrants at 30°C in the presence of 6% sucrose and kanamycin selects cells with the plasmid excised and lost from the cell due to the lethal effects of the sacB gene in the presence of sucrose. Kanamycin resistant cells are then transformed with pCP20 carrying the Flp recombinase resulting in the site-specific excision of the kan^r gene flanked by FRT sites (Flp recognition sequence). After Flp removal of the kan^r gene, a partial FRT sequence remains at the site of the deletion.


FRT
kanr
FRT
pCP15
A up
A down
pUAD
Clone FRT-kanr-FRT cassette between Aup-Adown in plasmid pUAD
A up
FRT
kanr
FRT
A down
ts-ori
pKO3
chloramphenicolr
sacB
Move the deletion allele into pKO3
A up
FRT
kanr
FRT
A down
sacB
ts-ori
chloramphenicolr
Transform REL strain with pKO3-delA
REL Genomic DNA
A up
16S
ITS
23S
A down
A up
FRT
kanr
FRT
A down
sacB
ts-ori
chloramphenicolr

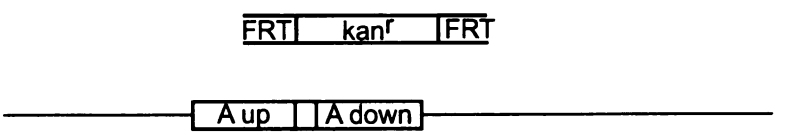
Grow transformants at 42° C with kanamycin and chloramphenicol to force cointegrate formation



Grow cells at 30° C in the presence of sucrose and kanamycin to force plasmid excision and curing and select for the desired excision type (kan^r)



Transform kan^r, sucrose^r cells with pCP20 at 30° C. Cure cells of the plasmid by growing at 42° C



Chapter 4

SUMMARY AND CONCLUSIONS

The initial aim of the research contained in this thesis was to complement strains of *E. coli* that had chromosomal deletions of *rrnA*. The low relative fitness phenotype in slow dilution chemostat seen in the *rrnA* deletion strains was expected to be reversed with the reinsertion of *rrnA* at its normal chromosomal position. The specific region of the *rrnA* operon that was required for the complementation was then to be determined by reinserting portions of the *rrnA* operon (Figure 3.1). However, the wild-type fitness was not rescued using two different strain constructions with a wild-type *rrnA* operon reinserted into the chromosome. It is likely that the use of the generalized transducing phage P1, which transfers ~80kb of DNA (39), was responsible for a portion of the fitness loss associated with the original *rrnA* deletion. Another source of variation in the relative fitness of the strain constructions is the high number of population bottlenecks that are encountered during the deletion and complementation procedure. An alternate procedure for generating chromosomal deletions of the rRNA operons of *E. coli* that does not use P1 transduction and limits the number of population bottlenecks is currently being pursued in the lab. Due to the lack of interpretable results from the complementation studies, I shifted the focus of my thesis research to the related topic of measuring the rate of and fitness effects of recombination between rRNA operons.

Recombination between the rRNA operons of *E. coli* had previously been measured (22, 24, 25, 27) but our system is unique in allowing the easy isolation

of recombinants with no other chromosomal alterations. We found that recombination events involving *rrsB* resulting in the removal of a 3.8 kb *sacB-neo* insertion occurred at a rate of ~$5x10^{-9}$ per cell division. The loss of the *sacB-neo* cassette from *rrnB* by recombination resulted in a 1% increase in fitness compared to the parental strain (BSV1). The increase in fitness of the recombinant isolates may be attributable to the loss of the *sacB-neo* cassette from the cell, regaining the full complement of rRNA operons, or a combination of both possibilities. The lack of any fitness cost associated with recombination also suggests that the sequence heterogeneity is not selectively maintained and alternate explanations for the presence of heterogeneity in the rRNA operons of *E. coli* should be explored.

Our findings regarding the fitness of the recombinant isolates were placed in the context of a conceptual model for the generation and maintenance of sequence heterogeneity in the rRNA operons of *E. coli*. In this model, mutation gives rise to heterogeneity and recombination results in the homogenization of the sequences. However, with such small fitness effects associated with the homogenization of the sequences, the fixation of such events in the population would be very rare. Thus, the relative rates of recombination and mutation and the rate and probability of fixation of the recombinant or mutant in the population will result in operons with either more or less heterogeneity. An additional variable in the model was added based on observations of the degree of sequence heterogeneity in 16S and 23S sequences of published genomes. A positive correlation exists between the number of pair-wise differences of 16S

and 23S sequences and the chromosomal distance between the operons (Figure 2.7). Thus, heterogeneity may also persist if an operon is located further away from other rRNA operons with the closely located operons showing higher degrees of sequence similarity.

