THE UNIVERSITY OF CALGARY

The Conservation of Genome Structure in *Salmonella typhi*

by

Rajrupa Sonali Majumdar

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN

PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES

CALGARY, ALBERTA

AUGUST, 1999

Canada

ABSTRACT

The genome of *Salmonella typhi* rearranges due to recombination between *rrn* operons; other enteric species like *Escherichia coli* and *Salmonella typhimurium* are more conservative. This variability in *S. typhi* often leads to unbalanced genomes (origin and terminus sites are no longer 180° separated). It appears that variability in *S. typhi* is tolerated, but not because of compensatory gain or loss of DNA or other recombinational events, which restore balance. Endonuclease digestion, pulsed-field gel electrophoresis and ^{32}P end-labeling showed that genome size in 28 *S. typhi* strains varies little (from 4500-4850 kb). Base pair changes are rare but genomic rearrangements occur within certain regions of the I-*Ceu*I A fragment in *S. typhi*, like the terminus region; however, these rearrangements encompass only small segments of DNA. Some *S. typhi* strains with unbalanced genomes have a significant reduction in growth rate (44%). Rearrangements, which rebalance the genome, restore the growth rate to near normal.

ACKNOWLEDGEMENTS

DEDICATION

To my parents, Amitabha and Gitali,

for their kindness, understanding and encouragement.

Also my loving friends, Farzana and Rupa,

for their constant support.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

α-^{32}P dCTP	Alpha-^{32}Phosphorus labeled deoxy-cytosine triphophate
BM	Boerhinger Manheim
BSA	Bovine serum albumin
dNTPs	Deoxy-nucleotide triphosphate
DTE	Dithioerythritol
DDT	Dithiothreitol
EtBr	Ethidium Bromide
EDTA	Ethylene-diaminetetra-acetate
g	Gram
Kb	Kilobase
LB	Luria-Bertani broth/ medium
μCi	Microcurie
μg	Microgram
μl	Microlitre
mg	Milligram
mm	Millimetre
ml	Millilitre
mM	Millimole
NEB	New England Biolabs
OD	Optical density
*ori*C	Origin of Replication

PMSF	Phenyl-Methyl-Sulfonyl Fluoride
PFGE	Pulsed Field Gel Electrophoresis
rpm	Revolution per minute
SGSC	*Salmonella* Genetic Stock Center
SARB	*Salmonella* Reference Collection B
s	Seconds
SSC	Sodium chloride and Sodium citrate
SDS	Sodium dodecyl sulfate
ter	Terminus
TBE	Tris-HCl-Boric acid-EDTA
V	Volt
V/cm	Volt per centimetre

CHAPTER 1: INTRODUCTION

The genome of most enteric bacteria is highly conserved. So much so, that even though *Escherichia coli* K-12 and *Salmonella typhimurium* LT2 had diverged over 100 million years ago, their genetic and physical maps are strikingly similar (Liu and Sanderson, 1995c; Weinstock and Lupski, 1997; Shimkets, 1997; Ochman and Wilson, 1987). This is especially surprising since in culture, rearrangements within the genome are common and occur at frequencies as high as 10^{-5} (Andersson and Roth, 1978). Such remarkable conservation may be a result of strong selective pressures that selectively remove rearranged genotypes.

One possible influence on gene conservation is genome balance, which refers to the symmetry between the origin of replication (*ori*C) and terminus sites (*ter*) on a circular bacterial chromosome that results from a 180° separation between the two regions (Sanderson et al., 1999). The *genome balance theory* states that such genomes are a requirement for optimal cell growth and survival. Unbalanced genomes, which are often a consequence of large insertion or inversion events in the chromosome, have their *ori*C site or *ter* sites displaced in such a way that a 180° symmetry no longer exists. Bacterial cells with unbalanced genomes have reduced growth rates and usually do not survive in nature (Hill et al., 1990; Hill and Gray, 1988). Therefore, chromosomal rearrangement events, which lead to changes in gene order may also give rise to changes in genome balance. Such imbalances in the genome often have negative effects on cell fitness and are usually selected against in nature. Thus, the strong need for cells to ensure

maximal fitness, through the maintenance of a balanced genome, probably plays a role in discouraging certain rearrangement events, and acts to conserve gene order in bacteria.

Surprisingly, unlike other *Salmonella* spp. and *Escherichia coli*, the gene order of *S. typhi* is not conserved. It appears that forces, which act to conserve the genome of other enteric bacteria, sometimes bypass *S. typhi*. Most of the variations seen in *S. typhi* are a result of inversions and translocations involving *rrn* operons (Sanderson et al., 1999; Liu and Sanderson, 1995a; Liu and Sanderson, 1996). In fact, due to these rearrangement events, the serovar of *S. typhi* is known to have 27 different genome orders, referred to as genome types. The entire *S. typhi* chromosome has not been studied in detail in regards to other rearrangements not mediated via *rrn* operons. Studies on one half of the chromosome indicate that these rearrangements are rare in *S. typhi* strains (Liu and Sanderson, 1995a). However, it is known that in other enteric bacteria like *E. coli*, *Salmonella typhimurium*, *Salmonella paratyphi* C and one well studied representative of *S. typhi*, *S. typhi* Ty2, rearrangements around the terminus region are common. The terminus region in *S. typhi* strains is found within the half of the chromosome of which not much is known.

In most of the 27 genome types, the location of the *ori* C site, found within the I-*Ceu* I C fragment, remains more or less unchanged; however, in some cases, for example genome types 25,26 and 27, the *ori* C site is grossly displaced, suggesting that balance may also be affected. There are two possible theories as to why the genome *of S. typhi* is so diverse. First, it could be that the genome balance theory does not apply to *S. typhi* as stringently as it does to other bacteria. For this reason, some strains of *S. typhi* are able

to survive in nature despite their unbalanced genomes. The second explanation pertains to the Adopt-Adapt Model of bacterial speciation. In this model, it is hypothesized that during the process of speciation, the bacteria often adopts foreign DNA, which may, as a consequence, disrupt its genome balance. To compensate for this, the bacteria undergo several changes via rearrangements, deletions or insertions in attempt to restore genome balance and maximize cell fitness. *S. typhi* is a fairly recent species and since humans are its only host, it is speculated that *S. typhi* can only be as old as the human race, around 2-3 millions years. All *S. typhi* strains have adopted 120 kb of foreign DNA which has also disrupted its genome balance (Liu and Sanderson, 1995c). Genomic reorganizations might be a mechanism to restore the 180° symmetry, between the origin and terminus, and may be what gives rise to the different genome types in *S. typhi*.

Recently, two opposing observations have been made about size conservation in *S. typhi*. Some researchers have claimed that the genome size of *S. typhi* strains is not conserved, with sizes ranging from 3980 to 4780 kb (Pang, 1998). However, in other enteric bacteria, this variation is normally only from 4600 to 4800 kb (Cole and Saint-Girons, 1999), though sizes up to 5400 kb have been reported for *E. coli* 0157 (Blattner, F., Personal communication.) Furthermore, the genome size of most enteric bacteria center around 4600-4700 kb; therefore, it is surprising that in *S. typhi* some strains are claimed to be as small as 3980 kb. This is especially surprising because *S. typhi* is a pathogenic organism and one would expect an increase, not a decrease, in genomic content since additional genes may be required for infection. However, preliminary studies on 144 wild-type *S. typhi* strains have shown that size is not significantly affected

in at least one half of the *S. typhi* chromosome; this implies that size is, in fact, conserved in *S. typhi* (Liu and Sanderson, 1995a).

S. typhi causes typhoid fever which is a serious health problem in many parts of the world (Thong et al., 1994). *S. typhi* infects humans exclusively (Edelman and Levine, 1986). Therefore, using animal models to investigate this organism is not feasible and most of the information known about this pathogen is from molecular evidence. Due to the uniqueness and diversity of the *S. typhi* genome, there is great interest in studying this organism at the genetic level.

CHAPTER 2: LITERATURE REVIEW

A. Genome conservation

1. Introduction

The genome of most enteric bacteria is conserved. In fact, that even though *Escherichia coli* K-12 and *Salmonella typhimurium* LT2 had diverged over 100 million years ago, their genetic and physical maps are near identical (Liu and Sanderson, 1995c; Shimkets, 1997; Weinstock and Lupski, 1997; Ochman and Wilson, 1987). This is surprising since over during evolution, there is on average a 15% divergence at the base pair level between homologous genes retaining identical gene location. In fact, comparison of 67 homologous protein-coding loci from *Escherichia coli* and *Salmonella typhimurium* reveals that individual genes varied up to 28% in sequence (Sharp, 1991). Furthermore, it has been observed that in culture, rearrangements like duplications, deletions and inversions, within the genome of enteric bacterial cells are common (Andersson and Roth, 1978). Despite all these mutational events and opportunities for change, the genomes of *E. coli* and *S. typhimurium*, along with *Salmonella enteritidis* and *Salmonella paratyphi* B remain stable (Liu and Sanderson, 1995a). This suggests that strong selective pressures must exist to maintain overall genome order within these enteric bacteria.

2. Theories on genome conservation

Several theories exist which attempt to explain the remarkable conservation in gene order seen within the enteric bacteria. Each theory introduces the notion of a selective force acting against genomic change. Moreover, it is suggested that the combination of all these forces helps to maintain gene order in bacteria.

a) Gene dosage theory-

The gene dosage theory deals with the origin of replication site *(oriC)* and how its position on the chromosome affects overall gene dosage. Bi-directional replication in bacteria initiates at the *oriC* site. This highly conserved region consists of several protein-binding sites, AT-rich direct repeats and GATC sites (Kogoma, 1997). Recognition of these sites by an ATP-bound DnaA protein is essential for the replication initiation process (Zyskind, 1990). Multiple replication forks originate from the *oriC* site and proceed towards the terminus region (*ter* site) at the opposite end of the chromosome. Unlike the *oriC* site, the terminus region is large and it encompasses about 7% of the *E. coli* chromosome. Replication inhibition takes place at the edges of the *ter* region and is mediated by the T1 and T2 sites. T1 and T2 consort together in a polar fashion to trap incoming replication forks. These proteins allow the forks to enter the *ter* region but prevent them from leaving (Kuempel et al., 1990). As the replication fork travels along the chromosome from the *oriC* site, genes are copied. The *gene dosage theory* states that genes located close to the *oriC* site are replicated more often and are present in higher copy numbers than those genes located close to the *ter* sites. This differential gene dosage results in a gradient of gene expression and may be dictated by the location of a

particular gene on the chromosome (Schmid and Roth, 1987). This observation has been confirmed in *S. typhimurium*, where transposition of *his* genes into several different sites on the bacterial chromosome resulted in a corresponding variation in the level of HisD (histidine dehydrogenase) production (Schmid and Roth, 1987). Thus, it is believed that the position of a gene, in relation to the *oriC* site, may be linked to its optimal expression which should be in accordance with the cell's need. Therefore, rearrangements in gene position resulting in altered gene expression, may be selected against in nature depending on how inappropriate these alterations are and how detrimental they are to the cell's survival. This type of negative selection is especially pronounced when dealing with relocation of essential genes on the chromosome. For example, rearrangements displacing a known segment of the *Salmonella typhi* genome, which harbour essential housekeeping genes, by more than 10% from either side of the *oriC* site is rarely seen (Liu et al., 1999). This, in accordance with the gene dosage theory, suggests that the above deleterious changes in the location of the I-*CeuI* F fragment result in incorrect gene expression that is disadvantageous to the cell's survival and are selected against in evolution (Liu et al., 1999). Therefore, the necessity for bacterial cells to maintain proper gene dosage may act as a force, which selects against changes in their chromosomal gene order.

b) Genome balance theory-

The genome balance theory refers to the symmetry between the *oriC* and *ter* sites on a circular bacterial chromosome, which results from a 180° separation of the two regions (Sanderson et al., 1999). The *genome balance theory* states that such genomes

are a requirement for optimal cell growth and survival. On the other hand, unbalanced genomes, which are a consequence of large insertion or inversion events in the chromosome, have their *oriC* site or *ter* sites displaced in such a way that a 180° symmetry no longer exists. Bacterial cells with unbalanced genomes usually have reduced growth rates, which prevent them from surviving in nature. Furthermore, these mutants are readily out-competed by their balanced counterparts (Sanderson et al., 1999; Liu et al., 1999; Hill et al., 1990; Hill and Gray, 1988). It is the degree to which the *oriC* or the *ter* sites are displaced from their original position, which determine how severely growth rate is affected. This has been experimentally observed using two *E.coli* K12 mutants, each harboring an inversion in their chromosome, and causing displacement of the *oriC* site to a certain extent. The first of these mutants had its *oriC* site displaced by 6% from the original location, due to an inversion event between *rrn* operons. Since this mutant has been present and undetected in *E. coli* sub-lines for over 30 years, it can be inferred that the effect of this 6% *oriC* displacement on cell growth is minimal (Hill and Harnish, 1981; Hill et al., 1990; Hill and Gray, 1988). The second mutant was experimentally derived and had its *oriC* displaced by 19%; this was again a result of an inversion event in the chromosome. This mutant had more obvious effects on cell growth in that it failed to grow on nutrient medium and showed limited growth on minimal medium (Hill and Gray, 1988). This suggested that these cells would be able to maintain themselves in a slow growing environment. As a result, these inverted mutants would probably be selected against quite rapidly in the course of evolution. To confirm this, extended growth analysis was done on the mutants which had their *oriC* site displaced by

19% and as predicted, the results showed that inverted mutants reverted back to their wild-type form, in which genome balance was reestablished, at a high frequency (Hill and Gray, 1988). There are at least two possible reasons as to why the above results were obtained. Firstly, gene dosage most likely plays a role in determining that an unbalanced genome has lower growth and survival rates. Displacement of the *oriC* site on a chromosome also results in the displacement of genes on the chromosome and could, in turn, disrupt the balance in gene products. Such changes could subsequently be detrimental to the cell (Hill and Harnish, 1981; Hill and Harnish, 1981). Secondly, balanced genomes have equal length chromosomal arms which initiate replication forks at the *oriC* site that move bi-directionally towards the terminus region, at a constant velocity. However, genomes which are unbalanced have unequal length chromosomal arms, due to the displacement of *oriC* or *ter* sites. As a result, their replication forks may observe differential velocities in certain regions of the chromosome, which can lead to incorrect or slow replication of the genome (Francois et al., 1990). This may also be a reason as to why cells with unbalanced genomes have such reduced growth rates. In summary, chromosomal rearrangement events, which lead to changes in gene order may also give rise to changes in genome balance. Such imbalances in the genome often have negative effects on cell fitness and are usually selected against in nature. Thus, the strong need for cells to ensure maximal fitness, through the maintenance of a balanced genome, probably plays a role in discouraging certain rearrangement events, and acts to conserve gene order in bacteria.

c) Chi sequences-

Chi sequences are octameric sequences, which are found, on the average, every 5 kb along the *E.coli* DNA. Surprisingly, chi sequences in *E. coli* are found in 60-fold excess to the frequency expected on the basis of base composition (Lloyd and Low, 1996; Masters, 1996; Sanderson et al., 1999). These sequences stimulate the action of RecBCD, which is a potent helicase and exonuclease. This enzyme complex encounters the chi sequence from the 3' end and acts on nicked double stranded DNA. As the enzyme moves along the chi sequence, it unwinds and displaces the DNA in this region to initiate recombination by invading other homologous DNA (Lloyd and Low, 1996). Thus, chi sequences function to stimulate recombination within the bacterial chromosome. What is unique about these sequences is that their orientation is nonrandom; 80% of all chi sequences in *E. coli* are oriented such that the RecBCD complex will bind and move towards the *oriC* (Sanderson et al., 1999). No reason is known as to why this particular orientation is preferred, however, it does suggest that there may be selective pressure against randomness in orientation of the chi sequences. Furthermore, this selection may also, indirectly, act as a force in conserving gene order in the vicinity of the chi sequences.

d) DNA and RNA polymerase orientation-

To prevent collisions between DNA and RNA polymerases on the bacterial chromosome, the orientation of these enzymes is often coordinated. Prokaryotic DNA and RNA polymerases are highly active enzymes. DNA replication of *E. coli* occurs in just 40 minutes and since both DNA and RNA polymerases move along the bacterial

chromosome concurrently, the chance that they may encounter each other is quite high. This is especially so because the replication fork moves approximately 10 times faster than transcriptional complexes (Brewer, 1988). Both DNA and RNA polymerase cannot occupy the same position on the chromosome and as a result, their functions must be coordinated in an orderly fashion to minimize collision. When transcription orientation is opposite to that of replication orientation, possibilities for head-on collisions arise, leading to stalled replication forks. However, when replication and transcription orientation are in the same direction, head on collisions are less likely to occur, although, at some point the replication fork will catch up to the transcription complex due to the quicker speed of replication. When confronted with such a situation, *E. coli* is speculated to do one of two things. It may bump the RNA polymerase off the DNA template, causing transcription to end prematurely, or DNA replication may slow down and proceed at the some rate as transcription until the RNA polymerase reaches its termination site and leaves the template (Brewer, 1988). A striking observation has been made in the organization of the *E. coli* genome which reveals that highly transcribed genes usually have their transcription orientation in the same direction as the replication orientation, as a mechanism to minimize polymerase collisions (Sanderson et al., 1999; Brewer, 1988). It has been suggested that rearrangements involving highly transcribed genes, which disrupt the orderly organization of replication and transcription orientation, are disadvantageous and therefore, do not exist in nature. For example, *rrn* operons are highly transcribed genes and although inversions between *rrn* operons are common in the *E. coli* and various *Salmonella* species, none which resulted in opposing replication and

transcription orientations are ever retained in wild-type strains (Sanderson et al., 1999; Hill et al., 1990). Thus, one can theorize that the need for maintaining similar orientations may aid in maintaining gene order by eliminating certain recombinational events in the bacterial chromosome. However, this theory is limited, in that it can only provide a mechanism by which highly transcribed genes remain conserved. Unfortunately, it cannot explain how genes, which are not highly active, also maintain gene order through evolution. Thus, this proposed mechanism cannot be the sole reason for how gene order is maintained in enteric bacteria.

e) Operon structure-

Lastly, operon structure also plays a role in the conservation of gene order. Operons are defined to contain a cluster of genes, which is controlled by promotors Thus, genes within an operon must remain close together to maintain proper function and regulation. One would expect that rearrangements, which result in the separation of genes in the same operon would be selected against in nature. Indeed, the operon structure of *E. coli* K12 and *S. typhimurium* LT2 is strikingly similar (Sanderson et al., 1999). This suggests that, order within operons in these bacteria and presumably other enteric bacteria, has been conserved. The maintenance of operon gene order may contribute to the overall conservation of gene order in enteric bacteria; however, this contribution would be small, since it does not explain the conservation in order of genes not present within operons.

In conclusion, enteric bacteria combine several selective mechanisms, including those which act to maintain gene dosage, genome balance, transcription and replication

orientation, chi sequence orientation and proper operon structure, to conserve overall gene order in their chromosome.

B. The genetic aspects of enteric bacteria

1. Causes of genetic change

Mutations in the genome, which lead to changes at the sequence level, are the ultimate source of genetic variability. Recombinational events, which then rearrange this variability and give rise to new combinations, is the foundation for evolution (Snustad et al., 1997). Repeated sequences like multigene families, *rrn* operons and IS elements, are found scattered throughout the genome. These sequences provide regions of homology that facilitate recombinational events, which give rise to inversions, deletions and duplications in the genome. An inversion occurs when the repeated sequences involved are in opposite orientation in respect to one another on the chromosome, while duplications and deletions occur when the repeated sequences are in the same orientation (Weinstock and Lupski, 1997). Strikingly, duplication and deletions are more common than inversions (Ornston et al., 1990). Possible reasons for the variety of inversions could be that gene dosage is being adversely affected, or replication functions are being disrupted. Furthermore, recombinational events often encompass large segments of the genome. For this reason, some inversions and deletions are often not recovered since too many essential genes are disrupted or erased (Weinstock and Lupski, 1997). Conversely, duplications appear at high frequencies in bacterial genomes. It is believed that

duplications provide bacteria with excess DNA, which help them adapt to stressful conditions in the nature, without irreversibly changing their genomes (Weinstock and Lupski, 1997).

2. Genetic variability in *E. coli* and *S. typhimurium*

E. coli and *S. typhimurium* have the ability to resist genetic drift despite the fact that their genomes have ample opportunity for change. At the sequence level, these two organisms have diverged up to 25% in regards to their homologous genes (Mahan et al., 1990). This reduces their ability to exchange genetic material and facilitates divergence; however, what is striking about these bacteria is that their genomes are remarkably conserved. What is especially surprising is that in culture, rearrangements occur frequently in *E. coli* and *S. typhimurium*. It has been observed that duplication of chromosomal segments occurs at frequencies as high as 10^{-3} to 10^{-5} (Shimkets, 1997; Weinstock and Lupski, 1997). Duplications, deletions, translocations and inversions are usually mediated via *rrn* operons. There are seven *rrn* operons in *E. coli* and *S. typhimurium*, which provide large, repetitive sequences for homologous recombination (Weinstock and Lupski, 1997). The postulated basis of these rearrangements are illustrated in Figure 2.1 (Hill and Combraito, 1973; Liu and Sanderson, 1996). The duplicated segments are fairly large in size, for example, in *E. coli* and *S. typhimurium*, the size ranges from 42 kb to 2100 kb (Weinstock and Lupski, 1997). It appears that duplications or deletions resulting in either gain or loss of genetic material are what predominantly distinguishes *E. coli* from *S. typhimurium* (Riley and Sanderson, 1990). In these bacteria, inversions and translocations are less frequent. Such recombination events

Figure 2.1: Postulated mechanisms for homologous recombination between the *rrn* operons (Hill and Combraito, 1973); (Liu and Sanderson, 1996).

The linear I-*Ceu*I cleavage map of *E. coli* and *S. typhimurium*, with the I-*Ceu*I fragment order of A-B-C-D-E-F-G, before any rearrangements have taken place (Liu and Sanderson, 1995b). a. Duplications can result from interchromosomal rearrangements. b. Deletions can arise from interchromosomal rearrangements. c. Transposition occurs if a deleted fragment (as seen in b) inserts back into the chromosome at a different *rrn* operon site via homologous recombination. d. Inversions result from homologous recombinations between rrn operons, which are opposite in their direction of transcription. The letters represent the I-*Ceu*I fragments between the *rrn* operons. The arrows represent the direction of transcription. The dot on the I-*Ceu*I C fragment represents the origin of replication (*oriC*).

A. Duplication
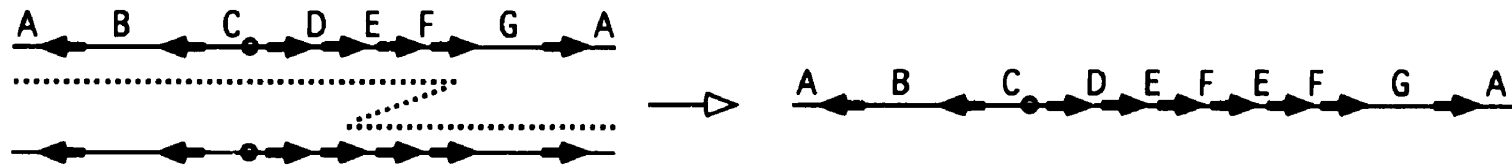
B. Deletion.
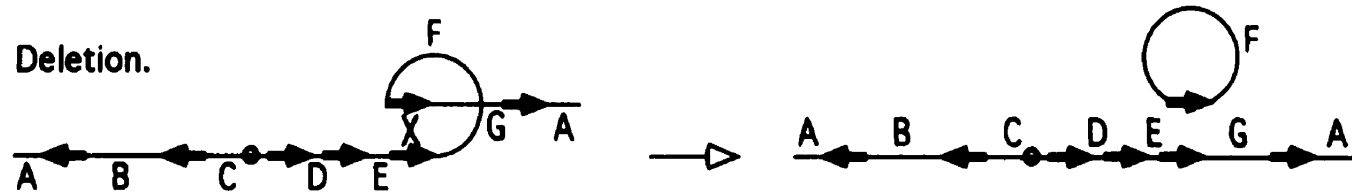
C. Transposition

D. Inversion

rrnD   *rrnE*
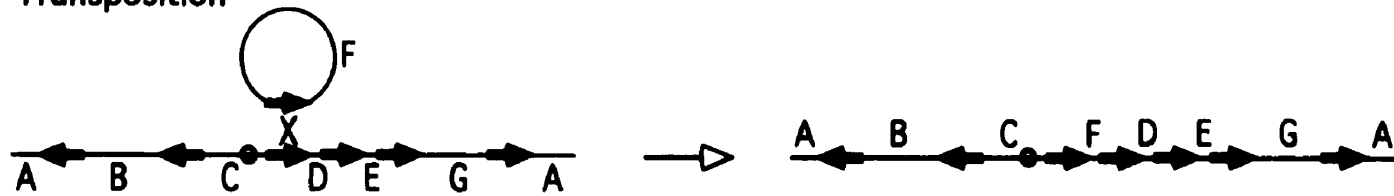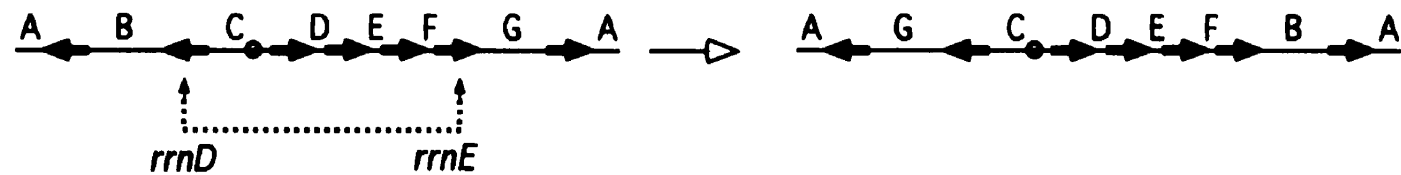
often lead to large tandem duplications and reshuffling of gene position. Nevertheless, despite all the potential *E. coli* and *S. typhimurium* have for change, their genomes are strikingly conserved, reiterating the notion that selective pressures must exist to negate genome variability in these bacteria.

3. Genetic variability in *S. typhi*

a) Distribution of *S. typhi*-

Of the 2,300 serovars of *Salmonella*, *Salmonella typhi* exclusively infects human hosts, where it causes enteric or typhoid fever (Thong et al., 1994) *S. typhi* infections are a serious cause of morbidity and mortality in many developing countries such as China, India and Malaysia. The estimated annual global incidence is 20 million cases and more than 700,000 deaths; this appears to be on the rise in many parts of the world due to the emergence of antibiotic-resistant *S. typhi* strains (Thong et al., 1994) ((Chee et al., 1992).

b) Genome of *S. typhi*-

The genome size of *S. typhi* Ty2, the laboratory strain, is known to be 4870 kb (Liu and Sanderson, 1995c). Like *E. coli* K12 and *S. typhimurium* LT2, *S. typhi* Ty2 has seven *rrn* operons; however, the order in which these operons are present on the chromosome is not similar to that of the first two bacteria. Moreover, after analyzing several *S. typhi* strains, the order of *rrn* operons within the species is also variable. Thus, it has been concluded that, although most enteric bacteria have highly conserved genomes, *S. typhi* is an exception to this rule (Liu and Sanderson, 1996; Liu and Sanderson, 1995c). In fact, chromosomal rearrangements caused by translocations and

inversions involving the *rrn* operons are commonly maintained in *S. typhi* strains, unlike most other enteric bacteria.

4. Rearrangements in the *ter* region of enteric bacteria

Although the genome of most enteric bacteria is conserved, the area encompassing the *ter* region is somewhat variable. Inversions are commonly not maintained in wild-type strains of *E. coli* and *Salmonella*, especially if the inversions do not involve *rrn* operons. However, this is not true within the *ter* region where inversions are surprisingly frequent. When using the *ter* region of *S. typhimurium* as a reference point, the same area in *E. coli* has a 480 kb block which has been inverted. Similarly, in *S. typhi* Ty2 this inverted block is 500 kb while in *Salmonella paratyphi* C it is 700 kb in length (Sanderson et al., 1999; Liu and Sanderson, 1995c). It is important to note that none of these inversions significantly displaces the position of the *ter* region and hence, should not have a major effect on genome balance. There are three possible reasons why this region is so susceptible to recombination. The first relies on the presence of the *dif* locus within the *ter* region (Weinstock and Lupski, 1997). This locus functions to facilitate chromosomal partitioning and is believed to act as a substrate for site-specific recombinases XerC and XerD (Hill, 1996). These recombinases resemble the lambda integrase family and are essential for recombination at the *dif* site (Hill, 1996; Sanderson et al., 1999). Secondly, some claim that hyperrecombination occur within the terminus region. This could be due to presence of several recombinational hotspots which all, interestingly, possess chi activity (Hill, 1996). These sequences are thought to promote homologous recombination. Lastly, DNA replication arrest occurs at the terminus site,

this produces stalled replication forks that increase the chance for illegitimate recombination to occur (Sharp, 1991).

C. Genomic cleavage maps

1. I-*Ceu*I genomic cleavage maps of *E. coli* K12, *S. typhimurium and S. typhi*

Endonuclease digestion combined with pulsed-field gel electrophoresis has proven to be a useful tool in deriving physical maps of genomes. *Xba*I, *Bln*I, *Spe*I and I-*Ceu*I are all rare cutting endonucleases, which have been used for this purpose. Of these enzymes, I-*Ceu*I cuts the bacterial chromosome least frequently, due to its large recognition site, and thus gives rise to the fewest number of DNA fragments. I-*Ceu*I is encoded by a class I mobile intron found in the chloroplast DNA of *Chlamydomonas eugamatos*. This intron is inserted in the *rrl* gene for the 23S rRNA and it specifically cuts in a 19 base pair sequence of the *rrl* gene (Liu and Sanderson, 1995b). Since RNA genes are highly conserved in bacteria, this 19 base pair sequence is found in all enteric bacteria, *Rhizobium meliloti*, and eukaryotic mitochondrial and chloroplast DNA (Liu and Sanderson, 1995b).

The I-*Ceu*I cleavage maps of *E. coli* K12 and 18 independent strains of *S. typhimurium*, including *S. typhimurium* LT2 have been determined. Since there are seven *rrn* operons in *E. coli* and *S. typhimurium*, there are seven I-*Ceu*I fragments and the order of these fragments, ABCDEFG, is the same in the two groups of bacteria (Liu and Sanderson, 1995b; Liu-Shu-Lin et al., 1993). The I-*Ceu*I cleavage map of *S.*

typhimurium LT2 is shown in Figure 2.2. The I-*CeuI* cleavage map was determined by partially digesting genomic DNA with I-*CeuI* restriction endonuclease and separating the restriction fragments by PFGE. The fragments, which appear to be a sum of other individual fragments on the pulse-field gel, reveal adjacent fragments.

2. Genome types of *S. typhi*

a) Balanced genome types-

S. typhi also has seven I-*CeuI* fragments, however, the order of these fragments is highly variable among individual strains. In fact, of the 144 wild-type *S. typhi* strains studied so far, there are 27 different I-*CeuI* fragment orders (Liu and Sanderson, 1995a). The 27 I-*CeuI* fragment orders are depicted in Figure 2.3a. Each of these genome types result from at least one translocation or inversion event involving *rrn* operons. Out of the 27 genome types, 24 appear to have balanced genomes, such that the distance between their *oriC* site and the *ter* region is separated by 180°. *S. typhi* Ty2, 4780 kb in size, falls into genome type 9 which has a I-*CeuI* fragment order of AGCEFDB(Liu and Sanderson, 1995a). A circular I-*CeuI* cleavage map of *S. typhi* Ty2 is shown in Figure 2.2. A detailed genome cleavage map depicting *XbaI*, I-*CeuI*, *AvrII* (*BlnI*) and *SpeI* cleavage sites in *S. typhi* Ty2 is shown in Figure 2.4. The order of I-*CeuI* restriction fragments was determined by partial I-*CeuI* digestion. Genome type 3 is the most common, with 57 out of the 144 strains studied so far having the I-*CeuI* fragment order of ABCEFDG (Liu and Sanderson, 1996). When comparing all these genome types, a few striking features are observed. Firstly, the position of the 502 kb I-*CeuI*C fragment, harboring the *oriC* site, rarely moves more than 1-2%, in relation to its position in *S. typhi* Ty2. *S. typhi* Ty2

Figure 2.2: Comparison of the I-*Ceu*I cleavage maps of *S. typhimurium* LT2 and *S. typhi* Ty2 (Liu and Sanderson, 1995a)

The I-*Ceu*I cleavage map of *S. typhimurium* LT2 illustrates an I-*Ceu*I fragment order of A-B-C-D-E-F-G and the I-*Ceu*I cleavage map of *S. typhi* Ty2 represents an I-*Ceu*I fragment order of A-G-C-E-F-D-B. The total genome size of *S. typhimurium* LT2 is 4600 kb and *S. typhi* Ty2 is 4780 kb. I-*Ceu*I restriction endonuclease cuts within the *rrl* genes for the 23s rRNA in the rrn operons. There are seven *rrn* operons in the genome of *S. typhimurium* LT2 and *S. typhi* Ty2 and they are designated alphabetically from A to G. The arrows besides the *rrn* operons indicate the postulated direction of replication, which is accordance with the position of the origin of replication (*oriC*) present in the I-*Ceu*I C fragment. The arc with arrowheads at both ends represents the segment of the *S. typhi* Ty2 genome that appears to be inverted in relation to *S. typhimurium* LT2. The open arrowheads indicate regions in *S. typhi* Ty2, predicted to have insertions, in relation to *S. typhimurium* LT2. The largest insertion, approximately 120 kb, is found within the I-*Ceu*I G fragment (Liu and Sanderson, 1995a).

