May 1998

The Influence of a Human Repetitive Dna on Genome Stability

Eugenia L. Posey
East Tennessee State University

Follow this and additional works at: https://dc.etsu.edu/etd
Part of the Biochemistry Commons, and the Molecular Biology Commons

Recommended Citation

This Dissertation - Open Access is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
THE INFLUENCE OF A HUMAN REPETITIVE DNA ON GENOME STABILITY

A Dissertation
Presented to
the Faculty of the Department of Biochemistry
and Molecular Biology
James H. Quillen College of Medicine
East Tennessee State University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Biomedical Science

by
Eugenia Lee Posey
May 1998
APPROVAL

This is to certify that the Graduate Committee of Eugenia Lee Posey met on the 24th day of February, 1998.

The committee read and examined her dissertation, supervised her defense of it in an oral examination, and decided to recommend that her study be submitted to the Graduate Council, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biomedical Science.

Signed on behalf of the Graduate Council

Interim Dean, School of Graduate Studies
ABSTRACT

THE INFLUENCE OF A HUMAN REPETITIVE DNA ON GENOME STABILITY

by

Eugenia Lee Posey

A uniquely human interspersed repetitive DNA sequence family, the L2Hs, are highly polymorphic in human genomes. Several features of interspersed repeated DNA may contribute to the instability observed. Certain motifs (direct repeats, palindromes, and inverted repeats) comprising L2Hs elements may adopt unusual secondary structures such as cruciforms or hairpins. These motifs have been associated with features of genome instability in recombination, insertions and deletions. The L2Hs elements also are AT-rich (76%) compared to the bulk of human DNA (52%). That their dynamic nature (i.e. polymorphisms) may arise from recombination, insertions and deletions has led to the hypothesis that the L2Hs element is intrinsically dynamic and may influence the stability of the surrounding genome. Thus, the stability of the L2Hs element was tested in a bacterial model system. A cloned 0.6 kb L2Hs element forms non-B-form structures in recombinant plasmids pN6 and pN2, which differ only in insert orientation. Instability of pN6 and pN2 plasmids was observed in serial propagation studies in which E.coli cells containing the plasmids were cultured every 24 hours for 28 days. The vector plasmid pTZ19U, as control, was found to be stable in all passages while the two L2Hs recombinants developed deletions of the L2Hs insert as well as adjacent vector sequences. The isolated deletion mutants have been characterized via restriction cleavage studies and sequencing to map the boundaries of the deletions. Direct repeats and potential stem-loop structures have been discovered at or within close proximity to the deletion boundaries. The data demonstrate that the L2Hs recombinants' unusual sequence features with potential for non-B-form secondary structures, influence genome stability via their involvement in generating errors during DNA replication and DNA repair.
ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to all my teachers and mentors along this quest for knowledge and truth, especially to my high school math teacher Mr. John Michael. I want to thank all the faculty and staff of the Biochemistry and Molecular Biology Department at ETSU for their support and encouragement.

To the members of my graduate advisory committee, Dr. Scott Champney, Dr. Lou Ernst-Fonberg, Dr. John Laffan, Dr. Foster Levy, and Dr. Phillip R. Musich, I extend my deepest appreciation.

To Dr. Phillip R. Musich, my advisor, I am eternally indebted for molding a scientist out of a formless piece of clay. Thank you. Dr. Musich provided me with guidance, skill, advice and discipline so that I would be able to think for myself.

I want to thank my parents and especially my sisters who encouraged me to take a leap of faith into a new endeavor and supported me along the way.

Special thanks to all my dear friends at ETSU and to Dumi, Harold, Sharon, Raymonde, and the two Karens. And thank you Shirl and Sally, whose antics, jokes, laughter, and food kept me going all those late nights of laboratory work. I’ll always have “precious memories.”

iv
## CONTENTS

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROVAL</td>
<td>ii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>xiii</td>
</tr>
</tbody>
</table>

## Chapter

1. **INTRODUCTION** ........................................ 1

   The Relationship of Genome Instability to Human Disease .................................. 2

   Repeated DNA Sequences Involved With Genome Instability .................................. 4

   Interspersed Repeated DNA and Non-B-form DNA Structures: Effects On Genome Instability ... 6

   DNA Slippage and Unusual Sequence Motifs .... 10

   DNA Repair Enzymes and Genome Instability .... 15

   Palindromes and Genomic Instability ........... 17

   Inviability Versus Instability ............... 19

   DNA Unwinding Elements and Genome Instability ................................................. 20

   Cell Strains and Genotypes That May Affect Genome Stability ................................ 20
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of the L2Hs Element</td>
<td>22</td>
</tr>
<tr>
<td>Hypotheses Concerning the L2Hs Element</td>
<td>23</td>
</tr>
<tr>
<td>Summary</td>
<td>24</td>
</tr>
<tr>
<td>2. MATERIALS AND METHODS</td>
<td>25</td>
</tr>
<tr>
<td>Materials</td>
<td>25</td>
</tr>
<tr>
<td>Methods</td>
<td>28</td>
</tr>
<tr>
<td>Bacterial Strains and Growth Conditions</td>
<td>28</td>
</tr>
<tr>
<td>Plasmids</td>
<td>28</td>
</tr>
<tr>
<td>Serial Passaging of Cells</td>
<td>30</td>
</tr>
<tr>
<td>Methods for Producing Competent <em>E. coli</em> Cells and Their Transformation</td>
<td>31</td>
</tr>
<tr>
<td>Electroporation Method</td>
<td>31</td>
</tr>
<tr>
<td>Transformation Using Heat Pulse</td>
<td>32</td>
</tr>
<tr>
<td>Two-layer Plating</td>
<td>33</td>
</tr>
<tr>
<td>Preparation of Plasmid DNA</td>
<td>33</td>
</tr>
<tr>
<td>Hot Alkaline Lysis Method</td>
<td>33</td>
</tr>
<tr>
<td>Magic™ Miniprep Method</td>
<td>34</td>
</tr>
<tr>
<td>Magic™ Maxiprep Method</td>
<td>36</td>
</tr>
<tr>
<td>Quick Boil Miniprep (Qbmp)</td>
<td>36</td>
</tr>
<tr>
<td>Alcohol Precipitation of DNA</td>
<td>38</td>
</tr>
<tr>
<td>Preparation of Intact Cellular DNAs for Analysis by Field Inversion Gel Electrophoresis</td>
<td>38</td>
</tr>
</tbody>
</table>
Chapter Page

Gel Electrophoresis of DNA ........ 40
Standard Agarose Gel Electrophoresis .................. 40
Polyacrylamide Gel Electrophoresis (PAGE) ........ 40
Sequencing Gel Electrophoresis .......... 41
Field Inversion Gel Electrophoresis ... 42
DNA Cleavage by Restriction Enzyme Digestion ........ 43
DNA Dideoxynucleotide Thermocycle Sequencing .......... 43
 Autoradiography .................. 45
DNA Primary Sequence Analysis ........ 45
Polymerase Chain Reaction Labeling of βLac and Ori Probes 46
Bidirectional Blotting and Hybridization of Field Inversion Gels 47
Chemiluminescent Detection of Hybridized Membranes ................ 48
Subcloning .......................... 49
Plating Experiment With Varying Amounts of Ampicillin ........ 49
Preparation of Colonies On Matrix Array With Varying Amounts of Ampicillin ... 49

3. RESULTS ................................................. 51
Construction of Plasmids .................. 51
Experimental Overview .................. 53
Initial Serial Passages .................. 54
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subcloning of pTZ19U</td>
<td>56</td>
</tr>
<tr>
<td>Linearization of Plasmid DNAs</td>
<td>56</td>
</tr>
<tr>
<td>Subcloning of Plasmid from JC7623 Cells</td>
<td>57</td>
</tr>
<tr>
<td>Sizing by Restriction Cleavage (SspI cuts)</td>
<td>60</td>
</tr>
<tr>
<td>Sizing by Restriction Cleavage (Acc65I)</td>
<td>60</td>
</tr>
<tr>
<td>Mapping Deletion Boundaries</td>
<td>63</td>
</tr>
<tr>
<td>\textit{HaeIII} Restriction Cleavage</td>
<td>63</td>
</tr>
<tr>
<td>\textit{HaeIII} + \textit{HinfI} Restriction Cleavage</td>
<td>63</td>
</tr>
<tr>
<td>Refinement of Deletion Boundaries Through Additional Restriction Mapping</td>
<td>68</td>
</tr>
<tr>
<td>Screening Additional Subclones for Deletions</td>
<td>69</td>
</tr>
<tr>
<td>Comparison of pN2, pN6, and pRh Plasmids</td>
<td>73</td>
</tr>
<tr>
<td>Screening JC7623 Cells for Reduced pN6 Plasmid Copy Number</td>
<td>75</td>
</tr>
<tr>
<td>DNA Sequencing Data</td>
<td>81</td>
</tr>
<tr>
<td>4. DISCUSSION</td>
<td>87</td>
</tr>
<tr>
<td>Retrospection and Additional Concerns</td>
<td>106</td>
</tr>
<tr>
<td>The Relationship of the Bacterial Model to the Human System</td>
<td>108</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>114</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>125</td>
</tr>
<tr>
<td>APPENDIX A, Buffers and solutions</td>
<td>126</td>
</tr>
<tr>
<td>APPENDIX B, pN2 Sequence</td>
<td>129</td>
</tr>
</tbody>
</table>
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Primers Used for PCR and DNA Sequencing</td>
<td>27</td>
</tr>
<tr>
<td>2.</td>
<td><em>Escherichia coli</em> Strain List</td>
<td>29</td>
</tr>
<tr>
<td>3.</td>
<td>Tally of Present, Missing or New Bands from PAG Assays of PN2 Plasmids Restricted with HaeIII</td>
<td>65</td>
</tr>
<tr>
<td>4.</td>
<td><em>AluI</em>, <em>MspI</em>, <em>HaeII</em>, and <em>DdeI</em> Restriction Fragments of P14JC3 and P28JC4 Subclones Corresponding to Figure 17</td>
<td>72</td>
</tr>
<tr>
<td>5.</td>
<td>Results of Gel Assays of Plasmid DNA Isolated from Serial-Passaged JC7623 Cells Containing PN2</td>
<td>74</td>
</tr>
<tr>
<td>6.</td>
<td>PN2 Subclone Deletion Boundaries and Sizes of Deletions</td>
<td>82</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Repeat motifs in the L2Hs element that may adopt unusual secondary structures and participate in replication slippage events</td>
<td>7</td>
</tr>
<tr>
<td>2.</td>
<td>Hairpin and cruciform formation</td>
<td>9</td>
</tr>
<tr>
<td>3.</td>
<td>Two pathways of cruciform formation</td>
<td>11</td>
</tr>
<tr>
<td>4.</td>
<td>Two adjacent direct repeats and the occurrence of slipped mispaired DNA</td>
<td>13</td>
</tr>
<tr>
<td>5.</td>
<td>Slipped mispairing of an interspersed direct repeat DNA during DNA replication</td>
<td>14</td>
</tr>
<tr>
<td>6.</td>
<td>A triplet repeat expansion model</td>
<td>16</td>
</tr>
<tr>
<td>7.</td>
<td>Vector plasmid pTZ19U and recombinants</td>
<td>52</td>
</tr>
<tr>
<td>8.</td>
<td>Gel analysis of pN2 and vector pTZ19U plasmid DNA passaged for 28 days in E. coli strain JC7623 cells</td>
<td>55</td>
</tr>
<tr>
<td>9.</td>
<td>Gel assay of passaged pTZ19U and pN2 plasmids cleaved with SspI that were isolated from host DH5αF'IQ cells</td>
<td>58</td>
</tr>
<tr>
<td>10.</td>
<td>Electrophoretic patterns of intact plasmid DNAs of pN2 subclones from passages p2, p5, p14 and p28 propagated in JC7623 cells</td>
<td>59</td>
</tr>
<tr>
<td>11.</td>
<td>Size analysis of linearized pN2 subclone DNAs</td>
<td>61</td>
</tr>
<tr>
<td>12.</td>
<td>Gel assay of insert sizes in p2 and p5 pN2 subclone DNAs</td>
<td>62</td>
</tr>
<tr>
<td>13.</td>
<td>5% PAG assay of p14 and p28 pN2 subclones digested with HaeIII</td>
<td>64</td>
</tr>
<tr>
<td>14.</td>
<td>Pictorial representation of p14 and p28 deletion subclones based on HaeIII restriction cleavage analysis</td>
<td>66</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>15. Analysis of 5% PAGs of passaged pN2 subclones digested with <em>Hae</em>III + <em>Hin</em>fI</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>16. Restriction maps of pN2 with <em>Alu</em>I, <em>Dde</em>I, <em>Hae</em>II, or <em>Msp</em>I</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>17. A 5% PAGE analysis of selected pl4JC3 and p28JC4 subclone DNAs digested with <em>Alu</em>I, <em>Dde</em>I, <em>Hae</em>II or <em>Msp</em>I</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>18. 1% agarose gel of four different p11 subclones: A1, B3, C2, D1</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>19. FIGE analysis of the distribution of ORI probe sequences</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>20. Collapse of a replication fork by sbcCD and the fork’s repair via the RecABCD-x-mediated recombination/repair</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>21. The deletion mechanism of mutant p5JC10</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>22. A possible multiplex stem-loop structure in the L2Hs element</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>23. Schematic of proposed deletion mechanism generating p14 and p28 subclones involving replication slippage and unusual non-B-form DNA structures</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
ABBREVIATIONS