The redundancy of the rRNA operons in *E. coli* allows the cell to adjust to the deletion of operons allowing near normal growth rates. This redundancy may also allow mutations within the coding sequences that disrupt the secondary or tertiary structure of the rRNA and persist until compensatory mutations occur. Thus, the rRNA operons provide an intriguing system to broaden our understanding of the physiological implications of multiple rRNA operons, the dynamics and interplay of coordinately evolving gene families, and, potentially, the conditions required to escape the homogenizing effects of recombination.

Literature Cited

1. **Abdulkarim, F., and D. Hughes.** 1996. Homologous recombination between the tuf genes of *Salmonella typhimurium*. J Mol Biol. **260:**506-22.

2. **Ammons, D., J. Rampersad, and G. E. Fox.** 1999. 5S rRNA gene deletions cause an unexpectedly high fitness loss in *Escherichia coli*. Nucleic Acids Res. **27:**637-42.

3. **Anton, A. I., A. J. Martinez-Murcia, and F. Rodriguez-Valera.** 1999. Intraspecific diversity of the 23S rRNA gene and the spacer region downstream in *Escherichia coli*. J Bacteriol. **181:**2703-9.

4. **Anton, A. I., A. J. Martinez-Murcia, and F. Rodriguez-Valera.** 1998. Sequence diversity in the 16S-23S intergenic spacer region (ISR) of the rRNA operons in representatives of the *Escherichia coli* ECOR collection. J Mol Evol. **47:**62-72.

5. **Asai, T., C. Condon, J. Voulgaris, D. Zaporojets, B. Shen, M. Al-Omar, C. Squires, and C. L. Squires.** 1999. Construction and initial characterization of *Escherichia coli* strains with few or no intact chromosomal rRNA operons. J Bacteriol. **181:**3803-9.

6. **Bachellier, S., E. Gilson, M. Hofnung, and C. W. Hill** 1996. Repeated Sequences, p. 2012-2040. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (eds), *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. American Society for Microbiology, Washington, D.C.

7. **Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science. **277:**1453-74.

8. **Blomfield, I. C., V. Vaughn, R. F. Rest, and B. I. Eisenstein.** 1991. Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. Mol Microbiol. **5:**1447-57.

9. **Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. M. Geoghagen, and J. C. Venter.** 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science. **273:**1058-73.

10. **Cilia, V., B. Lafay, and R. Christen.** 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. Mol Biol Evol. **13:**451-61.

11. **Condon, C., S. French, C. Squires, and C. L. Squires.** 1993. Depletion of functional ribosomal RNA operons in *Escherichia coli* causes increased expression of the remaining intact copies. Embo J. **12:**4305-15.

12. **Condon, C., D. Liveris, C. Squires, I. Schwartz, and C. L. Squires.** 1995. rRNA operon multiplicity in *Escherichia coli* and the physiological implications of *rrn* inactivation. J Bacteriol. **177:**4152-6.

13. **Condon, C., J. Philips, Z. Y. Fu, C. Squires, and C. L. Squires.** 1992. Comparison of the expression of the seven ribosomal RNA operons in *Escherichia coli*. Embo J. **11:**4175-85.

14. **Condon, C., C. Squires, and C. L. Squires.** 1995. Control of rRNA transcription in *Escherichia coli*. Microbiol Rev. **59:**623-45.

15. **Crane, G. J., S. M. Thomas, and M. E. Jones.** 1996. A modified Luria-Delbruck fluctuation assay for estimating and comparing mutation rates. Mutat Res. **354:**171-82.

16. **Deckert, G., P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen, and R. V. Swanson.** 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature. **392:**353-8.

17. **Dennis, P. P., S. Ziesche, and S. Mylvaganam.** 1998. Transcription analysis of two disparate rRNA operons in the halophilic archaeon *Haloarcula marismortui*. J Bacteriol. **180:**4804-13.

18. **Dvorak, J., D. Jue, and M. Lassner.** 1987. Homogenization of tandemly repeated nucleotide sequences by distance-dependent nucleotide sequence conversion. Genetics. **116:**487-98.

19. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, and et al.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science. **269:**496-512.