S. typhi Ty2

S. typhimurium LT2
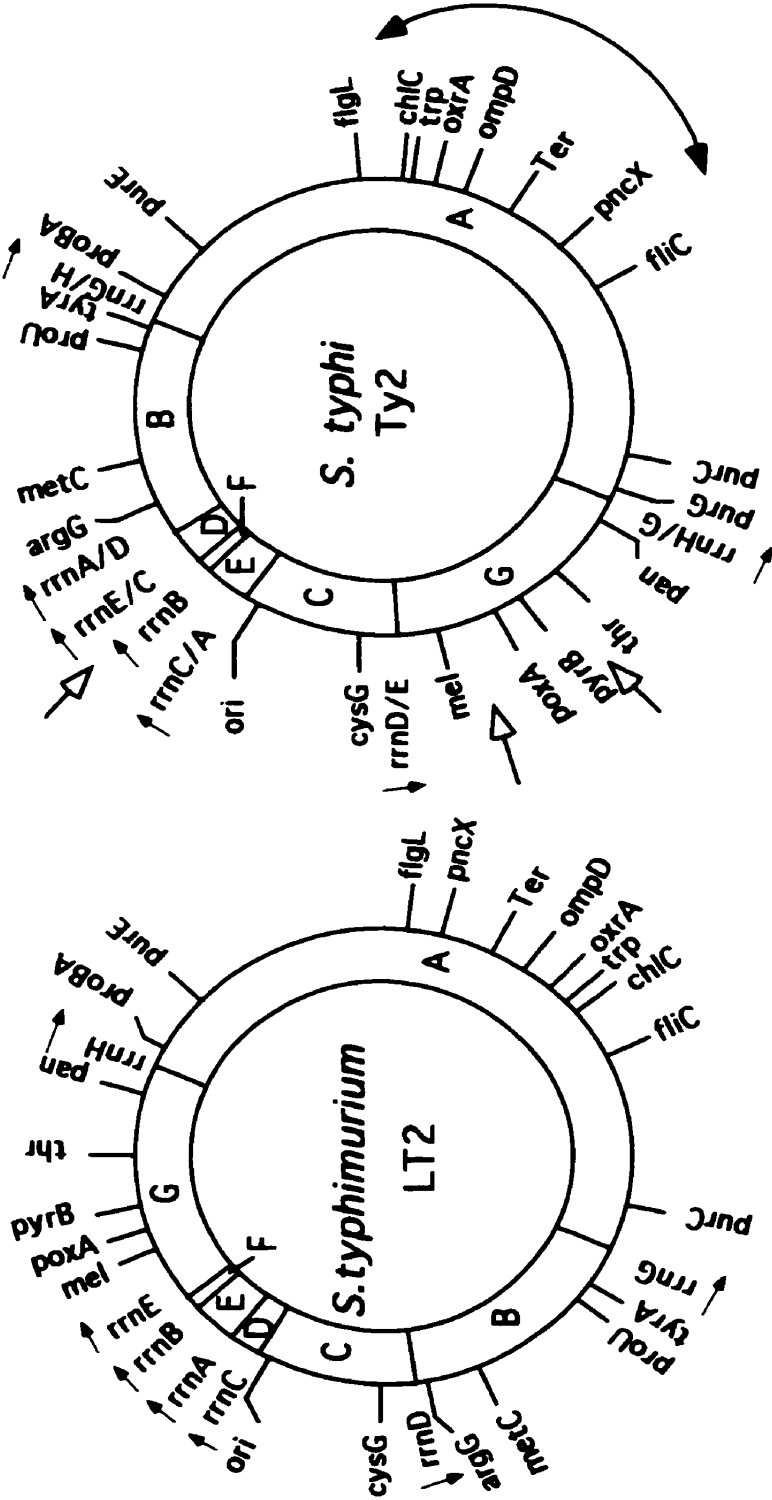
Figure 2.3: The genome types of wild-type and laboratory-derived *S. typhi* strains, determined by their order of I-*Ceu*I fragments (Liu and Sanderson, 1996)

a. 141 strains of wild-type *S. typhi* strains fall into 27 genome types determined by the order of their I-*Ceu*I fragments. Note that some genome types expected to be present were not detected (Liu and Sanderson, 1996). The order of the I-*Ceu*I fragments in *S. typhimurium* LT2 and the position and direction of its *rrn* operons are shown. Homologous recombination between *rrn* operons has rearranged the order of I-*Ceu*I fragments in *S. typhi* with respect to *S. typhimurium* LT2. The I-*Ceu*I A fragment is assumed to join the left end to the right end in the fragments shown, creating a circle; however, the orientation of the A fragment is not known in most of these strains. The genome types, in which the orientation of the I-*Ceu*I A fragment is known, are labeled with an asterisk; however, what is known is with respect to the orientation of the I-*Ceu*I A fragment in *S. typhi* Ty2, represented by genome type 9. The I-*Ceu*I C fragment carries the *oriC* site but this is only known for *S. typhimurium* LT2 and *S. typhi* Ty2. The orientation of the I-*Ceu*I C fragment cannot be predicted in these *S. typhi* strains since it is flanked by *rrn* operons with opposing transcription direction and the fragment can be present in an inverted form as well. **b.** The five classes of genome types, a to e, represent 20 laboratory-derived *S. typhi* SARB 63 revertant strains, in which the I-*Ceu*I C fragment was originally displaced by 15%. These strains now have restored the I-*Ceu*I C fragment close to the position maintained by *S. typhi* Ty2.
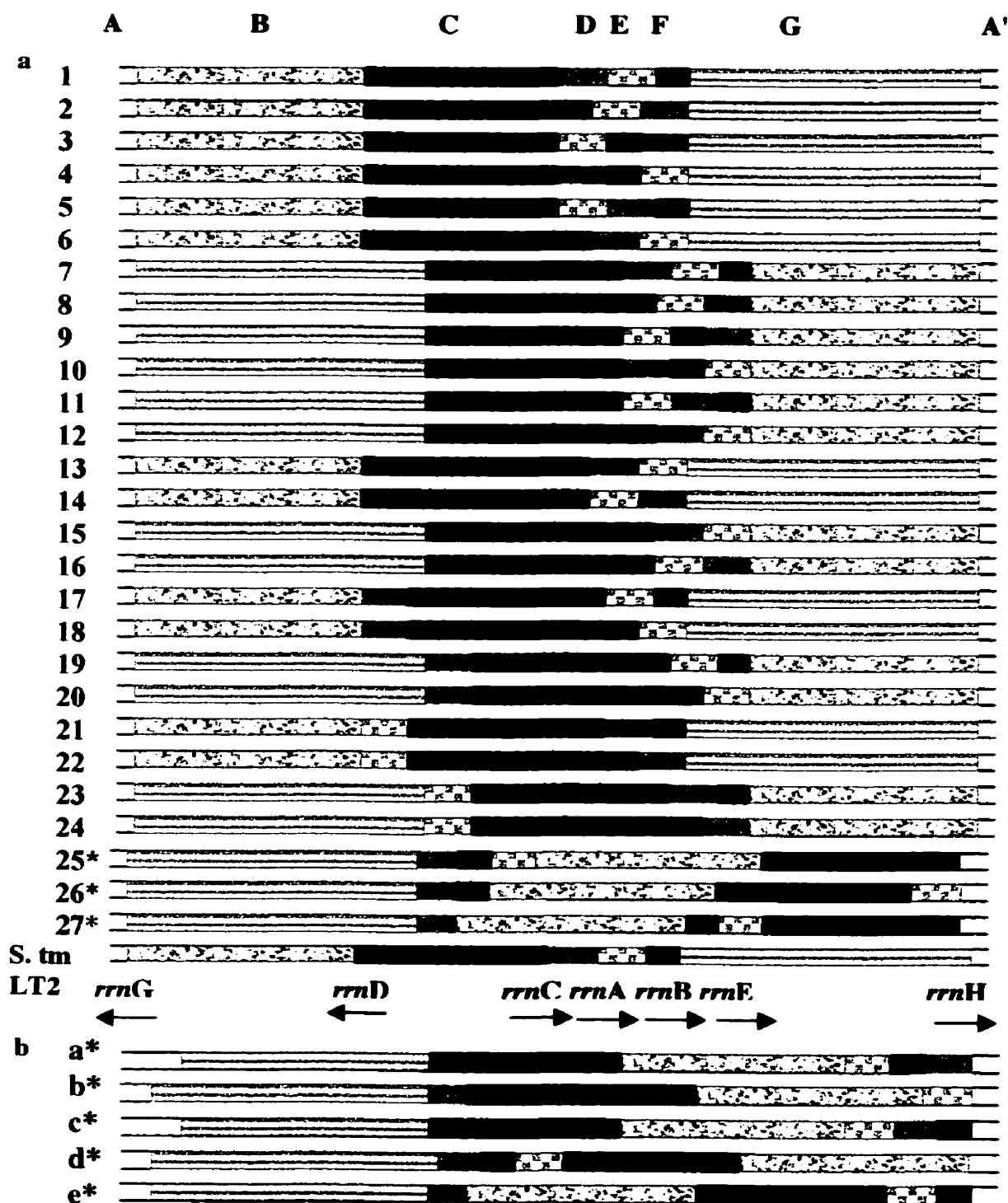
Figure 2.4: The genome cleavage map of _Salmonella typhi_ Ty2 containing _Xba_I, I-_Ceu_I, _Avr_II (_Bln_I) and _Spe_I cleavage sites (Liu and Sanderson, 1995c)

The total size of _S. typhi_ Ty2 is 4780 kb and the sizes are shown in fractions of 500 kb, going in a clock-wise direction. The _Xba_I, _Bln_I (_Avr_II) and _Spe_I cleavage fragments are presented in alphabetical order starting at 0 kb. For the I-_Ceu_I cleavage map of _S. typhi_ Ty2, the fragments are labeled in accordance with the seven I-_Ceu_I fragments present in _E. coli_ K12 and _S. typhimurium_ LT2; however, in Ty2, these fragments are found in the rearranged order of A-G-C-E-F-D-B. The structure and cleavage sites of the seven _rrn_ operons are shown outside the circle. The arrows represent the direction of transcription of each _rrn_ operon (Liu and Sanderson, 1995c).

Genome Map of

Salmonella typhi Ty2

4780 Kb
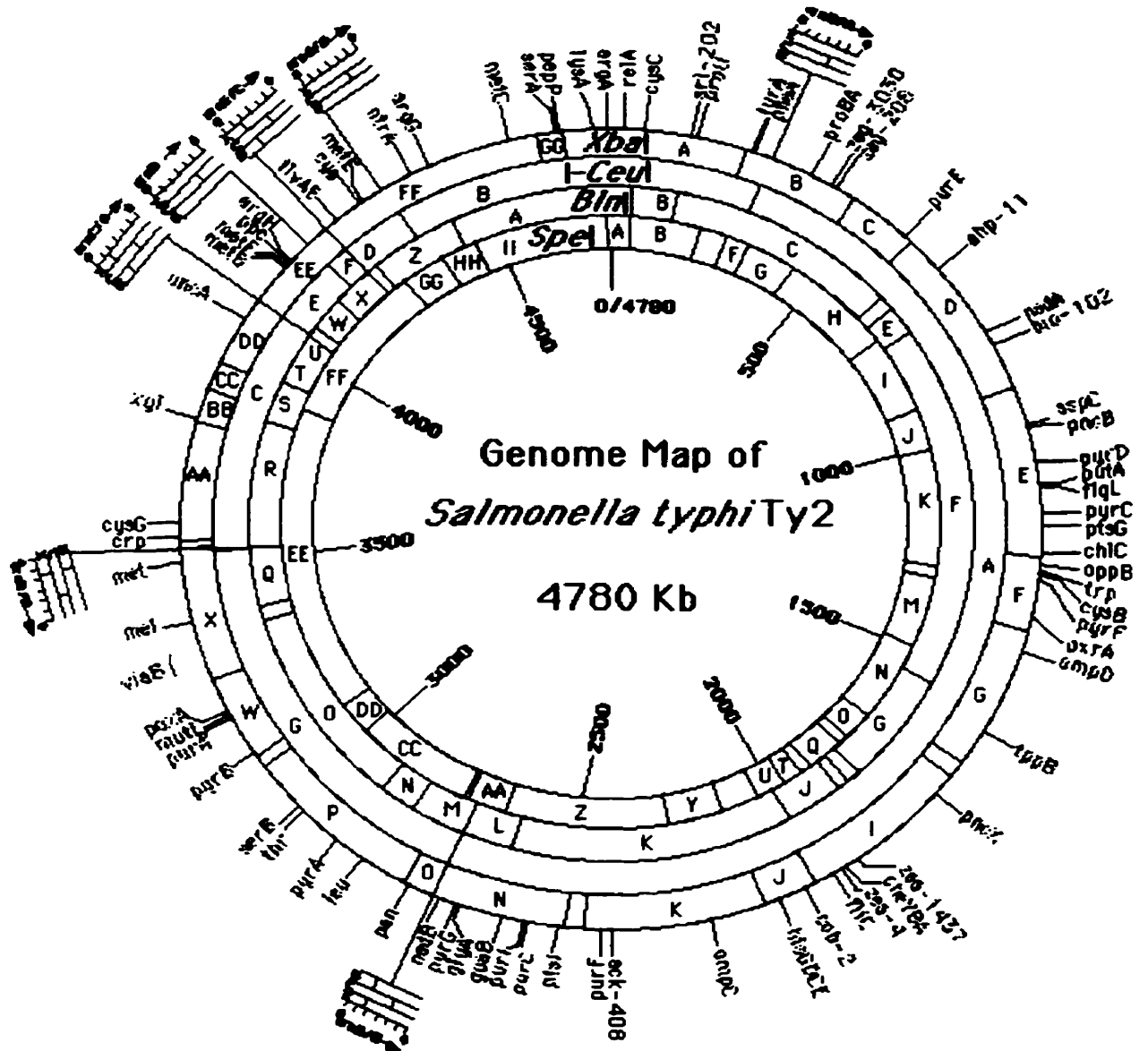
serves as an accurate reference point, since its *oriC* site, as well as its *ter* region (found in the I-*CeuI* A fragment), have been mapped (Liu and Sanderson, 1995c) and indeed, they are found to be approximately 180° apart. The fact that the location of the I-*CeuI* C fragment is conserved in most *S. typhi* strains suggests the genome balance theory is upheld; this is assuming that the *ter* region, in the I-*CeuI* A fragment, has not moved. Secondly, the I-*CeuI* F fragment houses essential housekeeping genes, and inversions which displace this fragment by more than 10% from either side of the *oriC* site are rarely seen (Liu et al., 1999). This suggests that the gene dosage theory is also upheld in most of the *S. typhi* strains.

b) Unbalanced genome types-

However, there are three *S. typhi* strains, SARB 63 (genome type 25), PL27566 (genome type 26) and 701Ty (genome type 27), in which neither genome balance nor gene dosage appear to be maintained (Liu and Sanderson, 1996). These genome types are illustrated in Figure 2.3a. The *oriC* site in SARB 63 is thought to be displaced by almost 15%, which in accordance with the genome balance theory should be gravely disadvantageous to the cell's fitness. However, it has been observed that the growth of SARB 63 is not as severely hindered as expected, instead only a small reduction in colony size is seen. It is a mystery as to how SARB 63 manages to survive with its genome being in such imbalance.

c) SARB 63 revertants-

Five classes of revertants of the *S. typhi* strain SARB 63 have been isolated (Ng, 1999). These classes are represented in Figure 2.3b. Initially, SARB 63 revertant has an

unbalanced genome (*ori* C was displaced by 15%). However, one or more inversion events, presumably mediated via homologous recombination between *rrn* operons, have restored the I-*Ceu*I C fragment, close to its original position such that the genomes are no longer unbalanced. This has resulted in revertant classes a to e as seen in Figure 2.3b.

D. Possible theories for the genome diversity in *S. typhi*

The vast diversity found in the *S. typhi* genome, represented by its multiple genome types, suggest that the serovar of *S. typhi* is tolerant to genetic variability, unlike its relatives, *E. coli* and *S. typhimurium*. There are two possible explanations for the plasticity observed in *S. typhi*. First, it could be that the genome balance theory does not apply to *S. typhi* as stringently as it does to other bacteria. For this reason, some strains of *S. typhi* are able to survive in nature despite their unbalanced genomes.

The second explanation pertains to the Adopt-Adapt Model of bacterial speciation. In this model, it is hypothesized that during the process of speciation, the bacteria often adopts foreign DNA, which may, as a consequence, disrupt its genome balance. To compensate for this, the bacteria undergo several changes via rearrangements, deletions or insertions in an attempt to restore genome balance and maximize cell fitness. *S. typhi* is a fairly recent species and since humans are its only host, it is speculated that *S. typhi* has only been around as old as the human race, which is approximately 2-3 millions years (Campbell, 1997). Thus, *S. typhi* is most likely still undergoing speciation. Secondly, all *S. typhi* strains have adopted 120 kb of foreign

DNA, which may have disrupted its genome balance (Liu and Sanderson, 1995a). Genomic reorganizations might be a mechanism to restore the 180° symmetry between the origin and terminus, and could be what gives rise to the different genome types in *S. typhi*.

E. Conservation of genome size in *S. typhi*

Genome size within enteric bacteria is relatively conserved. The size of *E. coli* K12 and its derivatives range from 4600 kb to 4800 kb (Brewer, 1988; Cole and Saint-Girons, 1999; Blattner et al., 1997). This variation is relatively small and somewhat expected since insertions and deletions are common in most bacteria. The genome size of *E. coli* 0157 is 5400 kb (Cole and Saint-Girons, 1999). This increase in size is not too surprising since this is a clinical strain *of E. coli* and its genome may harbor additional DNA to accommodate for its invasiveness. *S. typhimurium, S. enteritidis, S. paratyphi* A and *S. paratyphi* B have genome sizes of 4890 kb 4660 kb, 4595 kb and 4655 kb, respectively (Liu and Sanderson, 1995; Liu et al., 1993; Liu et al., 1994; Liu and Sanderson, 1995b). Thus, it appears that variation in size of these enteric bacteria is slight. Notably, most of these genome sizes center around 4600 kb to 4700 kb, suggesting that a genome smaller than this cannot exist or is selected against in enterics, possibly due to not enough essential genes being present.

S. typhi appears to be yet another exception in respect to size conservation. It has been claimed that the genome size of *S. typhi* isolates can vary as much as 20%,

representing approximately 960 kb; analysis of 17 independent isolates of *S. typhi* shows size to range from 3980 kb to 4840 kb (Pang, 1998). This is a strikingly large variation not seen in other enteric bacteria. Furthermore, the fact that these *S. typhi* strains have genome sizes which are as small as 3980 kb is surprising, especially since they represent clinical isolates. In addition, there have been no reports of enteric bacteria having genomes this small, which is yet another unique feature of these *S. typhi* strains.

Contrary to the above observation, size studies on 144 wild-type *S. typhi* strains has shown very little variation. I-*Ceu*I cleavage maps indicate that I-*Ceu*I fragments B, C, D, E, F and G are conserved in size(Liu et al., 1999). Although the size of the I-*Ceu*I A fragment, which makes up about half the genome size of *S. typhi*, was not accurately determined in these studies, conclusions drawn from the other I-*Ceu*I fragments strongly suggest overall size is conserved in *S. typhi*.

To summarize, the genome *S. typhi* appears to be highly plastic in respect to order, and maybe even size. These features, along with its capability to cause serious disease in many parts of the world, make *S. typhi* an important candidate for molecular and genetic study.

F. Hypothesies

In the process of speciation, *S. typhi* has acquired novel DNA, presumably through lateral transfer. Since this may result in asymmetry between the *ori* site and the *ter* region, the chromosome could be unbalanced. Consequently, this may give rise to

non-synchronous completion of bi-directional genomic replication and could adversely affect cell fitness. Genomic reorganizations may be a mechanism to restore the 180° separation.

In our previous studies, it was observed that half the genome of 144 wild-type *S. typhi* strains is relatively conserved in respect to size. It is therefore predicted that in *S. typhi*, reorganization is not through large-scale gain or loss of DNA but through genomic recombinations. Thus, the overall genome size of *S. typhi* is expected to be conserved.

It is known that rearrangements involving *rrn* operons are common in *S. typhi* strains. It is hypothesized that rearrangements not involving the *rrn* operons also occur but mainly around the *ter* region. This is anticipated to be most evident in strains like *S. typhi* SARB 63, whose *oriC* site is displaced. Rearrangements around the *ter* region may be a way of compensating for the changes in the position of the *oriC* site, so that overall genome balance in *S. typhi* is conserved.

Since a 180° symmetry is optimal for maximal cell fitness and growth, it is hypothesized that disrupting this symmetry will adversely affect cell fitness in *S. typhi*. Similarly, it is expected that restoring this 180° symmetry between the *oriC* site and *ter* region should improve cell fitness.

G. Objectives

In this study, the genome of several *S. typhi* strains was investigated to detect overall conservation of genome size and balance. Furthermore, the consequences of these genomic variations in *S. typhi* , was also examined in regards to its effect on growth and cell fitness.

1) To determine the size of the I-*Ceu* I A fragment to assess the overall size variation in the total genome size of *S. typhi* strains.

2) Identify rearrangements, other than those mediated by homologous recombination between *rrn*-operons, in *S. typhi* strains, specifically in the I-*Ceu* I A fragment.

3) To study the effect of genome imbalance on cell fitness through growth studies, on specific strains of wild-type and laboratory-derived *S. typhi.*

CHAPTER 3: MATERIALS AND METHODS

A. Bacterial strains

S. typhi strains were obtained from the following sources: Laboratory Centre for Disease Control, Ottawa, Canada; Centers for Disease Control and Prevention, Atlanta; Provincial Laboratory of Alberta, Calgary; Tikki Pang (University of Malaya, Malaysia); Robert Selander (Pennsylvania State University); Bruce Stocker (Stanford University); David Hone (University of Maryland, Baltimore). They were identified as *S. typhi* based on biochemical and antigenic characterizations, which were determined by the laboratories of origin, and confirmed by the Laboratory Centre for Disease Control, Ottawa. Strains were maintained in 15% glycerol at -70°C, and a single colony was isolated prior to use.

B. Media

1. Luria-Bertani (LB)

Bacterial strains were grown in LB-broth (10g tryptone, 5g yeast extract, 10g NaCl, 3.5 ml 1M NaOH- GibcoBRL) and in solid medium containing LB-broth with 1.5% bacto-agar (Difco).

2. Minimal medium

Bacterial cells were also grown in minimal salts broth containing 10.5g K_2HPO_4, 7.5g KH_2PO_4, 1g $(NH_4)_2SO_4$ and 0.5g $MgSO_4 \cdot H_2O$, supplemented with 0.2% glucose. Solid minimal medium was also used containing the above minimal salts broth with 1.5% bacto-agar (Difco). Both broth and solid medium were supplemented with 0.5% Caseamino acids and 0.33% of tryptophan, methionine and cysteine to insure proper growth.

C. Enzymes and chemicals

Restriction endonucleases were provided by New England Biolabs (I-*CeuI*, *AvrII* and *SpeI*) and Boehringer Mannheim (*XbaI*). α-^{32}P dCTP, Klenow, corresponding buffers and dNTPs were obtained from Amersham Pharmacia Biotech. Most other chemicals were from Sigma-Aldrich Chemical Company.

D. Genomic DNA isolation

Single colony isolates of *S. typhi* strains was obtained from cell collection frozen at –70°C. Cells were incubated in 2 ml of LB broth and shaken overnight at 37°C. Cells were then centrifuged at 1000xg for 5 minutes and the pellet was resuspended in 0.5 ml of Cell Suspension Solution (10 mM Tris-HCl pH 7.2, 20 mM NaCl, 100 mM EDTA). 0.5 ml of molten 1.4 % agarose was added to this suspension, mixed quickly and drawn

into a Gel syringe (NE Biolabs). This cell-agarose mixture was left to harden at room temperature for 5-10 minutes and subsequently cut into disks of 1mm thickness. The 1 ml sample of disks were treated with 5 ml of Lysozyme Buffer (10 mM Tris-HCl pH 7.2, 100 mM EDTA, 0.2% SDS, 0.5% N-laurylsarcosine) at 70°C for 1 hour with gentle shaking. The solution was discarded and disks were washed twice in 5 ml of Wash Solution (20 mM Tris-HCl, pH 8.0, 50 mM EDTA) at room temperature with gentle shaking for 15 minutes each. Digestion of disks was done with 3 ml of Proteinase K Solution (1.0 mg/ml Proteinase, 100 mM EDTA, 0.2% SDS, 0.5% N-laurylsarcosine) at 42°C, shaking for 18 hours. Sample was then washed once in Wash Solution for 15 min at room temperature with shaking. Incubation in 5 ml of PMSF Solution (Wash Solution containing 1 mM Phenylmethylsulfonyl Flouride) followed for 2 hours at room temperature with gentle shaking. The sample was then washed twice in Wash Solution and subsequently in Storage Solution (10X diluted Wash Solution) for 15 minutes each and shaking. The samples were stored in Storage solution at 4°C.

E. Restriction endonuclease digestions

1. I-*Ceu*I digestion

A single disk containing *S. typhi* genomic DNA was first equilibrated in 2X I-*Ceu*I NE buffer (20 mM Tris-HCl, 20 mM $MgCl_2$, 2 mM DTT) for 15 minutes at room temperature. Digestion was then done using 0.4 units of I-*Ceu*I enzyme in 1X I-*Ceu*I NE buffer and 1x BSA for 2-3 hours in a 37°C water bath.

2. *Avr*II, *Spe*I and *Xba*I digestions

When *Avr*II and *Spe*I digests were done, each block of agarose containing the *S. typhi* I-*Ceu*I A fragment DNA was first equilibrated in 2X NE buffer 2 (100 mM NaCl, 12 mM Tris-HCl, 20 mM MgCl$_2$, 2 mM DTT) for 15 minutes at room temperature. Digestion was then done using 50 units of *Spe*I in 1X NE buffer 2 and 1x BSA, or 4 units of *Avr*II in 1X NE buffer 2, for 4 hours in a 37°C water bath.

When digesting with *Xba*I, a block of agarose containing the *S. typhi* I-*Ceu*I A fragment DNA was equilibrated in 2X buffer H (100 mM of NaCl, 50 mM of Tris-HCl, 10 mM of MgCl$_2$, 1 mM DTE- BM) for 15 minutes at room temperature. Digestion was then done using 40 units of *Xba*I in 1X buffer H for 4 hours in a 37°C water bath.

3. *Spe*I and *Xba*I digestions to isolate probe DNA

Digestions were done on disks of *S. typhi* Ty2 genomic DNA, PFGE was used to separate specific *Spe*I and *Xba*I fragments. The digestions were done in the same manner as stated above.

F. End-labeling

End-labeling was done using the End-labeling Kit provided by Pharmacia Biotech. Agarose blocks of I-*Ceu*I A fragment digested with either *Xba*I, *Spe*I and *Avr*II were each end-labeled using 2.4 μCi of α-^{32}P dCTP and 0.5 units of Klenow enzyme diluted in Reagent mix (buffered aqueous solution of dNTPs) and dH$_2$O. The blocks

were labeled with α-^{32}P dCTP for 1 hour at 37°C. After electrophoresis of end-labeled

blocks, they were vacuum dried at 80°C for 3 hours and loaded into a cassette for

autoradiography.

G. Pulsed-Field Gel Electrophoresis

Each set of DNA digests was immersed into the wells of a 0.7% agarose gel

(GibcoBRL) containing 1µg of EtBr/ml and was separated by a CHEF MAPPER

electrophoresis system (Bio-Rad) at 6.0V/cm. When probe DNA was isolated from the

gel, low melting temperature agarose was used (SeaPlaque GTG agarose-FMC

BioProducts). Electrophoresis was performed in 0.5X TBE buffer (90 mM Tris Base, 90

mM boric acid and 2 mM EDTA, pH 7). Pulse times varied depending on the

endonuclease cleavage pattern being obtained and the size of fragments being separated.

Typically, pulse times are such that for every 10 kb, there is a 1 second pulse (e.g.

wanting to separate a 400 kb band, the pulse conditions would be 30s to 50s).

1. Conditions for end-labeling templates

Gel conditions for I-*Ceu*I digests involved pulse times of 30s - 150s for 12 hours

at 6V/cm² and 120°, followed by pulse times of 90s – 120s for 12 hours at 180V and

110°. Gel conditions for end-labeled *Xba*I digests involved pulse times of 10s – 60s for

12 hours at 6V/cm² and 120°, followed by pulse times of 10s – 30s for 6 hours at 6V/cm²

and 120°. Gel conditions for end-labeled *Spe*I digests involved pulse times of 5s – 20s

for 15 hours at 6V/cm² and 130°, followed by pulse times of 2s – 6s for 36 – 40 hours at

$6V/cm^2$ and $150°$. Gel conditions for end-labeled *Avr*II digests involved pulse times of $30s - 90s$ for 12 hours at $6V/cm^2$ and $120°$, followed by pulse times of $40s - 60s$ for $36 - 40$ hours at $6V/cm^2$ and $120°$.

2. Conditions for probing templates

Gel conditions for *Spe*I digests involved pulse times of $20s - 80s$ for 16 hours at $6V/cm^2$ and $120°$, $5s - 20s$ for 24 hours at $6V/cm^2$ and $140°$, and $6s - 8s$ for 20 hours at $6V/cm^2$ and $150°$. Gel conditions for *Xba*I digests are $10s - 60s$ for 12 hours $6V/cm^2$ and $120°$, $8s - 24s$ for 12 hours at $6V/cm^2$ and $130°$, and $22s - 26s$ for 16 hours at $6V/cm^2$ and $150°$.

3. Conditions for isolating probe

The *Spe*I AA fragment was isolated using the following pulse conditions: $10s - 40s$ for 10 hours at $6V/cm^2$ and $120°$ and $6s - 8s$ for 24 hours at $6V/cm^2$ and $150°$. The *Xba*I G, I and J fragments were isolated using the following pulse conditions: $10s - 60s$ for 12 hours $6V/cm^2$ and $120°$, $8s - 24s$ for 12 hours at $6V/cm^2$ and $130°$, and $22s - 26s$ for 16 hours at $6V/cm^2$ and $150°$.

H. Southern blotting

Gel containing DNA to be transferred onto membrane was trimmed and rinsed in freshly made 0.25 M HCl for 15 minutes with gentle shaking. The gel was washed twice

with ddH$_2$O for 2 minutes. The gels were then soaked in 0.4 M NaOH for 30 minutes with gentle shaking. Prior to loading on top of gel to initiate transfer, the Immobilon-P transfer membrane(Millipore cooperation) was immersed in methanol for 15 minutes. The membrane was rinsed in ddH$_2$O 6 times and finally soaked in 0.4 M NaOH briefly. The DNA transfer took 24-48 hours after which the membrane was air dried for 30 minutes and heat fixed at 80°C for 2 hours.

I. Probe-labeling

Blocks of probe DNA was washed in sterile ddH$_2$O two times, boiled for 8 minutes at 97°C and then cooled for 5 minutes at 37°C. The probe was labeled using 4 μCi of α-^{32}P dCTP and 0.5 units of Klenow enzyme (Oligolabelling Kit- Pharmacia) diluted in Reagent mix (buffered aqueous solution of dNTPs) (T7 Quick Prime kit- Pharmacia) and ddH$_2$O. The blocks were labeled with α-^{32}P dCTP for 45 minutes at 37°C.

J. Southern hybridization and washings

1. Pre-hybridization

The membrane was loaded into a hybridization cylinder between two nylon mesh and soaked in hybridization solution. The membrane was pre-hybridized for 24-48 hours within a 42°C hybridization oven.

2. Hybridization

Prior to adding label to membrane, the label was denatured at 97°C for 5 minutes. Hybridization solution was removed from the pre-hybridizing cylinder, the probe was added to this solution and the entire mixture was put back into the cylinder. Hybridization took place within a 42°C hybridization oven over 2 days.

3. Washings

Membrane removed from the hybridization oven was immediately washed in pre-warmed Solution 1 (2X SSC, 0.5%SDS) briefly. After this, the membrane was washed in Solution 1 twice for 15 minutes with gentle shaking. The membrane was then washed twice in Solution 2 (2X SSC and 0.1% SDS) for 15 minutes with gentle shaking. Washing twice with Solution 3 (0.1X SSC and 0.5%) followed for 1 hour each with gentle shaking. The membrane was finally washed in Solution 4 (0.1X SSC) a few times briefly and allowed to air dry before loading into a cassette for autoradiography.