$A_{600}$  absorbance at a wavelength of 600 nm
amp       ampicillin
amp$^3$   ampicillin resistance
amp$^s$   ampicillin sensitivity
βMCE      β-mercaptoethanol
bp        base pair(s)
cm        centimeter
dH$_2$O   deionized water
DNA       deoxyribonucleic acid
DNase     deoxyribonuclease
dsDNA     double-stranded DNA
DTT       dithiothreitol
DUE       DNA unwinding element
E. coli  Escherichia coli
EBr       ethidium bromide
EDTA      ethylenediaminetetraacetic acid
ETOH      ethanol
FIGE      field inversion gel electrophoresis
g         gravity
gm        gram(s)
kb        kilobase(s) or kilobase pair(s)
L         liter
LB        Luria-Bertani broth
SSPE  0.18 M NaCl, 10 mM NaH₂PO₄, 1 mM EDTA, pH 7.5

T:En.  10 mM Tris-HCl (pH 8.4), 1.0 mM EDTA

T:E:  1 mM Tris-HCl (pH 8.4), 0.1 mM EDTA

TAE  Tris-Acetate-EDTA

TB  Terrific broth

TBE  Tris-Boric acid-EDTA

TEMED  N,N,N',N'-tetramethylethylenediamine

Tris  Tris-(hydroxymethyl) aminoethane

UV  ultraviolet

U  unit

µg  microgram

µl  microliter

µM  micromolar

v  volume

V  volt(s)

w  weight
CHAPTER 1
INTRODUCTION

The concepts of DNA dynamics and genome instability are often used interchangeably. However, a distinction between these two notions should be discerned. DNA is an energetic molecule with a helix that can undergo conformational changes in its secondary structure (Sinden 1994) that may or may not lead to instability of the genome. DNA dynamics may be defined as an inherent potential for change within the DNA molecule. This applies especially to secondary structural changes from which alteration in the DNA sequence may or may not occur. Although both dynamics and instability broadly refer to DNA changes, DNA dynamics encompasses all aspects of change (transient, structural plus genetic) while genome instability manifests itself as a hereditary modification of DNA base sequences (i.e. the linear arrangement of bases). The latter may involve expansion, deletion or rearrangement of the genome (Richards and Sutherland 1992). Genome instability involves various structural processes including translocation, recombination, deletion and amplification (Cohen and others 1997). One example of instability, found in repeated sequences of humans and other eukaryotes, is copy number polymorphisms.
within or between individuals. A specific instance of nucleotide repeat polymorphisms is that associated with the human chromosomal fragile site FRA16B, which is an amplified AT-rich minisatellite repeat (Yu and others 1997). An example of DNA dynamics is cruciform formation in a superhelical plasmid. The plasmid’s secondary structure changes while the plasmid’s primary sequence remains stable.

The Relationship of Genome Instability to Human Disease

Genome instability may be responsible for both somatic and germline mutations implicated in the development of genetic diseases. For instance, many disorders have been correlated with genetic instability via repetitive DNA sequence polymorphisms. Several trinucleotide repeat disorders are examples of such polymorphisms (Dariow and Leach 1995, Timchenko and Caskey 1996, Gellibolian and others 1997). In the case of the fragile X syndrome, one of the most common causes of mental retardation, normally polymorphic exonic trinucleotide repeats expand beyond the normal size range and alter gene expression. Likewise, the disease mutation in myotonic dystrophy, an autosomal dominant disorder, results from expansion of a trinucleotide repeat array in the 3’ untranslated region of a gene mapped
to chromosome 19, resulting in changes in message stability (Singh 1995). Several neurodegenerative disorders (i.e. spinal and bulbar muscular atrophy, Huntington’s disease, dentatorubral-apallidoluysian atrophy and Machado-Joseph disease) are associated with expansion of CAG repeat arrays (Timchenko and Caskey 1996).

While instability can arise by expansion of certain repeat arrays, deletion events likewise are prominent in genome instability. One example is a form of β-thalassemia found in patients of Asian Indian origin. A deletion involving the 3' end of the β-globin gene reduces β-globin production, resulting in anemia and gross variation in shape and fragmentation of red blood cells. Such deletions account for nearly one third of the defective thalassemia genes in this ethnic group (Orkin 1987). Another example is Fanconi’s anemia, an autosomal recessive disorder in which small deletions are observed in repetitive DNA repeat arrays (Friedberg and others 1995). Several cancers also have been associated with instability of interspersed repeated and microsatellite sequences resulting from deletions: hereditary nonpolyposis colorectal cancer and sporadic colorectal cancers, endometrial and ovarian cancers (Risinger and others 1993, Orth and others 1994). Thus, genomic instability via deletion events also may represent
an important source of human genetic disease (Lovett and Feschenko 1996).

Repeeted DNA Sequences Involved With Genome Instability

The repeated DNA sequences involved in genome instability include tandem repeats, microsatellite DNAs, short interspersed repeat DNA (SINEs) and long interspersed repeat DNA (LINEs). Instability of repeated DNA sequences close to or even within genes may result in dire consequences for the organism. Because of their remoteness from sequences encoding genes, some mutations in these repetitive sequences may not lead to gross phenotypic changes. However, effects of instability may result in changes in genome organization and structure. Hence, instability in such sequences may have structural influences on genome function, including long-range (i.e. telesability) effects on neighboring DNA sequences.

The topological state of the DNA is one of the manifestations of DNA dynamics and a major contributor to genome instability. Particularly important to the topological state is torsional stress (i.e. supercoiling), which influences DNA replication, recombination and transcription, and may promote formation of unusual DNA structures (Kornberg and Baker 1992). In prokaryotes it is well documented that DNA topology is involved in regulating
gene expression (Cantor and others 1988). Torsional stress may influence certain DNA sequence motifs (i.e., palindromes, trinucleotide repeats, or certain AT-rich DNA sequences) to adopt unusual structures such as hairpins or cruciforms (Mizuuchi and others 1982). In turn, these non-B form DNA structures promote genome instability by increasing the frequency of mispairing reactions during DNA replication and recombination (Leach 1994, Darlow and Leach 1995). Torsional stress also provides an energy source which stably maintains non-B-form DNA conformations which are inherently unstable (Lilley and others 1988, Davison and Leach 1994, van Holde and Zlatanova 1994).

In contrast to prokaryotic systems, the role of torsional stress in DNA of eukaryotes is not well characterized. Nevertheless, several experiments have shown that supercoiled DNA in eukaryotic cells is more actively transcribed than linear DNA. However, torsional stress in eukaryotes is difficult to measure. Torsional stress may localize into domains and may have large global effects. The complexities of the eukaryotic system make interpretations regarding the roles of DNA topology and torsional stress in eukaryotes difficult to interpret (Cantor and others 1988). Thus, prokaryotic model systems are often chosen for studies concerning influences of
torsional stress on DNA structure and genome regulation.

**Interspersed Repeated DNA and Non-B-form DNA Structures: Effects on Genome Instability**

Several features of interspersed repeated DNA are influenced by the topological state of the DNA and may contribute to genome instability. Certain direct repeats, palindromes and inverted repeats may adopt unusual secondary structures such as cruciforms, hairpins or pseudo-hairpins under specific conditions including torsional stress (Lilley and others 1988, Ratnasinghe 1993, Musich 1996). Figure 1 illustrates actual direct repeats, interspersed direct repeats, palindromes and inverted repeats that are found in the human L2Hs repetitive element. Direct and interspersed direct repeats may contribute to genome instability via a mechanism of slip-mispairing in recombination and/or replication (Sinden 1994). Making up from 1% to 50% of a eukaryotic genome, direct repeats of 2 to 10 bp may be present $10^5$ to $10^7$ times in eukaryotes (Kornberg and Baker 1992). In comparison, inverted repetitive sequences, which can form "snapback" or "hairpin" sequences by intrastrand folding in single-stranded DNA (ssDNA), make up approximately 6% of the human genome (Kornberg and Baker 1992). The DNA palindrome, defined with respect to the
A Direct repeat (as found in the L2Hs element in plasmid pN2):

\[
\begin{align*}
&\text{bp 776-780} \quad \text{bp 781-785} \\
&5'-\text{TATAT TATAT-3'} \\
\end{align*}
\]

B Interspersed direct repeat (as found in the L2Hs element in plasmid pN6):

\[
\begin{align*}
&\text{bp 371} \quad \text{bp 382} \\
&5'-\text{AAATATATAT TTGATGTACT TTTATTTT TTATATTTT ATGTACAGCA TATAATATAT} \\
&\text{GCTTTGGGTA CTTTGATATT TTTTGTACAG TATGGAATAT ATACCTTGGG} \\
&TACTTTGATA TTTATGTGC ATGATATAAT ATATAGTTTG AGAACTTTGA \\
&\text{bp 539} \quad \text{bp 550} \\
&5'-\text{TATTTCCATGT ACAGTATAAA ATATATATT-3'}
\end{align*}
\]

C Palindrome (as found in the L2Hs element in plasmid pN2):

\[
\begin{align*}
&\text{bp 645} \quad \text{bp 656} \\
&5'-\text{AAATAT ATATTT-3'} \\
&\text{y-TTTATA TATAAA-y} \\
\end{align*}
\]

D Interspersed inverted repeat (as found in the L2Hs element of plasmid pN2):

\[
\begin{align*}
&\text{bp 353 - bp 358} \quad \text{bp 374 - bp 379} \\
&5'-\text{ATATTT. ATATTTTT. TTTTATA-y} \\
&3'-\text{TATAAA-5' TATAATATT-3'} \\
&\text{y-ATATTT. TTTATA-y} \\
\end{align*}
\]

Figure 1. Repeat motifs in the L2Hs element that may adopt unusual secondary structures and participate in replication slippage events. (A) Tandem direct repeats may be as small as two bases to thousands of bases. Here the repeat motif is 5 bases. (B) Interspersed direct repeats may be separated by a few bases to hundreds of bases. (C) Palindromes exhibit symmetry so that each DNA strand reads the same way in the 5' to 3' direction and, if single stranded, each strand would be self-complementary. (D) An interspersed inverted repeat is separated by many bases. If denatured, the motif at the 5' end could base pair with the complementary motif at the 3' end of the same strand. The bases in the middle could remain in a single-stranded denaturation bubble.
double-stranded form, is a pair of inverted repeats with two-fold rotational symmetry (Leach 1996) and may be a site of cruciform formation. A cruciform structure consists of two complementary hairpins that are completely base-paired in a B-form helix, except for the tip of each stem which consists of a loop containing at least three unpaired bases (Sinden 1994). The four-way junction at the base of the cruciform is similar to a Holliday junction in genetic recombination. Hairpin and cruciform formation are illustrated in Figure 2.