20. **Fraser, C. M., S. J. Norris, G. M. Weinstock, O. White, G. G. Sutton, R. Dodson, M. Gwinn, E. K. Hickey, R. Clayton, K. A. Ketchum, E. Sodergren, J. M. Hardham, M. P. McLeod, S. Salzberg, J. Peterson, H. Khalak, D. Richardson, J. K. Howell, M. Chidambaram, T. Utterback, L. McDonald, P. Artiach, C. Bowman, M. D. Cotton, J. C. Venter, and et al.** 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science. **281:**375-88.

21. **Hamilton, C. M., M. Aldea, B. K. Washburn, P. Babitzke, and S. R. Kushner.** 1989. New method for generating deletions and gene replacements in *Escherichia coli*. J Bacteriol. **171:**4617-22.

22. **Harvey, S., and C. W. Hill.** 1990. Exchange of spacer regions between rRNA operons in *Escherichia coli*. Genetics. **125:**683-90.

23. **Harvey, S., C. W. Hill, C. Squires, and C. L. Squires.** 1988. Loss of the spacer loop sequence from the *rrnB* operon in the *Escherichia coli* K-12 subline that bears the *relA*1 mutation. J Bacteriol. **170:**1235-8.

24. **Hill, C. W., R. H. Grafstrom, B. W. Harnish, and B. S. Hillman.** 1977. Tandem duplications resulting from recombination between ribosomal RNA genes in *Escherichia coli*. J Mol Biol. **116:**407-28.

25. **Hill, C. W., and B. W. Harnish.** 1981. Inversions between ribosomal RNA genes of *Escherichia coli*. Proc Natl Acad Sci U S A. **78:**7069-72.

26. **Hill, C. W., and B. W. Harnish.** 1982. Transposition of a chromosomal segment bounded by redundant rRNA genes into other rRNA genes in *Escherichia coli*. J Bacteriol. **149:**449-57.

27. **Hill, C. W., S. Harvey, and J. A. Gray** 1990. Recombination between rRNA genes in *Escherichia coli* and *Salmonella typhimurium*, p. 335-340. *In* K. Drlica, and M. Riley (eds), The Bacterial Chromosome. American Society of Microbiology, Washington, D.C.

28. **Hughes, D.** 1999. Impact of homologous recombination on genome organization and stability. *In* R. L. Charlebois (ed.), Organization of the Prokaryotic Genome. American Society for Microbiology, Washington, D.C.

29. **Jones, M. E., S. M. Thomas, and A. Rogers.** 1994. Luria-Delbruck fluctuation experiments: design and analysis. Genetics. **136:**1209-16.

30. **Klappenbach, J. A., J. M. Dunbar, and T. M. Schmidt.** 2000. rRNA operon copy number reflects ecological strategies of bacteria. Appl Environ Microbiol. **66:**1328-33.

31. **Koch, A. L.** 1971. The adaptive responses of *Escherichia coli* to a feast and famine existence. Adv Microb Physiol. **6:**147-217.

32. **Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin, and et al.** 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature. **390:**249-56.

33. **Lenski, R. E., M. R. Rose, S. C. Simpson, and S. C. Tadler.** 1991. Long-term experimental evolution in *Escherichia coli*: I. Adaptation and divergence during 2,000 generations. The American Naturalist. **138:**1315-1341.

34. **Liao, D.** 2000. Gene conversion drives within genic sequences: Concerted evolution of ribosomal RNA genes in Bacteria and Archaea. J Mol Evol. **51:**305-317.

35. **Link, A. J., D. Phillips, and G. M. Church.** 1997. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. J Bacteriol. **179:**6228-37.

36. **Lloyd, R. G., and K. B. Low** 1996. Homologous Recombination, p. 2236-2255. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (eds), *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. American Society for Microbiology, Washington, D.C.

37. **Lovett, S. T., T. J. Gluckman, P. J. Simon, V. A. Sutera, Jr., and P. T. Drapkin.** 1994. Recombination between repeats in *Escherichia coli* by a *recA*-independent, proximity-sensitive mechanism. Mol Gen Genet. **245:**294-300.

38. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics. **28:**491-511.

39. **Milkman, R.** 1999. Gene transfer in *Escherichia coli*, p. 291-309. *In* R. L. Charlebois (ed.), Organization of the Prokaryotic Chromosome. American Society for Microbiology, Washington, D.C.