K. Growth curves

1. In LB medium

Cells from *S. typhi* Ty2, SARB 63, SARB 63 revertants, 701Ty and PL27566 were obtained from cell collection frozen at -70°C. Cells were inoculated in LB broth aerated overnight at 37°C. A 0.05% inoculum of each overnight culture was pipetted into Klett flasks containing 10 ml of LB broth. The cells were grown for an hour or until they reached in between 0.2 – 0.3 OD. The cells were again diluted to a 1.0 % inoculum and

readings were taken every 10 minutes for 2 hours or until stationary phase was reached. Cell growth was measured as Klett units, where each Klett unit is equivalent to 1.0×10^7 cells (Curry and Ross, 1998). Growth curves for each strain was obtained in duplicate.

2. In minimal medium

Bacterial cells from *S. typhi* Ty2, SARB 63, SARB 63 revertants, 701Ty and PL27566 were obtained from cell collection frozen at −70°C, inoculated in minimal broth and aerated overnight at 37°C. A 0.1% inoculum of each overnight culture was pipetted into Klett flasks containing 10 ml of minimal broth. These cells were then grown for approximately two hours or until they reached an value in between 0.2 -0.3 OD. The cells were then inoculated to a 1.0% dilution and klett readings were taken over a four hour period. In the first hour, klett readings were taken in 20 minute intervals, in the 2nd and 3rd hours they were taken at10 minute intervals and the last hour klett readings were taken every 20-30 minutes.

L. Statistical tests

The Tukey's multiple comparison test was done to compare the generation times of balanced genomes, unbalanced genomes and revertant strains grown in LB medium and minimal medium. The Tukey's test conducted pair wise comparisons between strains found within each of the balanced, unbalanced and revertants groups (Linton and Harder,

1998). It also compared all strains between the three groups in a pair-wise fashion. In the

Tukey's test, if statistical comparison between two groups (ex. average generation times

of strains with balanced and unbalanced genomes) show that the probability of the two

groups being the same is more than 0.05 (5%), then there is no statistical difference. If

this probability for being the same is less than 0.05 or 5% then there is statistical

difference. Furthermore, the smaller the value is from 0.05, the higher the significance

because the lower the probability is of the two groups being the same.

M. Illustrations and Graphs

Illustrations were done using various programs. Polaroid pictures of gels were

scanned and modified in Adobe Photoshop 4.0, they were then imported to Microsoft

Word 98. Illustrations were made in Adobe Photoshop 4.0 and again imported to

Microsoft Word 98. The graphs were created in Microsoft Excel 98.

CHAPTER 4: RESULTS- Conservation of genome size in *S. typhi*

A. Isolation of the I-*Ceu*I A fragment from *Salmonella typhi* strains

In previously reported work, the genomic DNA of 144 wild-type *S. typhi* strains was isolated, digested with I-*Ceu*I restriction endonuclease and run on PFGE (Liu et al., 1999). Since I-*Ceu*I cuts once within the seven *rrn* operons of *S. typhi*, seven I-*Ceu*I restriction fragments were obtained from each of these strains. These seven I-*Ceu*I fragments along with some partially digested I-*Ceu*I fragments are shown in Figure 4.1a. The size of the restriction fragments was determined using a Lambda ladder PFG marker (NE Biolabs), which ranges from 48.5 kb to 1018.5 kb, with 48.5 kb divisions. It can be observed in Figure 4.1a that although the sizes of the I-*Ceu*I B, C, D, E, F and G can be determined using the Lambda ladder, the same cannot be easily done for the I-*Ceu*I A fragment. This is because the sizes of the I-*Ceu*I B, C, D, E, F and G fragments fall within the range of the ladder, whereas the I-*Ceu*I A fragment does not. In addition, it is known that the I-*Ceu*I A fragment for *S. typhi* Ty2 is 2400kb (Liu and Sanderson, 1995c), and this size is too large for accurate separation of DNA fragments with the PFGE equipment available. Consequently, to determine the size of the I-*Ceu*I A fragment, it was first separated (Figure 4.1b) and then agarose with the bands was excised from a pulsed field gel (each strain is ran in three adjacent lanes, to provide enough DNA for subsequent analysis), re-digested with either *Xba*I, *Avr*II or *Spe*I, end-labeled with α-^{32}P

Figure 4.1: A pulsed field gel containing I-*Ceu*I digested fragments from the genome of several *S. typhi* strains.

a. *S. typhi* Ty2 genomic DNA partially digested with I-*Ceu*I restriction endonuclease and separated on PFGE. A 0.7% agarose gel containing ethidium bromide was used so that the I-*Ceu*I fragments could be visualized in UV light. The running conditions were as follows: 50s-90s for 18 hours at 120° and 180V, 8s-10s for 11hours at 120° and 180V, 50s-60s for 10 hours at 120° and 180V, 100s-152s for 10.5 hours at 120° and 180V and 80s-180s for 13 hours at 120° and 180V. The Lambda ladder was run along side so that the fragment sizes of *S. typhi* Ty2 could be determined. The I-*Ceu*I fragments, and their respective sizes are given beside the lane for *S. typhi* Ty2. The sizes for the Lambda ladder fragments are also given. Sizes are in kilobases (kb). b. Genomic DNA from *S. typhi* strains completely digested with I-*Ceu*I restriction endonuclease and separated on PFGE so that the I-*Ceu*I A fragment could be isolated and excised. The run conditions were as follows: 30s-150s for 12 hours at 120° and 180V and 90s-120s for 12 hours at 110° and 180V. Other I-*Ceu*I fragments, including the I-*Ceu*I A fragment, on the gel are also shown and the sizes for these fragments are given.

a
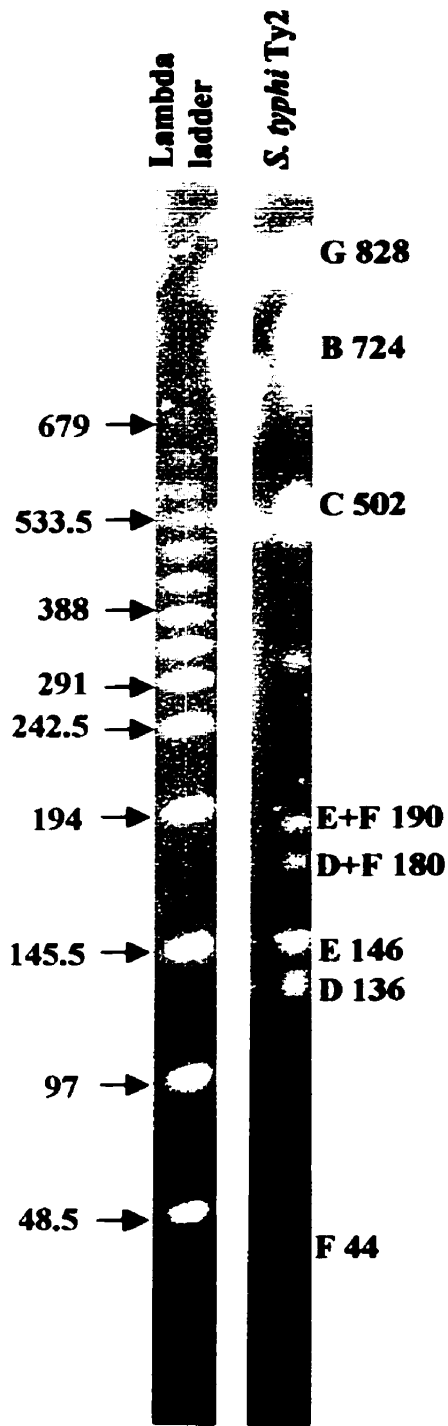
b
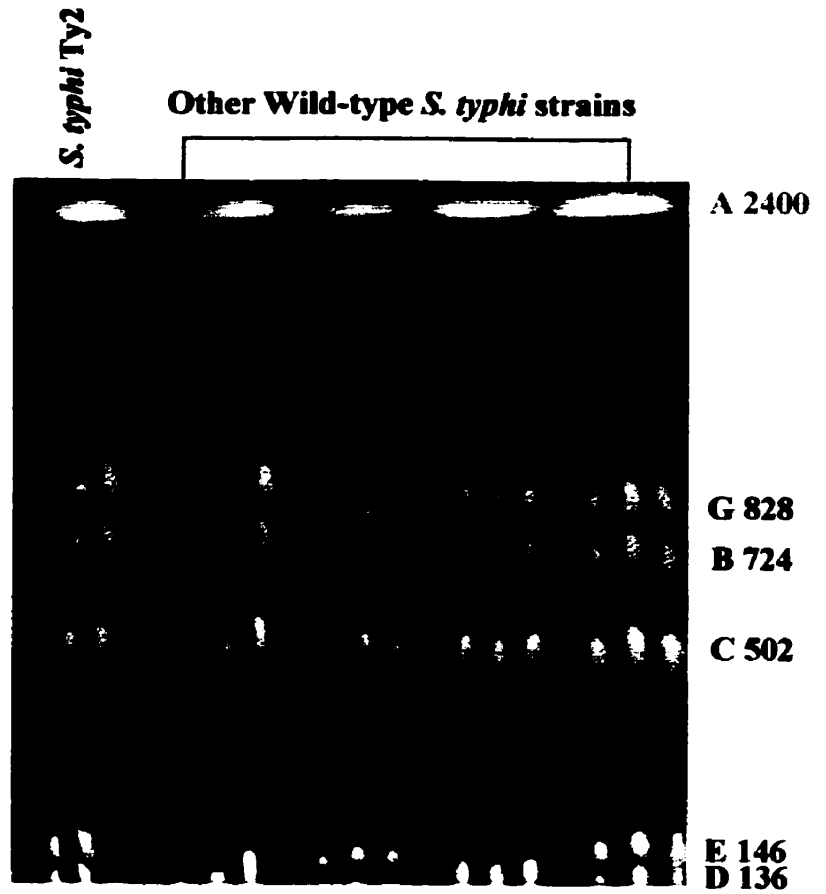
dCTP and separated on PFGE. The purpose of the re-digestion was to obtain smaller DNA pieces, so that sizes can be easily resolved and the total size of the I-*Ceu*I A fragment can be determined. Autoradiographs of *Xba*I, *Avr*II and *Spe*I digested, and α-^{32}P dCTP end-labeled, I-*Ceu*I A fragment of *S. typhi* strains are depicted in Figures 4.2, 4.3 and 4.4. It was necessary to obtain all three restriction endonuclease cleavage patterns because often a single endonuclease restriction pattern will have areas in which more than one band is present. In such a case, it would difficult to resolve exactly how many bands are actually present within this area and as a consequence, large additions or deletions of DNA may be easily overlooked. The following three sections (B, C and D) describe the size of the I-*Ceu*I A fragment in three representative wild-type strains, ST 1002, ST 24A and PNG 31, and compare them to *S. typhi* Ty2.

B. A comparison of the size of the I-*Ceu*I A fragment between *S. typhi* Ty2 and ST 1002

ST 1002 is an example of a wild-type *S. typhi* strain, which appears to have an I-*Ceu*I A fragment that is identical in size to *S. typhi* Ty2. When comparing the end-labeled *Avr*II fragments of ST 1002 to *S. typhi* Ty2, as depicted in Figure 4.2a, it can be clearly seen that there are no differences in the banding pattern of ST 1002 and the size of all the *Avr*II fragments remain the same. Similar results are observed when comparing the *Spe*I cleavage patterns of ST 1002 and *S. typhi* Ty2, as seen in Figure 4.2b. The banding patterns for these two strains are exactly alike, suggesting that their I-*Ceu*I A

Figure 4.2: *AvrII* and *SpeI* digested, and ^{32}P end-labeled I-*Ceu* I A fragments of *S. typhi* Ty2 and strain ST 1002.

a.The I-*CeuI* A fragments of *S. typhi* Ty2 and ST 1002 were excised from the gel following PFGE, digested with *AvrII*, end-labeled with α-[32]P dCTP and run on PFGE for a 30-35 hours. Pulse times were set such that fragments between 300 kb to 900 kb could be separated. **b.** The I-*CeuI* A fragments of *S. typhi* Ty2 and ST 1002 were digested with *SpeI*, end-labeled with α-[32]P dCTP and run on PFGE for a 50-55 hours. Pulse times were set such that fragments between 20 kb to 200 kb could be separated. The I-*CeuI* A fragment of *S. typhi* Ty2 was run to act as a standard to determine sizes in other *S. typhi* strains, since the sizes of its own *AvrII* fragments are known (Liu and Sanderson, 1995c). Sizes are represented in kilobases (kb) and the letters corresponding to each band are taken from the *AvrII* fragments within the I-*CeuI* A fragment of *S. typhi* Ty2 as illustrated in Figure 2.3 (Liu and Sanderson, 1995c). Letters with an apostrophe are fragments, which have been truncated, due to the presence of a *rrn* operon within the fragment.

a

ST 1002

S. typhi Ty2

F 830

K 600

C' 335

G 265

J 115
L 100
E 68
D, H 32
I 23

b

ST 1002 S. typhi Ty2

Z 380

I 195 K 276
N 193 H 270
M 180

Y 140

J 121

G 104

Q 80
U 70 O 77 U 70

AA' 60
V 52

T 45
F' 40
L 33

P 31 R 30 P 31
W 25 W 25

X 12
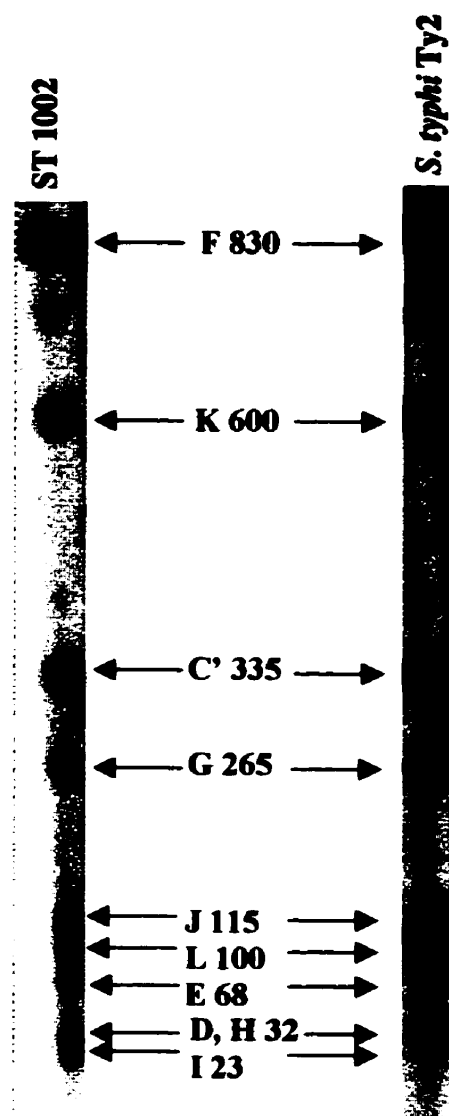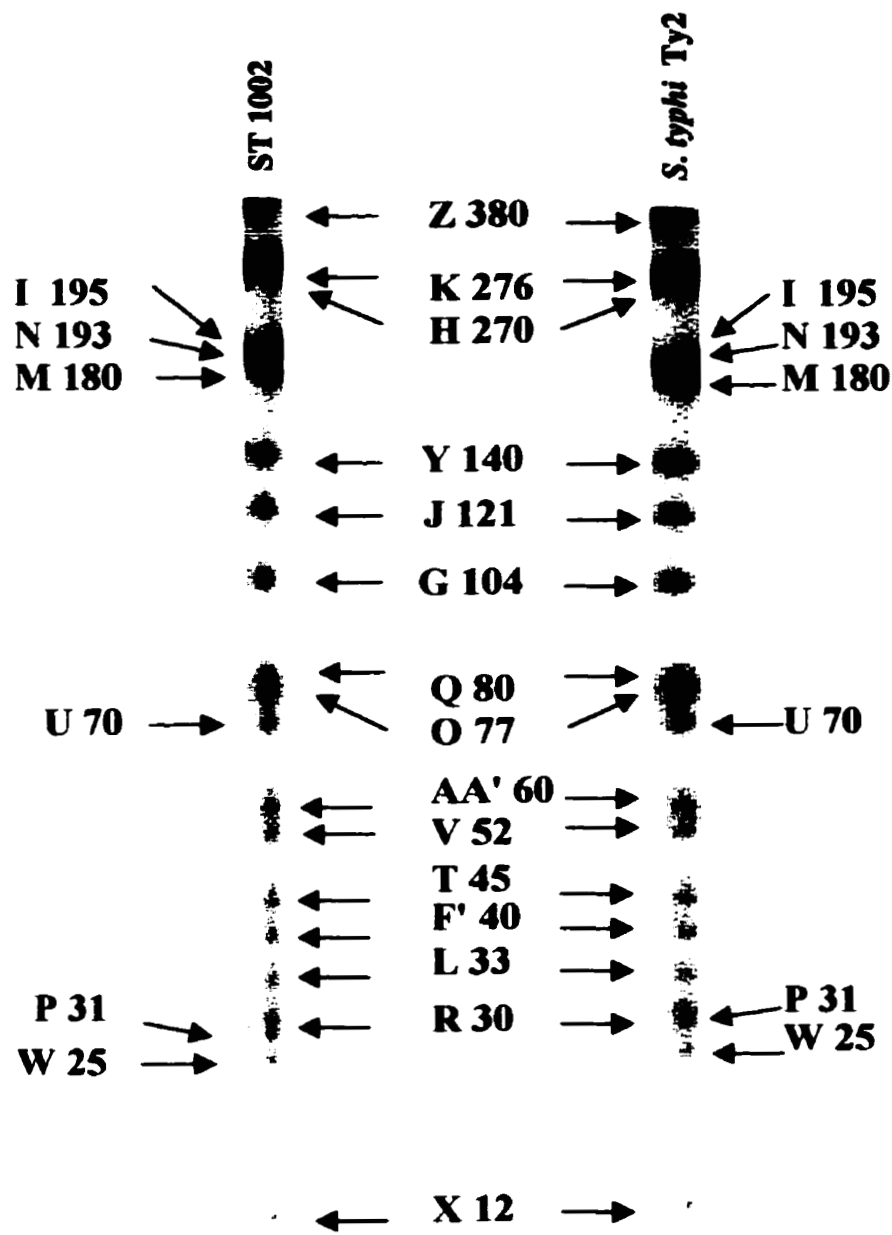
Figure 4.3: *Xba*I and *Avr*II digested, and ³²P end-labeled I-*Ceu* I A fragments of *S. typhi* Ty2 and strain ST 24A.

a. The I-*Ceu*I A fragments of *S. typhi* Ty2 and ST 24A were excised from the gel following PFGE, digested with *Xba*I, end-labeled with α-³²P dCTP and run on PFGE for a 30-35 hours. Pulse times were set such that fragments between 250 kb to 600 kb could be separated. b. The I-*Ceu*I A fragment of *S. typhi* Ty2 and ST 24A was digested with *Avr*II, end-labeled with α-³²P dCTP and run on PFGE for a 30-35 hours. Pulse times were set such that fragments between 300 kb to 900 kb could be separated. The I-*Ceu*I A fragment of *S. typhi* Ty2 was run to act as a standard to determine sizes in other *S. typhi* strains, since the sizes of its own *Xba*I and *Avr*II fragments are known (Liu and Sanderson, 1995c). Sizes are represented in kilobases (kb) and the letters corresponding to each band are taken from the *Xba*I and *Avr*II fragments within the I-*Ceu*I A fragment of *S. typhi* Ty2 as illustrated in Figure 2.3 (Liu and Sanderson, 1995c). Letters with an apostrophe are fragments, which have been truncated, due to the presence of a *rrn* operon within the fragment.

a

b

ST 24A

S. typhi Ty2

← F 830 →

K (600) is
smaller
by 20 kb

K 600 →

← C' 335 →

Extra
175 kb
fragment

← G 265 →

Missing
J (115)

J 115 →
← L 100 →
← E 68 →

Missing
D, H (32)
and I (23)

D, H 32 →
← I 23
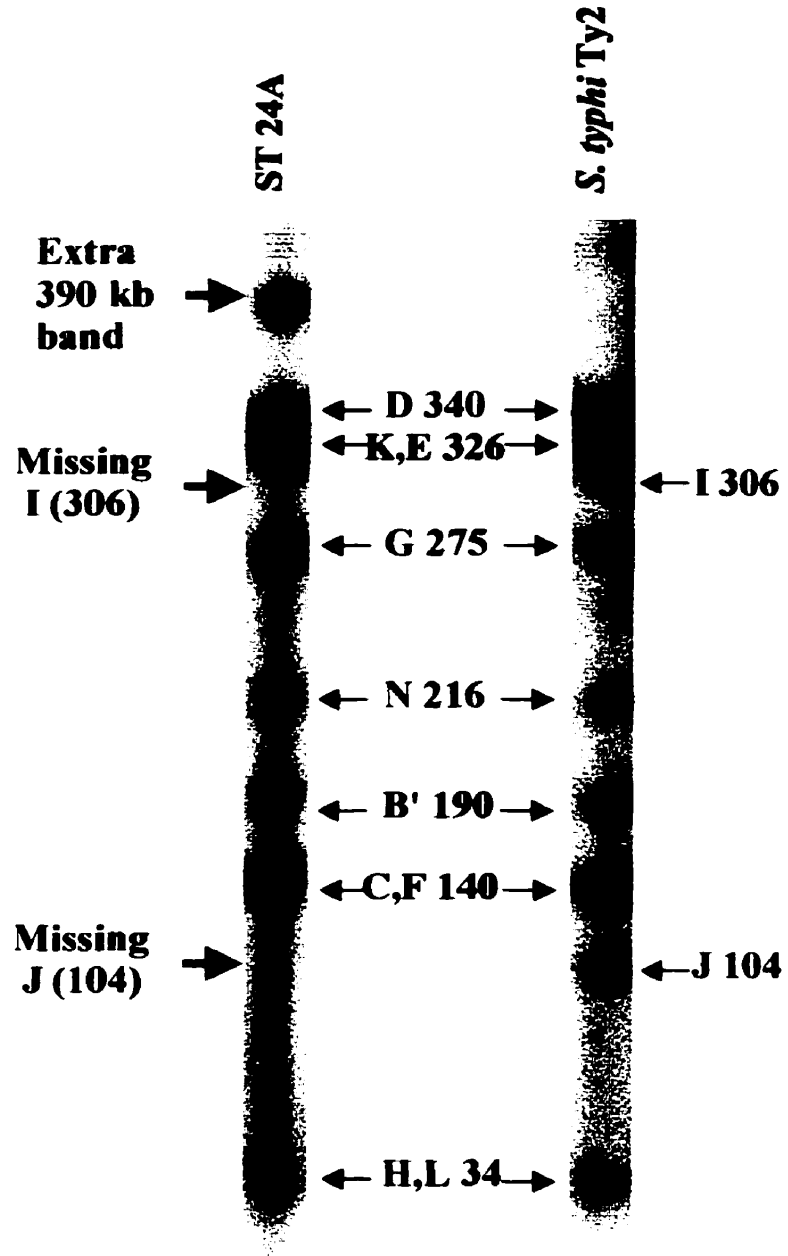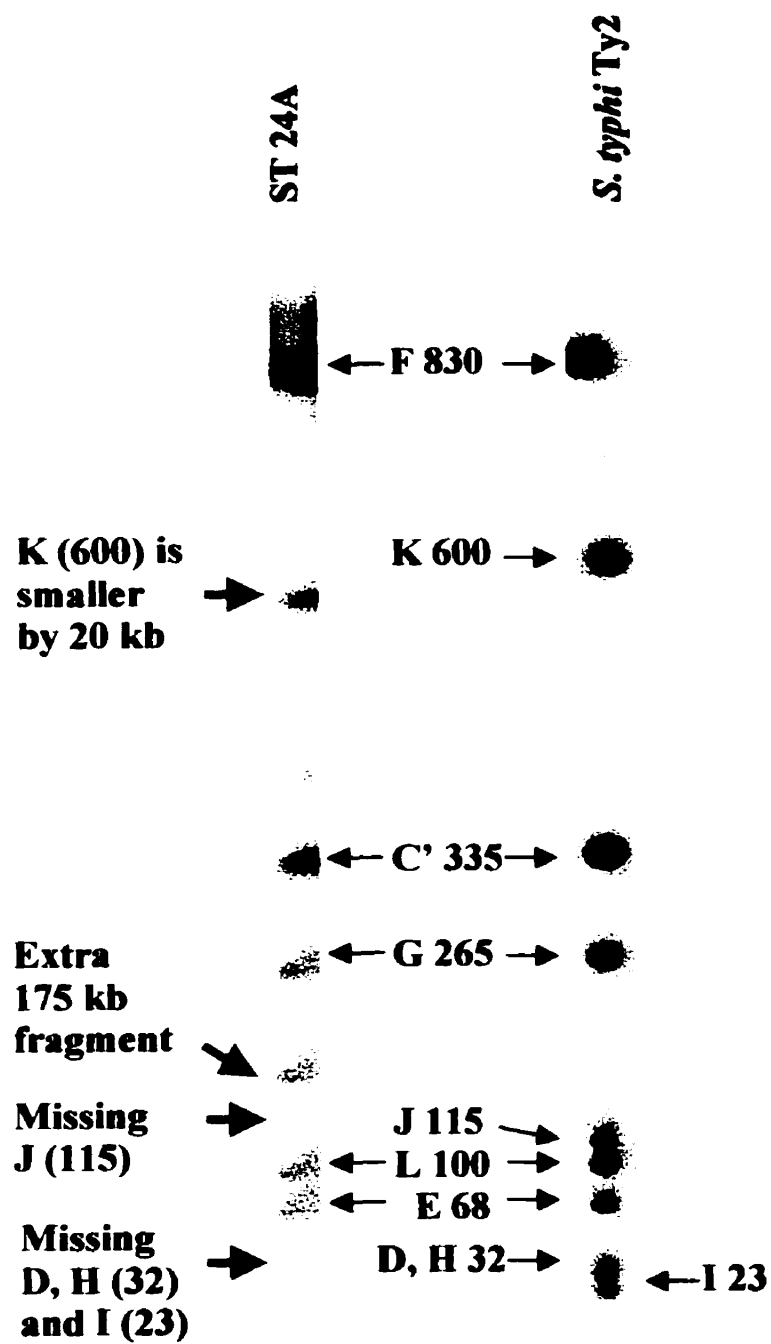
Figure 4.4: *Avr*II and *Spe*I digested, and ^{32}P end-labeled I-*Ceu* I A fragments of *S. typhi* Ty2 and PNG31.

a. The I-*Ceu*I A fragments of *S. typhi* Ty2 and PNG31 were excised from the gel following PFGE, digested with *Avr*II, end-labeled with α-[32]P dCTP and run on PFGE for a 30-35 hours. Pulse times were set such that fragments between 300 kb to 900 kb could be separated. b. The I-*Ceu*I A fragment of *S. typhi* Ty2 and PNG31 was digested with *Spe*I, end-labeled with α-[32]P dCTP and run on PFGE for a 50-55 hours. Pulse times were set such that fragments between 20 kb to 200 kb could be separated. The I-*Ceu*I A fragment of *S. typhi* Ty2 was run to act as a standard to determine sizes in other *S. typhi* strains, since the sizes of its own *Avr*II and *Spe*I fragments are known (Liu and Sanderson, 1995c). Sizes are represented in kilobases (kb) and the letters corresponding to each band are taken from the *Avr*II and *Spe*I fragments within the I-*Ceu*I A fragment of *S. typhi* Ty2 as illustrated in Figure 2.3 (Liu and Sanderson, 1995c). Letters with an apostrophe are fragments, which have been truncated, due to the presence of a *rrn* operon within the fragment.

a

PNG 31

S. typhi Ty2

← F 830 → F 830

K (600) is bigger by 20 kb → K 600 → K 600

← C' 335 → C' 335

← G 265 → G 265

Extra 175 kb fragment →

Missing J (115) → J 115 → J 115
L 100 → L 100
E 68 → E 68

Missing H,D (32) and I (23) → D,H 32 → D,H 32
← I 23

b

Z (380) is
missing

Extra 210
kb band

Extra 185
kb band

U (70)
smaller
by 5 kb

V (52)
smaller
by 5 kb

P 31
W 25

PNG 31

S. typhi Ty2

Z 380

K 276
H 270

I 195
N 193
M 180

Y 140

J 121

G 104

Q 80
O 77

U 70

AA' 60
V 52

T 45
F' 40
L 33

R 30

P 31
W 25

X 12
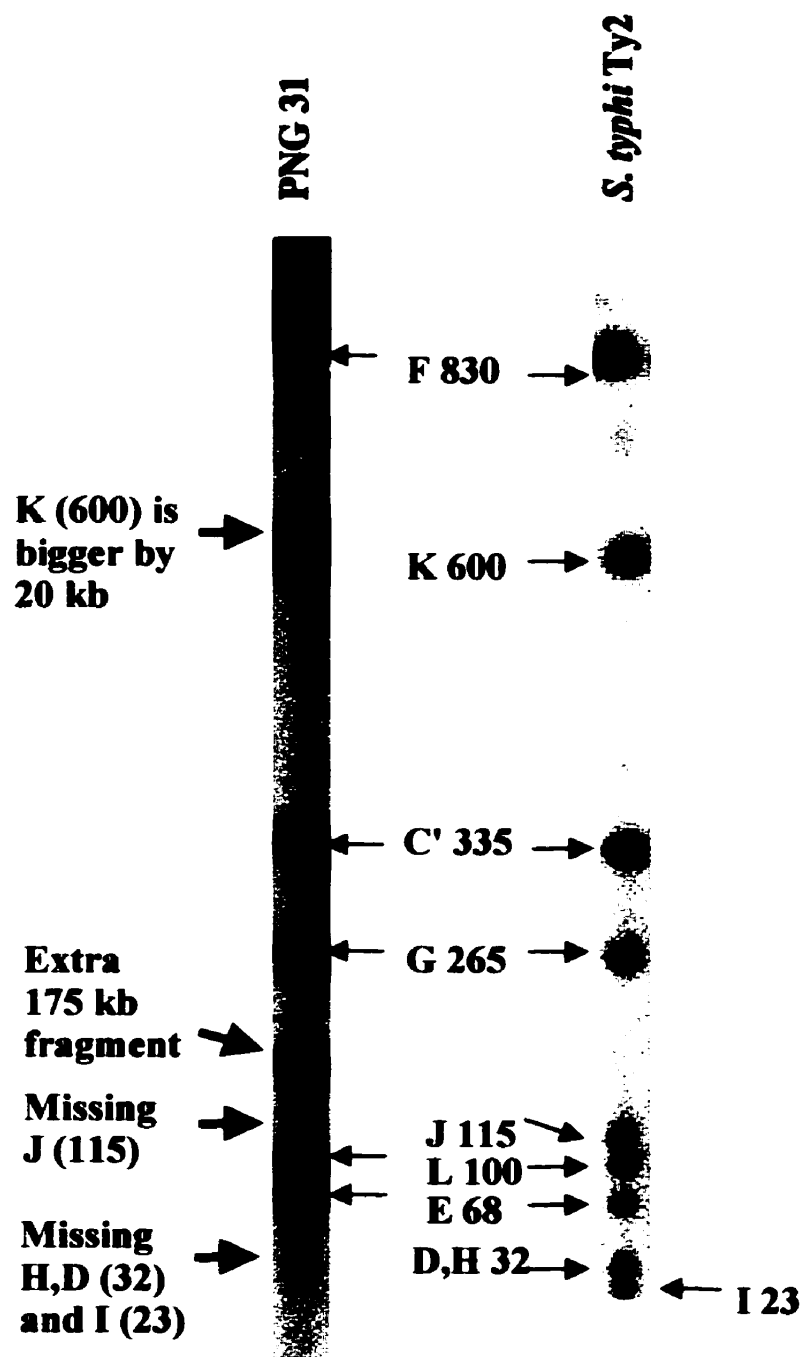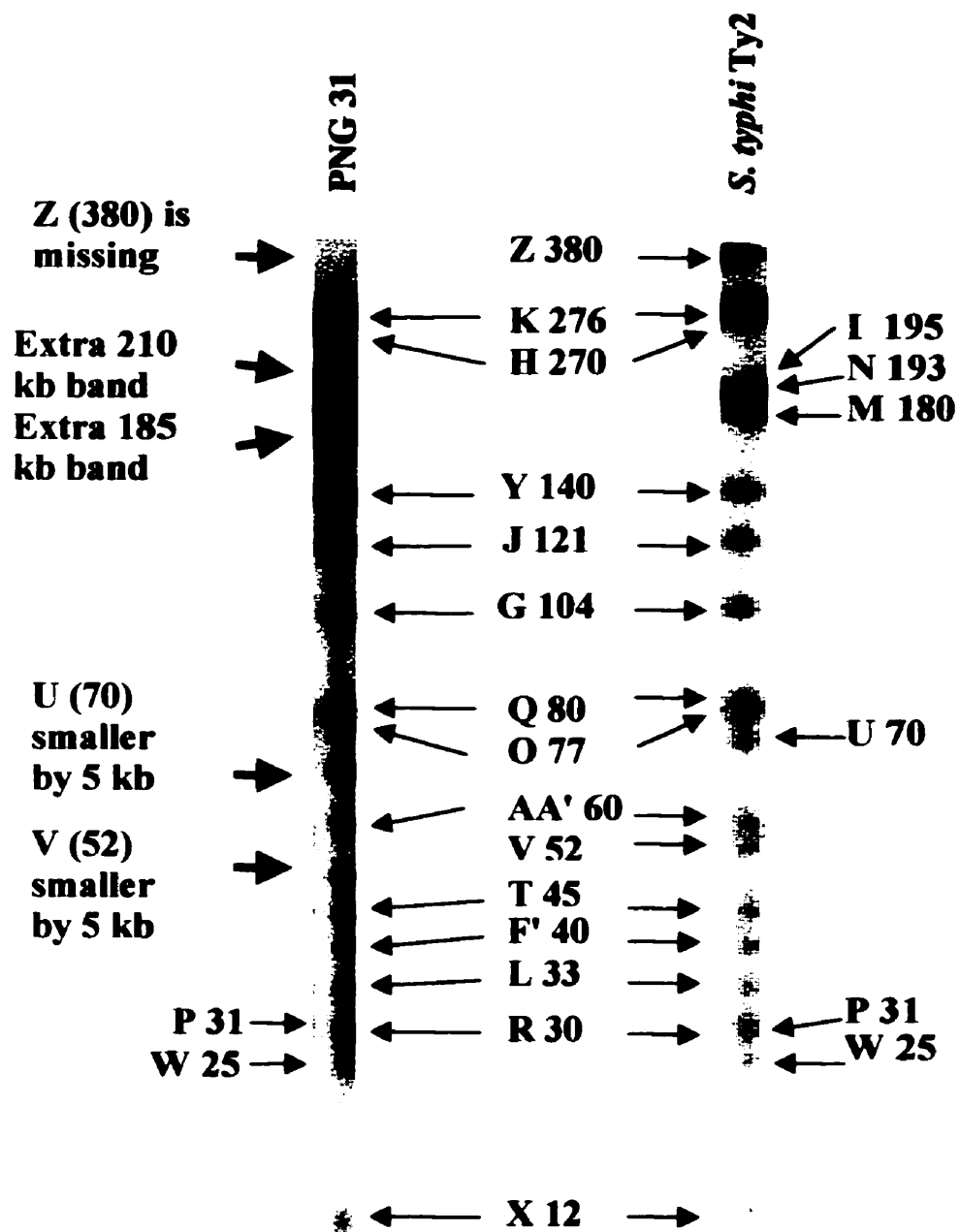
fragment are similar in size. The *Xba*I cleavage pattern for the I-*Ceu*I A fragment of ST 1002 is also unchanged from that of *S. typhi* Ty2, however, the data are not shown . Thus, it can be concluded that the size of the I-*Ceu*I A fragment of ST 1002 is 2400 kb, which is similar to that of *S. typhi* Ty2.