Two pathways for cruciform extrusion have been postulated: S-type formation is dependent on super-coiling, temperature and ionic conditions and C-type formation is dependent on AT-richness (Sinden 1994). C-type cruciform formation is of particular interest because of possible telestability effects. An AT-rich region flanking an inverted repeat opens and forms a denaturation bubble which is enlarged and finally encompasses the inverted repeat. Within the region of the inverted repeat, two hairpin structures may form which result in a cruciform structure. Thus, the process of C-type extrusion is an example of short-range cis-effects (i.e. telestability) of neighboring DNA on an adjacent site. Such cruciform structures and their adjacent, possibly single-stranded regions may be more
Figure 2. Hairpin and cruciform formation. (A) An inverted repeat in a single-stranded region of DNA is shown forming a hairpin. The inherent thermodynamic stability of double-stranded DNA relative to single-stranded DNA drives the formation of the hairpin. (B) This cartoon depicts two hairpins forming a cruciform structure from the existing inverted repeat in a double-stranded DNA region. Source: Sinden, R. R. (1994). DNA Structure and Function. (San Diego: Academic Press), 136.
susceptible to attack from nucleases, metabolic reaction by-products (i.e. free radicals) and/or environmental agents (Sinden and others 1991). Moreover, in replication, stalling of the replication fork occurs when a cruciform structure is encountered, leading to loss of processivity of DNA polymerase and premature termination of replication (Mytelka and Chamberlin 1996). Alternatively, the unusual secondary structure might be skipped over by the DNA polymerase (Kang and others 1995), leading to deletion of bases that made up the cruciform. Figure 3 illustrates both S- and C-type cruciform extrusion. The growing denaturation bubble that influences flanking regions of the inverted repeats should be noted in the C-type extrusion model.

DNA Slippage and Unusual Sequence Motifs

Other hypothetical mechanisms for the role of unusual sequence motifs and potential non-B-form DNA structures in genome instability have been proposed. The DNA polymerase slippage model (Streisinger and others 1966) may account for many incidences of genome instability associated with non-B-form structures and their effects on replication fidelity. Certain DNA sequences may adopt secondary structures that
Figure 3. Two pathways of cruciform formation. (A) The S-type (salt-dependent type) mechanism of cruciform formation refers to the reaction that occurs at physiological ionic strengths. Before nucleation and cruciform extrusion occur, 10 bp at the center of symmetry must melt. The rate of cruciform formation depends on the base composition at the center of symmetry. (B) The C-type mechanism of cruciform formation occurs in solutions with little-to-no salt and is dependent on an AT-rich region of DNA that flanks inverted repeats. The AT-rich region breathes, forming a denaturation bubble that can expand and encompass the inverted repeat. This stable open region that includes the inverted repeats is stabilized by DNA negative supercoiling. Complementary hairpins form which may result in cruciform formation. Source: Sinden, R.R. (1994). DNA Structure and Function. (San Diego: Academic Press), 148.
promote misalignment of complementary DNA strands or promote intrastrand base pairing (Trinh and Sinden 1993). In the slippage model, the strands dissociate during stalled DNA replication, followed by misaligned reassociation of the strands with resultant unpaired repeats (Wierdl and others 1996). Slipped mispaired DNA can exist when two repeated sequences are adjacent (Sinden 1994) as depicted in Figure 4. The direct repeat found in one strand can pair with the second direct repeat found on the complementary strand. A looping out of the skipped direct repeat on both strands occurs with the final outcome being extrusion of these loops from the DNA. Alternatively, slipped mispairing can occur between interspersed direct repeats (Sinden 1994) as illustrated in Figure 5. If backward slippage in the template strand occurs, the next round of replication generates in a deletion of the intervening sequence and one copy of the direct repeat (Fig. 5B-E). A duplication event occurs if the progeny strand slips back after replication of both repeats and the second copy base pairs with the first copy in the template strand (Fig. 5B'-E'). If the intervening DNA between direct repeats are palindromic, the ensuing hairpin structures may increase the frequency of deletion by stabilizing the misalignment (Sinden 1994). A replication misalignment mechanism specific to deletions may
Figure 4. Two adjacent direct repeats and the occurrence of slipped mispaired DNA. (A) Adjacent direct repeats are labeled 1 and 2, and are indicated by the arrows. On the complementary strand, the direct repeats are labeled 1' and 2'. (B) A slipped mispaired structure is shown, the result of direct repeat 2 pairing with complementary direct repeat 1'. Two single-strand loops are created, one on each strand. (C) Direct repeat 1 pairs with complementary direct repeat 2', the result of slippage of the complementary strand (bottom 3'-5' strand). Two single-strand loops are created, one on each strand. Source: Adapted from Sinden, R.R. (1994). DNA Structure and Function. (San Diego: Academic Press), 261.
Figure 5. Slipped mispairing of an interspersed direct repeat DNA during DNA replication. (A) Double-stranded DNA is shown with two interspersed direct repeats. (B) Replication of the first direct repeat (DR1) is depicted. (C) Slippage in the template strand results in the alignment of DR1 with its complement DR2'. (D) Replication of the top strand continues. (E) Subsequent replication of the top strand in (D) results in a deletion of DR2. Replication of the bottom strand would result in DNA like that in (A). (B') Replication proceeds through the second direct repeat (DR2). (C') Slippage in the nascent strand occurs resulting in DR2 aligning with DR1'. (D') Replication of the top strand continues resulting in an amplification of one direct repeat (DR3). (E') Subsequent replication of the top strand in (D') results in three interspersed direct repeats. Replication of the bottom strand would result in two direct repeats DR1 and DR2. Source: Sinden, R.R. (1994). DNA Structure and Function. (San Diego: Academic Press), 263.
require inverted repeats that can form hairpin structures. These hairpins may juxtapose flanking direct repeats so that slipped misalignment could occur (Lovett and Feschenko 1996). Furthermore, evidence from mutagenesis suggests that slipped mispairing may occur during DNA replication (Sinden 1994). A more current model of slipped mispairing by Gellibolian and others (1997) takes into account the influence of DNA supercoiling during replication and the possible outcomes when stalling occurs. This model (Fig. 6) shows how stalling of the replication fork influences genome stability and reveals how trinucleotide repeat array expansion or reduction might occur. It also is an important model concerning the potential of pausing during lagging strand synthesis.

**DNA Repair Enzymes and Genome Instability**

Another source of genomic instability that can be coupled to slippage events and non-B-form DNA structures involves defective DNA repair. Examples include disease states as observed in Fanconi's anemia, xeroderma pigmentosum and ataxia telangiectasia. In these examples the genomic instability observed precedes cancer (Cohen and others 1997). A particular type of repair called "arm-directed secondary-structure repair" leads to repair of a damaged replication fork (Leach 1994). In *E. coli*, this
Figure 6. A triplet repeat expansion model. The top illustration indicates a DNA triplet repeat array flanked by non-triplet repeat regions. This triplet repeat could involve other direct repeat sequences greater than three bases. (A) In step A, the replication fork is shown with accompanying replication machinery (primase, helicase, polymerase). Negative supercoiling (neg σ) occurs behind the replication machinery while positive supercoiling (pos σ) occurs downstream of the replication fork. (B) In step B, the accumulation of negative supercoils behind the replication machinery leads to hyperwrithing of the helix in front of the complex. Topoisomerases may not remove all the positive supercoils generated. (C) Thus, stalling of the replication machinery can occur as shown in C. This allows time for one of the lagging strands to form a hairpin structure in the repeat array. Also the replication machinery decreases its processivity along with the rate of reiterative DNA synthesis. (D) Another outcome, as shown in step D, is that the replication machinery can come off the DNA, reform a complex, and cause expansion of the repeat array through replication slippage. Source: Gellibolian, R., Bacolia, A. and Wells, R.D. (1997). Triplet repeat instability and DNA topology: An expansion model based on statistical mechanics. J. Biol. Chem. 272, 16793-16797.
process results from an unusual secondary structure (generated from a palindromic motif) that forms on one arm of a replication fork, stalling replication. The multimeric protein SbcCD recognizes this structure and endonucleolytically cleaves the DNA. Part or all of the hairpin arm is degraded by SbcCD (now acting as an exonuclease) or by RecBCD. The DNA (chromosome) is repaired to its unreplicated state and reinitiation of replication can occur (Leach 1994). If the SbcCD protein is defective, cleavage and degradation of the stem-loop structure would not occur. The replication fork would not have a second chance to initiate replication. Rather, the palindrome-induced secondary structures would stall DNA polymerase resulting in truncation of replication, or the secondary structure could be skipped over by the polymerase and not be replicated. The latter would result in deletion of the palindromic sequence or inviability of the progeny.

**Palindromes and Genomic Instability**

Chromosome instability has been observed for palindromes as short as 22 bp and increases with the length of a palindrome. It remains unclear whether palindromes cause instability problems in eukaryotic genomes. However, palindrome-mediated instability has been observed in *E. coli*, *B. subtilis*, *Streptococcus*, *Streptomyces* and
Saccharomyces cerevisiae (Leach 1996). Although wild-type E. coli do not retain large palindromes (>150 bases), some mutant strains are able to propagate palindromic sequences. Palindrome-mediated instability is a function of the relative position of the inverted repeats. Using a plasmid model system, identical palindromes have been inserted into plasmids at different nucleotide positions. A difference in relative location of only one base has been shown to have a thousand-fold difference in deletion frequencies (Leach 1996). Such positional effects must be taken into account when cloning DNA with palindromic motifs like the human L2Hs element into a plasmid vector. Orientation differences may play a significant role in genome stability.

Palindromic and repeated sequences have been suggested to jointly participate in the formation of deletions (Glickman and Ripley 1984). The formation of hairpins or cruciforms, capable of juxtaposing otherwise distant bases, have been observed to acquire double-strand breaks (Glickman and Ripley 1984, Salganik and Dianov 1992). Instability via deletions may arise from structures that involve DNA repeats. One deletion model postulated by Salganik and Dianov (1992) requires two repeat sequences that flank an intervening sequence. This model (Dianov and others 1991, Salganik and Dianov 1992) has shown that deletions develop
when the intervening sequence acquires a double-strand break, followed by nucleolytic cleavage of both strands of DNA to repair the strand break. In the process one of the repeats is lost along with the intervening sequence. Repeat sequences as small as six bp exhibit this behavior.

Another deletion model describes direct repeat motifs flanking intervening sequences that contain indirect repeat (or palindromic) motifs. Mukaihara and Enomoto (1997) suggest three factors that lead to deletion formation: the length of the direct repeats, the distance between them, and the non-B-form DNA structures between them.

**Inviability Versus Instability**

Long palindromic sequences or inverted repeats, when cloned into a plasmid vector in host wild-type *E. coli*, are either inviable or unstable when propagated (Allers and Leach 1995, Connelly and Leach 1996, Leach 1996). Inviability means that the cloned sequence cannot be propagated while instability refers to possible DNA sequence changes, whether by deletion, expansion or rearrangement, that can be passed down to subsequent generations. Usually, a palindromic DNA sequence must be 150-200 base pairs (bp) in length before inviability is conferred (Leach 1996). In this instance, inviability refers to the loss of the vector plus its palindromic or inverted repeat DNA and disallows
any further propagation of this construct. Instability, on the other hand, refers to changes in the palindromic DNA construct which allows replication. Instability may be detected within palindromes as short as 22 bp (Leach 1996).