40. **Mitchell, M. B.** 1955. Aberrant recombination of pyridooxine mutants of *Neurospora*. Proc. Natl Acad. Sci. USA. **41:**215-220.

41. **Mitchell, M. B.** 1955. Further evidence of aberrant recombination in *Neurospora*. Proc. Natl Acad. Sci. USA. **41:**935-937.

42. **Mylvaganam, S., and P. P. Dennis.** 1992. Sequence heterogeneity between the two genes encoding 16S rRNA from the halophilic archaebacterium *Haloarcula marismortui*. Genetics. **130:**399-410.

43. **Nakamura, Y., T. Kaneko, M. Hirosawa, N. Miyajima, and S. Tabata.** 1998. CyanoBase, a www database containing the complete nucleotide sequence of the genome of *Synechocystis sp.* strain PCC6803. Nucleic Acids Res. **26:**63-7.

44. **Parkhill, J., B. W. Wren, K. Mungall, J. M. Ketley, C. Churcher, D. Basham, T. Chillingworth, R. M. Davies, T. Feltwell, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Moule, M. J. Pallen, C. W. Penn, M. A. Quail, M. A. Rajandream, K. M. Rutherford, A. H. van Vliet, S. Whitehead, and B. G. Barrell.** 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature. **403:**665-8.

45. **Segall, A. M., and J. R. Roth.** 1994. Approaches to half-tetrad analysis in bacteria: recombination between repeated, inverse-order chromosomal sequences. Genetics. **136:**27-39.

46. **Silvestri, M. L., W. J. Siqueira, A. A. de Souza, A. P. de Souza, M. F. Terenzi, D. Truffi, S. M. Tsai, M. H. Tsuhako, H. Vallada, M. A. Van Sluys, S. Verjovski-Almeida, A. L. Vettore, M. A. Zago, M. Zatz, J. Meidanis, and J. C. Setubai.** 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature. **406:**151-7.

47. **Singer, B. S.** 1988. On the role of homologous sequences in chromosomal rearrangements. Genes Dev. **2:**1800-11.

48. **Smith, D. R., L. A. Doucette-Stamm, C. Deloughery, H. Lee, J. Dubois, T. Aldredge, R. Bashirzadeh, D. Blakely, R. Cook, K. Gilbert, D. Harrison, L. Hoang, P. Keagle, W. Lumm, B. Pothier, D. Qiu, R. Spadafora, R. Vicaire, Y. Wang, J. Wierzbowski, R. Gibson, N. Jiwani, A. Caruso, D. Bush, J. N. Reeve, and et al.** 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. J Bacteriol. **179:**7135-55.

49. **Stevenson, B. S.** 2000. Doctoral dissertation. Life history implications of ribosomal RNA gene copy number in *Escherichia coli*. Michigan State University.

50. **Stevenson, B. S., and T. M. Schmidt.** 1998. Growth rate-dependent accumulation of RNA from plasmid-borne rRNA operons in *Escherichia coli*. J Bacteriol. **180:**1970-2.

51. **Tettelin, H., N. J. Saunders, J. Heidelberg, A. C. Jeffries, K. E. Nelson, J. A. Eisen, K. A. Ketchum, D. W. Hood, J. F. Peden, R. J. Dodson, W. C. Nelson, M. L. Gwinn, R. DeBoy, J. D. Peterson, E. K. Hickey, D. H. Haft, S. L. Salzberg, O. White, R. D. Fleischmann, B. A. Dougherty, T. Mason, A. Ciecko, D. S. Parksey, E. Blair, H. Cittone, E. B. Clark, M. D. Cotton, T. R. Utterback, H. Khouri, H. Qin, J. Vamathevan, J. Gill, V. Scarlato, V. Masignani, M. Pizza, G. Grandi, L. Sun, H. O. Smith, C. M. Fraser, E. R. Moxon, R. Rappuoli, and J. C. Venter.** 2000. Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. Science. **287:**1809-15.

52. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzegerald, N. Lee, M. D. Adams, J. C. Venter, and et al.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature. **388:**539-47.

53. **Wang, Y., Z. Zhang, and N. Ramanan.** 1997. The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. J Bacteriol. **179:**3270-6.

54. **Yap, W. H., Z. Zhang, and Y. Wang.** 1999. Distinct types of rRNA operons exist in the genome of the Actinomycete *Thermomonospora chromogena* and evidence for horizontal transfer of an entire rRNA operon. J Bacteriol. **181:**5201-5209.