C. A comparison of the size of the I-*Ceu*I A fragment between *S. typhi* Ty2 and ST 24A

1. *Xba*I cleavage pattern of ST 24A

*Xba*I fragments within the I-*Ceu*I A fragment of *S. typhi* Ty2, previously studied in detail (Liu and Sanderson, 1995c), and ST24A are fairly comparable in size, however there are some changes. The *Xba*I cleavage pattern shows that the I-*Ceu*I A fragment of *S. typhi* Ty2 is 2400 kb, while that of ST 24A is 2380 kb (Figure 4.3a and Table 4.1); this is based on the following data. In ST 24a the J fragment (104 kb) and the I fragment (306 kb) are missing, suggesting that 410 kb of DNA has been deleted. However, there is an additional 390 kb *Xba*I fragment, which approximately makes up for the loss of the 410 kb. It is possible that the unaccounted 20 kb is within the cluster of *Xba*I fragments D, K and E, since exact size cannot be determined for these three fragments by looking at the *Xba*I cleavage pattern alone. As seen in Figure 4.3a, these bands appear as a single dark band on the gel but in fact there are three fragments within this region ranging in size from 325 kb to 340 kb. Thus, it is difficult to ascertain exact size of these bands within this area and for this reason, more than one endonuclease cleavage pattern is obtained for

**Table 4.1: A comparison of the *Xba*I restriction endonuclease fragments of the I-
*Ceu*I A fragment between *S. typhi* Ty2 and strain ST 24A**

*Xba*I FRAGMENTS	SIZE in *S. typhi* Ty2 (kb)	SIZE in ST 24A (kb)	NET GAIN (+) or LOSS (-) of DNA (kb)
D[a]	340	340	0
K[a]	326	326	0
E[a]	325	325	0
I	306	Not present	- 306
G	275	275	0
N	216	216	0
B	190	190	0
C	140	140	0
F	140	140	0
J	104	Not present	-104
H	34	34	0
L	34	34	0
Extra bands		390	+390
Total size (kb)	2400	2380	-20

[a] The exact sizes of the *Xba*I fragments D, K and E are difficult to determine. As well, evaluating changes within this area is also difficult to ascertain. This is because on an end-labeled, autoradiograph of a *Xba*I cleavage pattern, these three fragments appear as a single band even though its sizes range from 325 kb to 340 kb. Thus, it is difficult to determine accurately the size of each single band.

each *S. typhi* strain in order to accurately determine size of the I-*Ceu*I A fragment. It should be noted that sizes given in Figure 4.3a and Table 4.1 are not exact since commonly used standards like Lambda ladder cannot be successfully end-labeled, due to technical difficulties. Instead, sizes are determined by using known *S. typhi* Ty2 cleavage fragments.

2. *Avr*II cleavage pattern of ST 24A

The *Avr*II cleavage pattern of the I-*Ceu*I A fragment of ST 24A also shows changes, when compared to *S. typhi* Ty2. The *Avr*II cleavage pattern depicted in Figure 4.3b and Table 4.2 illustrates that the size of I-*Ceu*I A in *S. typhi* Ty2 is 2400 kb and in ST24A is 2353 kb. It can be observed that *Avr*II fragments D and H (32 kb), I (23 kb) and J (115 kb) are missing. In addition, the *Avr*II K fragment, which is adjacent to *Avr*II J on the chromosome, is approximately 20 kb smaller. Thus, it appears that there is a total loss of a 222 kb of DNA, but this is counteracted by the presence of an extra 175 kb fragment. The *Avr*II cleavage pattern is clearer to analyze than that of *Xba*I because bands of similar size are not present. Despite this, the missing 47 kb cannot be accounted for, which suggests that approximately 47 kb of DNA, or if not at least the 20 kb that could not be explained by the *Xba*I cleavage pattern either, is missing from ST 24A.

3. *Spe*I cleavage pattern of ST 24A

Table 4.3 compares the *Spe*I restriction fragments of *S. typhi* Ty2 and ST 24A, which shows that the I-*Ceu*I A fragment of *S. typhi* Ty2 is 2400 kb and ST 24A is 2488 kb. The gel from which these conclusions are drawn is not shown. The *Spe*I cleavage pattern is harder to interpret since there is a greater number of bands to contend with and bands of

Table 4.2: A comparison of the *Avr*II restriction endonuclease fragments of the I-*Ceu*I A fragment between *S. typhi* Ty2 and strain ST 24A

*Avr*II FRAGMENTS	SIZE in *S. typhi* Ty2 (kb)	SIZE in ST 24A (kb)	NET GAIN (+) or LOSS (-) of DNA (kb)
F	830	830	0
K	600	580	-20
C	335	335	0
G	265	265	0
J	115	Not present	-115
L	100	100	0
E	68	68	0
D[a]	32	Not present	-32
H	32	Not present	-32
I	23	Not present	-23
Extra bands		175	+175
TOTAL SIZE (kb)	2400	2353	-47

[a] The position of the *Avr*II D in the I-*Ceu*I A fragment is not known

Table 4.3: A comparison of the *Spe*I restriction endonuclease fragments of the I-*Ceu*I A fragment between *S. typhi* Ty2 and strain ST 24A

*Spe*I FRAGMENTS	SIZE in *S. typhi* Ty2 (kb)	SIZE in ST 24A (kb)	NET GAIN (+) or LOSS (-) of DNA (kb)
Z	380	380	0
K[a]	276	276	0
H[a]	270	270	0
I[a]	195	195	0
N[a]	193	193	0
M[a]	180	180	0
Y	140	140	0
J	121	121	0
G	104	104	0
Q	80	80	0
O	77	77	0
U	70	Not present	-70
AA'	60	60	0
V[b]	52	Not present	-52
T	45	45	0
F'	40	40	0
L	33	33	0
R[b]	31	31	0
P[b]	29	29	0
W[b]	25	25	0
Extra bands		110	+110
Total size (kb)	2400	2388	-12

[a] The exact sizes of the *Spe*I fragments K, H and I, N, M are difficult to determine. As well, evaluating changes within this area is also difficult to ascertain. This is because on an end-labeled, autoradiograph of a *Spe*I cleavage pattern, these fragments appears as single bands. One set harboring the K and H bands in the 270-280 kb region and the other set containing the I, N and M bands within the 180-200 kb region. Thus, it is difficult to determine accurately what size each single band is within each cluster.

[b] The exact positions of the *Spe*I P,R, V and W are not known.

similar size are common. However, it appears that only minor changes occur in the *Spe*I cleavage pattern of ST 24A. According to Table 4.3, in ST 24A the *Spe*I U (70 kb) and the V (52 kb) bands have disappeared, this adds up to 122 kb of lost DNA. As well, there is the appearance of a 110 kb band, which approximately makes up for the above missing DNA. The remaining 10 kb could be in the fragments adjacent to either *Spe*I U or V. However, the *S. typhi* Ty2 cleavage map in Figure 2.4 shows that the *Spe*I fragments U and V are within a region of the map in which the restriction fragments are not well defined; this is due to the clutter of numerous small fragments in this area. Therefore, it is difficult to ascertain exact changes in size within this region with endonuclease digestion and end-labeling method alone. All that can be said is that according to the SpeI cleavage pattern of ST 24A, 12 kb of DNA is lost from the I-*Ceu*I A fragment.

When comparing the I-*Ceu*I A fragment of ST 24A with *S. typhi* Ty2, a small variation in size is observed. The I-*Ceu*I A fragment of ST 24A appears to be on average approximately 26 kb smaller than that of *S. typhi* Ty2. It is consistent in all three cleavage patterns that at least 10 kb is lost, with the *Avr*II cleavage pattern showing the greatest loss of about 43 kb. These sizes are shown for *S. typhi* Ty2 (Table 4.4, row 1) and strain ST 24A (Table 4.4, row 4), for each of the three enzymes; the overall size of the I-CeuI A fragment in ST 24A is set at 2374 kb.

63

Table 4.4: Comparison of total genome size for various *S. typhi* strains as assessed by the sizes of their I-*Ceu*I fragments

No.	*S. typhi* Strain No.	SGSC No.	Source[a]	SIZE OF I-*Ceu*I FRAGMENTS[c] (kb)						A — Total fragment size using:				Total Genome size (kb)
				B	C	D	E	F	G	*Xba* I	*Avr* I	*Spe* I	Total Size (avg.)[b]	
1	Ty2	2408	Home	724	502	136	146	44	828	2400	2400	2400	2400	4780
2	PNG31	3194	Pang	724	502	136	146	44	828	2410	2393	2405	2403	4783
3	PNG32	3195	Pang	724	502	136	146	44	828	2410	2493	2405	2403	4783
4	ST 24A	2771	Pang	724	502	136	146	44	828	2380	2353	2388	2374	4754
5	ST308	2775	Pang	724	502	136	146	44	828	2400	2400	2400	2400	4780
6	ST318	2778	Pang	724	502	136	146	44	828	2400	2380	2415	2398	4778
7	ST309	2779	Pang	739	502	136	146	44	828	2400	2400	2400	2400	4795
8	ST495	2781	Pang	724	502	136	146	44	828	2370	2380	2390	2380	4760
9	In15	2783	Pang	724	502	136	146	44	828	2410	2493	2390	2403	4783

SIZE OF I-CeuI FRAGMENTS (kb)

	S. typhi Strain No.	SGSC No.	Source	B	C	D	E	F	G	Total fragment size using: A				Total Genome size (kb)
										Xba I	Avr II	Spe I	Total Size (avg.)	
10	ST1006	2780	Pang	724	502	136	186	44	828	2380	2353	2388	2374	4754
11	ST143	2773	Pang	724	502	136	146	44	282	2410	2403	2405	2406	4786
12	CC6	3198	Pang	724	502	136	146	44	828	2380	2343	2388	2370	4750
13	ST1002	2776	Pang	724	502	136	146	44	828	2400	2400	2400	2400	4780
14	In20	3191	Pang	724	502	136	186	44	828	2380	2390	2390	2387	4807
15	T189	3187	Pang	724	502	136	146	44	828	2380	2353	2390	2374	4757
16	ST1	2728	Pang	724	502	136	146	44	698	2400	2400	2400	2400	4650
17	In4	3190	Pang	804	502	136	146	44	828	2380	2360	2388	2376	4836
18	3123	3184	Pang	724	502	136	146	44	828	2410	2385	2400	2398	4778
19	3125	3185	Pang	724	502	136	146	44	828	2470	2475	2440	2462	4842
20	In24	3192	Pang	724	502	136	146	44	828	2380	2363	2400	2381	4761

	S. typhi Strain No.	SGSC No.	Source	SIZE OF I-*CeuI* FRAGMENTS (kb)						A — Total fragment size using:				Total Genome size (kb)
				B	C	D	E	F	G	*Xba* I	*Avr* II	*Spe* I	Total Size (avg.)	
21	SARB 63	2520	RKS	724	502	136	146	44	828	2400	2363	2400	2388	4768
22	PL257 66	2990	Prov. eL	744	502	136	146	44	828	2385	2253	2390	2343	4743
23	PNG 30	3193	Pang	724	502	136	146	44	828	2410	2393	2405	2403	4783
24	SA 4865	2684	Hone	724	502	136	146	44	578	2380	2353	2390	2374	4504
25	ISP 1820	2406	Hone	724	502	136	146	44	828	2380	2353	2390	2374	4754
26	26T20	3143	LCDC	724	502	136	146	44	828	2380	2353	2390	2374	4754
27	26T38	3161	LCDC	724	502	136	146	44	828	2380	2395	2385	2387	4767
28	26T50	3173	LCDC	804	502	136	146	44	828	2360	2370	2310	2347	4807

Malaysia; Dr. Robert K. Selander at the University of Pennsylvania, PA; Southern Alberta Provincial Lab, Edmonton, AB; Dr. Bruce Stocker at Stanford University and the LCDC (Laboratory Center for Disease Control) Ottawa, ON..

[b]The size of the I-*Ceu*I A fragment as determined by the different restriction endonuclease digests, was resolved by comparing corresponding *S. typhi* Ty2 fragments, whose size was previously known. Since it was not possible to run a standard lambda ladder to determine accurate size of the restriction endonuclease fragments (because end-labeling was used), the above given sizes for the 1-*Ceu*I A fragment should not be accepted as absolute values.

[c] The size of the I-*Ceu*I B, C, D, E, F and G fragments was determined by comparison to lambda standards, and to I-*Ceu*I fragments from S. typhi Ty2

D. A comparison of the size of the I-*Ceu*I A fragment between *S. typhi* Ty2 and strain PNG31

1. *Xba*I cleavage pattern of PNG 31

Only minor changes are observed in the *Xba*I cleavage pattern of the PNG31 I-*Ceu*I A fragment, as seen in Table 4.5. The gel from which these conclusions are drawn is not shown. It appears that the *Xba*I J fragment of PNG31 is 10 kb smaller than in *S. typhi* Ty2, while the *Xba*I K fragment, adjacent to J, is 20 kb larger in PNG31 than in *S. typhi* Ty2. Thus, according to the *Xba*I cleavage pattern, the I-*Ceu*I A fragment of PNG31 is 2410 kb, which is a total gain of 10 kb. As noted in the previous section, the *Xba*I K fragment is part of a three band cluster within the 320-340 kb region. Thus, the size the K fragment cannot be as accurately resolved as other single bands present in the *Xba*I cleavage pattern.

2. *Avr*II cleavage pattern of PNG31

The *Avr*II cleavage pattern for PNG31 indicates that the I-*Ceu*I A fragment is around 2393 kb, as represented in Figure 4.4a and Table 4.6. As in the *Avr*II cleavage pattern of ST 24A described above, *Avr*II fragments D and H (32 kb), I (23 kb) and J (115 kb) are missing in PNG31 and there is an additional 175 kb fragment present. However, the difference is that the *Avr*II K fragment, adjacent to J on the chromosome, is now 20 kb larger instead of smaller. Therefore, according to this *Avr*II cleavage pattern, PNG31 only appears to have lost 5-10 kb.

Table 4.5: A comparison of the *Xba*I restriction endonuclease fragments of the I-

CeuI A fragment between *S. typhi* Ty2 and strain PNG31

*Xba*I FRAGMENTS	SIZE in *S. typhi* Ty2 (kb)	SIZE in PNG31 (kb)	NET GAIN (+) or LOSS (-) of DNA (kb)
D[a]	340	340	0
K[a]	326	346	+20
E[a]	325	325	0
I	306	306	0
G	275	275	0
N	216	216	0
B	190	190	0
C	140	140	0
F	140	140	0
J	104	94	-10
H	34	34	0
L	34	34	0
Extra bands			
Total size (kb)	2400	2410	+10

[a] The exact sizes of the *Xba*I fragments D, K and E are difficult to determine. As well, evaluating changes within this area is also difficult to ascertain. This is because on an end-labeled, autoradiograph of a *Xba*I cleavage pattern, these three fragments appear as a single band even though its sizes range from 325 kb to 340 kb. Thus, it is difficult to determine accurately what the size of each single band.

Table 4.6: A comparison of the *Avr*II restriction endonuclease fragments of the I-

***Ceu*I A fragment between *S. typhi* Ty2 and strain PNG31**

*Avr*II FRAGMENTS	SIZE in *S. typhi* Ty2 (kb)	SIZE in PNG31 (kb)	NET GAIN (+) or LOSS (-) of DNA (kb)
F	830	830	0
K	600	620	+20
C	335	335	0
G	265	265	0
J	115	Not present	-115
L	100	100	0
E	68	68	0
D[a]	32	Not present	-32
H	32	Not present	-32
I	23	Not present	-23
Extra bands			+175
TOTAL SIZE (kb)	2400	2393	-7

[a] The position of the *Avr*II D in the I-*Ceu*I A fragment is not known in *S. typhi* strains.

3. *Spe*I cleavage pattern of PNG31

The size difference in the I-*Ceu*I A fragment between *S. typhi* Ty2 and PNG31 is negligible, since according to the *Spe*I cleavage pattern, the size of the I-*Ceu*I A fragment of PNG31 is 2405 kb, instead of 2400 kb. Despite this, there are a few changes observed in the *Spe*I restriction cleavage pattern of PNG31, when compared to *S. typhi* Ty2, as seen in Figure 4.4b and Table 4.7. The *Spe*I Z fragment (380 kb) in PNG31 is missing but in its place there are two additional bands which are 210 kb and 185 kb in size. Furthermore, the *Spe*I U and V fragments are both around 5-10 kb smaller. In total, it appears that the I-*Ceu*I A fragment of PNG31 is at most only 5 kb larger than *S. typhi* Ty2.

In light of the information obtained from the *Xba*I, *Spe*I and *Avr*II cleavage patterns of PNG31, it can be concluded that on average, the size of its I-*Ceu*I A fragment is between 2400-2405 kb. This is consistent with the size of the I-*Ceu*I A of *S. typhi* Ty2, which is 2400 kb. Consequently, the size of the I-*Ceu*I A fragment of PNG31 is arbitrarily determined to be 2403 kb in size (Table 4.4, row 2).

E. Changes in the size of the I-*Ceu*I A fragment and total genome size in wild-type strains of *S. typhi*

1. Size of the I-*Ceu*I A fragment

Using the methods of I-*Ceu*I A excision, re-digestion, α ^{32}P-dCTP end-labeling and PFGE separation described above, 24 other wild-type *S. typhi* strains (in addition to

Table 4.7: A comparison of the *Spe*I restriction endonuclease fragments of the I-

*Ceu*I A fragment between *S. typhi* Ty2 and strain PNG31

*Spe*I FRAGMENTS	SIZE in *S. typhi* Ty2 (kb)	SIZE in PNG31 (kb)	NET GAIN (+) or LOSS (-) of DNA (kb)
Z	380	Not present	**-380**
K[a]	276	276	**0**
H[a]	270	270	**0**
I[a]	195	195	**0**
N[a]	193	193	**0**
M[a]	180	180	**0**
Y	140	140	**0**
J	121	121	**0**
G	104	104	**0**
Q	80	80	**0**
O	77	77	**0**
U	70	65	**-5**
AA'	60	60	**0**
V[b]	52	47	**-5**
T	45	45	**0**
F'	40	40	**0**
L	33	33	**0**
R[b]	30	30	**0**
P[b]	30	30	**0**
W[b]	25	25	**0**
Extra bands (kb)		210 185	**+395**
Total size (kb)	2400	2405	**+5**

[a] The exact sizes of the *Spe*I fragments K, H and I, N, M are difficult to determine. As well, evaluating changes within this area is also difficult to ascertain. This is because on an end-labeled, autoradiograph of a *Spe*I cleavage pattern, these fragments appears as single bands. One set harboring the K and H bands in the 270-280 kb region and the other set containing the I, N and M bands within the 180-200 kb region. Thus, it is difficult to determine accurately what size each single band is within each cluster.

[b] The exact positions of the *Spe*I P,R, V and W are not known.

S. typhi Ty2, ST 1002, ST 24A and PNG 31 already described) were analyzed to determine the average size of the I-*Ceu*I A fragment. Table 4.4 lists sizes of the I-*Ceu*I A fragment obtained from all three *Xba*I, *Avr*II and *Spe*I cleavage patterns, for each wild-type *S. typhi* strain. The gels from which these conclusions are derived are not shown. These sizes were then averaged to give a single number for the size of the I-*Ceu*I A fragment for each strain. It must be noted that the average size recorded for the I-*Ceu*I A fragment cannot be taken as a exact number but shows the size difference from *S. typhi* Ty2. In analyzing this table, it can be seen that the size of the I-*Ceu*I A fragment varies little. Four of the *S. typhi* strains in Table 4.4, (ST 308, row 5; ST 309, row 7; ST 1002 and ST 1, row 16), have *Xba*I, *Avr*II and *Spe*I fragments which are all indistinguishable from *S. typhi* Ty2, indicating a total size of 2400 kb. In the whole set of 28 wild-type strains listed in Table 4.4, the size range is in between 2343 kb - 2462 kb. PL25766 and 26T50 [Table 4.4 (row 22 and 28)] appear to have an average I-*Ceu*I A fragment sizes of 2343 kb and 2347 kb, respectively, while 3125 [Table 4.4 (row 19)] has a larger I-*Ceu*I fragment, which is 2462 kb in size. The rest of the strains have I-*Ceu*I A fragments which fall within 25 kb, either smaller or larger, of the 2400 kb I-*Ceu*I A fragment size, known for *S. typhi* Ty2.

2. Total genome size of *S typhi* strains

In previous studies, the sizes of I-*Ceu*I fragments B, C, D, E, F and G of all 28 wild-type *S. typhi* strains, were determined (Liu et al., 1999). This was done by digesting *S. typhi* genomic DNA with I-*Ceu*I and separating it on PFGE. The I-*Ceu*I cleavage patterns of these wild-type strains were then compared to that of *S. typhi* Ty2, whose I-

*Ceu*I fragment sizes were known (Liu and Sanderson, 1995c). Size of the I-*Ceu*I fragments were also confirmed by comparing them with known fragment sizes of a Lambda ladder. Along with the size of the I-*Ceu*I A fragment, Table 4.4 lists sizes for the other six I-*Ceu*I fragments. It is apparent from Table 4.4 that the I-*Ceu*I C (502 kb), D (136 kb) and F (44 kb) fragments are highly conserved in all the 28 wild-type *S. typhi* strains. The total genome size of *S. typhi* Ty2 is 4780 kb and most of the other *S. typhi* strains analyzed range between 4740 kb to 4840 kb; thus, showing a total range of variation of less than 100 kb.

There are a few exceptions to this, for example, *S. typhi* strains ST 1006 and In20 [Table 4.4 (row 10 and 14)] have a 40 kb larger I-*Ceu*I E fragment that is normally 146 kb in size. ST 309 and PL25766 [Table 4.4 (row 7 and 22)] have I-*Ceu*I B fragments (normally 704 kb), which are 15 and 20 kb larger, respectively. Despite this extra DNA, the total genome size of ST 1006, In20, ST 309 and PL25766 fall within the range mentioned above. Two other *S. typhi* strains have I-*Ceu*I B fragments which are 80 kb larger, these are In4 and 26T50 [Table 4.4 (17 and 28)]. In these strains, the increase in size of the I-*Ceu*I B fragment appears to be compensated by a decrease in size of the I-*Ceu*I A fragment. In In4, the average size of the I-*Ceu*I A fragment is 2376kb, which allows its total genome size to be 4836 kb. Similarly, in 26T50, the average size of the I-*Ceu*I A fragment is 2347 kb but the total genome size is 4807 kb. In both these cases , the total genome size still remains within the 4750 to 4840 kb range.

The most significant changes in total genome size are observed when the I-*Ceu*I G fragment is altered. In ST1 [Table 4.4 (row 16)], the size of the I-*Ceu*I G fragment is

reduced by 130 kb, whereas the I-*Ceu*I A fragment remains unchanged. As a consequence, the total genome size of ST1 is also reduced by 130 kb, making it 4650 kb. A greater change is seen in SA 4865 [Table 4.4 (24)] because its I-*Ceu*I G fragment is smaller by 250 kb. In addition, the I-*Ceu*I A fragment is smaller by about 25 kb in this strain. As a result, the total genome size of SA 4865 is 4504 kb, which is approximately 275 kb smaller than *S. typhi* Ty2.

In summary, the total genome size of the 28 wild-type *S. typhi* strains studied varies from 4504 kb to 4840 kb, a total range of 330 kb. However, in most of these strains the variation is less than 100 kb.

F. Changes in the size of the I-*Ceu*I A fragment and total genome size in *S. typhi* SARB 63 revertants

As mentioned earlier, SARB 63 is a wild-type *S. typhi* strain in which the I-*Ceu*I C fragment is displaced such that the *oriC* site has moved from its original position by about 15%. SARB 63 revertants are spontaneously arising, laboratory-derived strains in which the I-*Ceu*I C fragment has close to its original position such that the *oriC* site is no longer displaced. These strains were randomly isolated and the location of their I-Ceu1 fragments were determined (Ng, 1999).

Table 4.8 allows for a size comparison of all seven I-*Ceu*I fragments and total genome size, between *S. typhi* Ty2, SARB 63 and 20 independent SARB 63 revertants. As done previously, the size of the I-*Ceu*I A fragment for each strain was determined by

averaging the size obtained from all three *Xba*I, *Avr*II and *Spe*I cleavage patterns. These cleavage patterns were determined by isolation of the I-*Ceu*I A fragment, re-digestion, end-labeling with α-^{32}P dCTPs, PFGE and autoradiography. The sizes of the I-*Ceu*I B to G fragments were determined by separating I-*Ceu*I digested genomic DNA on PFGE and comparing the fragments to known I-*Ceu*I fragment sizes of *S. typhi* Ty2.

SARB 63 [Table 4.8 (2)] has a total genome size of 4768 kb. Its I-*Ceu*I B, C, D, E, F and G fragments are unchanged when compared with *S. typhi* Ty2, but the average size of its I-*Ceu*I A fragment is smaller than in *S. typhi* Ty2 by about 12 kb. The discrepancy in size of the I-*Ceu*I A fragment is what contributes to the smaller genome size of SARB 63. In comparing the total genome size and I-*Ceu*I A fragment size in the revertants, there appears to be two groups. SARB 63 revertant strains IV152, 264, 58, 85, 90, 228, 117 and 156 [Table 4.8 (rows15-22)] are identical to SARB 63, in that all 7 of their I-*Ceu*I fragments are unchanged in size. This conclusion is firmly based, since the size of I-*Ceu*I fragments, and of *Xba*I, *Avr*II and *Spe*I fragments of I-*Ceu*I A, were all clearly shown to be indistinguishable from those of SARB 63 (data not shown). However, the remaining 12 SARB 63 revertants IV 52, 57, 60, 64, 71, 122, 132, 145, 150, 154, 155 and 292 [Table 4.8 (rows 3-14)] are different from SARB 63, although total genome size remains approximately the same. The total genome size of these revertants is 4774 kb while in SARB 63 it is 4768 kb. In this set of revertants, the I-*Ceu*I G fragment is smaller by 200 kb; however, this decrease in size is offset by an increase in size of the I-*Ceu*I A fragment by approximately the same amount. As a result, in all the

Table 4.8: Comparison of total genome size for various *S. typhi* strains, including *S. typhi* SARB 63 and revertants, as assessed by the sizes of their I-*Ceu*I fragments

	S. typhi Strain No.	SGSC No.	SIZE OF I-*Ceu*I FRAGMENTS[b] (kb)						A Total fragment size using:				Total Genome size (kb)
			B	C	D	E	F	G	*Xba* I	*Avr* II	*Spe* I	Total Size (avg.)[a]	
1	Ty2	2408	724	502	136	146	44	828	2400	2400	2400	2400	4780
2	SARB 63	2520	724	502	136	146	44	828	2400	2363	2400	2388	4768
3	IV 52	3793	724	502	136	146	44	628	2575	2548	2660	2594	4774
4	IV 57	3794	724	502	136	146	44	628	2575	2548	2660	2594	4774
5	IV 60	3796	724	502	136	146	44	628	2575	2548	2660	2594	4774
6	IV 64	3797	724	502	136	146	44	628	2575	2548	2660	2594	4774
7	IV 71	3798	724	502	136	146	44	628	2575	2548	2660	2594	4774
8	IV122	3802	724	502	136	146	44	628	2575	2548	2660	2594	4774
9	IV132	3803	724	502	136	146	44	628	2575	2548	2660	2594	4774

	S. typhi Strain No.	SGSC No.	SIZE OF I-CeuI FRAGMENTS (kb)						A Total fragment size using:				Total Genome size (kb)
			B	C	D	E	F	G	Xba I	Avr II	Spe I	Total Size	
10	IV145	3804	724	502	136	146	44	628	2575	2548	2660	2594	4774
11	IV150	3805	724	502	136	146	44	628	2575	2548	2660	2594	4774
12	IV154	3807	724	502	136	146	44	628	2575	2548	2660	2594	4774
13	IV155	3809	724	502	136	146	44	628	2575	2548	2660	2594	4774
14	IV292	3813	724	502	136	146	44	628	2575	2548	2660	2594	4774
15	IV152	3806	724	502	136	146	44	828	2400	2363	2400	2388	4768
16	IV264	3812	724	502	136	146	44	828	2400	2363	2400	2388	4768
17	IV 58	3795	724	502	136	146	44	828	2400	2363	2400	2388	4768
18	IV 85	3799	724	502	136	146	44	828	2400	2363	2400	2388	4768
19	IV 90	3800	724	502	136	146	44	828	2400	2363	2400	2388	4768
20	IV228	3811	724	502	136	146	44	828	2400	2363	2400	2388	4768
21	IV117	3801	724	502	136	146	44	828	2400	2363	2400	2388	4768
22	IV156	3810	724	502	136	146	44	828	2400	2363	2400	2388	4768

[a] The size of the I-*Ceu*I A fragment as determined by the different restriction endonuclease digests, was resolved by comparing corresponding *S. typhi* Ty2 fragments, whose size was previously known. Since it was not possible to run a standard lambda ladder to determine accurate size of the restriction endonuclease fragments (because end-labeling was used), the above given sizes for the I-*Ceu*I A fragment should not be accepted as absolute values.

[b] The size of the I-*Ceu*I B, C, D, E, F and G fragments was determined by comparison to lambda standards, and to I-*Ceu*I fragments from S. typhi Ty2.

SARB 63 revertants, it appears that overall genome size is not significantly altered with respect to SARB 63and *S. typhi* Ty2. This group of SARB 63 revertants deviate in total genome size from SARB 63 by 5-10 kb and from *S. typhi* Ty2 by only 6 kb.

CHAPTER 5: RESULTS- Rearrangements in the I-*Ceu*I A fragment of *S. typhi* strains

A. Evaluating the restriction endonuclease cleavage patterns of the I-*Ceu*I A fragment of various wild-type *S. typhi* strains

The *Xba*I, *Avr*II and *Spe*I restriction endonuclease cleavage patterns of 28 *S. typhi* strains were examined. This was done to identify areas within the *S. typhi* genome where change in the restriction fragment pattern is common. Previous analysis revealed very few changes in the half of the chromosome covered by I-*Ceu*I fragments B to G, so this study concentrated on the I-*Ceu*I A fragment. The analysis was done in the same way as the size comparison study involving the I-*Ceu*I A fragment of *S. typhi*, detailed in chapter four. An I-*Ceu*I restriction digest of *S. typhi* genomic DNA run on PFGE separates the DNA into seven I-*Ceu*I fragments. A gel showing the seven I-*Ceu*I fragments is seen in Figure 4.1. From this, the I-*Ceu*I A fragment was isolated and re-digested with either *Xba*I, *Avr*II or *Spe*I. The digests were end-labeled with α ^{32}P-dCTPs and run on PGFE. An autoradiograph of the gel depicts a restriction cleavage pattern, which is then compared to the known restriction pattern of *S. typhi* Ty2 for that particular enzyme (Liu and Sanderson, 1995c). Examples of such autoradiographs were shown previously in Figures 4.2 and 4.3.