**DNA Unwinding Elements and Genome Instability**

Another DNA motif that can influence genome stability is the DNA unwinding element (DUE). DUEs have been found in both prokaryotes and eukaryotes. DUEs are AT-rich regions of DNA associated with origins of replication and chromosomal DNA-nuclear membrane/matrix attachment sites. These sequences range from 30 to more than 100 bp in length. In supercoiled DNA, these AT-rich regions are the first to unwind. DUEs also are required for initiation of DNA replication at certain origins (Umek and Kowalski 1987, Umek and Kowalski 1988, Kohwi-Shigematsu and Kohwi 1990, Bode and others 1992). Moreover, putative DUEs in a plasmid could result in stably unwound regions that might be involved in recombination, rearrangements or promote the formation of non-B-form DNA structures.

**Cell Strains and Genotypes That May Affect Genome Stability**

Because the experimental methods in this dissertation involve cloning an element rich in inverted repeats and palindromic sequences into *E. coli* hosts, it is worthwhile
to mention aspects of the host which could influence the system dynamics. Although little is known about the character and stability of long inverted repeats (150 bp or larger) in eukaryotes, such repeats present cloning and maintenance problems in plasmids of *E. coli* (Lilley 1981, Mizuuchi and others 1982, Leach and Stahl 1983, Sinden 1994). To prevent homologous recombination, cloning often is done in recombination-deficient host cells.

In *E. coli*, the stability of clones with long perfect palindromes and inverted repeats is increased in hosts containing mutants of the *sbcC* and *sbcD* genes (Leach 1996). The *sbcC* gene is co-transcribed with the *sbcD* gene. Together these genes encode a multimeric protein that has an ATP-dependent double-stranded DNA (dsDNA) exonuclease and single-stranded DNA (ssDNA) endonuclease activities. Mutations in these genes confer a recombination co-suppressor phenotype such that the viability and stability of palindromes can be maintained (Connelly and Leach 1996). SbcCD appears to cleave secondary structures such as hairpin loops formed during DNA replication (Connelly and Leach 1996) that leads to the "arm-directed secondary-structure repair" previously discussed. In the JC7623 cell strain, the genotype (mutant *recBrecCreC* *sbcBsbccsbcD*) may contribute to plasmid instability via defective arm-directed
secondary-structure repair. Also, these mutant strains form linear and circular multimers which contribute to plasmid instability (Biek and Cohen 1986, Cohen and Clark 1986, Leach 1996). Other strains (C600, CES201, CES201 and DH5αF'IQ cells) do not have a mutant $sbcC$sbc$D$ genotype and their plasmids are more stably propagated as will be shown.

Overview of the L2Hs Element

The laboratory of Dr. P.R. Musich has been studying moderately repetitive, long interspersed repeated DNA sequences called the L2Hs (Line 2 Homo sapiens) family. The discovery of the L2Hs family came about by studying the nuclear organization of human LINE 1 elements. Further studies have shown that the L2Hs family exhibits polymorphisms within and between individual genomes. For instance, DNA fingerprinting of sperm and blood cells from the same individual exhibited distinct L2Hs fingerprint differences (Musich and Dykes 1986, Musich 1996). This suggests that during various developmental stages genomic differentiation may occur.

Two recombinant plasmids, pN6.4.39 (pN6) and pN6.4.39-2 (pN2) were constructed by inserting the same 596-bp L2Hs element into the plasmid vector pTZ19U, but in opposite orientations (Fig. 7, page 54). Sequence analysis revealed this cloned L2Hs element to be AT-rich (76%) relative to the
bulk of human DNA (56%) and to contain an unusually high density of palindromic, inverted and directly repeated DNA sequences. This AT-rich element has the potential to become single-stranded and to form cruciform structures with a C-type extrusion activity (Ratnasinghe 1993, Musich 1996). These structural motifs and activities suggest that the L2Hs family may be associated with replication or recombination events in the human genome.

**Hypotheses Concerning the L2Hs Element**

These key hypotheses concerning individual L2Hs elements have developed: a single L2Hs is intrinsically dynamic, has the potential to influence adjacent sequences and, thus, contributes to genome instability. By considering the dynamic nature of the L2Hs family implied by its polymorphic features, the stability of a single L2Hs element is investigated in a simple prokaryotic model system.

To test the hypothesis that the L2Hs element contributes to genomic instability, the dynamics of the L2Hs element were monitored over between several hundred cell generations in a simple bacterial molecule. Plasmids were assayed for evidence of instability (i.e. deletion or rearrangements of the chromosome construct) and whether other factors besides the L2Hs element were involved. In
the bacterial model system, orientation of the L2Hs element in the vector may have also contributed to the dynamics observed. The approaches taken to resolve these issues have served as the foundation of this work.

Summary

It appears that the stability of a genome depends on several features. First, the potential to assume secondary non-B-form DNA structures may be a determinant of instability. Intrinsic to the DNA, though, are certain sequence motifs which may form these unusual DNA structures under specific conditions of supercoiling or strand separation. Finally, extrinsic factors such as host cell strain may influence stability, especially when certain sequence motifs are present in the DNA. These three factors together may promote instability of genomic elements. In addition, their impact is witnessed throughout this investigation.
CHAPTER 2

MATERIALS AND METHODS

Materials

Ampicillin (amp), β-mercaptoethanol (βMCE), ethidium bromide (EBr), lysozyme, N, N, N',N'-tetramethylethelenediamine (TEMED), proteinase K, ribonuclease A (RNase A), and Tris-(hydroxymethyl) aminomethane (Tris) were purchased from Sigma Chemical Company (St. Louis, MI). Bactotryptone and yeast extract were purchased from Sigma Chemical Company and Difco Laboratories (Detroit, MI). Formamide was purchased from Sigma Chemical Company and Fisher Scientific (Pittsburgh, PA).

Acrylamide, agarose, ammonium acetate, ammonium sulfate, boric acid, bromophenol blue, ethylenediaminetetraacetic acid (EDTA), polyethylene glycol 8000 (PEG), potassium phosphate monobasic, potassium phosphate dibasic, sodium chloride, sodium hydroxide, and urea were purchased from Fisher Scientific. InCert® agarose was FMC Bioproducts (Rockland, ME).

Deoxyadenosine 5′-triphosphate [α-33P] (specific activity >3000 Ci/mmol) and deoxyadenosine [α-35S] 5′-triphosphate (specific activity = 1000 Ci/mmol) were purchased from New England Nuclear (Boston, MA). Biotin-14-dCTP and biotin-14-dATP were obtained from GibcoBRL.

25
Bovine serum albumin, the Magic™ Miniprep and Maxiprep kits, the fmol™ kit for thermocycle sequencing and restriction enzymes with 10x buffers were purchased from Promega Corporation (Madison, WI), New England Biolabs (NEB; Beverly, MA), or Boehringer Mannheim (Indianapolis, IN).

Filter paper for blotting was supplied by Whatman. Immobilon-S membrane was obtained from Millipore Corporation (Bedford, MA). Hydrolink Long Ranger Gel Solution was obtained from AT Biochem (Malvern, PA) and the FMC Bioproducts (Rockland, ME). SeaKem™ agarose was obtained from the FMC Bioproducts. Chill-out 14™ liquid wax was obtained from MJ Research, Inc. (Watertown, MA).

Primers for the polymerase chain reaction (PCR), hybridization probes and DNA sequencing were obtained from Millipore Corporation, NEB or Integrated DNA Technologies, Inc. (Coralville, IA). Table 1 lists the primers used. DNA sequencing products were either radioactively labeled using direct incorporation via an extension/termination reaction with [α-33P]dATP or [α-35S]dATP for sequencing reactions or non-radioactively labeled using a 5' biotin-labeled primer. Hybridization probes were nonradioactively labeled by using biotinylated primers, deoxynucleotides, and/or random primers in the synthetic reactions.

The composition of solutions and buffers used are
<table>
<thead>
<tr>
<th>Name</th>
<th>Length</th>
<th>Sequence</th>
<th>5' primer position in pN2(pTZ)</th>
<th>5' → 3' orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>puc/M13 Universal Reverse¹</td>
<td>16-mer</td>
<td>5'-AAC AGO TAT GAC CAT G-3'</td>
<td>210 (210)</td>
<td>→</td>
</tr>
<tr>
<td>puc/M13 Universal Reverse²</td>
<td>22-mer</td>
<td>5'-TCA CAC AGG AAA CAG CTA TGA C-3'</td>
<td>200 (200)</td>
<td>→</td>
</tr>
<tr>
<td>puc/M13 Universal Forward¹ ³</td>
<td>24-mer</td>
<td>5'-CGC CAG GGT TTT CCC AGT CAC GAC-3'</td>
<td>951 (383)</td>
<td>↑</td>
</tr>
<tr>
<td>pN1343R²</td>
<td>20-mer</td>
<td>5' AAC GTC AAA GGG CGA AAA AC 3'</td>
<td>1343 (753)</td>
<td>←</td>
</tr>
<tr>
<td>pNAMP1678²</td>
<td>20-mer</td>
<td>5' GCG ACA CGG AAA TGT TGA AT 3'</td>
<td>1678 (1088)</td>
<td>←</td>
</tr>
<tr>
<td>OriReverse (O)²</td>
<td>23-mer</td>
<td>5' ACG ACC TAC ACC GAA CTG AGA TA 3'</td>
<td>3039 (2443)</td>
<td>→</td>
</tr>
<tr>
<td>OriForward (O²)²</td>
<td>23-mer</td>
<td>5' TAT CTC AGT TCG GTG TAG GTC GT 3'</td>
<td>3087 (2471)</td>
<td>→</td>
</tr>
<tr>
<td>β-Lac Upper²</td>
<td>20-mer</td>
<td>5' TTT CCG TGT CGC CCT TAT TC 3'</td>
<td>1672 (1077)</td>
<td>→</td>
</tr>
<tr>
<td>β-Lac Lower²</td>
<td>21-mer</td>
<td>5' GCT CAG TGG AAC GAA AAC TCA 3'</td>
<td>2635 (2040)</td>
<td>←</td>
</tr>
<tr>
<td>ORI Upper²</td>
<td>20-mer</td>
<td>5' AGC GTC AGA CCC CGT AGA AA 3</td>
<td>2653 (2058)</td>
<td>→</td>
</tr>
<tr>
<td>ORI Lower²</td>
<td>23-mer</td>
<td>5' TAT CTC AGT TCG GTG TAG GTC GT 3'</td>
<td>2449 (3044)</td>
<td>←</td>
</tr>
</tbody>
</table>

¹New England Biolabs
²Integrated DNA Technologies
³Millipore
Methods

Bacterial Strains and Growth Conditions

The bacterial strains employed are listed in Table 2. Plasmid-carrying derivatives were produced by transformation using either electroporation or the heat-pulse method described by Nishimura and others (1990). Cultures were grown in Luria-Bertani broth (LB) (Sambrook and others 1989) or Terrific Broth (TB) (Tartof and Hobbs 1987) with amp (150 µg/ml) at 37°C to an absorbance at 600 nm (A₆ₐ₀) of approximately 4.0. Five or 20 ml cultures were grown in 14 ml glass culture tubes (1.5 cm diameter) or in 50 ml plastic centrifuge tubes (3.0 cm diameter), respectively, by inoculating amp medium with 50 µl or 200 µl, respectively, of overnight culture. Alternately, single colonies on agar plates were selected and placed into individual wells of 24-well microtiter plates (1.0 - 1.5 ml amp medium per well) or sterile 1.5 ml microcentrifuge tubes containing 0.1 - 1.0 ml amp medium.

Plasmids

Plasmids used in this study were derived from the vector pTZ19U (US Biochemicals; Cleveland, OH). Recombinants containing the L2Hs element are denoted pN6.
<table>
<thead>
<tr>
<th>STRAIN NAME</th>
<th>GENOTYPE</th>
<th>Phenotype and Effects</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C600</td>
<td>F' thr-1 leuB6 thi-1 <strong>supE44</strong> lacY1 tonA21 λ'</td>
<td>recA', recBCD' recombination proficient</td>
<td>Appleyard 1954 Bachmann 1987</td>
</tr>
<tr>
<td>CES200</td>
<td>argE3 ara-14 Δ (gpt-proA)62 galK2 hisG4 hsdR kdgK51 lacY1 λ' leuB6 mtl-1 rac' recB21 recC22 rfbD1 rpsL31 sbcB15 thi-1 tsx-33 xyl-5 thr-34::Tn10</td>
<td>recombination proficient recF pathway</td>
<td>Wyman et al. 1985 Wertman et al. 1986</td>
</tr>
<tr>
<td>CES201</td>
<td>argE3 Δ(gpt-proA)62 galK2 hisG4 hsdR kdgK51 lacY1 λ' mtl-1 rac' recB21 recC22 rfbD1 rpsL31 sbcB15 thi-1 tsx-33 xyl-5 Δ (rec-srl)306 srlR::TN10-84</td>
<td>recombination proficient recF pathway</td>
<td>Wyman et al. 1985</td>
</tr>
<tr>
<td>DH5αF'IQ</td>
<td>byfΔ (lacZYA-argR)U169F' deoR dncZ1 endA1 gyr gyrA96 hsdR17 lacAΔM15 lacI Q λ Φ80dlacZΔM15 proAB (rmk) recA1 relA1 <strong>supE44</strong> thi-1 Tn5(KmR)</td>
<td>recombination deficient recBCD pathway direct repeats are stably propagated</td>
<td>Hanahan 1983 Jessee and Blodgett 1988</td>
</tr>
<tr>
<td>JC7623</td>
<td>argE3 ara-14 Δ (gpt-proA)62 galK2 hisG4 kdgK51 lacY1 λ' leuB6 mtl-1 rac' recB21 recC22 rfbD1 rpsL31 sbcB15 sbcC201 <strong>supE44</strong> thi-1 thr-1 tsx-33 xyl-5</td>
<td>recombination proficient recF pathway &quot;arm-directed secondary-structure repair&quot; inhibited with sbcCD mutation</td>
<td>Horii and Clark 1973 Lloyd and Buckman 1985</td>
</tr>
</tbody>
</table>

* The underlined genotypes are involved in recombination
(pN6.4.39) or, with the L2Hs element inserted in reverse orientation, as pN2 (pN6.4.39.2). Recombinant plasmid pRh contains a 680 bp tetrameric rhesus alphoid monkey DNA element inserted into the BamHI site of pTZ19U (Pike and others 1986).

**Serial Passage of Cells**

An initial culture was prepared by inoculating 7 ml of amp medium with a single colony (0.5 -1.0 mm diameter) grown from cells transformed with pTZ19U or the recombinant plasmids mentioned above. Denoted as passage zero (p0), this culture was grown for 12 hours in TB plus amp or for 24 hours in LB plus amp at 37°C in an orbital shaker set at 200 rpm. After the appropriate incubation time, an aliquot of the p0 culture (50 μl) was transferred to 7 ml of fresh medium plus amp and grown as above to generate p1. After the set incubation time, an aliquot of 50 μl of p1 culture was used to inoculate fresh medium plus amp for the next passage. This serial propagation continued for a total of 28 passages. A second serial propagation experiment was done for 14 passages with 12 hour growth intervals using TB plus amp.

After each growth incubation period, 0.2-0.5 ml of cell culture was reserved in sterile tubes and re-fed with 0.5 to 1.0 ml of freezing medium (TB with 70% glycerol) with amp;
50 μl of culture was used to inoculate fresh medium for the next passage, and the remainder was used for plasmid DNA isolation.

**Methods for Producing Competent E. coli cells and Their Transformation**

**Electroporation Method.** To prepare competent cells for electroporation using the BTX Transporator (BTX Company, San Diego, CA), E. coli cells were grown to mid-log phase (A600 of ~0.5) in LB containing 50% of the standard salt (5 gm/L rather than 10 gm/L NaCl) in order to prevent electrical arcing during the electroporation. Cells were chilled on ice for 20 minutes (min) followed by centrifugation at 4000 x g, 4°C for 15 min. The pellet was resuspended in sterile deionized water (dH2O) to wash the cells, centrifuged and again resuspended in dH2O. After a third centrifugation, the pellet was resuspended in 10% glycerol and saved in 50 μl aliquots, either for immediate use or frozen at -80°C. Throughout the centrifugation and wash steps, cells were kept between 2-4°C. One μl of the transformant DNA was pipetted into a pre-cooled electroporation cuvette and 49 μl of competent cells added. After electroporation with a one-second pulse at 1000 volts, the cell sample was immediately transferred into one ml of LB. The resulting cell culture
was then incubated for one hour at 37°C with aeration before aliquots were plated on LB-agar plates containing 150 μg/ml amp.

Transformation Using Heat Pulse. A variation of a protocol by Nishimura and others (1990) was used for preparing competent cells. Fifty μl of an overnight *E. coli* stationary phase culture (*A₆₅₀ > 4.0*) was used to inoculate 24 ml of medium A (LB broth supplemented to a final concentration of 10 mM of MgSO₄ and 0.2% glucose). Cells were grown to mid-logarithmic phase (*A₆₅₀ ~0.5*) at 37°C with aeration in an orbital shaker set at 200 rpm. The culture was placed on ice for 10 min before centrifugation for 10 min at 1500 x g. The supernatant fraction was decanted and the cell pellet resuspended in 0.5 ml of medium A; 2.5 ml of storage solution B (LB broth supplemented to 12 mM MgSO₄, 36% glycerin and 12% PEG) was added. Fifty μl of the cell suspension was aliquoted into sterile microcentrifuge tubes for immediate use or for storage at -80°C. For transformation, 1.0 to 3.5 μl of DNA (5 to 20 ng/μl) were added to 50 μl of thawed competent cells. The mixture of cells and DNA was kept on ice for 30 min, subjected to a heat pulse at 42°C for exactly one min, then cooled on ice for five min. Warmed LB (0.5 ml) was added, and the transformed cells were incubated for one hour at 37°C with
aeration before aliquots were plated on LB-agar containing 150 μg/ml amp.

Two-layer Plating

In some instances freshly transformed cells were plated directly onto a "two-layer" LB-agar plate without prior incubation (Fleischmann and others 1995). The "two layer" plate consisted of a bottom LB-agar layer containing 5 ml of LB-agar with 150 μg/ml amp. A top LB-agar layer of 15 ml of LB-agar only was poured within ten min before cell plating. Transformed cells (100-200 μl) were spread onto the hardened top layer. The plates were incubated 8 to 10 hours at 37°C. Final amp concentration was 37.5 μg/ml per plate.

Preparation of Plasmid DNA

Hot Alkaline Lysis Method. A hot alkaline lysis procedure for plasmid DNA isolation described by Musich and Chu (1993) was used. Cells containing plasmid were grown to stationary phase and collected by centrifuging for 10 min at 1000-1500 x g. After decanting the supernatant fraction, the cell pellet was vortexed until creamy before dilution with 1.3 ml SET (20% sucrose, 50 mM Tris-HCl (pH 8.0) and 50 mM EDTA) buffer and transfer to a microcentrifuge tube. The cells were collected by a 20 sec centrifugation at 14,000 x g and the supernatant fraction discarded. The soft cell
pellet was vortexed until creamy before resuspension in 250 μl of SET buffer. Five hundred μl of freshly made 0.2 N NaOH-1% sodium dodecyl sulfate (SDS) was added and mixed by gentle tube inversion. Cell lysis was completed at 65°C for 30 min, inverting at 15 min. Next, 400 μl of 5 M potassium acetate [(3 M K⁺, 5 M acetate (pH 4.8-4.9)] was added, mixed by inversion and the samples placed on ice for 20 min to aggregate the SDS-protein-chromosomal DNA complex. Precipitate was collected at 14,000 x g for 15 min. The supernatant fraction was transferred to a clean tube and spun again for 10 min at 14,000 x g. The plasmid DNA was precipitated from the cleared supernatant fraction by adding 300 μl of 27% PEG in 3.3 M NaCl and chilling on ice for at least two hours. The DNA precipitate was collected by centrifugation at 14,000 x g for 15 min. Residual PEG and salt were removed by rinsing the precipitate with one ml of 70% ethanol. The DNA pellets were dried in vacuo and resuspended in T: cE:

**Magic™ Miniprep Method.** Plasmid DNA was isolated using the Magic™ Miniprep Kit from Promega. Plasmid-carrying E. coli cells were grown to stationary phase and collected by centrifugation for 10 min at 1000 x g. Cell pellets were resuspended in 200 μl cell resuspension solution (Magic™ Miniprep Kit) and transferred to a microcentrifuge tube.
Four hundred μl of cell lysis solution (Magic™ Miniprep Kit) was added, and the contents were mixed by inverting the tube until the cell suspension was clear. The mixture was neutralized with 400 μl neutralization solution (Magic™ Miniprep Kit), mixing the contents by inverting the tube several times. The cell lysate was spun at 14,000 x g in a microcentrifuge for five min and the cleared supernatant fraction decanted to a new tube. One ml of DNA purification resin was mixed into the supernatant fraction. For each miniprep, one Magic™ minicolumn was prepared with a 3 ml syringe barrel attached to a column connected to a vacuum manifold. The resin/DNA mix was transferred into the syringe barrels while a vacuum was applied to pull the resin/DNA mix into the minicolumn. The resin was washed under vacuum using column wash solution (Magic™ Miniprep Kit) containing ethanol. To elute DNA, the column was placed inside a 1.5 ml microcentrifuge tube (without an attached lid), 50 μl of preheated (65°C) TEₐ⁻ added to each column, and tubes spun in the microcentrifuge for 20 seconds (sec) after waiting one min. A second aliquot of preheated TEₐ⁻ was added to each column for a second elution. Eluents were pooled and the eluted DNA stored at 4°C.
**Magic™ Maxiprep Method.** Large amounts of plasmid DNA were isolated using the Magic™ Maxiprep Kit by Promega. Two-to-five hundred ml of culture were grown to stationary phase and cells collected by centrifugation for 10 min at 1000 x g. Cell pellets were resuspended in 15 ml cell resuspension solution (Magic™ Maxiprep Kit) and lysed using 15 ml of cell lysis solution (Magic™ Maxiprep Kit). Fifteen ml of neutralization solution (Magic™ Maxiprep Kit) were added, mixing by inversion and spun at 14,000 x g for 15 min. Ten ml of DNA purification resin suspension (Magic™ Maxiprep Kit) were mixed into the supernatant fraction. For each maxiprep, one Magic™ maxicolumn was prepared and inserted into a vacuum manifold. The resin/DNA mix was transferred to the column and a vacuum applied to pull the DNA solution through the resin. The resin was washed under vacuum using column wash solution containing ethanol. To elute DNA, the column was placed inside a 50 ml reservoir tube and 1.5 ml of preheated (65°C) T:E buffer was added to each column. After allowing the maxicolumn to soak for one min, tubes were spun in a clinical centrifuge for 5 min at 1000 x g. The eluted DNA was stored at 4°C.

**Quick Boil Miniprep (Obmp).** Two adaptations of the boiling methods described by Rajeevan and Bassett (1994) and Liu and Mishra (1995) were used for quick plasmid isolation.
In the Rajeevan and Bassett method (1994), 0.75 ml of overnight cell culture was transferred into microcentrifuge tubes and 15 μl 2% (v/v) Triton X-100 detergent (or 0.33% Triton X-100 final concentration) was added to each. The mixtures were vortexed and tubes placed on ice for one min, boiled for one min and then subjected to centrifugation at 14,000 x g for 10 min. The resulting viscous pellets were removed with a toothpick and 13 μl of each cleared lysate was analyzed by gel electrophoresis. In the Liu and Mishra method (1995), 0.75 ml of cell culture was placed in a microcentrifuge tube and spun for 15 sec. The supernatant fraction was decanted and the cell pellet vortexed until creamy. Lysozyme was dissolved into STT buffer [8% sucrose; 5mM Tris-HCl (pH 8); 5% Triton X-100] to a final concentration of 100 μg/ml]. Seventy-five μl of the Lysozyme-STT buffer was added to each cell sample and mixed by inversion. Samples were incubated on ice for at least ten min before placing the samples into boiling water for exactly one min. Samples were returned to the ice for five min and subjected to microcentrifugation for five min. Seven μl of each cleared cell lysate was assayed on an agarose gel.
Alcohol Precipitation of DNA. To concentrate DNA or to remove salts and/or unincorporated nucleotides from DNA samples, alcohol precipitation was used. Ammonium acetate was added to a final concentration of 2.0-2.5 M in the DNA sample. Next, ethanol or isopropanol was added. If ethanol was used, 2.5 volumes were added and the mixed solution placed in an ethanol-dry ice bath for at least ten min. A one-to-one volume ratio was used when isopropanol precipitation was done and the solution placed on ice for at least 20 min. DNA precipitates were collected at 14,000 x g for 15 min. The pellets of DNA obtained were rinsed with 70% ethanol and dried in vacuo.