B. Changes in the restriction endonuclease cleavage pattern of the I-*Ceu*I A fragment within the genome of various *S. typhi* strains

Autoradiographs of *Xba*I, *Avr*II and *Spe*I cleavage patterns of the I-*Ceu*I A fragment for various *S. typhi* strains were compared to that of *S. typhi* Ty2. Restriction fragments, which were altered with respect to that of *S. typhi* Ty2, were noted for each strain. Strains with similar changes and hence, similar restriction patterns for all the three enzymes were grouped and are presented as separate rearrangement classes in Table 5.1. Representatives of these classes are shown in Figure 5.1 and will be discussed below in detail. 17 of the 28 the wild-type *S. typhi* strains studied fell within one of these four classes. The remaining strains were altered such that they were similar to at least one of the above classes but they had additional changes as well. However, the some general conclusions can be drawn from all of the *S. typhi* strains studied above in regards to changes within the I-*Ceu*I A fragment.

1. Characteristics of *S. typhi* strains belonging in Class I

Figure 5.1 shows the *S. typhi* Ty2 genomic map and its I-*Ceu*I, *Xba*I, *Avr*II and *Spe*I restriction fragments. The I-*Ceu*I A fragment is shown in color so that differences in the *Xba*I, *Avr*II and *Spe*I restriction patterns for this fragment can be observed when comparing *S. typhi* Ty2 to other strains. Different colors on the I-*Ceu*I A fragment represent the types of changes observed in doing this comparison. Restriction fragments colored green represent no changes with respect to the fragments produced by *S. typhi*

Table 5.1: A list of changes in the *Xba*I, *Avr*II and *Spe*I restriction endonuclease cleavage patterns in comparison to *S. typhi* Ty2 for various *S. typhi* strains. These strains are grouped according to the area of the genome in which changes are being seen.

Strain No.#	Rearrange-ment Class	Changes in the endonuclease cleavage pattern of[c]:			Other *S. typhi* members in each class listed by strain no.
		*Xba*I	*Avr*II	*Spe*I	
ST 308	I	– All fragments similar to *S. typhi* Ty2	– All fragments similar to *S. typhi* Ty2	– All fragments similar to *S. typhi* Ty2	– ST 309 – ST 1 – ST 1002
ST 24A[b]	II	– No I (306 kb) – No J (104 kb) – Extra 390 kb	– No D (32 kb) – No H (32 kb) – No I (23 kb) – No J (115 kb) – K-20 kb smaller – Extra 175 kb band	– No U (70 kb) – No V (52 kb) – Extra 110 kb band	– 26T20 – SA 4865 – ISP 1820 – ST 1006
SARB 63	IIa	– Similar to class 2 with the exception of: – G-20 kb larger	– Similar to class 2 with the exception of: – F-10 kb larger	– Similar to class 2 with the exception of : – No M (180 kb) – Extra 145 kb band – Extra 46 kb band	– In 24

Strain No.#	Rearrange-ment Class	Changes in the endonuclease cleavage pattern of[a]:			Other *S. typhi* members in each class listed by strain no.
		*Xba*I	*Avr*II	*Spe*I	
26T50	IIb	– Similar to class 2 with the exception of: – G-20 kb smaller	– Similar to class 2 with the exception of: – F-20 kb larger	– Similar to class 2 with the exception of: – K-80 kb smaller	
PNG 31[c]	III	– J-10 kb smaller – K-20 kb larger	– No D (32 kb) – No H (32 kb) – No I (23 kb) – No J (115 kb) – K-20 kb larger – Extra 175 kb band	– No Z (380 kb) – U-5 kb smaller – V-5 kb smaller – Extra 210 kb band – Extra 185 kb band	– PNG 30 – PNG 32
ST 3123	IV	– G-20 kb smaller – F-30 kb larger	– K-5 kb smaller – F-10 kb smaller	– No M (180 kb) – No N (193 kb) – Extra 310 kb band – Extra 80 kb band	

[a]The changes are determined by comparing to the known endonuclease cleavage pattern of *S. typhi* Ty2 {Liu & Sanderson

1995 }. It should be noted that sizes are not exact, as they are determined by comparing to known sizes of the same fragment

form *S. typhi* Ty2 and a size standard was not available.

[b]*Xba*I and *Avr*II endonuclease cleavage pattern of the I-*Ceu*I A fragment for ST 24A is shown in Figure 4.2.

[c]*Avr*II and *Spe*I endonuclease cleavage pattern of the I-*Ceu*I A fragment for PNG 31 is shown in Figure 4.3.

Figure 5.1: A diagrammatic depiction of the changes in the *Xba*I, *Avr*II and *Spe*I cleavage patterns of wild-type *S. typhi* strains with respect to *S. typhi* Ty2

The known I-*Ceu*I, *Xba*I, *Avr*II and *Spe*I endonuclease restriction cleavage maps of *S. typhi* Ty2 are presented (Liu and Sanderson, 1995c). Focus is put on the restriction fragments of the I-*Ceu*I A fragment. The I-*Ceu*I A fragment of *S. typhi* Ty2 is used as a base model to compare the *Xba*I, *Avr*II and *Spe*I cleavage patterns of the I-*Ceu*I A fragment of several wild-type *S. typhi* strains. The changes in the cleavage patterns of these strains are diagrammatically depicted by color in the following manner: in a given strain, if a particular restriction fragment is identical to its counterpart in *S typhi* Ty2 it is colored in green. If the fragment is missing with respect to *S. typhi* Ty2 then it is colored in pink. Finally, if the fragment is either smaller or larger than its counterpart in *S. typhi* Ty2, then the fragment is colored yellow. The actual changes in the fragments that are occurring in each diagram from a to f is presented in Table 5.1. **a.** Depicts *S. typhi* strain ST 308, which is an example of rearrangement class I. **b.** Illustrates *S. typhi* strain ST 24A, which is an example of rearrangement class II. **c.** Presents *S. typhi* strain SARB 63, which is an example of rearrangement class IIa. **d.** Shows *S. typhi* strain 26T50, which is an example of rearrangement class IIb. **e.** Is *S. typhi* strain PNG 31, which is a representative of rearrangement class III and **f.** ST 3123 is an example of a *S. typhi* strain which belongs to rearrangement class IV.

a. CLASS I: ST 308
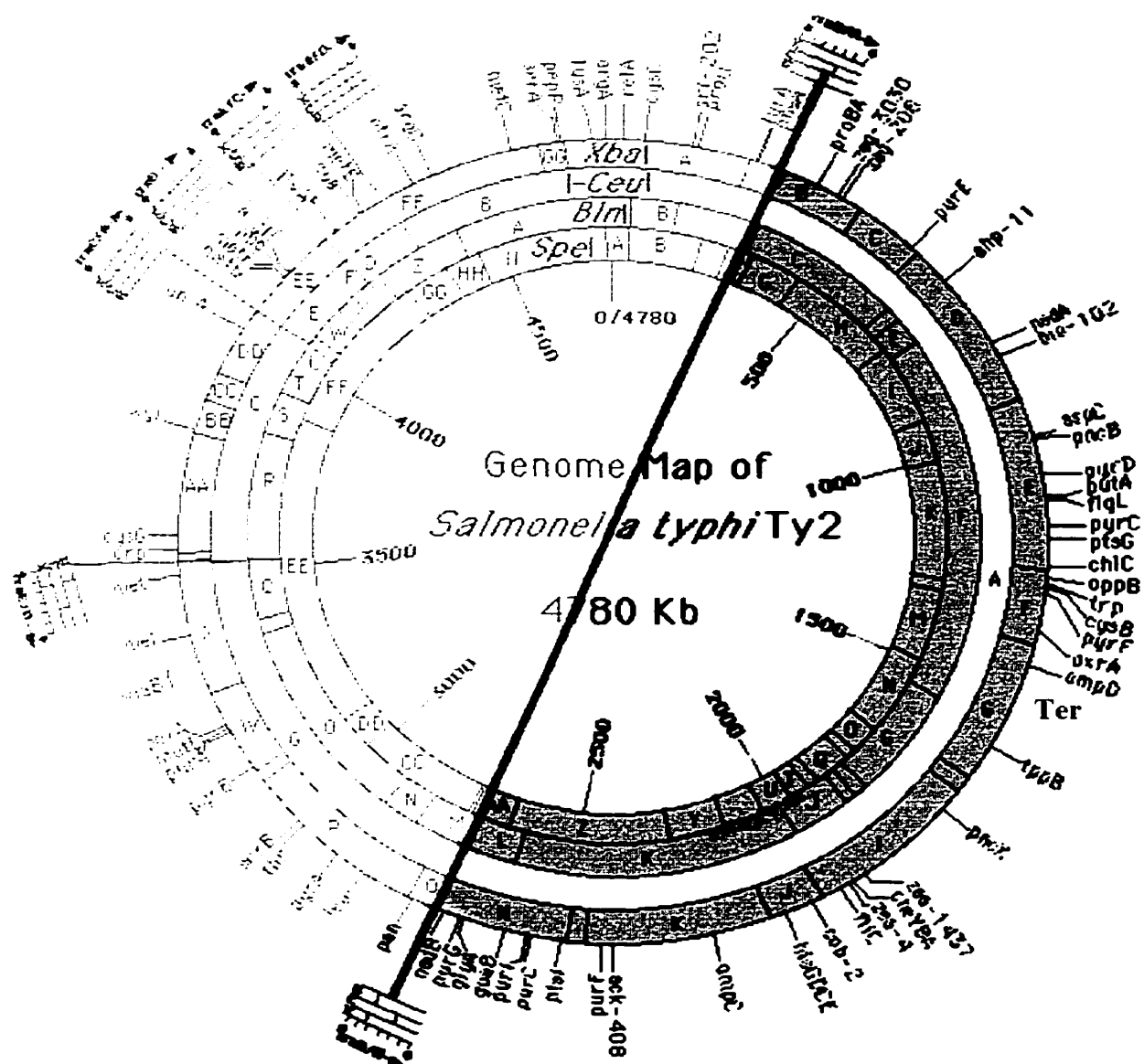
Genome Map of
Salmonella typhi Ty2
4780 Kb

b. CLASS II: ST 24A

Genome Map of *Salmonella typhi* 4780 Kb

c. CLASS IIa: SARB 63

Genome Map of Salmonella typhi 4780 Kb

d. CLASS IIb: 26T50

Genome Map of *Salmonella typhi* 4780 Kb

e. CLASS III: PNG 31

Genome Map of
Salmonella typhi Ty2
4780 Kb

f. CLASS IV: 3123

Genome Map of *Salmonella typhi* Ty2
4780 Kb
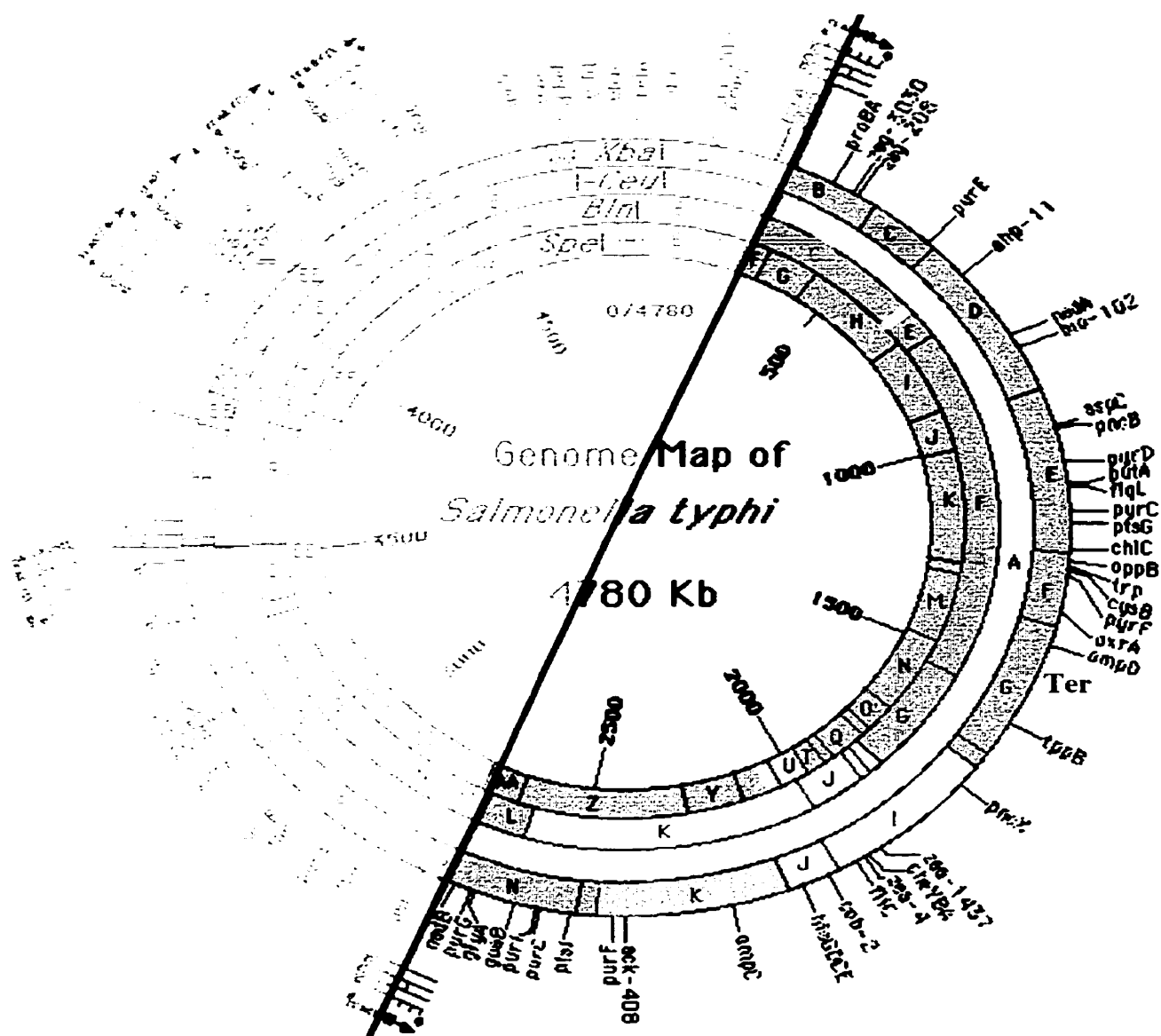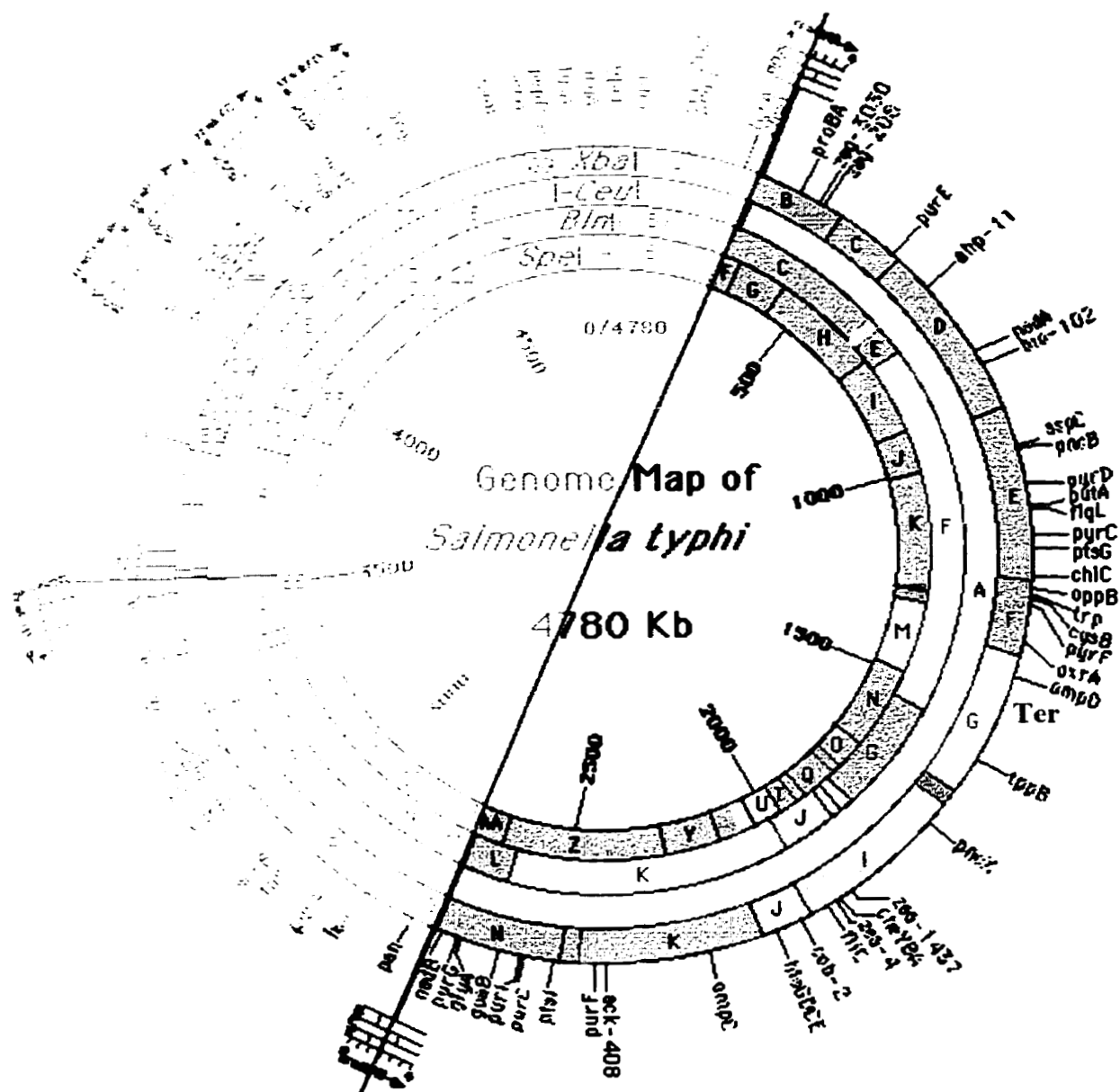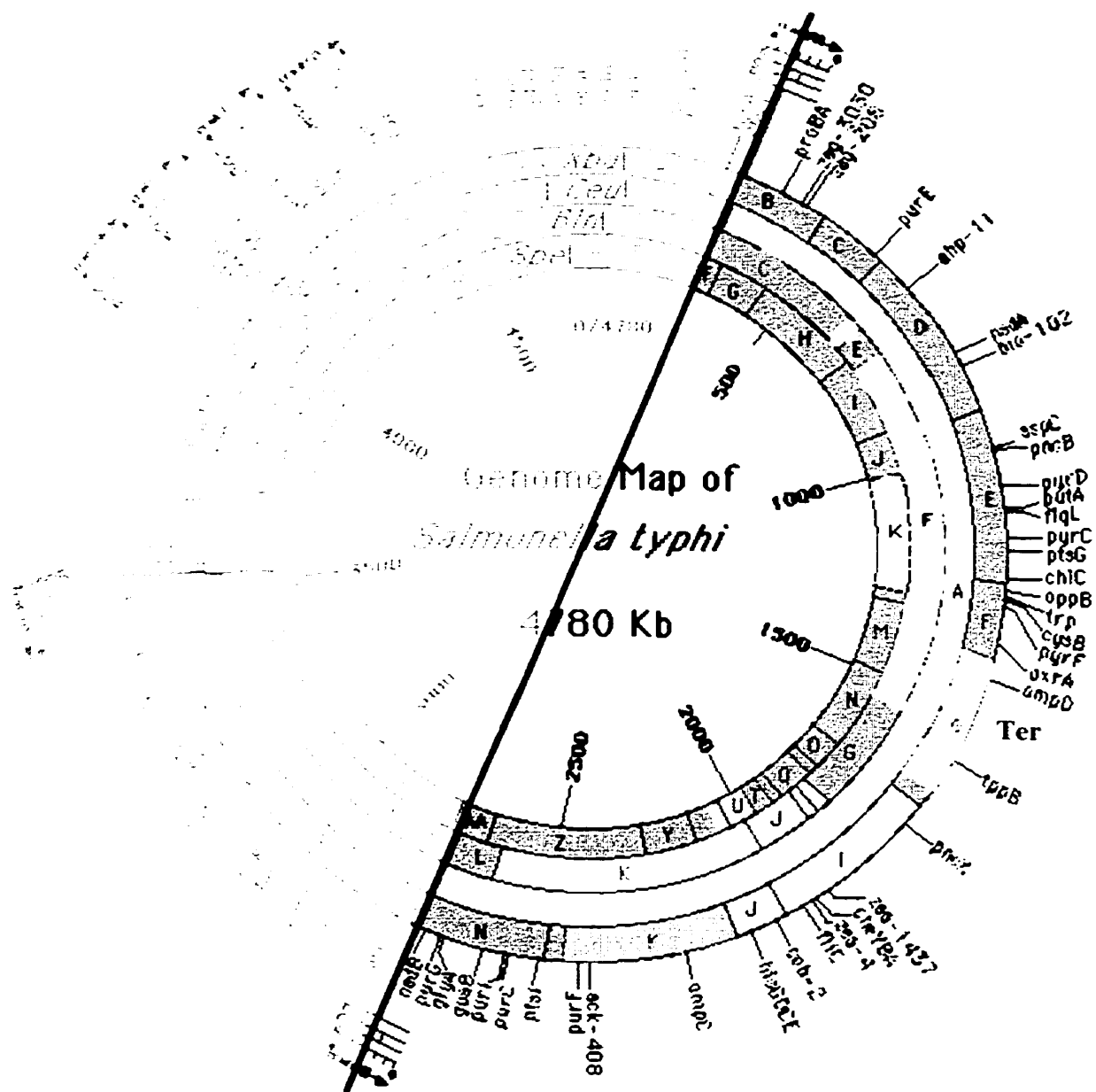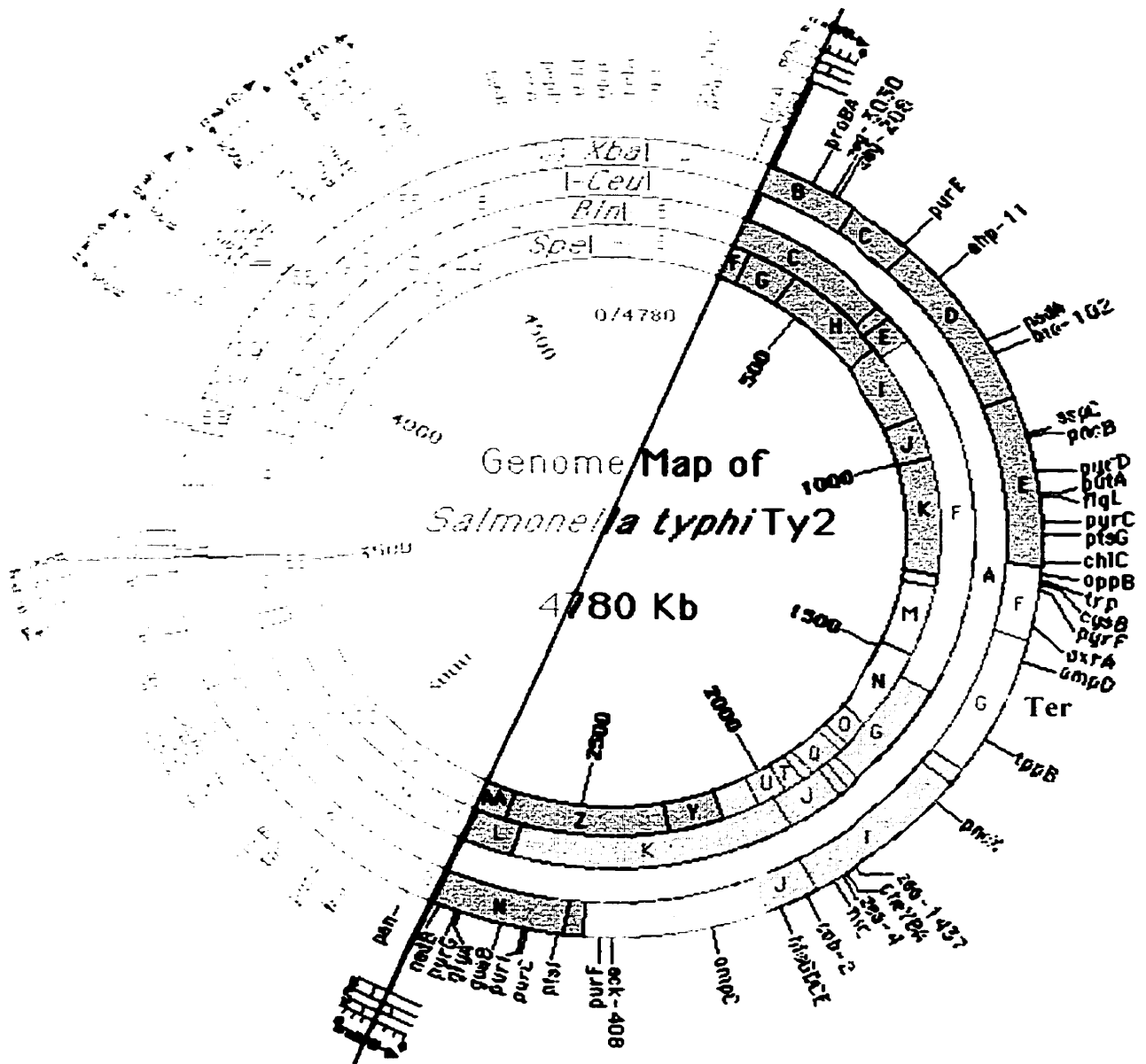
Ty2. Restriction fragments colored in pink indicates these fragments are missing when compared to the restriction pattern for *S. typhi* Ty2. Lastly, fragments colored in yellow indicate a change in size, the fragment can be either smaller or larger with respect to the same fragment in *S. typhi* Ty2. Table 5.1 lists each class, indicates which strains belong to each class and what changes, for the *Xba*I, *Avr*II and *Spe*I cleavage patterns, are actually occurring. Part of the data to support these conclusions is in figure 5.1 and other data are not shown.

S. typhi strains belonging in Class I, like ST 308, ST 309, ST 1 and ST 1002, are indistinguishable from *S. typhi* Ty2 (Table 5.1, Figure 5.1a) in their *Xba*I, *Avr*II and *Spe*I cleavage patterns. Therefore, strains belonging to Class I appear to have no changes in the I-*Ceu*I A fragment, with respect to *S. typhi* Ty2. The *Avr*II and *Spe*I cleavage patterns for the I-*Ceu*I A fragment of ST 1002 were shown earlier in Figure 4.2.

2. Characteristics of *S. typhi* strains belonging in Class II

S. typhi strains belonging in Class II, such as ST 24A and four other strains, have a specific set of changes in the *Xba*I, *Avr*II and *Spe*I restriction fragments in the I-*Ceu*I A fragment. All of these altered restriction fragments are localized within a small area of the I-*Ceu*I A fragment, around the 2000 kb region of the chromosome. Autoradiographs of end-labeled pulse-field gels, comparing the *Xba*I and *Avr*II cleavage patterns of the I-*Ceu*I A fragment, for *S. typhi* Ty2 and ST24A were shown earlier in Figure 4.3 and changes are listed in Table 5.1. In the *Xba*I cleavage pattern, fragments I and J, which are adjacent and around 2000 kb on the chromosome, are missing (Figure 5.1b). An extra fragment was detected, whose size (390 kb), indicates that it results from the addition of I

and J, with a 20 kb loss. This suggests there may be an alteration, perhaps a deletion of the *Xba*I cleavage site between these two fragments, causing them to join. The *Spe*I cleavage pattern for ST 24A is similar to the *Xba*I cleavage pattern (Table 5.1). The *Spe*I U (70 kb) and V (52 kb) fragments, which are adjacent to each other in the 2000 kb region of the chromosome, are missing; a new fragment of 110 kb is detected. The loss of 122 kb (70kb and 52 kb), and the gain of 110 kb, suggests a deletion of 12 kb containing the *Spe*I cleavage site between *Spe*I U and V, causing them to join to form a new 110 kb fragment. Interestingly, the *Xba*I cleavage site between I and J and the *Spe*I cleavage site between U and V are close together, supporting the idea of a deletion within this area. The *Avr*II cleavage pattern also indicates changes within the 2000 kb region of the chromosome. In strain ST 24A, *Avr*II D, H, I and J fragments are missing and K is 20 kb smaller (Table 5.1). These fragments are all adjacent to one another (Figure 5.1b) with the exception of D. It should be noted here that the *Avr*II D fragment is often missing in other *S. typhi* strains besides those belonging to class II. However, according to Figure 5.1b, the *Avr*II D fragment is within a region of the chromosome where there are no other changes present. In the *Avr*II cleavage pattern for the I-*Ceu*I A fragment of *S. typhi* Ty2 (Figure 5.1), it can be seen that the *Avr*II D, H and I fragments are not labeled. This indicates that the position of these fragments is not known for certain. It may be that the *Avr*II D fragment is not in the area indicated but is elsewhere on the chromosome where changes are more common. Due to the uncertainty of the position of the *Avr*II D fragment, it will not be discussed further. Nevertheless, although the *Avr*II H, I and J fragments are missing, whose sum is 170 kb, an extra fragment which is 175 kb in

size is seen. The surplus 5 kb may be coming from the 20 kb missing from the K fragment. The above modifications cannot be explained by simple base pair changes in *Avr*II cleavage sites alone, since more than one cleavage site would have to altered in a random manner within a small area. It is more likely that complex rearrangements are a cause for the modifications observed.

In examining the *Xba*I, *Avr*II and *Spe*I cleavage patterns of the I-*Ceu*I A fragment of ST 24A, it appears that two independent events, a deletion and a rearrangement, are occurring within the 2000 kb region of this strain and other members of class II.

3. Characteristics of *S. typhi* strains belonging in Class IIa

Only two strains fall into Class IIa, and they are SARB 63 and In 24. As mentioned earlier, SARB 63 is unusual in the order of I-*Ceu*I B to G fragments, such that it has an unbalanced genome. The *ori*C site, in the I-*Ceu*I C fragment of SARB 63, is displaced by 15% such that the origin and terminus are no longer 180° apart. Class IIa is a subset of class II in that all the changes detected in class II, at the 2000 kb region of the chromosome in I-*Ceu*I A, are present along with additional modifications in the 1300 kb to 1500 kb region of the chromosome (Figure 5.1c).

Table 5.1 shows that in the *Xba*I cleavage pattern, the G fragment is 20 kb larger and in the *Avr*II cleavage pattern, the F fragment is 10 kb larger. In the *Spe*I cleavage pattern, the M fragment, also within the 1300 - 1500 kb region, is missing. However, two extra bands in the *Spe*I cleavage pattern, equaling the M fragment are found.

Changes within this region of the chromosome are difficult to explain. It is unlikely that random base pair changes alone would alter all three *Xba*I, *Avr*II and *Spe*I

cleavage sites found in different positions within the 1300 kb to 1500 kb region of the chromosome. Furthermore, a insertion within this area would not explain why according to the AvrII and XbaI patterns there is a gain of 10 kb and 20 kb of DNA, respectively, but the SpeI digest reports no gain or loss of DNA. Changes within 1300 - 1500 kb region of the chromosome may be associated with the changes in the 2000 kb region through a large inversion event. However, probing experiments below (section D. 2) show no evidence of an inversion.

4. Characteristics of *S. typhi* strain 26T50, belonging in Class IIb

26T50 is another *S. typhi* strain whose *XbaI, AvrII* and *SpeI* cleavage patterns for the I-*CeuI* A fragment are similar to that of class II, but also harbors additional changes. Therefore, 26T50 is listed as a second subset class of class II (Table 5.1); however, it is its only representative.

26T50 illustrated in Figure 5.1d, has a *XbaI* cleavage pattern similar to class II, except that its G fragment is 20 kb smaller. Similarly, the *AvrII* cleavage pattern is identical to class II also, however, the F fragments appears to be 20 kb larger. Finally, the *SpeI* cleavage pattern, shows that its K fragment is 80 kb smaller; the missing 80 kb may be present as an extra band, this is difficult to ascertain since there is a cluster of bands near 80 kb in the *SpeI* cleavage pattern, making it difficult to enumerate the exact number of bands within this area.

In examining the *XbaI, AvrII* and *SpeI* restriction pattern of 26T50, it appears that two areas in the I-*CeuI* A fragment is affected. As in class II, the 2000 kb region of the

chromosome is affected. In addition, the 1000 kb - 1500 kb region also appears to be altered in 26T50.

5. Characteristics of *S. typhi* strains belonging in Class III

S. typhi strains which fall into class III all have changes within the 2000-2500 kb region of the chromosome. Figure 5.1e depicts the chromosome of PNG 31 and Table 5.1 lists the altered restriction fragments. In addition, autoradiographs of end-labeled pulse-field gels, comparing the *Spe*I and *Avr*II cleavage patterns of the I-*Ceu*I A fragment for *S. typhi* Ty2 and PNG 31 were shown earlier in Figure 4.4.

In the *Xba*I digest, the *Xba*I J fragment is 10 kb smaller, while the K fragment, adjacent to J, is 20 kb larger. This could be due to a rearrangement event, which has moved the *Xba*I cleavage site between the J and the K towards the *Xba*I K fragment. The *Avr*II cleavage pattern of PNG 31 resembles that of ST 24A (rearrangement class II), in that the *Avr*II D, H, I and J (total size 170 kb) are missing, but a new fragment of 175 kb is detected. The *Avr*II K fragment is now 20 kb larger, but this may be associated with the changes further down in the chromosome, around the 2500 kb mark. The *Spe*I cleavage pattern of PNG 31 shows that the U and V fragments are each 5 kb smaller and the Z fragment, which is 380 kb in size, appears to be split into two fragments, 180 kb and 310 kb in size. These changes could be explained by an inversion event between the *Spe*I U and Z, however, this is not proven.

In evaluating the *Xba*I, *Avr*II and *Spe*I cleavage patterns of PNG 31, it can be concluded that changes are occurring within the 2000 kb - 2500 kb region of the chromosome and these changes are seen in all three cleavage patterns. As a result,

random mutations are not the likely cause for these changes, instead the modifications are probably due to rearrangement events within the 2000 kb - 2500 kb region of the chromosome.

6. Characteristics of *S. typhi* strain 3123, belonging in Class IV

S. typhi strain 3123, the only *S. typhi* strain found within rearrangement class IV (Figure 5.1f, Table 5.1), has changes within the 1300 kb to 1500 kb region of the chromosome in its I-*Ceu*I A fragment. In its *Xba*I cleavage pattern it can be seen that the *Xba*I G is 20 kb smaller, while the F is 40 kb larger. Since these fragments are adjacent to one another, as seen in Figure 5.1f, the changes could have resulted from an inversion event. In the *Avr*II digest, the F is 10 kb smaller and the G is 5 kb smaller. The *Spe*I cleavage pattern of ST 3123 suggests that the M and N fragments are missing, which totals a loss of 373 kb of DNA. However, these fragments are replaced by two other fragments, a 310 kb and a 80 kb fragment, which more or less makes up for the above loss. The changes in the *Spe*I digest could also be explained by a rearrangement event. Since the M and N fragments are adjacent to each other, the restriction site between them could have moved such that one fragment has become larger, while the other has become smaller.

All of the modifications seen in the *Xba*I, *Avr*II and *Spe*I cleavage patterns of ST 3123 suggest changes in the 1500 kb region of the chromosome. Since these alterations are seen in the same location for all three of the digests, a possible rearrangement event is more likely to have occurred within this area. However, the loss of 5 kb from the *Avr*II K fragment cannot be explained in this way, since a change in the 2500 kb region is only

seen in the *Avr*II digest. It could be possible this change is not seen in the *Spe*I and *Xba*I

cleavage patterns due to the crowding of bands on the gels. The loss of the 5 kb in the

*Avr*II K fragment is most likely a result of a deletion event within this 2000 kb region of

chromosome.

7. Variations in the other wild-type *S. typhi* strains studied

Seventeen out of the 28 wild-type *S. typhi* strains studied have been described

above; the other eleven follow similar trends as shown in the above four classes;

however, they each represent their own class since the changes in their *Xba*I, *Avr*II and

*Spe*I restriction patterns are slightly different from one another. Like the four classes

discussed above, these 11 strains have changes in their I-*Ceu*I A around the 1000 kb to

1500 kb and 2000 kb to 2500 kb regions of the chromosome. Since changes within these

regions are seen for all three *Xba*I, *Avr*II and *Spe*I restriction cleavage patterns with

respect to *S. typhi* Ty2, they are likely a result of rearrangement events and deletion

events, not base pair changes which are randomly affecting the cleavage sites. If the

latter reasoning were true, changes caused by an altered *Xba*I cleavage site in a particular

region of the I-*Ceu*I A fragment, would not necessarily affect the *Avr*II cleavage pattern

within the same area of the chromosome; however, changes in the same region in all

these restriction digestions are most commonly seen.

In summary, it appears that changes within the I-*Ceu*I A fragment of wild-type *S.*

typhi strains are a result of rearrangement and deletion events, instead of base pair

changes. In addition, these changes are found in certain regions of the chromosome.

For examples, many strains commonly had alterations within the 2000 kb mark of the

chromosome. Modifications were also seen in the 1000 kb to 1500 kb and the 2000 kb to

2500 kb region. What is particularly interesting about the 1000 kb to1500 kb region is

that the *ter* sites are situated here and rearrangements within this area affects the position

of the *ter* region.

C. Types of rearrangements within the I-*CeuI* A fragment of *S. typhi* strains

In the previous section, it was determined that deletions and rearrangements are

occurring within the 1000 kb to 1500 kb and 2000 kb to 2500 kb region of the I-*CeuI* A

fragment of *S. typhi* strains. For example, in rearrangement class II, as shown in Table 5.1

and Figure 5.1b., the *XbaI*, *AvrII* and *SpeI* cleavage patterns vary around the 2000 kb

region of the genome map. In this situation, it is predicted that a possible rearrangement

and a deletion event are involved. The rearrangement most likely encompasses a small

area within the 2000 kb and 2100 kb region. The deletion appears to be occurring around

the *XbaI* cleavage site between the *XbaI* I and J fragments, which then causes these

fragments to join.

Rearrangements within the 1000 kb - 1500 kb region of the chromosome are

interesting because the terminus sites are within this area. In subclass IIa, 26T50 (Figure

5.1c and e), and other wild type *S. typhi* strains like ST 318, PL25766, CC6 and ST 495

(cleavage patterns not shown), cleavage patterns in both the 1000 kb - 1500 kb and the

2000 kb - 2500 kb regions vary. The changes in the *XbaI*, *AvrII* and *SpeI* cleavage

patterns of these strains within the 2000 kb - 2500 kb region are identical to those strains

in class II, therefore, it is suspected that they result from the same type of rearrangement and deletion events. Cleavage pattern changes in the 1000 kb - 1500 kb region are also suspected to result from rearrangement events. However, it is not known if the rearrangement is restricted within the 1000 kb - 1500 kb region or if it is affiliated with the changes in the 2000 kb region of the chromosome. If the latter is true, the inverted piece of DNA would be large and it would affect the position of the terminus site.

D. Hybridization experiments to determine the size of the rearrangements found within the I-*Ceu*I A fragment of *S. typhi* strains.

End-labeling experiments done in the previous section determined the presence of possible rearrangements within certain regions of the I-*Ceu*I A fragment but could not indicate the end-points of these rearrangements. To ascertain if the rearrangements occur within small areas or encompass large portions of the I-*Ceu*I A fragment, probing was done.

Probing experiments were done using *Xba*I restriction fragments J and G of *S. typhi* Ty2 as probes. Probes were labeled using α ^{32}P-dCTP. The templates contained genomic DNA from *S. typhi* strains digested with *Xba*I. The digests were run on PFGE and the DNA was transferred onto membranes. The presence of a rearrangement was ascertained in the following manner: if the *Xba*I J fragment was used as a probe on a strain and one signal appeared on the autoradiograph, this would indicate that the *Xba*I J fragment of this strain has remained intact. Changes could still be occurring within the J

fragment, but none that would disrupt the structure of this fragment. If a rearrangement is suspected in this case, this result would indicate that the rearrangement is small and has remained within the J fragment. If more than one signal appears, this would imply that the *Xba*I J fragment of this strain has split in some way. It could be a rearrangement, in which case the DNA from the J fragment has broken and joined other fragments, giving rise to more than one signal. This would indicate that the rearrangement is large enough to affect other *Xba*I bands in the I-*Ceu*I A fragment. The size of this rearrangement would be determined by which other *Xba*I band is being affected. One must be aware that more than one signal could also be explained by the introduction of a cleavage site, but in this case, the signals should sum up to the original size of the J fragment.

1. Hybridization with the *Xba*I J fragment

The genomic DNA from several wild-type *S. typhi* strains was digested with *Xba*I and run on PFGE, transferred onto a membrane and probed with a α ^{32}P-dCTP labeled *Xba*I J fragment isolated from *S. typhi* Ty2. Figure 5.2a represents the *Xba*I cleavage pattern of several *S. typhi* strains, stained with ethidium bromide, and Figure 5.2b shows the probing results obtained using this digest as the template for probing. *S. typhi* Ty2 (lane 1) is the positive control, since its *Xba*I J fragment is the same fragment used as the *Xba*I J probe.

The *Xba*I J fragment is found around the 2000 kb region of the chromosome. Thus, strains which vary in their *Xba*I cleavage pattern within this area of the chromosome were studied. As example is ISP 1820, which is a member of rearrangement class II and is represented in Figure 5.2b, lane 8. The signal for the *Xba*I J

Figure 5.2: The detection of rearrangements involving the _XbaI_ J fragment of various _S. typhi_ strains

_Xba_I digests of the total genomic DNA of various _S. typhi_ strains were probed with a α ^{32}P-dCTP labeled _Xba_I J fragment isolated from _S. typhi_ Ty2 to detect the presence of an intact or a disrupted _Xba_I J fragment in each strain. **a.** Genomic DNA isolated from _S. typhi_ strains were digested with _Xba_I and run on PFGE using the following pulse conditions: 10s - 60s at 120°and 180V for 12 hours, 8s - 24s at 130°and 180V for 12 hours and 22s - 26s at 150°and 180V for 16 hours. The sizes of some _Xba_I restriction fragments of _S. typhi_ Ty2 are presented along side of the XbaI cleavage pattern of _S. typhi_ Ty2 in lane 1. **b.** The DNA from these digests were then transferred to a membrane and probed with a α ^{32}P-dCTP labeled _Xba_I J fragment isolated from _S. typhi_ Ty2. The X-ray film was exposed to radiation for 12 hours.

a.

(1) *S. typhi* Ty2

(7) Pl25766
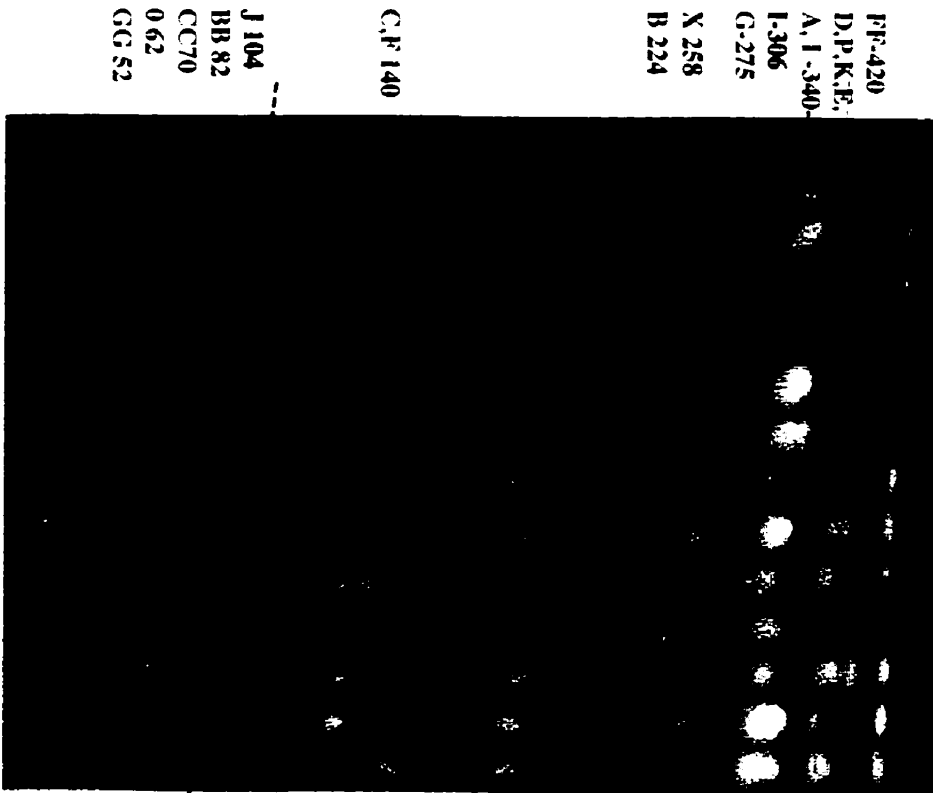
(8) ISP 1820

(9) ST 318

(10) 26T50

(12) SARB 63

(13) CC6
(14) ST 495

b.

(1) *S. typhi* Ty2

(7) Pl25766
(8) ISP 1820

(9) ST 318

(10) 26T50

(12) SARB 63

(13) CC6
(14) ST 495

GG 52
O 62
CC70
BB 82
J 104

C.F 140

B 224
X 258
G-275
I-306
A.I -340
D.P.K.E.
FF-420
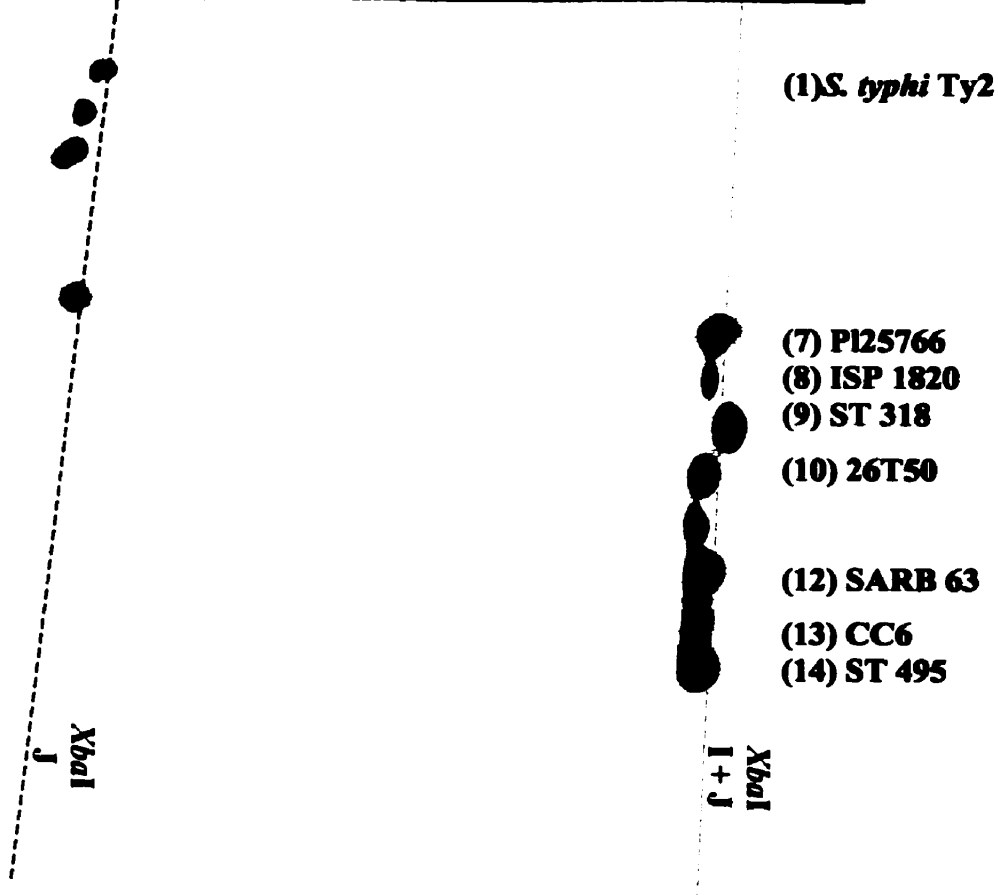
*Xba*I
J

*Xba*I
I+J

fragment appears as a larger fragment, whose size is more or less the sum of the *Xba*I I and J fragments. This confirms previous results, which indicated that a deletion had removed the *Xba*I site between these two fragments and caused them to join. The rearrangement, which is also predicted to be present in this area, cannot be seen in these results because the *Xba*I I and J fragments cannot be distinguished here and the rearrangement, being small, most likely only involves these two fragments.

As mentioned earlier, it is known that rearrangements occur in certain areas within the I-*Ceu*I A fragment of *S. typhi* strains. Most of these rearrangements are restricted within small regions, for example members of rearrangement class II, III and *S. typhi* 3123 (Figure 5.1b, d and f). However, in some cases like rearrangement class IIa, 26T50 (Figure 5.1c and e), ST 318, PL25766, CC6 and ST 495 (cleavage patterns not shown), two regions on the I-*Ceu*I A fragment, some distance apart appear to harbor changes. It could be possible that these two regions are endpoints for one large inversion event or are two independent rearrangement events. If the former were true, the position of the terminus site would be affected and consequently, so would the 180° symmetry between the origin of replication and terminus sites. If such a large inversion were to occur and if one of the endpoints was in the *Xba*I J fragment, then the J fragment would be split in the above strains, such that some of its DNA would join the *Xba*I G fragment. As a result, probing with an intact *Xba*I J fragment would result in two bands. However, when examining SARB 63 (lane 12), 26T50 (lane 10), ST 318 (lane 9), PL25766 (lane 7), CC6 (lane 13) and ST 495 (lane 14) in Figure 5.2, none appear to have two bands. Instead, the only signal given by all these strains is similar to that of rearrangement class

II. This result shows that the only change occurring in the 2000 kb region of these strains is identical to that which occurs in strains of rearrangement class II. If a large rearrangement were to exist in combination with this one, the fragment size of the signal would change but this does not happen either. Thus, SARB 63, 26T50 and the other *S. typhi* strains mentioned above, have a small rearrangement within the 2000 kb region of the chromosome, which is independent of the changes in the 1000 kb - 1500 kb region.

2. Probing with the *Xba*I G fragment

To confirm that the changes in the 1000 kb - 1500 kb region of the chromosome of SARB 63, 26T50 and similar *S. typhi* strains are independent of the changes in the 2000 kb region, probing experiments using the *Xba*I G fragment were also done. The approach was similar to that of when the *Xba*I J fragment was used as a probe. Figure 5.3a represents the *Xba*I cleavage pattern of several *S. typhi* strains and Figure 5.2b shows the probing results obtained using this digest as the template for probing. *S. typhi* Ty2 (lane 1) is the positive control, since its *Xba*I G fragment is the same fragment used as the *Xba*I G probe.

The *Xba*I G fragment is found around the 1500 kb mark of the *S. typhi* chromosome; thus, probing with this fragment detects changes within this area of the chromosome. *S. typhi* strain 3123 harbors changes solely within the 1000 kb - 1500 kb region of the chromosome. As seen in Figure 5.3b, probing 3123's *Xba*I digest of the I-*Ceu*I A fragment, with the *Xba*I G fragments shows the presence of a rearrangement

Figure 5.3: The detection of rearrangements involving the *Xba*I G fragment of various *S. typhi* strains

*Xba*I digests of the total genomic DNA of various *S. typhi* strains were probed with a α

³²P-dCTP labeled *Xba*I G fragment isolated from *S. typhi* Ty2 to detect the presence of an

intact or a disrupted *Xba*I G fragment in each strain. **a.** Genomic DNA isolated from *S.*

typhi strains was digested with *Xba*I and run on PFGE using the following pulse

conditions: 10s - 60s at 120°and 180V for 12 hours, 8s - 24s at 130°and 180V for 12

hours and 22s - 26s at 150°and 180V for 16 hours. The sizes of some *Xba*I restriction

fragments of *S. typhi* Ty2 are presented alongside of the *Xba*I cleavage pattern of *S. typhi*

Ty2 in lane 1. **b.** The DNA from these digests were then transferred to a membrane and

probed with a α ³²P-dCTP labeled *Xba*I G fragment isolated from *S. typhi* Ty2. The X-

ray film was exposed to radiation for 5 days.

a

GG 52
0 62
CC70
BB 82
J 104

C,F 140

X 258
B 224

G-225
I-306
A,1 -340
D,P,K,E
FF-420

25

(1) *S. typhi* Ty2

(2) 3125

(3) In4

(6) In24

(7) SARB 63

(11) 3123

b

(1) *S. typhi*
Ty2

(6) In24

(7) SARB 63

(11) 3123

within the 1500 kb region of this strain. In 3123, the *Xba*I G fragment (275 kb) splits into two smaller fragments; one which is a smaller version of the *Xba*I G and the other, appears to be a part of *Xba*I G combined with the *Xba*I F fragment. Table 5.1 lists which *Xba*I fragments in the I-*Ceu*I A fragment of 3123 have changed. It has been observed that 3123's *Xba*I G fragment has decreased in size by 20 kb while the *Xba*I F fragment, adjacent to G, has increased in size by about 30 kb. The increase and decrease in size do not balance exactly, but as mentioned earlier, the sizes are derived through comparisons made with known sizes from similar fragments of *S. typhi* Ty2, thus sizes are not exact.

In combining the information from Table 5.1 and Figure 5.3b, it can be inferred that a rearrangement involving the G and F fragments has moved roughly 20-30 kb of DNA from the XbaI G fragment into the XbaI F fragment. This rearrangement is not a large one as it only involves about 40 kb of DNA.

SARB 63 and In24 belong to rearrangement class IIa (Figure 5.1c) and the 1300 kb to 1500 kb region and the 2000 kb region of their chromosomes appear to be altered. As mentioned earlier, the *Xba*I G fragment is within the 1300 kb to 1500 kb region of the chromosome and according to Figure 5.3b, SARB 63 and In24 (lanes 3 and 6), continue to have an intact *Xba*I G fragment. This indicates that a large inversion event, disrupting the *Xba*I G fragment, is not present within these strains, suggesting that the changes in the 1300 kb to1500 kb and 2000 kb regions are not associated and therefore are not caused by a single large inversion event. However, since all three *Xba*I, *Avr*II and *Spe*I cleavage patterns are altered within the 1500 kb region of SARB 63 and In24, it is

possible that there is a rearrangement event within the *XbaI* G fragment which does not disrupt the overall structure of the fragment.

To summarize, in the *S. typhi* strains studied, the probing results indicate that rearrangements occur within small areas of the I-*CeuI* A fragment. No large inversions are seen which invert the 1500 kb and 2000 kb regions of the chromosome and changes within these areas are independent of each other; therefore, the position of the terminus site remains more or less unaltered. In *S. typhi* strain 3123, there appears to be an inversion involving the terminus sites; however, this inversion is small and does not significantly displace the terminus sites to the extent that their 180° symmetry with the *oriC* is modified.

E. The position of the I-*CeuI* C, in relation to the I-*CeuI* A fragment in SARB 63

It has been previously determined that SARB 63 has a displaced I-*CeuI* C fragment, in relation to *S. typhi* Ty2 (Liu and Sanderson, 1996). This is especially important because the I-*CeuI* C fragment harbors the *oriC* site. Displacing the I-*CeuI* C fragment changes the position of the *oriC* site and disrupts the 180° symmetry between the *oriC* and the *ter* region. In SARB 63, the I-*CeuI* C fragment and consequently the *oriC* site is displaced by 15%. As seen in Figure 2.3, SARB 63, represented by genome type 25, has its I-*CeuI* C fragment adjacent to the I-*CeuI* A fragment. This is unlike *S. typhi* Ty2, represented by genome type 9, whose I-*CeuI* C fragment is found between the

Figure 5.4: An illustration of the I-*Ceu*I cleavage map of *S. typhi* Ty2 and SARB 63 with its two possible positions for the I-*Ceu*I C fragment in relation to the I-*Ceu*I A fragment

I. The I-*Ceu*I cleavage map of *S. typhi* Ty2 representing the order of its I-*Ceu*I fragments (Liu and Sanderson, 1995c). The position of the *Spe*I AA fragment on the chromosome is given. **IIa.** One possibility for the order of I-*Ceu*I fragments in SARB 63, in which the I-*Ceu*I C fragment is adjacent to the "A" end of the I-*Ceu*I A fragment. **IIb.** A second possibility for the order of I-*Ceu*I fragments in SARB 63, in which the I-*Ceu*I C fragment is adjacent to the "A'" end of the I-*Ceu*I A fragment.

I
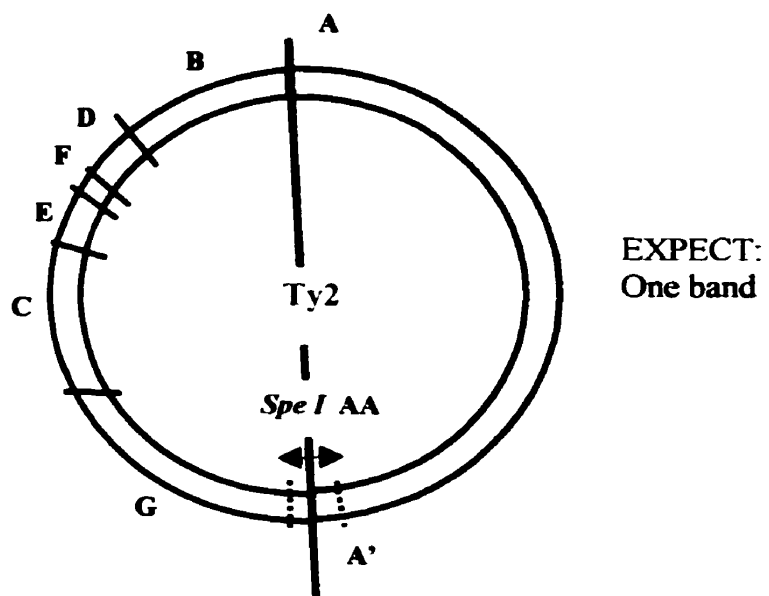
EXPECT:
One band

IIa

IIb

SARB 63

or
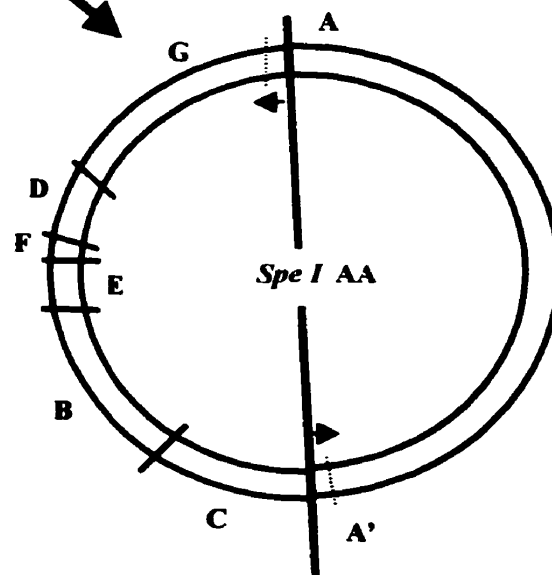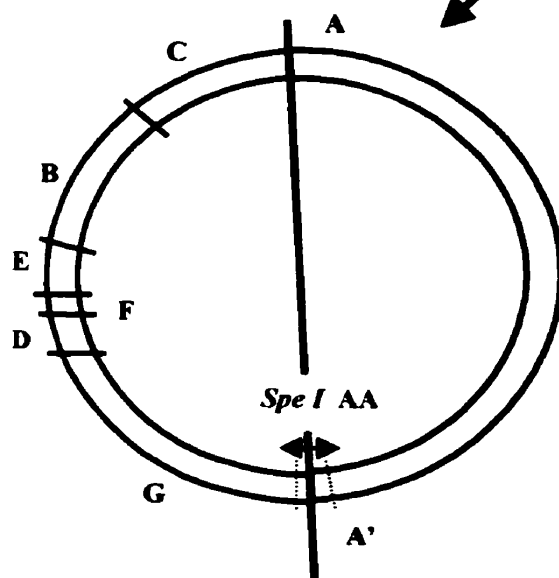
EXPECT:
One band

EXPECT:
Two bands

G and E fragments. However, the side of the I-CeuI A fragment to which the I-*Ceu*I C

fragment of SARB 63 is adjacent to is not known. Figure 5.4, diagrammatically depicts

the two possible positions of the I-*Ceu*I C fragment for SARB 63. When comparing the

first position (Figure 5.4, IIa) to *S. typhi* Ty2 (Figure 5.4, I), it can be seen that the I-*Ceu*I

C fragment of SARB 63, has taken the position of the I-*Ceu*I B fragment in *S. typhi* Ty2.

In this situation the other end of the I-*Ceu*I A fragment, abbreviated as A', is joined to the

I-*Ceu*I G fragment for both SARB 63 and *S. typhi* Ty2. The other possible position

(Figure 5.4, IIb) has the I-*Ceu*I A fragment of SARB 63 inverted such that the A' is now

adjacent to I-*Ceu*I C and the A is beside I-*Ceu*I G. In this case, the I-*Ceu*I G fragment in

SARB 63 and is now in a different position than in *S. typhi* Ty2. An inversion, causing

the I-*Ceu*I A fragment to flip, could have readily occurred due to homologous

recombination between the *rrn* operons flanking this fragment.

To determine which of the above two orientations SARB 63 adopts, probing was done.

The approach was similar to that undertaken in the previous section when the *Xba*I J and

G fragments were used as probes to study rearrangements within the I-*Ceu*I A fragment.

However, in this case, the *Spe*I AA fragment is used as a probe on a *Spe*I digest of

genomic DNA isolated from SARB 63. As seen in Figure 2.4 and in lesser detail in

Figure 5.4 (I), the *Spe*I AA fragment spans the I-*Ceu*I G and the A' end of the I-*Ceu*I A

fragment, in *S. typhi* Ty2. If in SARB 63, the A' end of the I-*Ceu*I A fragment is also

adjacent to the G fragment, then the *Spe*I AA will remain intact and probing with the

same fragment from *S. typhi* Ty2 will reveal one signal. However, if in SARB 63 the I-

*Ceu*I G is not adjacent to the A' end as in *S. typhi* Ty2, but instead the A end of the I-

Figure 5.5: The condition of the *Spe*I AA fragment within the genome of SARB 63 to determine the position of its I-*Ceu*I C fragment on its I-*Ceu*I cleavage map

*Spe*I digests of the total genomic DNA of *S. typhi* Ty2 and SARB 63 were probed with a α ^{32}P-dCTP labeled *Spe*I AA fragment isolated from *S. typhi* Ty2 to detect the presence of an intact or a disrupted *Spe*I AA fragment in each strain. **a.** Genomic DNA was digested with *Spe*I and run on PFGE using the following pulse conditions: 20s - 80s at 120°and 180V for 16 hours, 5s - 20s at 140°and 180V for 24 hours and 6s - 8s at 150°and 180V for 24 hours. The sizes of some relevant *Spe*I restriction fragments of *S. typhi* Ty2 are presented alongside of the *Spe*I cleavage pattern of *S. typhi* Ty2 in lane 1. **b.** The DNA from these digests were then transferred to a membrane and probed with a α ^{32}P-dCTP labeled *Spe*I AA fragment isolated from *S. typhi* Ty2. The X-ray film was exposed to radiation for 6 days.

a

b

Ty2

SARB 63

G 104
DD 99
HH 95
AA 91
Q 80
O 77

F 56
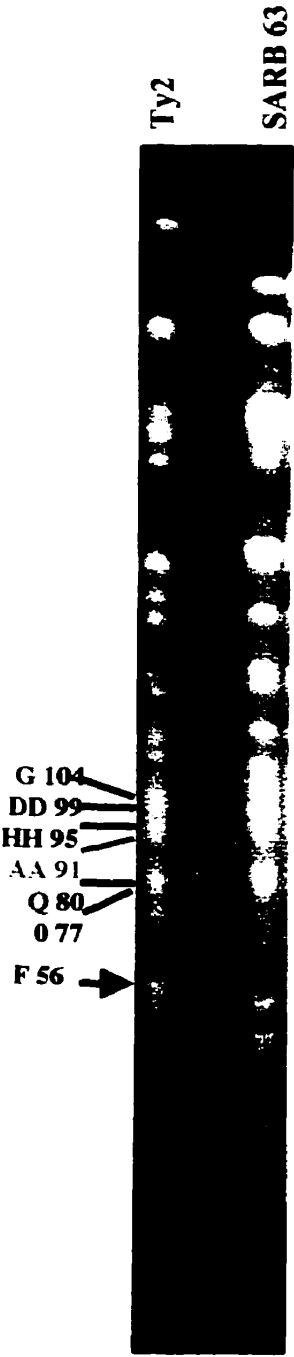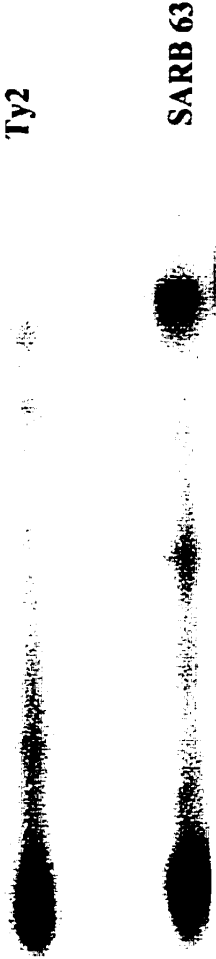
CeuI A fragment, then the *SpeI* AA will be split into two and once probed, will result in two signals. Figure 5.5a compares the *SpeI* genomic cleavage pattern of *S. typhi* Ty2 and SARB 63 and Figure 5.5b shows the probing results after the above cleavage patterns were probed with the *S. typhi* Ty2 *SpeI* AA fragment. From this figure, it can be concluded that in SARB 63, the *SpeI* AA fragment remains intact. This indicates that the I-*CeuI* G fragment in SARB 63 and *S. typhi* Ty2, is in the same position. Thus, in SARB 63 the I-*CeuI* G fragment is adjacent to the A' end of the I-*CeuI* A fragment and the I-*CeuI* C fragment is adjacent to the A end of the I-*CeuI* A fragment, as represented in Figure 5.4, IIa.