Preparation of Intact Cellular DNAs for Analysis by Field Inversion Gel Electrophoresis

For total DNA, cells were formed into gel plugs for lysis, deproteinization and RNA removal. Twenty-eight colonies of DH5αF’IQ cells containing p11 pN6 subclones were selected and cultured overnight in cell culture plates (24 wells, one ml per well) in TB. In addition cells representing plasmid-negative (DH5αF’IQ or JC7623 host cells without plasmids) and plasmid-positive (p0 JC7623 cells and DH5αF’IQ cells containing pN6 plasmids) were analyzed also. Cells cultured overnight in TB were aliquoted into 1.5 ml
microcentrifuge tubes and spun for 15 sec at 14,000 x g.
The supernatant fraction was decanted and cell pellet washed
by resuspending in 200 µl wash solution (200 mM NaCl, 10 mM
Tris-HCl (pH 7.2), 100 mM EDTA). Cells were pelleted again
at 14,000 x g for 15 sec. The supernatant fraction was
discarded and 200 µl of fresh wash solution [200 mM NaCl,
10mM Tris-HCl (pH 7.2), 100 mM EDTA] was added. The cell
pellet was resuspended by vortexing. Cells were mixed with
an equal volume of 1.5% InCert® agarose at 40°C. The cell-
agarose mixture was pipetted into a 0.5 ml syringe for gel
formation. The polymerized "noodle" was ejected into a 5 ml
plastic tube and covered with bacterial lysis solution [10mM
Tris-HCl (pH 7.5), 50 mM NaCl, 100 mM EDTA, 0.2% Na
deoxycholate, 0.5% Sarcosyl (Na salt) and lysozyme (1
mg/ml)] for eight hours at 37°C. The lysis solution was
decanted and noodles were covered with a 5 mM proteinase K-
1% sarcosyl-0.5 M EDTA solution and incubated overnight at
50°C. The proteinase K solution was decanted and gel plugs
incubated three times for 30 min at 50°C in 5 ml T:E: to
change the buffer. Gel plugs were ready for FIGE.
Gel Electrophoresis of DNA

**Standard Agarose Gel Electrophoresis.** Agarose gels were prepared and electrophoresed in 50 mM TBE with EBr at 0.25 µg/ml. Final agarose concentrations ranged between 1% - 1.4% depending on the size of the DNA fragments to be separated. Minigels (7.5 cm x 5.0 cm) were electrophoresed at a constant voltage of 3-5 V/cm. Longer gels (10 cm x 10 cm) were electrophoresed at a constant voltage of 6-10 V/cm. Following electrophoresis, all gels were destained in dH₂O. DNA was visualized by UV transillumination and photographed on Polaroid positive/negative type 55 film or positive type 667 or type 52 film using a Polaroid MP-4 Land camera, f-stop 4.5, with a #23A filter.

**Polyacrylamide Gel Electrophoresis (PAGE).** To effectively separate small DNA fragments and for better resolution of short PCR products, nondenaturing polyacrylamide gel electrophoresis (PAGE) was done. For PCR products ranging from 0.5-2 kb, 4% polyacrylamide gel solutions were used. Resolution of restriction fragments between 80 and 500 bp required 5% or 6.5% polyacrylamide gels. The ratio of acrylamide to bisacrylamide was 19 to 1 and 50 mM TBE was the electrophoresis buffer. Immediately before pouring, ammonium persulfate and TEMED were added to 0.07% (w/v) and 0.03% (v/v), respectively. Gels were pre-
electrophoresed for 20 min before sample application. After sample loading, gels of 8 cm x 10 cm (1 x w) were run at 6.25 V/cm for 0.5-1.5 hours; gels of 16.5 cm x 19.0 cm were run at 12.12 V/cm for 2-3 hours. Polyacrylamide gels were post-stained in TBE containing EBr (0.05 µg/ml) for at least 30 min and destained in dH₂O. DNA was visualized by UV transillumination using a 312 nm Variable Intensity Transilluminator (Fisher Scientific).

**Sequencing Gel Electrophoresis.** An International Biotechnologies, Inc. model STS 45 apparatus with 64- or 96-well sharks-tooth combs was used to display sequencing products. The sequencing gels contained 5-6% polyacrylamide (acrylamide/bis: 19/1) and 42% urea in 50 mM TBE buffer or 5-6% Long Ranger™ gel solution (formerly Hydrolink® Long Ranger gel solution, AT Biochem) and 42% urea in 89 mM TBE buffer. The gel solutions were filtered and vacuum degassed. Immediately before pouring, ammonium persulfate and TEMED were added to 0.07% (w/v) and 0.03% (v/v), respectively. Long Ranger™ gels were 0.4 mm thick; 0.4-0.8 mm wedge spacers were used for acrylamide gels. Both types of gels were pre-electrophoresed for 20 min at 55 W. An aliquot of 3-6 µl of each sequencing reaction was loaded and electrophoresis continued for an appropriate time.

Two procedures were used to process the gels. If
nonradioactive chemiluminescent sequencing with biotin end-labeled primers was used, the gel was processed according to Musich and others (1995). The gel was lifted onto filter paper containing 89 mM TBE plus 20% methanol. The DNA was transferred using a Gene Sweep™ electroblotting apparatus (Hoefer Scientific Instruments, San Francisco, CA) to an Immobilon-S membrane. After drying the membrane, the DNA was cross-linked to the membrane with UV irradiation using a Fisher UV Crosslinker apparatus (FB-UVXL-1000) on the optimum setting (120,000 J/cm²). The membrane then was processed using a NEBlot Phototype kit (NEB) for chemiluminescent detection of DNA bands. Alternatively, when radioactive sequencing was used, following electrophoresis, the sequencing gel was transferred onto filter paper and vacuum dried at 80°C. DNA was detected by autoradiography.

Field Inversion Gel Electrophoresis. Field inversion gel electrophoresis (FIGE), a type of pulsed field gel electrophoresis, was performed to separate plasmid DNA from the bacterial chromosomal DNA. For FIGE analysis, a 1% agarose gel (14.0 x 20.2 cm², 200 ml SeaKem™ agarose gel) was prepared in 45 mM TBE. A 40-well or a 36-well comb (each 1.0 mm thick) was used to make a row of wells. Before pouring the gel, 1.0 mm thick slices were cut off the DNA
gel plugs and positioned on each tooth of the comb which lay horizontally. After loading all the gel plugs, the comb was placed vertically in the gel running tray. The molten agarose was added at the bottom end of the gel mold. After 60 min the solidified gel was transferred to the SuperSub apparatus (Amersham Pharmacia Biotech, Piscataway, New Jersey) and TBE buffer (45 mM) added to the buffer chamber. The system was cooled to 12°C. The electrodes were connected to a Hoefer Scientific PC750 Pulse Controller which was connected to a power supply. The ramp was set for zero. The forward pulse was set for 10.0 sec and the reverse pulse set for 3.3 sec. The gel was run at 12°C and 3V/cm for 12 hours. After completion, the gel was stained in 45 mM TBE containing EBr (0.25 µg/ml) for 30 min.

**DNA Cleavage by Restriction Enzyme Digestion**

DNA used for restriction enzyme cleavage was diluted into the appropriate enzyme buffer provided by the enzyme supplier. Nuclease-free BSA and 8MCE were added, as appropriate, to 100 µg/ml and 6 mM, respectively. Restriction enzymes were added to 1-4 U/µg of DNA. Samples were incubated at 37°C for 2-12 hours.

**DNA Dideoxynucleotide Thermocycle Sequencing**

DNA sequencing was performed by the dideoxynucleotide sequencing method (Sanger and others 1982) employing
Promega's fmol™ DNA sequencing kit. The sequencing primers used are described in Table 1.

DNA (500 fmol) was mixed with a sequencing primer in a total volume of 16 µl containing 1X fmol™ sequencing buffer. This reaction contained 3.0 pmol of primer, 5 µCi of [α-35S]dATP or [α-33P]dATP and 5 units of sequencing grade Taq-DNA polymerase. For DNA templates that were high in A and T, 4 µl of a solution of 20 µM each in dATP and dTTP was added to extend the length of the sequencing ladders for each reaction volume. For nonradioactive chemiluminescent sequencing, biotinylated primers were used in lieu of radioactive [α-35S]dATP or [α-33P]dATP. An aliquot of 4 µl of this reaction mix was placed into each of four tubes containing the appropriate deoxynucleotide-dideoxynucleotide mix (G,A,T or C). Each reaction was covered with 5 µl of Chill-out 14™ liquid wax or one drop of mineral oil. The tubes were placed in a thermal cycler (GL Applied Research, model GTC-2) preheated to 94°C. The reaction mixes were thermocycled in the following manner:

For the first cycle

94°C for 2 min, (initial denaturation)
55°C for 45 sec (annealing)
70°C for 2 min (extension)

After the first cycle, 60 more cycles were done as follows:

94°C for 45 sec (denaturation)
55°C for 45 sec (annealing)
70°C for 2 min (extension)

When the thermocycling program was complete, the reaction mixes were cooled to 4°C and 3 µl of fmol™ sequencing stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol) was added. The DNA samples were denatured at 80°C for 2 min and chilled on ice before loading on sequencing gels.

**Autoradiography**

Gels with 32P- or 35S-labeled DNA were dried onto Whatman 3 MM paper before exposure to Kodak (XAR) or Fuji (RX) films at room temperature from 12 hours to 2 months. All films were developed for up to 7 min in GBX developer, fixed for 10 min in GBX fixer, and rinsed in running water for 15 min before air drying.