CHAPTER 6: RESULTS- Growth studies on *S. typhi* strains with balanced and unbalanced genomes

A. Growth studies to monitor the level of cell fitness in various *S. typhi* strains

Salmonella typhi strains fall into several genome types depending on their I-*Ceu*I restriction fragment order. As mentioned earlier, these genome types are depicted in Figure 2.3. Most of these genome types (1-24) are wild-type *S. typhi* strains which typically carry a balanced genome. A balanced genome maintains close to 180° symmetry between the *oriC* site and the *ter* region. It is known that the *oriC* site is found in the I-*Ceu*I C fragment of *S. typhi* Ty2 and it is assumed to be in the same position in other *S. typhi* strains (Liu and Sanderson, 1996). Similarly, it is known that the *ter* region is found within the I-*Ceu*I A fragment of *S. typhi* Ty2. It is expected that the *ter* region has remained in the I-*Ceu*I A fragment in other *S. typhi* strains, however its position may have varied. Of the genome types shown in Figure 2.3, there are a few genome types (25-27) in which the I-*Ceu*I C fragment has been displaced (Liu and Sanderson, 1996). Since the *oriC* site is found within the I-*Ceu*I C fragment, it is speculated that *S. typhi* strains found within these genome types have their *oriC* site displaced, in relation to *S. typhi* Ty2. It has been previously determined in *E. coli* that displacing the *oriC* site and disrupting the 180° symmetry between origin and terminus reduces cell fitness significantly; returning the *oriC* site back to its original position maximizes cell fitness once again (Hill and Gray, 1988). To determine if the position of the *oriC* site also affects cell fitness in *S. typhi* strains, growth curves were completed to derive generation

time. Generation times of strains with balanced genomes, *S. typhi* Ty2 and ST 308, were compared with those of unbalanced genomes, SARB 63 and 701Ty, and those of several revertant strains. These revertant strains are randomly selected, laboratory-derived variants of SARB 63, in which the I-*Ceu*I C fragment, and consequently the *oriC* site, has returned to its original position. At least two representatives from each class of revertants shown in Figure 2.3b were studied (Ng, 1999). Generation times were used as a measure of cell growth and in this way the level of cell fitness for strains with balanced genomes, unbalanced genomes and revertants were compared.

B. Growth curves and generation times for *S. typhi* strains grown in LB medium

Growth curves were determined by first growing *S. typhi* bacterial cells to saturation and subsequently diluting them and growing them again until early log phase is attained. Lastly, the cells were diluted once again and growth was measured in terms of Klett units. One klett unit is equivalent to 1.0×10^6 bacterial cells (Curry and Ross, 1998).

1. The generation times for S. typhi Ty2, SARB 63 and revertant strain, IV 90

Figure 6.1 illustrates the growth curves for *S. typhi* Ty2, SARB 63 and IV 90. Klett readings were taken every 10 minutes from time zero to 200 minutes. The growth curve of *S. typhi* Ty2 does not show a distinct lag phase. However, there are probably two factors contributing to this: the small number of cells present in the sample at time

Figure 6.1: The growth curves of *S. typhi* Ty2, SARB 63 and IV 90 grown in LB medium.

Growth studies of *S. typhi* Ty2, SARB 63 and IV 90 were conducted in the following manner: Cells were obtained from -70 °C freezer and grown in LB broth overnight. Cells were then diluted to a 0.05% inoculum and grown till log phase, after which the cells were diluted again to a 1.0% inoculum and a growth analysis was done. Klett readings were taken every 10 min. for 3 -3 1/2 hours. Klett readings were taken as a measure of the amount of turbidity in the sample and 1 Klett unit is equivalent to 1×10^6 cells (Curry and Ross, 1998). The time in minutes vs. klett units was plotted on a semi-logarithmic graph. The generation time was calculated using points taken from the exponential portion of the graph representing the log phase. The generation time was calculated using the formula $0.301(T_2-T_1)/$ (log # of cells$_2$- log # of cells$_1$) (Lim, 1998) after individual klett readings were converted into # of cells. The time points, which were taken for each strain, to calculate generation time are shown on the graph. A sample calculation is shown in the results. The generation time of *S. typhi* Ty2 is 28.7 minutes, SARB 63 is 39.6 minutes and IV 90 is 31.1 minutes.

SARB 63
—— S. typhi Ty2
—— IV 90

Time (min.)

Log (Klett units)
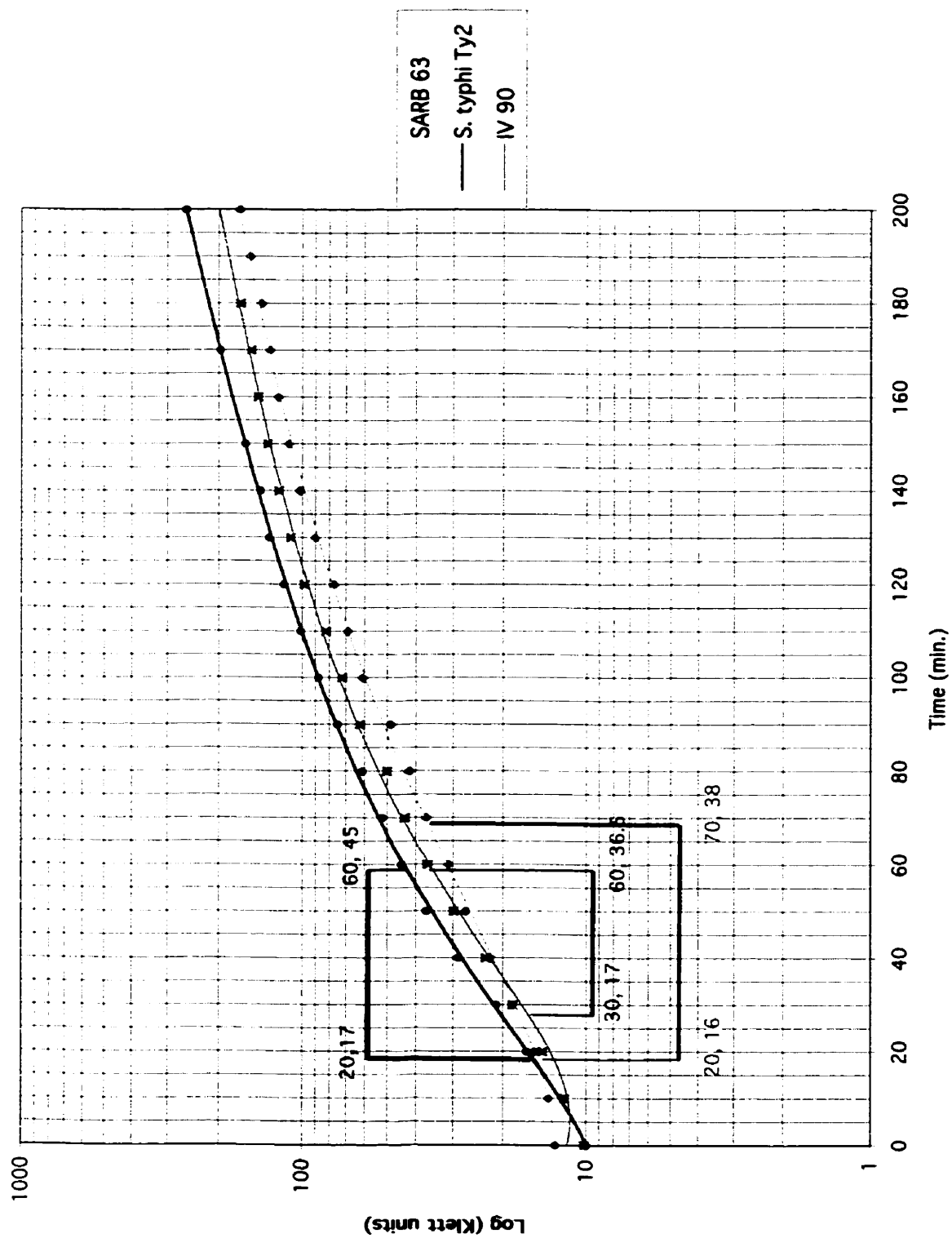
zero and the fact that growth rate is high. The exponential phase appears to be in between 20 minutes and 60 minutes while the growth rate tapers off after this point. SARB 63 has a lag phase, extending from time zero to 20 minutes. The log phase starts at 30 minutes and lasts till about 100 minutes, after which growth rate decreases. Lastly, the revertant strain IV90 follows a similar growth pattern as *S typhi* Ty2. IV90 appears to have a distinct lag phase, unlike *S typhi* Ty2, this may be a result of a larger sample size at time zero. The lag phase of IV90 extends from time 0 to 15 minutes. The log phase extends from 20 minutes to 60 minutes and after 60 minutes growth rate declines.

After comparing the three growth curves in Figure 6.1 visually, it appears that *S. typhi* Ty2 and IV90 have a steeper slope during log phase than SARB 63, this suggests that the growth rate of *S. typhi* Ty2 and IV90 is faster than SARB 63. The generation times calculated for these strains confirm the above observation. The generation time for each strain was calculated using the following formula (Lim, 1998):

$$\text{Generation time} = 0.301 \ (T_2 - T_1)/ \ \text{Log \# cells}_1 - \text{Log \# cells2}$$

$$\text{NB- 1 klett unit} = 1 \times 10^6 \text{ cells}$$

$T_1 = 1^{st}$ time point (min.)   \# cells$_1$= number of cells at 1st time point

$T_2 = 2^{nd}$ time point (min.)   \# cells$_2$ = number of cells at 2nd time point

A sample calculation of generation time using the above formula for *S. typhi* Ty2 is shown below:

$T_1 = 20$ min.   \# cells$_1$ = \# klett units X 1 x10^6

                                 = 1.7 x 10^7

                                 = 1.7 x 10^7 cells

$T_2 = 60$ min.   \# cells$_2$ = 4.5 x 10^7

$$\text{Generation time} = 0.301 \ (60 - 20)/ \ \text{Log } 4.5 \times 10^7 - \text{Log } 1.7 \times 10^7$$

$$= 0.301 \ (40)/ \ 7.65 - 7.23$$

$$= 12.04/ \ 0.42$$

$$= 28.7 \text{ min.}$$

As shown above, the generation time calculated for *S. typhi* Ty2 is 28.7 minutes. Not shown are the calculations for SARB 63 and IV90 but their generation times are 39.3 min. and 31.1 min., respectively. The generation times show that SARB 63 has a 38% slower growth rate than *S. typhi* Ty2, while the growth rate of IV90 is only 8% slower than *S. typhi* Ty2. Consequently, the growth rate of SARB 63 is also slower than IV90 by 27%. These results imply that there is a significant difference between the growth rates of *S. typhi* Ty2 and SARB 63, and IV90 and SARB 63. In contrast, there appears to be little difference in the growth rates of *S. typhi* Ty2 and IV90. Statistical analyses on generation time of these strains confirm the above conclusions; however, the statistical results from these strains and others will be discussed in greater detail later.

2. Variation in the generation times of SARB 63 revertants

A sample growth curve representing each class of revertant is shown in Figure 6.2. At first glance, the growth curves appear fairly similar and the region of the curve representing the log phase seems parallel. However, it can be seen that IV 64 and IV 264 have a slightly steeper slope within the log phase region. Generation times were calculated using Klett readings taken within 20 min. to 60 min. because it is within this range that a clear log phase is occurring in the revertant strains. The generation times ranged between 26.2 minutes and 32 minutes These generation times are listed in Table 6.1. This table will be discussed in detail below.

Table 6.1: A list of generation times for various *S. typhi* strains grown in LB medium: those with balanced genomes, unbalanced genomes and SARB 63 revertant strains.

Classification	*S. typhi* strain No.	Generation Time trial #1 (min.)	Generation Time trial #2 (min.)	Avg. Generation Time (min.)	Avg. Generation Time for each classification (min.)
Balanced Genomes	Ty2	26.2	28.7	27.5	
	ST 308	26.5	27.4	27.0	27.3
Unbalanced Genomes	SARB 63	38.8	39.3	39.2	
	701Ty	39.3	39.3	39.3	39.3
Revertants:					
Group A	IV 52	32.3	33.4	32.9	
	IV 64	26.2	28.2	27.2	
	IV 71	28.2	25.1	26.7	
	IV 122	29.4	26.2	27.8	
	IV 145	33.4	30.1	31.8	29.3
Group B	IV 117	32	33.4	32.7	
	IV 156	32.3	34.7	33.5	33.1
Group C	IV 155	36.9	36.1	36.5	
	IV 292	28.7	26.8	27.8	32.2
Group D	IV 152	27.4	25.8	26.6	
	IV 264	29.5	32.3	30.9	28.8
Group E	IV 58	29.1	28	28.6	
	IV 90	29.2	30.1	30.2	
	IV 228	25.1	25.8	25.5	28.1
Revertant Gen. Time (avg.)					30.3

Figure 6.2: A comparison of the growth curves of SARB 63 revertant strains grown in LB medium

The growth curves of IV 64, IV 117, IV 58, IV 264 and IV 292 are shown. Each strain belongs to a different revertant class. IV 64 belongs to class A, IV 117 belongs to class B, IV 292 belongs to class C, IV 264 belongs to class D and IV 58 belongs to class E. Growth studies were conducted in the same manner as described previously for Figure 6.1.

Figure legend: IV 292, IV 264, IV 64, IV117, IV 58

Y-axis: Log (Klett units)

X-axis: Time (min.)

Annotations: 20 min., 60 min.
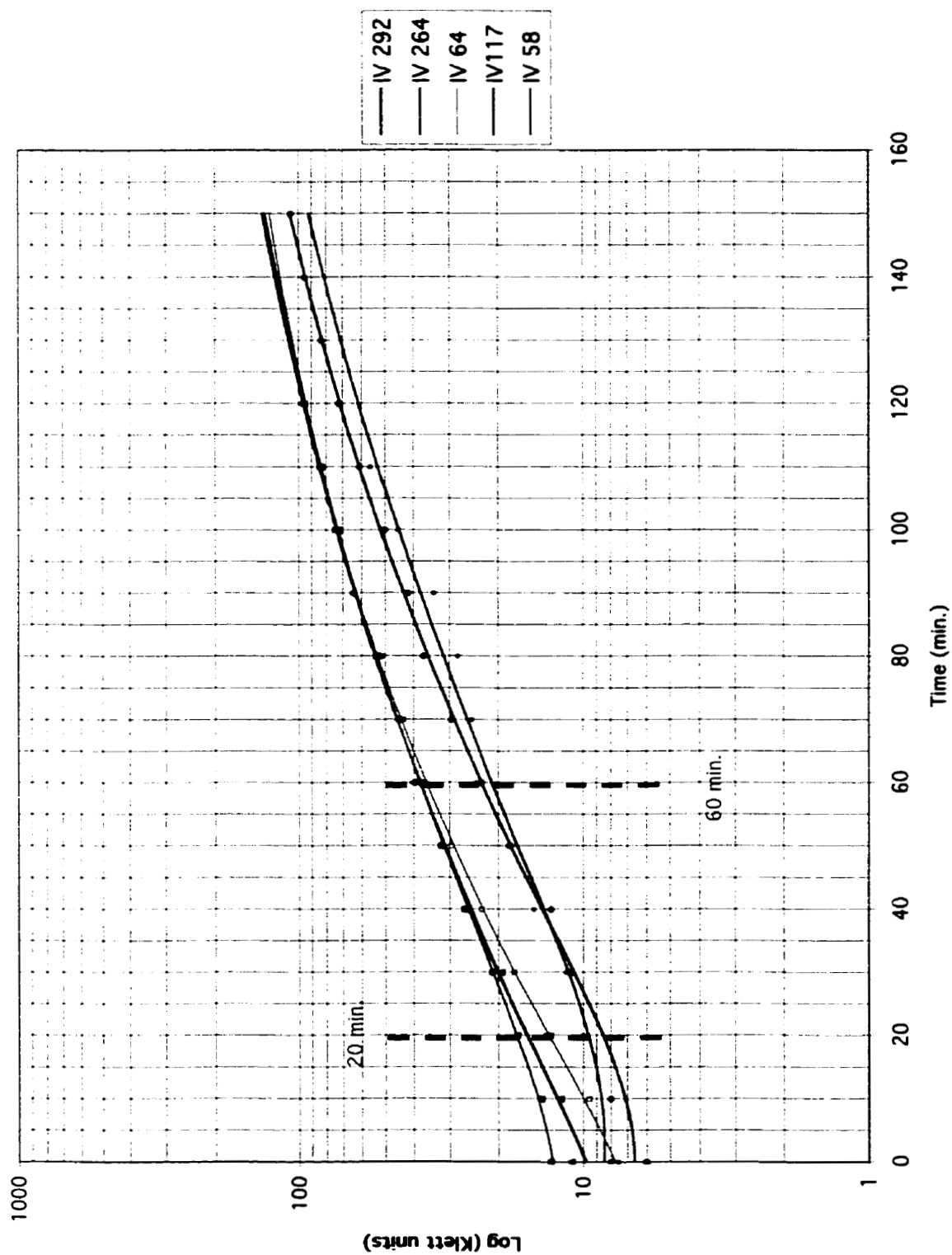
3. An overall comparison of generation time of *S. typhi* strains

The generation times for several *S. typhi* strains were obtained in duplicate and averaged. In addition an average generation time for balanced, unbalanced and revertant genomes was obtained. An average generation time for each class of revertants was also calculated. These are presented in Table 6.1. Two representative strains with balanced genomes were studied, *S. typhi* Ty2 and ST 308. They belong to different genome types, genome type 9 (*S. typhi* Ty2) and genome type 3 (ST 308) and their I-*CeuI* C fragment is roughly in the same position, suggesting that the *oriC* site should also be in the same position (Figure 2.3). The average generation times for these strains are almost identical, for *S. typhi* Ty2 has an average generation time of 27.5 min. and ST 308 has a generation time of 27 min.. Thus, the overall generation time for balanced genomes is around 27.3 min..

To analyze generation time in unbalanced genomes, SARB 63 and 701Ty were studied. In both these strains the I-*CeuI* C fragment is now adjacent to the I-*CeuI* A, causing the *oriC* site to be displaced by about 15% (Figure 2.3). Both SARB 63 and 701Ty have similar generation times, for SARB 63 has an average generation time of 39.2 min. and 701Ty has a generation time of 39.3 minutes. Thus, unbalanced genomes have a generation time which is 44% longer than the generation of balanced genomes. Furthermore, statistical analysis using the Tukey's Multiple Comparison test clearly shows that there is significant difference in the two generation times of balanced and unbalanced genomes. In the Tukey's test, if statistical comparison between two groups (e.g. average generation times of strains with balanced and unbalanced genomes) show

that the probability of the two groups being the same is more than 0.05 (5%), then there is no statistical significance. If this probability for being the same is less than 0.05 or 5% then there is statistical difference. Furthermore, the smaller the value is from 0.05, the higher the significance because the lower the probability of the two groups being the same. The probability of the generation times of balanced and unbalanced genomes being the same is null. This shows that there is a significant difference between the generation times of these two genome types and it is not just random variability between independent generation times.

As mentioned above, the *S. typhi* revertant strains were originally SARB 63 strains with unbalanced genomes. However, these revertant strains have undergone rearrangements which have moved their I-*Ceu*I C fragment, and consequently their *oriC* site, back to a position such that they have balanced genomes. In Table 6.1, the average generation times of the five strains in revertant class A vary from 26.7 minutes to 32.9 minutes; however, statistical results, using the Tukeys test, calculate this difference to be insignificant. The overall average of the generation time in revertant class A is 29.3 minutes. This is similar to the total generation time calculated for all the revertants, which is 30.3 minutes. In group D, the generation time varies from 26.6 minutes to 30.9 minutes and in group E, the generation time varies from 25.5 minutes to 30.2 minutes. There is an even larger difference in generation time is seen within group C; IV292 has a generation time of 27.8 minutes and a IV155 has a generation time of 36.5 minutes. The average generation time for groups C, D and E is 32.2 minutes, 28.8 minutes and 28.1 minutes, respectively. This is again not significantly different from the overall average

generation time for the revertants, which is 30.3 minutes. The average generation time of revertant strains in group B is similar: 32.7 minutes and 33.5 minutes, and the overall generation time is 33.1 minutes. It should be noted that the sample size in groups B, C and D is small, since only two *S. typhi* revertant strains, fall into each group.

The Tukey's test was used to statistically compare the revertant strains to one another individually. The results of this test suggested that there is no significant difference between these revertant strains. When comparing the generation times of the revertant strains to *S. typhi* Ty2 using the Tukeys test, there also appears to be no significant difference, as the probability value is 0.15. This is expected since the average generation time for the revertants is only 10% slower than for *S. typhi* Ty2. However, there is significant difference between the generation times of the revertant strains to SARB 63. This was also expected since the generation time of SARB 63 is 29% slower than that of the revertants. In summary, the above results imply that the generation time of unbalanced genomes is significantly slower than balanced genomes and genomes of revertant strains.

C. Growth curves and generation times for *S. typhi* strains grown in minimal medium

1. The generation times for Ty2, SARB 63 and revertant strain, IV 117

Figure 6.3 illustrates the growth curves for *S. typhi* Ty2, SARB 63 and IV 117. Klett readings were taken every 1/2 hour between time zero and 120 minutes. The log

Figure 6.3: The growth curves of *S. typhi* Ty2, SARB 63 and IV 117 grown in minimal medium.

Growth studies of *S. typhi* Ty2, SARB 63 and IV 117 were conducted in the following manner: Cells were obtained from -70 °C freezer and grown in Minimal broth supplemented with Cas amino acids, Tryptophan, Methione and Cystiene overnight. Cells were then diluted to a 0.1% inoculum and grown till log phase, after which the cells were diluted again to a 1.0% inoculum and a growth analysis was done. Klett readings were taken every 1/2 hour for 2 hours so that log phase could be reached. Klett readings were taken as a measure of the amount of turbidity in the sample and 1 Klett unit is equivalent to 1 x 10^6 cells (Curry and Ross, 1998). The time in minutes vs. klett units was plotted on a semi-logarithmic graph. The generation time was calculated using points taken from the exponential portion of the graph representing the log phase. The generation time was calculated using the formula 0.301(T$_2$-T$_1$)/ (log # of cells$_2$- log # of cells$_1$) (Lim, 1998) after individual klett readings were converted into # of cells. The time points, which were taken for each strain, to calculate generation time are shown on the graph. A sample calculation is shown in the results. The generation time of *S. typhi* Ty2 is 53.2 minutes, SARB 63 is 62.3 minutes and IV117 is 60.3 minutes.

Legend:
— S. typhi Ty2
— SARB 63
— IV 117

Y-axis: Log (Klett units)

X-axis: Time (min.)
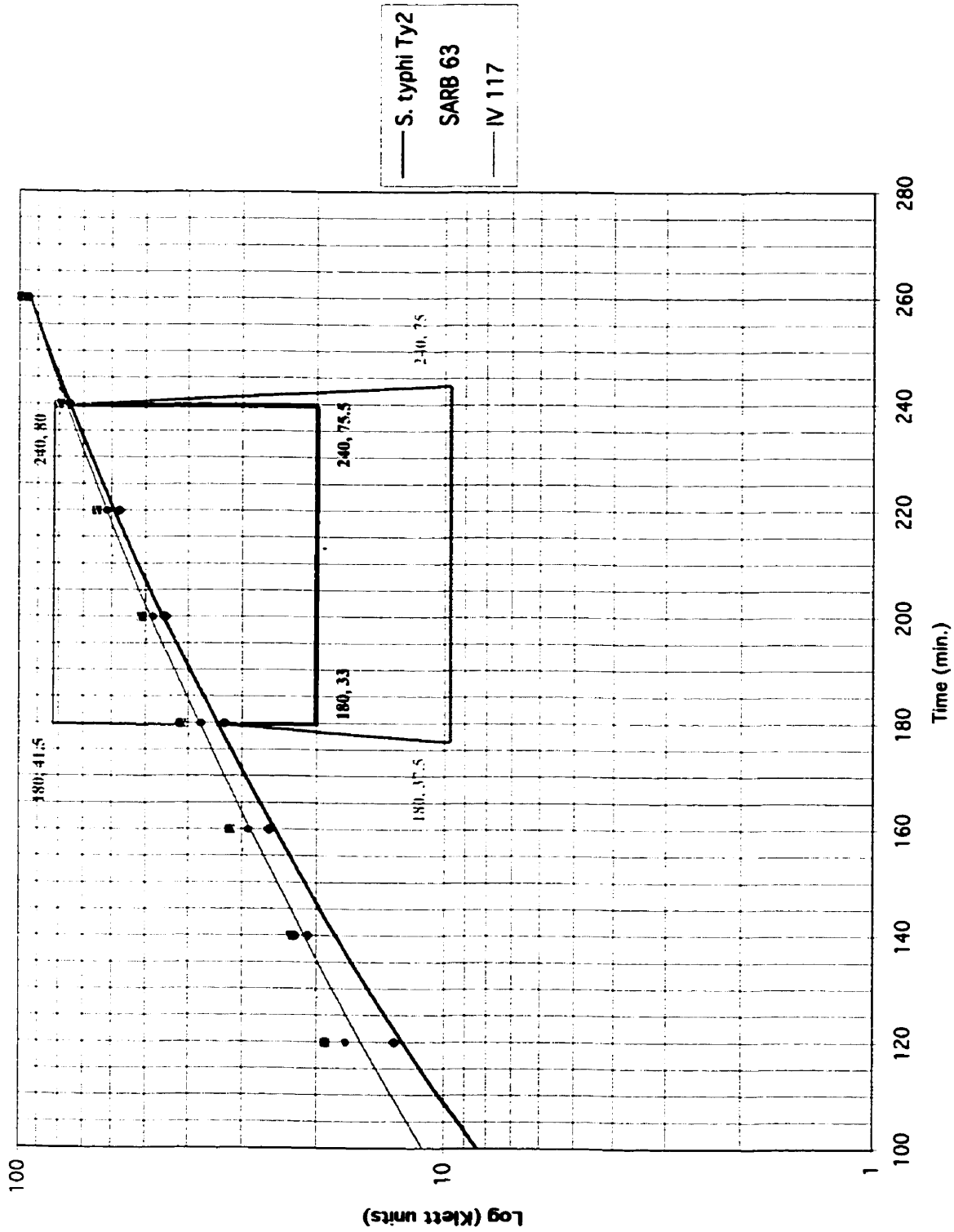
phase of *S. typhi* strains grown in minimal media begins after 120 minutes, as a result, klett readings were taken every 20 minutes for the following two hours. Figure 6.3 graphs time vs. log of # of cells starting at 100 min because prior to that, klett reading were erratic due to a small sample size at time zero. Klett readings were not taken after four hours because the growth rate declines for all the strains studied. Since the purpose of these growth curves was to derive generation time during log phase, only the log phase of these strains is portrayed in Figure 6.3. The generation time was calculated for *S. typhi* Ty2, SARB 63 and IV 117 using klett readings taken from 180 minutes and 240 minutes. *S. typhi* Ty2 does appear to have a shorter generation time than SARB 63 and IV 117; whereas SARB 63 and IV117 have very similar generation times. The generation times and statistical tests will be discussed in detail in the following section.

2. An overall comparison of generation time of *S. typhi* strains

The generation times for several *S. typhi* strains were obtained in duplicate and averaged. In addition an average generation time for balanced, unbalanced and revertant genomes was obtained. This information is presented in Table 6.2 and Figure 6.3. To determine the generation time for balanced genomes, *S. typhi* Ty2 and ST 308 were analyzed. As mentioned before, *S. typhi* Ty2 and ST 308 belong to different genome types, however their I-*CeuI* C fragment and hence their *oriC* site are found roughly in the same position. As expected, the generation times for *S. typhi* Ty2 and ST 308 are similar. *S. typhi* Ty2 has an average generation time of 51.7 minutes and ST 308 has a generation time of 54.6 minutes. Thus, the average generation time for balanced genomes is 53.2 minutes.

Table 6.2: A list of generation times for various *S. typhi* strains grown in minimal medium: those with balanced genomes, unbalanced genomes and SARB 63 revertant strains.

Classification	*S. typhi* strain No.	Generation Time [trial #1] (min.)	Generation Time [trial #2] (min.)	Avg. Generation Time (min.)	Avg. Generation Time for each classification (min.)
Balanced Genomes	Ty2	53.2	50.2	51.7	
	ST 308	53.2	56.0	54.6	53.2
Unbalanced Genomes	SARB 63	62.3	58.3	60.3	60.3
Revertants:					
	IV 52	56	58.8	57.2	
	IV 117	60.3	55.8	58.1	57.7

SARB 63 was the only strain studied to determine the generation time of unbalanced genomes. SARB 63 is the most important member of the balanced genomes since the revertant strains are variants of this strain. Therefore, to conduct proper comparison studies between unbalanced genomes and revertants, SARB 63 should be used as an example of the former. The average generation time for SARB 63 is 60.3 minutes. The Tukey's test compares the generation times of balanced and unbalanced genomes. Significant difference is measured as a function of probability for similarity between these two values. The probability of these generation times being the same is 0.009, this is below the 0.05 threshold value. Therefore, it can be claimed that the generation times of balanced and unbalanced genomes are significantly different and the generation time of unbalanced genomes is 13% longer than the generation time of balanced genomes. *S. typhi* revertant strains IV52and IV117 were analyzed to derive the average generation time for the revertants. IV 52 belongs to revertant class a and IV 117 belongs to revertant class b. Other revertant classes were not studied and the sample size was kept at a minimum since growth studies in LB medium verified that there is no significant difference between members of different classes. The generation time for IV 52 is 57.2 minutes and for IV 117 is 58.1 minutes. The average generation time for the revertant strains is 57.7 minutes.

Statistical analysis using the Tukey's test shows that there is significant different between the revertant strains and balanced genomes but not between revertant strains and unbalanced genomes. The probability for similarity between revertants and balanced genomes is 0.01, which is below the threshold value of 0.05, suggesting that there is

significant difference between these two groups. Conversely, the probability for similarity between revertants and unbalanced genomes (SARB 63) is 0.712, this is above the 0.05 threshold, which implies there is no significant difference revertants and unbalanced genomes.

In summary, it appears that there is significant difference between the generation times of strains with balanced and unbalanced genomes; this result was expected and is similar to that of the comparison between strains with balanced genomes and unbalanced genomes grown in LB medium. However, there is statistical difference between strains with balanced genomes and revertants. In addition, there is no significant difference between strains with unbalanced genomes and revertants. This result is unlike what was observed in the growth studies done using LB medium for similar strains. In the growth studies done using LB medium, there was no significant difference between generation times of balanced genomes and revertants but there was significant difference between generation times of unbalanced genomes and revertants. This was somewhat unexpected and will be discussed later.

CHAPTER 7: DISCUSSION

A. Conservation of genome size in *S. typhi*

In general, the genome of most enteric bacteria is relatively conserved. Variation in genome size is small but common since insertions and deletions occur frequently. Upon analyzing the genome size of *E. coli* K-12, its derivatives and several *Salmonella* species, it can be seen that genome size in these enteric bacteria range from 4500 kb to 4900 kb (Brewer, 1988; Cole and Saint-Girons, 1999). In 14 independent natural isolates of *E. coli* strains studied, it was seen that variation is large, 4660 kb to 5300 kb (Bergthorsson U. and Ochman, 1997; Bergthorsson and Ochman H., 1995). The clinical strain, *E. coli* O157, has an even larger genome, which is 5400 kb in size (Blattner, F., per. comm.). However, variations in natural isolates are generally larger than those seen in laboratory strains. This increase in size is expected since *E. coli* O157 may have acquired additional genes to promote its invasiveness. Therefore, it appears that in enteric bacteria, genome size can be as large as 5400 kb but cannot be less than 4550 kb, suggesting that a genome smaller than this cannot exist or is selected against, possibly due to not enough essential genes being present.

Studies done by Tikki Pang on several independent clinical isolates of *S. typhi*, suggest that *S. typhi* does not follow the above rules on size conservation, as seen for other enteric bacteria. In fact, according to Pang's study, the genome size of *S. typhi* ranges in size from 3980 kb to 4840 kb (Pang, 1998). This variation is striking because it

is not seen in other enteric bacteria. In addition, since the strains studied were clinical isolates of *S. typhi* and were as small as 3980 kb, Pang's results are especially surprising because no other enteric bacteria have genomes this small. However, contrary to the Tikki Pang's observations, preliminary studies conducted on 144 wild-type *S. typhi* strains, suggest very little size variation (Liu et al., 1999). These studies were done by digesting *S. typhi* genomic DNA with I-*Ceu*I restriction endonuclease and separating the DNA on PFGE. The sizes of six out of the seven I-*Ceu*I fragments, the I-*Ceu*I B, C, D, E, F and G, were determined. I-*Ceu*I A fragment was too large to ascertain size in this manner. Conclusions drawn from the total sum of the I-*Ceu*I B, C, D, E, F and G fragments suggest that size is conserved in *S. typhi*. The sum of the I-*Ceu*I B, C, D, E, F and G fragments in the 144 *S. typhi* strains studied vary from 2130 kb to 2460 kb (Liu et al., 1999). The size of the I-*Ceu*I A fragment in *S. typhi* Ty2 is known to be 2400 kb. If it can be assumed that the size of the I-*Ceu*I A is conserved as well, then the total genome size of the *S. typhi* strains studied by Liu would fall into the 4500 kb to 4900 kb range seen in other enteric bacteria.