**DNA Primary Sequence Analysis**

Gel autoradiograms containing DNA sequence patterns were scanned, compiled and analyzed using a Sun Microsystems SparcStation IPX computer, Howtek scanner and Bioimage sequencing software (Millipore Corp. version 2.1). Harr plots, sequence motif searches and maximum homology matching analyses were done using DNASIS (Hitachi America).
Polymerase Chain Reaction Labeling of β-lac and Ori Probes

Probes were synthesized from an oligonucleotide primer set that flanked the β-lactamase gene region of the pN6 plasmid or a primer set that flanked the ori region of pN6. Each 100 μl PCR reaction contained 0.05 μg of each biotinylated primer, 0.8 ng pN6 DNA, 0.2 mM dNTPs and biotinylated dATP, 1X Thermo buffer (Promega), 1.5 mM MgCl₂, and 5 u of Taq polymerase. After mixing, the PCR samples were placed in an MJR thermocycler for 30 cycles using the following specifications:

For the first step

94°C for 1.5 min (initial denaturation)

After the first step, 30 cycles were done as follows:

94°C for 30 sec (denaturation)
57°C for 30 sec (annealing)
72°C for 1.5 min (extension)

To ensure completion of the extension process an additional step was done as follows:

72°C for 5 min

After these cycles, then the reaction mixes were cooled to 4°C. One μl of each reaction mix was assayed on a 1.0 % agarose gel. One hundred μl of dH₂O was added to each of the remaining PCR reaction mixes.
Bidirectional Blotting and Hybridization of Field Inversion Gels

DNA from FIGE gels was transferred to two Immobilon-S membranes by bidirectional transfer (Smith and Summers 1980). Each FIGE gel was incubated in two volumes 0.25 M HCl for 30 min and briefly rinsed in dH2O. The gel was incubated in two volumes of 0.5 M NaOH, 1.5 M NaCl for 30 min each to denature the DNA. Finally, the gel was incubated in two volumes of 1 M ammonium acetate, 20 mM NaOH for one hour to neutralize the gel. The gel was placed on top of a moist Immobilon-S membrane located on top of moist filter paper and dry absorbent paper towels. A second membrane was placed on top of the gel, followed by filter paper and another stack of paper towels. A weight was placed on top to compress the whole stack. Transfer to both membranes was accomplished in less than two hours. After the transfer, the membranes were dried at 80°C and the DNA was cross-linked to the membrane at the optimum setting (120,000 J/cm²) in a UV Crosslinker apparatus (Fisher). Each membrane was incubated at 65°C for 30 min in 20 ml 3X SSPE-0.1% SDS twice. Each membrane was incubated at 65°C for one hour in prehybridization buffer (3X SSPE-3% SDS-5X Denhardt's reagent plus 100 µg/ml single-stranded herring sperm DNA (ssHsDNA). Hybridization solution was prepared by aliquoting 5 µl of biotinylated probe to 15 ml fresh
prehybridization buffer. The solution was denatured by boiling for 6-10 min. Membranes were incubated in the hybridization solution overnight, approximately 10-12 hours at 60°C. After hybridization and in order to prepare the hybridized membranes for chemiluminescent detection (via the Phototope™ Star Detection system), filters were rinsed in 3X SSPE-0.1% SDS for 2-3 min, washed three times in 20 ml 3X SSPE-0.1% SDS at 60°C for 20 min each and finally incubated in 3X SSPE-0.1% SDS at 20°C for 30 min.

Chemiluminescent Detection of Hybridized Membranes

The Phototope™ Star Detection process is summarized as follows. After the hybridization procedure, the membrane was incubated in blocking solution [5% SDS, 25 mM Phosphate (pH 7.2)], incubated in streptavidin (1 μg/ml) in blocking solution for five min, and washed twice in Wash Solution I (Appendix A). This was followed by a five min incubation in biotinylated alkaline phosphatase (0.5 μg/ml) and washing twice in Wash Solution II (Appendix A). After discarding and draining this solution, the Lumigen-PPD reagent (1X dilution of the supplied 100X stock solution) was added to the hybridization bag and incubated for five min. The majority of the Lumigen-PPD reagent was drained off the membrane which then was sealed in a plastic bag and exposed to X-ray film.
Subcloning

Two kinds of subcloning were done: DNA cloning and cell cloning. In DNA cloning, isolated plasmid DNA was used to transform DH5αF’IQ cells. The transformed cells were plated on LB-agar containing amp (150 μg/ml). In cell cloning, cells from an individual culture passages were plated. Colonies derived from these plates were grown up in medium with amp and the plasmid DNA isolated for further study.

Plating Experiment With Varying Amounts of Ampicillin

To determine if plasmid copy numbers were affected by the amount of amp used in growing colonies, plating was carried out using different amounts of amp. Four plates each were prepared with the following amounts of amp: 0, 1, 2, 8, 20 and 150 μg/ml. E. coli JC7623 host cells containing pN6 plasmid subclones of passage p11C2 were assayed on these plates. To determine the level of amp sensitivity of JC7623 cells without plasmid, these cells were plated on the following amounts of amp: 0, 0.25, 0.5 and 4 μg/ml.

Preparation of Colonies On Matrix Array With Varying Amounts of Ampicillin

A subcloned culture of JC7623 cells containing pN6 was selected for further study because some cell clones contained less plasmid DNA than other subclones, yet
maintained amp^3. An overnight culture was prepared by inoculating TB with frozen subclone cells from pN6 passage 11. After dilution, an aliquot of 100 μl overnight culture was spread on each of eight LB-agar plates containing amp at 0 (two plates), 0.25, 1.0, 2, 20 or 150 μg/ml. Two hundred eight-eight individual colonies which grew on the 0.25 μg/ml amp plate were transferred to LB-agar plates containing either 0.25 or 150 μg/ml amp. The plates were incubated at 37°C overnight. The 117 colonies that grew on both low and high amp plates were inoculated into 1.5 ml TB containing 0.25 μg/ml amp for overnight growth. A Qbmp was used to isolate plasmid DNA. Aliquots of these cultures were also reserved and used to inoculate additional TB for cultures used in making gel plugs for FIGE analysis.
CHAPTER 3

RESULTS

Construction of Plasmids

Recombinant plasmids pN6.4.39 (pN6) and pN6.4.39-2 (pN2) were constructed by ligation of identical copies of a 0.6 kb L2Hs segment into the KpnI site (bp position 305) of pTZ19U, but in opposite orientations. Also, a 0.68 kb segment of Rhesus monkey tandemly repetitive alphoid DNA (Pike and others 1986) was ligated into the BamHI site of pTZ19U plasmid for use as a control recombinant plasmid (pRh). These plasmids were transformed into E. coli host strains C600, CES200, CES201, DH5αF′IQ, and JC7623. Several features of the 2.86 kb pTZ19U vector should be noted. The replication control region (that which transcribes RNA I and RNA II) is located between bp positions 2101 to 2675. A ColEl replication ori site is located between bp positions 2683 to 2688 in pTZ19U. The β-lactamase gene (bp positions 1063 to 1923) confers ampR to host cells. Thus, cells grown in media containing amp survive only if they carry the β-lactamase gene, whether on the plasmid or via recombination of this gene into the bacterial chromosome. These features are illustrated in Figure 7.
Figure 7. Vector plasmid pTZ19U and recombinants. All constructs in this study are derived from plasmid vector pTZ19U. This vector has two essential regions which must be maintained in all plasmids: The coding regions for replication RNAs I and II are needed for plasmid DNA replication while the ß-lactamase (bla) gene confers ampicillin resistance on its host cell. Two recombinants, pN6 and pN2, have been constructed by inserting identical copies of a 596 bp human L2Hs element into the KpnI site of plasmid vector pTZ19U, but in opposite orientations. The pRh plasmid, used as a control, has a 680 bp rhesus monkey alphoid DNA inserted into the BamHI site of pTZ19U.

*Beta-lactamase gene
Experimental Overview

To test the hypothesis that the observed instability of the L2Hs element is intrinsic to the L2Hs element, this study investigated the stability of a single L2Hs element in a model prokaryotic chromosome. The recombinant L2Hs plasmids pN6 and pN2 and plasmid vector pTZ19U were serially propagated in various E. coli strains for 28 passages. To monitor L2Hs chromosome stability, the plasmid DNA was isolated and characterized after each passage by analyzing the plasmid DNA via agarose gel electrophoresis. In addition, isolated plasmids were cleaved with SspI restriction endonuclease into two linear fragments before electrophoretic analysis to test whether mobility differences between intact plasmids were topological or due to size differences in the DNA. To examine the stabilities of the recombinants more closely and to map deletion boundaries of mutants, individual plasmid molecules were isolated by DNA subcloning and characterized by restriction mapping, PCR analysis, and DNA sequencing. A more extensive population survey was performed to determine at what passage the deletion mutants were first detectable. Also, differences between pN6 and pN2 deletion patterns were compared to test for orientation affects of the L2Hs element. Finally, a reduced yield of plasmid DNA was observed from some cells in later passages while amp^R
remained high. To determine whether intermolecular recombination occurred between plasmids and the host chromosome, bacterial genomic DNA was probed with plasmid-containing sequences.

**Initial Serial Passages**

The initial serial propagation of CES200, CES201, DH5αF’IQ, and JC7623 *E. coli* cells carrying the pTZ19U vector or the L2Hs-recombinant pN2 plasmid was performed by passaging these cells for 28 days from a series of 24-hour cultures. Plasmid DNAs were isolated by the alkaline lysis miniprep procedure and assayed by electrophoresis through 1% agarose gels. The pTZ19U vector plasmid was stable in all cell strains and in all passages as illustrated in Figure 8B. Marker DNAs included the initial passage pN2 and pTZ19U plasmid DNAs (p0 pN2 and p0 pTZ19U, respectively). The multiple banding patterns in the markers represent form I supercoiled or form II nicked circular plasmid DNAs. An additional less mobile band may be a pN2 plasmid dimer. The complex banding patterns for late passage L2Hs-containing plasmids from JC7623 cells indicated either topological differences and/or possible deletions by pl2 (Fig. 8A). As the JC7623 cell strain passages progressed, increased heterogeneity within the plasmid population became more
Culture passage of recombinant pN2 plasmid in JC7623 cells

Culture passage of pTZ19U plasmid vector in JC7623 cells

This passage was chosen for subcloning

Figure 8. Gel analysis of pN2 and vector pTZ19U plasmid DNA passaged for 28 days in *E. coli* strain JC7623 cells. Plasmid DNA was prepared from 24-hr cultures of JC7623 cells carrying pTZ19U vector (Panel B) or the L2Hs recombinant, pN2 (Panel A). These DNAs were analyzed by electrophoresis through a 1% agarose gel in 50 mM TBE buffer with 0.5 μg/ml EBr. The first lane of each row contains pTZ19U as a marker; the last lane of each row contains pN2 DNA as a marker. The labels on the side of the gel photo indicate Form I supercoiled (I) or Form II nicked circular (II) plasmid DNA. The arrow points out the faster moving bands that suggest deletions in the pN2 plasmid.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
evident. Thus, it appeared that the plasmid DNA from mid-to-late passages was a mixture of intact plasmid and deletion mutants. The electrophoretic mobility of plasmids from cell strains CES200, CES201 and DH5αF′IQ was similar to the pN2 control.

Subcloning of pTZ19U

To verify that the pTZ19U vector plasmid itself was stable in all passages, DH5αF′IQ cells were transformed with unpassaged pTZ19U and plasmid DNA isolated from p0, p7 or p28. Seventy-eight individual colonies from the representative passages were selected from LB-agar plus amp plates for culturing in TB with amp. Plasmid DNA, isolated by a Qbmp procedure, was analyzed by agarose gel electrophoresis. There were 75 intact plasmid samples except for three dimers, one each from p0, p7, and p28. HaeIII cleavage patterns, which gave seven distinct bands, confirmed that all subclone pTZ19U plasmids were intact (data not shown).

Linearization of Plasmid DNAs

In order to determine whether plasmid mobility differences observed in Figure 8 were the consequence of changes in topology or in molecular size, plasmid DNAs were cleaved with SspI. No deletions were observed in linearized pTZ19U plasmids at any passage in any host cell strain used.
Also, no deletions were detected in linearized pN2 DNAs passaged in *E. coli* strains C600, CES200, CES201 and DH5αF′IQ, indicating that the complex gel patterns of these intact plasmids represented differences in topoisomeric forms (Fig. 9). However, pN2 DNA passaged in strain JC7623 cells showed deletion forms in middle-to-late passages when linearized (data not shown).