1. Size conservation in respects to the I-*Ceu*I A fragment of *S. typhi* strains

In this study the size of the I-*Ceu*I A fragment was determined for 28 wild-type *S. typhi* strains. Many of these strains were obtained from Tikki Pang, Institute for Advanced Studies, University of Malaysia. This was accomplished by first digesting *S. typhi* genomic DNA with restriction enzyme I-*Ceu*I and separating the DNA on PFGE. The I-*Ceu*I A fragment was then excised, re-digested by *Xba*I, *Avr*II and *Spe*I restriction endonucleases, α-^{32}P dCTP end-labeled and separated on PFGE. This cut the I-*Ceu*I A

fragment into smaller fragments, whose sizes could then be determined by comparing to

known sizes of *S. typhi* Ty2. The sum of these sizes allowed the total size of the I-*Ceu*I

A fragment to be ascertained. The results indicate that the size of the I-*Ceu*I A fragment

of 27 wild-type *S. typhi* strains range in between 2340 kb to 2460 kb, while the size in *S.*

typhi Ty2, the standard used for comparison, is 2400 kb. This range of about 120 kb is

small considering the I-*Ceu*I A fragment accounts for more than half the genome size of

S. typhi. One would expect this size variation to be larger since in the 144 *S. typhi* strains

studied by Liu, the total sum of I-*Ceu*I B, C, D, E, F and G fragments, whose size makes

up the other half of the genome, ranges a total of 330 kb. If insertions and deletions

occur randomly within a chromosome, there should not be a bias in the size variations of

one half of the chromosome in comparison to the other. It has been previously

determined that the *S. typhi* genome is highly variable (Liu and Sanderson, 1996), for

rearrangements mediated by *rrn* operons occur frequently in *S. typhi*; this phenomenon is

rare in other enteric bacteria like *S. typhimurium* or *E. coli*. In *Salmonella typhi*, *rrn*

operons are found within the half of the chromosome which harbors the I-*Ceu*I B, C, D,

E, F and G fragments and not within the I-*Ceu*I A fragment. It could be possible that

heightened activity caused by the predominant presence of *rrn* operons in one half of the

chromosome is promoting other changes, which affect size, within the same part of the

chromosome. As a result, half of the *S. typhi* chromosome, the half which contains the I-

*Ceu*I B, C, D, E, F and G fragments, has a larger variation in size.

2. Variations in the overall genome size of *S. typhi*

The sizes of the I-*Ceu*I A fragment were combined with the known sizes of the I-*Ceu*I B, C, D, E. F and G fragments for each of the 27 *S. typhi* strains studied. In this way, the total genome sizes of the 27 *S. typhi* strains were determined and these sizes were compared to that of *S. typhi* Ty2, whose total genome size is 4780 kb. Overall, the genome size of these wild-type strains varies from 4504 kb to 4840 kb; however, most of the sizes remain in between 4740 kb to 4840 kb. Only two out of the 27 strains studied have genome sizes smaller than 4740 kb, these are ST1 and SA 4865. In these strains, the size of the I-*Ceu*I A fragment remains fairly constant; however, the size of the I-*Ceu*I G fragments decreases significantly. Thus, it appears that the I-*Ceu*I A fragment contributes very little to the variation in total genome size in *S. typhi*; instead changes in the size of I-*Ceu*I B and G seem to have the greatest affect on overall genome size.

3. Conservation in the overall genome size of SARB 63 revertant strains

SARB 63 is a wild-type *S. typhi* strain in which the I-*Ceu*I C fragment is displaced such that the *oriC* site has moved from its original position by about 15%. SARB 63 revertants are spontaneously arising, laboratory-derived strains in which the I-*Ceu*I C fragment has returned close to its original position such that the *oriC* site is no longer displaced (Ng, 1999). The movement of the I-*Ceu*I C fragment close to its original location results from rearrangements mediated via *rrn* operons.

It was previously determined that over half of the SARB 63 revertant strains had I-*Ceu*I G fragments which were smaller than the I-*Ceu*I G fragments of SARB 63 *and S. typhi* Ty2 by about 200 kb (Ng, 1999). However, analysis of the I-*Ceu*I A fragment of

these revertant strains shows that this size discrepancy is compensated by a 194 kb larger I-*Ceu*I A fragment. As a result, the overall genome size remains conserved in these SARB 63 revertant strains.

In summary, these results indicate that the overall genome size of *S. typhi* is fairly conserved. The size of the I-*Ceu*I A fragment in *S. typhi* changes little and contributes little to the variation in overall genome in the *S. typhi* strains studied. Most of the larger modifications in genome size seen in *S. typhi* are caused by changes in the I-*Ceu*I B and G fragments. However, the variation in total genome size seen in the *S. typhi* strains fall within the range seen in other enteric bacteria. Therefore, although the *S. typhi* genome reorganizes frequently, it is not through large-scale gain or loss of DNA but through genomic recombinations.

B. Rearrangements within the I-*Ceu*I A fragment of wild-type *S. typhi* strains

1. Rearrangements in *S. typhimurium* and within the I-*Ceu*I B to G fragments of *S. typhi* strains

The genome of most enteric bacteria is conserved even though in culture, rearrangements involving *rrn* operons are common in *E. coli* and *S. typhimurium* (Liu and Sanderson, 1995c). To identify rearrangements other than those mediated by *rrn* operons, earlier studies on different wild-type strains of *S. typhimurium* were conducted. Using *Xba*I cleavage patterns of the entire genome as a fingerprinting method, it was observed that variations in cleavage patterns exist between genomes of independent *S.*

typhimurium isolates. However, fingerprinting using I-*Ceu*I restriction endonuclease showed the cleavage patterns to be were very similar in these independent isolates (Liu and Sanderson, 1995b). This suggested that the overall genome structure in *S typhimurium* is conserved; but base pair changes and small rearrangement events may also be common in wild-types. In previous studies, the genome of 144 wild-type *S. typhi* strains were analyzed in regards to the size and arrangement of their I-*Ceu*I fragments (Liu and Sanderson, 1996; Liu and Sanderson, 1995a). These fragments are obtained by digesting the genome with restriction endonuclease I-*Ceu*I, which cuts within the seven *rrn* operons in the *S. typhi* genome, thus produces seven I-*Ceu*I fragments. The size and order of these seven I-*Ceu*I fragments were determined through partial digestion and separation using PFGE. From these studies, it was determined that rearrangements, such as inversions and translocations, mediated via *rrn* operons are common in *S. typhi*. Results showed that many of the *S. typhi* strains had different I-*Ceu*I fragment orders and this gave rise to the 27 genome types identified in *S. typhi* (Liu and Sanderson, 1996). Size studies on I-*Ceu*I B to G fragments were also done in the same study. Results showed that size remained fairly constant in these fragments. In some strains, the I-*Ceu*I fragments have gained or lost DNA but this was a result of insertion and deletion events, not rearrangements, since no reciprocal changes were seen in other fragments of the same chromosome. Therefore, in the 144 *S. typhi* strains within which half of the chromosome, the half which contains the I-CeuI B to G fragments, has been studied in detail, it appears that rearrangements other than those mediated by *rrn* operons are uncommon (Liu and Sanderson, 1996; Liu and Sanderson, 1995a).

2. Rearrangements within the the I-*Ceu*I A fragment of *S. typhi* strains

In regards to the other half of the chromosome in *S. typhi*, the I-*Ceu*I A fragment, not much is known about rearrangements within this fragment. However, the I-*Ceu*I A fragment of *S. typhi* contains the terminus region and rearrangements are known to occur frequently within this region in *S. typhi* Ty2 and other enteric bacteria. When using the terminus region of *Salmonella typhimurium* LT2 as a reference point, it can be seen that the same area within *E. coli* has a 480 kb inverted block. Similarly, in *S. typhi* Ty2 this inverted block is 500 kb while in *Salmonella paratyphi* C it is 700 kb in length (Sanderson et al., 1999; Liu and Sanderson, 1995c). There are three reasons as to why this region is susceptible to recombination. First, the *dif* locus, which is a substrate for site-specific recombinases XerC and XerD, is present within the *ter* region of enteric bacteria (Hill, 1996; Weinstock and Lupski, 1997). Secondly, stalled replication forks are common within the terminus region since within this region DNA replication arrest occurs. These stalled replication forks increase the chances of illegitimate recombination to occur (Sharp, 1991). Lastly, the terminus region is believed to have several recombinational hotspots, which could induce hyperrecombination within this region (Hill, 1996). It should be noted that the inversions observed within the terminus region of *E. coli*, *S. typhi* Ty2, and *S. paratyphi* C did not significantly displace the position of the terminus region and hence, should not have a major effect on genome balance (Liu and Sanderson, 1995c). In other *S. typhi* strains, it is suspected that rearrangements in the I-*Ceu*I A fragment, not mediated via *rrn* operons, will occur but mostly within the area of terminus region.

Rearrangement studies on the I-*Ceu*I A fragment of 28 wild-type *S. typhi* strains proved the following: first, deletions and rearrangements are occurring within the I-*Ceu*I A fragment of these *S. typhi* strains. Second, rearrangements are localized within certain areas of the I-*Ceu*I A fragment, these are the 1000 to 1500 kb region (terminus region) and the 2000 kb to 2500 kb region of the chromosome. Lastly, these rearrangements are small and do not disrupt the integrity of the I-Ceu I A fragment. To identify rearrangements in the *S. typhi* strains studied, all three *Xba*I, *Avr*II and *Spe*I cleavage patterns of the I-*Ceu*I A fragment were compared to that of known cleavage patterns of *S. typhi* Ty2. Variations in the cleavage patterns seen in each strain indicated the region of the I-*Ceu*I A fragment being affected. In the 28 *S. typhi* strains, base pair changes were rarely detected. In these strains, modifications in one of the three enzymatic cleavage patterns, suggesting changes in a particular region of the I-*Ceu*I A fragment, corresponded with changes in the other two cleavage patterns which also pointed to the same region of the chromosome. Base pair changes alone would not be able to explain these patterns since, for example, knocking out an *Xba*I cleavage pattern should only affect the *Xba*I restriction fragments, not the *Avr*II and *Spe*I fragments as well. Furthermore, if base pair changes were more common, the effect would be randomly distributed throughout the I-*Ceu*I A fragment, not restricted to certain regions, as seen in the 28 *S. typhi* strains studied. Therefore, changes caused by deletions and rearrangements are most commonly seen within the I-*Ceu*I A fragment of these strains, since only these kinds of events should affect all three cleavage patterns.

It was expected that rearrangements around the terminus region in the 1000 kb to 1500 kb region of the I-*Ceu*I A fragment would be common in the *S. typhi* strains studied. 12 out of 28 of these strains had modifications within the 1000 kb to 1500 kb region of the chromosome. However, these rearrangements could only be identified in two strains, 3123 (Figure 5.1f) and ST 495 (data not shown), through probing experiments. It is expected that in the other 10 strains, rearrangements are occurring within this region but they are small and do not disrupt the structure of any of the relevant restriction fragments. In 3123 and ST 495, the rearrangement event does encompass the terminus region (Figure 5.1f) but the inversion is small and does not significantly displace the position of the terminus sites. In the other 10 *S. typhi* strains, the recombinational event is expected to be even smaller and hence, should not affect the position of the terminus region either. These results are in accordance with similar changes within the terminis region seen in *E. coli*, *S. typhimurium* and. *S. paratyphi* C, in that rearrangements within the terminus region are common in enteric bacteria but these rearrangements do not displace the position of the terminus sites.

Interestingly, changes within another region of the I-*Ceu*I A fragment, the 2000 to 2500 kb region, were also common in the 28 *S. typhi* strains, even more so than changes seen within the 1000 kb to 1500 kb region. 24 out of the 28 strains had modifications within the 2000 kb – 2500 kb region. The most common ones were centered around the 2000 kb region of the chromosome and these were identified as class II rearrangements (Figure 5.1b). Many strains harboring class II rearrangements had alterations within the 1000 kb to 1500 kb region of the chromosome as well (Figure 5.1c and d). Changes

within the 2000 kb region were not restricted to rearrangements events only since deletions were also identified within this area. Five of the strains carried modifications closer to the 2500 kb region of the chromosome, these probably result from either insertions, deletions or rearrangements but the exact nature is not known since probing experiments were not carried out for this area. The possible presence of IS 200 elements could be what is promoting modifications within this area, since *S. typhi* is thought to have an increased number of IS 200 elements, 20 to 25 in number (Stanley et al., 1994), whereas *S. typhimurium* only has six (Sanderson et al., 1993). It is important to note that the 28 *S. typhi* strains studied were all independent isolates. For this reason, the variation in their cleavage patterns is large when compared to one another. As a result, these strains fell into several rearrangement classes of which only four were shown.

Lastly, in the 28 *S. typhi* strains studied, rearrangements identified within the I-*CeuI* A fragment appear to involve only small regions of the chromosome. The recombinational events, which have been identified in this study, only encompass about 50 kb DNA and the ones which are suspected to occur, but could not be identified, most likely involve less DNA. Thus, in the 28 *S. typhi* strains analyzed, it appears that no large inversion events are occurring, especially none that are involving both the 1300 kb and the 2000 kb region of the chromosome, as was hypothesized prior to this study. Instead, results indicate that modifications within the 1000 kb to 1500 kb region and 2000 kb to 2500 kb region are independent of each other.

3. Rearrangements within the terminus region of SARB 63

The above observations which indicate that changes in the 1000 kb to 1500 kb region and the 2000 kb region of the chromosome are not related, do not fit well with predictions made about the *S. typhi* strain SARB 63. As mentioned earlier, the *oriC* site in SARB 63 is displaced by 15%, this significantly disrupts the 180° symmetry between the *oriC* and terminus sites and consequently produces an unbalanced genome. Unbalanced genomes adversely affect cell fitness and survival (Hill and Gray, 1988) as seen in studies done on *E.coli*. In these studies, laboratory-derived *E. coli* strains which had their *oriC* site displaced by 19% had severely retarded growth, since they were unable to grow on nutrient agar but could grow on minimal medium (Hill et al., 1990). When the *oriC* site was returned to its original position in these strains, the growth rate improved dramatically. In this current study, the position of the *oriC* site in SARB 63 was further refined. The *oriC* site is found within the I-*Ceu*I C fragment, and previously it was determined that the I-*Ceu*I C fragment in SARB 63 is adjacent to the I-*Ceu*I A fragment; however, it was not known which end of the I-*Ceu*I A fragment it was beside (Liu and Sanderson, 1995a; Liu and Sanderson, 1995c). Probing experiments done in this study indicates that in SARB 63, the I-*Ceu*I C fragment takes the position of the I-*Ceu*I B fragment in *S. typhi* Ty2 (Figure5.4I and IIa). This movement from its original position causes the 15% displacement of the *oriC* site. After analyzing the I-*Ceu*I A fragment of SARB 63, it can be concluded that changes within the 1300 kb to 1500 kb region and 2000 kb region are found; however, these changes are independent and are not associated with each other. These results are contradictory to earlier hypothesis, which

predicted that a large inversion within the terminus region of the chromosome of SARB 63, would reposition the terminus region such that the 180° symmetry between the origin and terminus is once again be obtained. This would suggest that SARB 63, does in fact have a balanced genome, not an unbalanced one, which would explain why SARB 63 is a wild-type *S. typhi* strain and a clinical isolate seemingly able to survive and compete in nature. However, the above results opposed this hypothesis and indicated that SARB 63 does indeed harbor an unbalanced genome, in which the origin and terminus sites do not maintain the preferred 180° symmetry.

To summarize, in the 28 *S. typhi* strains studied, rearrangements not mediated via *rrn* operons do occur within the I-*Ceu*I A fragment. Rearrangements around the terminus region, which is within the 1000 kb to 1500 kb region of the chromosome, are common but they are more common within the 2000 kb to 2500 kb region. These rearrangements are believed to be small and independent within each of the above two regions. Large inversions were not seen suggesting that the integrity of the I-*Ceu*I A fragment is conserved in *S. typhi*. Lastly, results also show that *S. typhi* strain SARB 63 does indeed harbor an unbalanced genome and the effect this has on its growth will be discussed in the following section.

C. Growth effects on balanced and unbalanced genomes

1. The balanced genome theory and its affect of cell fitness

The *oriC* and *ter* sites on a circular bacterial chromosome are separated by 180° in a balanced genome; this is thought to be a requirement for optimal cell growth and survival (Hill and Harnish, 1981). Unbalanced genomes, which are a consequence of large insertion or inversion events in the chromosome, have their *oriC* site or *ter* sites displaced in such a way that a 180° symmetry no longer exists. Bacterial cells with unbalanced genomes usually have reduced growth rates, which prevent them from surviving in nature and are usually out-competed by their balanced counterparts (Sanderson et al., 1999; Liu et al., 1999; Hill et al., 1990; Hill and Gray, 1988). In addition, the extent of displacement in the position of either the *oriC* and *ter* sites, determines how severely cell growth is hindered; the further the displacement, the worse the effect is on growth. There are two possible reasons why balanced genomes are required for optimal cell growth; these depend on the importance of gene dosage, and on coordinated replication velocity. Firstly, displacement of the *oriC* site on a chromosome also results in the displacement of genes on the chromosome and could, in turn, disrupt the balance in the amount of gene products, which could be detrimental to the cell (Hill and Harnish, 1981). Secondly, balanced genomes have equal length chromosomal arms, which initiate replication forks at the *oriC* site that move bi-directionally towards the terminus region, at a constant velocity. However, genomes which are unbalanced have unequal length chromosomal arms and as a result, their replication forks may observe

differential velocities in certain regions of the chromosome, which can lead to incorrect or slow replication of the genome (Francois et al., 1990). This may also be a reason why cells with unbalanced genomes have such reduced growth rates.

2. Balanced and unbalanced genomes in *S. typhi* strains and their affect on cell growth

To determine if similar patterns for cell growth and fitness applies for *S. typhi*, in this study the cell growth of wild-type *S. typhi* strains with balanced and unbalanced genomes were compared. In addition, laboratory-derived strains, in which the unbalanced genomes have reverted back to its balanced form, were also studied in regards to growth. Two representatives of balanced genomes (*S. typhi* Ty2 and ST 308) and unbalanced genomes (SARB 63 and 701 Ty) were used. The strains with their *ori*C site displaced by 15%. Spontaneously derived laboratory strains of SARB 63, in which reversions have moved the *ori*C site close to its original position, were also studied. Growth in nutrient medium indicate that there is a significant growth difference between balanced and unbalanced genomes, since unbalanced genomes (generation time 39.3 minutes) grow 44% slower than balanced ones (generation time 27.3 minutes). The growth rate of the revertants (*S. typhi* strains with genomes that have converted from unbalanced to balanced) have an generation time of 30.3 minutes, which represents a significant difference from the generation times of unbalanced genomes but not from balanced genomes. This suggests that replacing the *ori*C site back to its original position improves cell growth and fitness.

Similar but not identical results were obtained in earlier studies done on *E.coli* K12 mutants with the *oriC* unbalanced by 19%. The unbalanced strains failed to grow at all in nutrient medium; however, growth was restored once balance was reestablished (Hill and Gray, 1988). Interestingly, this growth reduction in unbalanced genome is far more severe in *E. coli* than in *S. typhi*, since in *E.coli* no growth was seen whereas in *S. typhi* growth was reduced by only 44%. The Adopt-Adapt model has proposed to explain genomic rearrangements and bacterial speciation in *S. typhi*. Previous studies have shown that *S. typhi* has acquired novel blocks of DNA, up to 120 kb each, presumably through lateral transfer (Liu and Sanderson, 1995c). This may result in asymmetry between the *oriC* site and the *ter* region, causing the chromosome to be unbalanced, as seen in SARB 63 and 701Ty. It was thought that genomic reorganizations, through recombinational events, was a way of restoring the 180° separation. It was expected that in SARB 63 and 701Ty, this reorganization would occur in the terminus region such that the genome would be almost balanced. This would explain the less dramatic reduction in cell growth of unbalanced genomes of *S. typhi* has when compared to *E. coli*. However, in SARB 63 at least, there appears to be no evidence of significant changes around the terminus region which would allow the *oriC* and terminus sites to align in the proper 180° symmetry. Therefore, although *S. typhi* strains with unbalanced genomes grow slower than those with balanced genomes, growth is not prevented on nutrient medium, as seen *E. coli* strains. In addition, it was expected that the revertant strains, with newly restored balanced genomes, should have the same generation time as the strains with pre-existing balanced genomes, but their generation times are still 10% slower. However, it

should be noted that statistical tests deemed this 10% difference to be not significant. These results indicate that in *S. typhi* balanced genomes are not as stringently required for cell growth. For this reason, genome structure is not as well conserved within *S. typhi* strains as seen in other enteric bacteria. Growth studies using balanced, unbalanced and revertant *S. typhi* strains were also conducted in lesser detail in minimal media. Results show that differences in the generation times between the three groups of strains is not as dramatic as seen for growth in nutrient media. This is similar to the studies done in *E. coli*, where strains with unbalanced genomes grown on minimal medium show limited growth, whereas in rich medium they show no growth (Hill and Gray, 1988). In *S. typhi*, strains with balanced and unbalanced genomes and revertants all had longer generation times in minimal media than in rich medium. Strains with balanced genomes, like *S. typhi* Ty2, and revertant strains grow only half as fast in minimal medium as in nutrient medium. Strains with unbalanced genomes also grow slower in minimal medium than in nutrient medium. This indicates *that S. typhi* grows slower in minimal medium than in rich medium, which is expected since less available nutrients in minimal medium limits the growth rate of bacterial cells. The difference in generation time between balanced and unbalanced genomes is much less. SARB 63 (unbalanced genome) has a generation time which is only 13% longer than S. typhi Ty2 (balanced genome); this is not as significant as the 44% longer generation time seen between SARB 63 and *S. typhi* Ty2 in nutrient medium. It could be possible that genome imbalance is more tolerated in slow growing conditions than in fast ones; since there is more time for completion of DNA replication prior to cell division. This may be beneficial for strains with unbalanced

genomes because it would allow more time for the uneven replication fork to reach the terminus regions from both ends. When comparing the generation times of the revertant strains with balanced genomes and unbalanced genomes the following observations are made: the generation times of revertant strains compared to strains with balanced genomes are significantly different, in fact, the generation times of revertant strains and strains with unbalanced genomes are similar. This is opposite to what was seen when the same comparison made with these groups of strains grown in nutrient medium. Thus, the results indicate that restoring genome balance, as done in the *S. typhi* revertants strains, does not improve cell growth and restore cell fitness in these strains when grown in minimal medium.

In summary, in *S. typhi* it appears that genome balance does affect cell fitness but not as stringently as seen in *E. coli*. For this reason, *S. typhi* strains with unbalanced genomes still grow, although much more slowly, in nutrient medium whereas *E. coli* strains with unbalanced genomes grown, fail to grow at all in nutrient medium. In addition, restoring genome balance in previously unbalanced strains of *S. typhi* (revertant strains) does improve growth but still not as efficiently as the growth seen in strains with pre-existing balanced genomes. Furthermore, these revertant strains when grown in minimal media show no significant improvement in growth rate. These results indicate that in *S. typhi* balanced genomes are not as stringently required for cell survival and growth. This may explain why genome structure is not as well conserved within *S. typhi* strains, as in *E. coli*.

CHAPTER 8: LITERATURE CITED

Andersson, R.P. and Roth, J.R. (1978). Gene duplication in bacteria: alteration of gene dosage by sister chromosome exchange. Cold Spring Harbor Symp. Quantit. Biol *43*, 1083-1087.

Bergthorsson, U. and Ochman H. (1995). Heterogeneity of genome sizes among natural isolates of *Escherichia coli*. Journal of Bacteriology *177 (20)*, 5784-5789.

Bergthorsson U. and Ochman, H. (1997). Evolution of the *E. coli* genome. *In* Bacterial Genomes: Physical Sturcture and Analysis. de Bruijn Frans J., J.R. Lupski, and G.M. Weinstock, eds. (Florence, KY: International Thomson Publishing), pp. 177-186.

Blattner, F., Plunkett, G., Bloch, C., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J., Rode, C., Mayhew, G., Greyor, J., Davis, N., Kirkpatrick, H., Goeden, M., Rose, D., Mau, B., and Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. Science *277*, 1453-1462.

Brewer, B.J. (1988). When polymerases collide: replication and the transcriptional organization of the *E. coli* chromosome. Cell *53*, 679-686.

Campbell, N. (1997). The vertebrate genealogy. *In* Biology. Menlo Park, CA: The Benjamin/Cummings Publishing Company), pp. 628-665.

Chee, C.S., Noordin, N., and Ibrahim, L. (1992). Epidemiology and control of typhoid in

Malasyia. *In* Typhoid fever: strategy for the 90's. T. Pang, C.L. Koh, and C.L.

Puthucheary, eds. (Singapore: World Scientific Publishing), pp. 3-10.

Cole, S. and Saint-Girons, I. (1999). Bacterial genomes-all shapes and sizes. *In*

Organization of the Prokarytic Genome. R.L. Charlebois, ed. (Washington, D.C.:

American Society of Microbiology) (In press).

Curry, C.L. and Ross, H. (1998). Biology 231 Laboratory Manual. University of Calgary:

pp. 11.2-11.3

Edelman, R. and Levine, M.M. (1986). Summary of an international workshop on

typhoid fever. Rev. Infect. Dis. *8*, 329-349.

Francois, V., Louarn, J., Rebollo, J.-E., and Louarn, J.-M. (1990). Replication

termination, nondivisible zones, and structure of the *Escherichia coli* chromosome. *In*

The Bacterial Chromosome. K. Drlica and M. Riley, eds. (Washington, D.C.: American

Society of Microbiology), pp. 351-359.

Hill, C.W. and Combraito, G. (1973). Genetic duplications induced at very high

frequency by ultraviolet irradiation in *Escherichia coli*. Mol. Gen. Genet. *127*, 197-214.

Hill, C.W. and Gray, J.A. (1988). Effects of chromosomal inversion on cell fitness in *Escherichia coli* K-12. Genetics *119*, 771-778.

Hill, C.W. and Hamish, B.W. (1981). Inversions between ribosomal RNA genes of *Escherichia coli*. Proceedings of the National Academy of Science *78*, 7069-7072.

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

Hill, T.M. (1996). Features of the chromosomal terminus region. *In Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. F. Neidhardt, ed. (Washington, D.C.: American Society of Microbiology), pp. 1603-1613.

Kogoma, T. (1997). Origins of chromosome replication. *In* Bacterial Genomes: Physical Sturcture and Analysis. de Bruijn Frans J., J.R. Lupski, and G.M. Weinstock, eds. (Florence, KY: International Thomson Publishing), pp. 67-77.

Kuempel, P.L., Pelletier, A.J., and Hill, T.M. (1990). *tus* and the terminators: inhibition of replication in *Escherichia coli*. *In* The Bacterial Chromosome. K. Drilica and M. Riley, eds. (Washington D.C.: American Society of Microbiology), pp. 297-306.

Lim, D. (1998). Growth and control of growth. In Microbiology, (Boston: WCB McGraw-Hill Companies), pp. 106-107.

Linton, L. and Harder, L.D. (1998). Biology 315 Laboratory Manual. University of Calgary: pp. 13.1-13.2

Liu, S.-L., Hessel, A., Cheng, Y.M., and Sanderson K. (1994). The *XbaI-BlnI-I-CeuI* genomic cleavage map of *Salmonella paratyphi* B. Journal of Bacteriology *4*, 1014-1024.

Liu, S.-L., Hessel, A., and Sanderson, K. (1993). The *Xba* I-*Bln* I-I *Ceu* I genomic cleavage map of *Salmonella enteritidis* shows an inversion relative to *Salmonella typhimurium* LT2. Molecular Microbiology *10*, 655-664.

Liu, S.-L. and Sanderson, K. (1995). The chromosome of *Salmonella paratyphi* A is inverted by recombination between *rrnH* and *rrnG*. Journal of Bacteriology *177*, 6585-6592.

Liu-Shu-Lin, Hessel, A., and Sanderson, K. (1993). The *Xba* I-*Bln* I-I *Ceu* I genomic cleavage map of *Salmonella typhimurium* LT2 determined by double digestion, end labeling, and pulsed-field gel electrophoresis. Journal of Bacteriology *175*, 4104-4120.

Liu, S.-L., Khakhria, R., Woodward, D., Johnson, W., and Sanderson, K. (1999). Diversity of genome structure of *Salmonella typhi* populations and characterization of strains of different gemone types. (In preparation).

Liu, S.-L. and Sanderson, K. (1995a). Rearrangements in the genome of the bacterium *Salmonella typhi*. Proceedings of the National Academy of Science *92*, 1018-1022.

Liu, S.-L. and Sanderson, K. (1995b). I-*Ceu* I reveals conservation of the genome of independent strains of *Salmonella typhimurium*. Journal of Bacteriology *177*, 3355-3357.

Liu, S.-L. and Sanderson, K. (1995c). Genomic cleavage map of *Salmonella typhi* Ty2. Journal of Bacteriology *177*, 5099-5107.

Liu, S. and Sanderson, K. (1996). Highly plastic chromosomal organization in *Salmonella typhi*. Proceedings of the National Academy of Science *93*, 10303-10308.

Lloyd, R. and Low, K.B. (1996). Homologous recombination. *In Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. F. Neidhardt, ed. (Washington, D.C.: American Society of Microbiology), pp. 2244-2245.

Mahan, M., Segall, A., and Roth, J. (1990). Recombination events that rearrange the chromosome: barriers to inversion. *In* The Bacterial Chromosome. K. Drilica and M. Riley, eds. (Washington D.C.: American Society of Microbiology), pp. 341-349.

Masters, M. (1996). Generalized transduction. *In Escherichia coli* and *Salmonella*

typhimurium: cellular and molecular biology. F. Neidhardt, ed. (Washington, D.C.:

American Society of Microbiology), pp. 2432-2433.

Ng, Ivy. Genomic rearrangement in *Salmonella typhi*. MSc. Thesis. 1999. University of

Calgary.

Ochman, H. and Wilson, A.C. (1987). Evolutionary history of enteric bacteria. *In*

Escherichia coli and *Salmonella typhimurium*: cellular and molecular biology. F.

Neidhardt, ed. (Washington, D.C.: American Society of Microbiology), pp. 1649-1654.

Ornston, N., Neidle, E., and Houghton, J. (1990). Gene rearrangements, a force for

evolutionary change; DNA sequence arrangements, a source of genetic constancy. *In* The

Bacterial Chromosome. K. Drilica and M. Riley, eds. (Washington D.C.: American

Society of Microbiology), pp. 325-334.

Pang, T. (1998). Genetic dynamics of *Salmonella typhi*-diversity in clonality. Trends in

Microbiology *6*, 339-342.

Riley, M. and Sanderson, K. (1990). Comparative genetics of *Escherichia coli* and

Salmonella typhimurium. *In* The Bacterial Chromosome. K. Drlica and M. Riley, eds.

(Washington, D.C.: American Society for Microbiology), pp. 85-95.

Sanderson, K., McClelland, M., and Liu, S.-L. (1999). Stable genomes. *In* Organization of the Prokaryotic Genome. R.L. Charlebois, ed. (Washington, D.C.: American Society for Microbiology), (In press).

Sanderson, K., Sciore, P., Liu, S.-L., and Hessel, A. (1993). Location of IS 200 on the genomic cleavage map of *Salmonella typhimurium* LT2. Journal of Bacteriology *175* *(23)*, 7624-7628.

Schmid, M.B. and Roth, J. (1987). Gene location affects expression level in *Salmonella typhimurium*. Journal of Bacteriology *169*, 2872-2875.

Sharp, P.M. (1991). Determinants of DNA sequence divergence between *Escherichia coli* and *Salmonella typhimurium:* codon usage, map position, and concerted evolution. Journal of Molecular Evolution *33*, 23-33.

Shimkets, L.J. (1997). Structure and sizes of genomes of the archaea and bacteria. *In* Bacterial Genomes: Physical Structure and Analysis. de Bruijn Frans J., J.R. Lupski, and G.M. Weinstock, eds. (Florence, KY: International Thomson Publishing), pp. 5-11.

Snustad, P., Simmons, M., and Jenkins, J. (1997). Mutation, DNA repair and recombination. *In* Principles of Genetics. AnonymousToronto, ON: John Wiley and Sons, Inc.), pp. 311-347.

Stanley, J., Powell, N., Jones, C., and Burnens, A. (1994). A framework for IS 200, 16S rRNA gene and plasmid-profile analysis in *Salmonella* serogroup D1. Journal of Medical Microbiology *41*, 112-119.

Thong, K.-L., Cheong, Y.-M., Puthucheary, S., Koh, C.-L., and Pang Tikki (1994). Epidemiologic analysis of sporadic *Salmonella typhi* isolates and those from outbreaks by pulsed-field gel electrophoresis. Journal of Clinical Microbiology *32*, 1135-1141.

Weinstock, G.M. and Lupski, J.R. (1997). Chromosomal rearrangements. *In* Bacterial Genomes: Physical Sturcture and Analysis. de Bruijn Frans J., J.R. Lupski, and G.M. Weinstock, eds. (Florence, KY: International Thomson Publishing), pp. 112-118.

Zyskind, J. (1990). Priming and growth rate regulation: questions concerning initiation of DNA replication in *Escherichia coli*. *In* The Bacterial Chromosome. K. Drlica and M. Riley, eds. (Washington, D.C.: American Society for Microbiology), pp. 269-278.

.