**Subcloning of Plasmid From JC7623 Cells**

Passaged pN2 plasmids from JC7623 cells were chosen for subcloning in order to sort out the plasmid mixtures. Early-passage plasmid DNAs from p2 and p5 were chosen because they had no observable change relative to control pN2 plasmid. DNA from p14 was chosen because some deleted species were evident (Fig. 8) while p28 plasmid was selected because it was from the final passage. These DNAs were transformed into the DH5αF′IQ host strain and eight individual colonies selected for plasmid DNA isolation by the hot alkaline lysis miniprep procedure. Banding patterns of intact p2 and p5 subclone DNAs appeared like the control DNA (p0 pN2) except for p5 colony 10 (p5JC10), which was more mobile than the pN2 control DNA (Fig. 10 A and B). The banding patterns of all p14 subclones except one (p14JC5) (Fig. 10C) and all p28 subclones (Fig. 10D) indicated DNAs with greater mobilities than the pN2 control.
Figure 9. Gel assay of passaged pTZ19U and pN2 plasmids cleaved with SspI that were isolated from host DH5αF’IQ cells. “M” is a marker of SspI-linearized pTZ19U plus pN2 DNAs with fragment sizes indicated on the sides of the gels. SspI-linearized pTZ19U DNAs (p0 through p28) are shown in Panel A. SspI-linearized pN2 DNAs (p0 through p28) are shown in Panel B. Slight differences in band mobility appear to be caused by loading differences or contamination by host genomic DNA (p6, p14 and p23 of pTZ19U; p3 and p23 of pN2).
Figure 10. Electrophoretic patterns of intact plasmid DNAs of pN2 subclones from passages p2, p5, p14, and p28 propagated in JC7623 cells. Gel A is the p2 pN2 subclones 2 through 8; Gel B is p5 pN2 subclones 1 through 4, 7 through 10; Gel C is p14 pN2 subclones 1 through 8; Gel D is the p28 subclones 1 through 8. The p2 pN2 subclones and p5 subclones are like the control pN2 except for p5 subclone 10 (p5JC10) which seems to have a more mobile band than the other p2's and p5's. Subclone p14JC5 appears like control pN2. The lowest bands of p14 (excluding p14JC5) and p28 subclones run with or are more mobile than the 2.0 kb band.
**Sizing by Restriction Cleavage (SspI cuts)**

Subcloned plasmid DNAs were linearized by SspI cleavage in order to size the DNA. All p2 (Fig. 11A) and all but one p5 (Fig. 11B) exhibited the normal 3.32 kb SspI band. The exception, p5JC10, was slightly shorter, indicating a small deletion. The p14JC5 plasmid subclone appeared similar to control p0 pN2 DNA (Fig. 11C). The other p14 and p28 subclones (Fig. 11D) exhibit fragments of 1.8 to 2.0 kb indicating deletions ranging from 1.3 to 1.6 kb. p28JC1 exhibits a larger deletion by its faster gel mobility, compared to the other p28 subclones.

**Sizing by Restriction Cleavage (Acc65I)**

In order to size the small deletion found in p5JC10 and to detect any deletions in the other subclones that appeared to be the same size as the pN2 control, restriction cleavage was performed on p2JC and p5JC DNAs, using Acc65I (an isoschizomer of KpnI). Except for p5JC10, all plasmids had the vector band of 2.8 kb and the insert band of 595 bp as in the p0 pN2 control (Fig. 12). In p5JC10, the size of the L2Hs fragment was approximately 520 bp, representing a deletion of approximately 80 bp.
Figure 11. Size analysis of linearized pN2 subclone DNAs. Subclone plasmid DNAs were linearized by SspI cleavage and analyzed on 1% agarose gels. SspI produces fragments 3.32 kb and 131 bp (too small to resolve on 1% agarose gels) from intact pN2. Gel A is p2 pN2 subclones 2 through 7. Gel B is the p5 pN2 subclones. Gel C is p14 pN2 subclones. Gel D is p28 pN2 subclones. The sizes of the marker DNA fragments are indicated alongside the gels.
Acc65I Restriction Cuts

M1 = pTZ/Acc65I
M2 = pN2/ Acc65I

Figure 12. Gel assay of insert sizes in p2 and p5 pN2 subclone DNAs. Acc65I, an isoschizomer of KpnI, cleaves pN2 DNA into a 2863 bp vector band and a 595 bp L2Hs insert band. Gel A is Acc65I-cleaved p2 pN2 subclones. Gel B is Acc65I-cleaved p5 pN2 subclones 1 through 4, 7, 8 and 10. All p2 and p5 subclones have the 2863 bp and 595 bp bands except for p5JC10 which has a more mobile insert band. Faint bands seen in most of the lanes are due to incomplete Acc65I digestion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Mapping Deletion Boundaries

HaeIII Restriction Cleavage

Subclones from p2, p5, p14, and p28 were cleaved with HaeIII restriction endonuclease and assayed on 5% polyacrylamide gels (PAG). Every p2 and p5 subclone had all the normal pN2 bands present except for p5JC10 which had a more mobile (smaller) insert-containing band (data not shown). Except for the normal p14JC5 subclone, all other p14 and p28 subclones were missing the insert-containing plus other bands and had new bands which represented fusion junctions of partially deleted fragments. Figure 13 and Table 3 summarize this HaeIII analysis. Numerous fragments were missing from the cleavage patterns of the p14 and p28 clones, indicating that vector sequences adjacent to the insert fragment were deleted. Figure 14 displays a pictorial representation of the deletions.

HaeIII + Hinfl Restriction Cleavage

To map and size deletions that might be within the insert DNA, p2 and p5 subclones were cleaved with HaeIII + Hinfl and assayed on 5% PAGs (Fig. 15). All subclones had a banding pattern similar to the pN2 control except for p5JC10. This exception was missing the 652 bp insert-containing band but had a doublet band at ~600 bp. The
Figure 13. 5% PAG assay of pl4 and p28 pN2 subclones digested with HaeIII. Gel A is pl4 pN2 subclones 1 through 8 cleaved with HaeIII; Gel B is p28 pN2 subclones 1 through 8. The insert-containing 874 bp band found in M2 and pl4JC5 is missing in all other subclones. Several bands are missing in the subclones that are present in HaeIII-digested intact pN2 DNA.
**TABLE 3. TALLY OF PRESENT, MISSING OR NEW BANDS FROM PAG ASSAYS OF PASSAGED PN2 PLASMIDS RESTRICTED WITH HAEIII**

<table>
<thead>
<tr>
<th>HaeIII restriction fragment number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>12</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (bp):</td>
<td>80</td>
<td>458</td>
<td>434</td>
<td>174</td>
<td>874</td>
<td>102</td>
<td>290</td>
<td>142</td>
<td>603</td>
<td>267</td>
</tr>
<tr>
<td>Estimated Deletion Size (kb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p2JC's</td>
<td>0.00</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p5JC's</td>
<td>0.00</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p5JC10&lt;0.10&lt;sup&gt;b&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC1</td>
<td>1.44</td>
<td>+&lt;sup&gt;d&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC2</td>
<td>1.29</td>
<td>+</td>
<td>+</td>
<td>(590)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(590)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC3</td>
<td>1.16</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>(290)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>(290)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC4</td>
<td>1.29</td>
<td>+</td>
<td>+</td>
<td>(590)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(590)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC5</td>
<td>0.00</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC6</td>
<td>1.29</td>
<td>+</td>
<td>+</td>
<td>(590)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(590)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC7</td>
<td>1.38</td>
<td>+</td>
<td>+</td>
<td>(500)&lt;sup&gt;n&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(500)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p14JC8</td>
<td>1.44</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC1</td>
<td>1.44</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC2</td>
<td>1.33</td>
<td>+</td>
<td>+</td>
<td>(120)&lt;sup&gt;i&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(120)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC3</td>
<td>1.33</td>
<td>+</td>
<td>+</td>
<td>(120)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(120)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC4</td>
<td>1.33</td>
<td>+</td>
<td>+</td>
<td>(120)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(120)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC5</td>
<td>1.33</td>
<td>+</td>
<td>+</td>
<td>(120)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(120)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC6</td>
<td>1.33</td>
<td>+</td>
<td>+</td>
<td>(120)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(120)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC7</td>
<td>1.33</td>
<td>+</td>
<td>+</td>
<td>(120)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(120)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>p28JC8</td>
<td>1.38</td>
<td>+</td>
<td>+</td>
<td>(500)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(500)</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Footnotes:

<sup>a</sup>represents the HaeIII restriction fragment number according to size (smallest # = largest fragment). The fragments left-to-right are arranged in clockwise alignment.
<sup>b</sup>The deletion size is less than 100 bp.
<sup>c</sup>The i indicates that the fragment is reduced in size.
<sup>d</sup>The + means that the fragment is present in the HaeIII digest pattern.
<sup>e</sup>The - means that the fragment is absent.
<sup>f</sup>The 590 bp fragment is formed from the fusion products of partially deleted fragments 9 and 3.
<sup>g</sup>The 290 bp fragment is formed from the fusion products of partially deleted fragments 12 and 3.
<sup>n</sup>The 500 bp fragment is formed from the fusion products of partially deleted fragments 9 and 3.
<sup>i</sup>The 120 bp fragment is formed from the fusion products of partially deleted fragments 12 and 3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 14. Pictorial representation of p14 and p28 deletion subclones based on HaeIII restriction cleavage analysis. Deletion boundaries are mapped using HaeIII cleavage data presented in Figure 13 and Table 3.
Figure 15. Analysis of 5% PAGs of passaged pN2 subclones digested with *HaeIII* + *HinfI*. Gel A is p2 subclones 1 through 8. The subclones run like the M3 marker and are intact pN2's. Gel B is p5 subclones 1 through 4, 7, 8, and 10. In p5JC10 an intense doublet band is shown at or below the 603 bp while a 652 bp band is missing. Gel C is p14 subclones 1, 3, and 8 digested with *HaeIII* + *HinfI*. The absent 652 bp band suggests that subclones have deletions of the L2Hs insert as well as flanking regions.
A

\[ \text{p2 pN2JC subclone #s} \]

\[ \begin{array}{cccccccc}
M1 & M2 & 2 & 3 & 3 & 4 & 5 & 6 & 8 & M3 \\
1075 & & & & & & & & & \\
630 & & & & & & & & & \\
603 & & & & & & & & & \\
517 & & & & & & & & & \\
472 & & & & & & & & & \\
413 & & & & & & & & & \\
396 & & & & & & & & & \\
330 & & & & & & & & & \\
\end{array} \]

B

\[ \text{p5 pN2JC subclone #s} \]

\[ \begin{array}{cccccccc}
M3 & 10 & 10 & 9 & 8 & 7 & 4 & 2 & 1 & M2 \\
630 & & & & & & & & & \\
603 & & & & & & & & & \\
413 & & & & & & & & & \\
330 & & & & & & & & & \\
\end{array} \]

C

\[ \text{p14 pN2JC subclone #s} \]

\[ \begin{array}{ccccccc}
M4 & 8 & 3 & 1 & M5 & M1 \\
874 & & & & & & \\
603 & & & & & & \\
517 & & & & & & \\
468 & & & & & & \\
434 & & & & & & \\
396 & & & & & & \\
290 & & & & & & \\
\end{array} \]

\text{HaeIII + Hinfl Restriction Digests}

M1=\text{pTZ19U/Hinfl}
M2=\text{pTZ19U/HaeIII + Hinfl}
M3=\text{pN2/HaeIII + Hinfl}
M4=\text{pN2/HaeIII}
M5=\text{pTZ19U/HaeIII}

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
doublet band represented a vector band plus another segment that contained a partially deleted L2Hs element. The deletion was at least 50 bp and no larger than 80 bp, based on these data and that of the Acc65I restriction cleavage data (Fig. 12). HaeIII + HinfI digestion of pl4 subclones (Fig. 15C) shows that the 652 bp band is missing and indicates that subclones have deletions of the L2Hs insert as well as flanking regions.

Refinement of Deletion Boundaries Through Additional Restriction Mapping

Four additional restriction cleavage studies were done to further refine the deletion boundaries of pl4 and p28 pN2 subclones. Restriction cut sites for the enzymes used for the digests (AluI, DdeI, HaeII, and MspI) are displayed in Figure 16. Each pl4JC3 and p28JC4 pN2 subclone was digested separately with each of the mentioned enzymes. Comparisons were made with pTZ19U and pN2 standards that also had undergone cleavage. These restricted DNAs were assayed on 5% PAGs (Fig. 17). A list of fragments present or absent for each pl4 and p28 subclone DNA are summarized in Table 4.

Subclone pl4JC5 had the same AluI and MspI fragments as intact pN2. Two AluI bands were missing from each of the other subclones, suggesting a large deletion with boundaries between bp positions 3396-to-3453 and 1215-to-1472. MspI
cleavage patterns of p14 and p28 subclones indicated deletions from bp position 161-to-1192, with flanking fragments ranging from bp position 3123-to-160 and 1193-to 1903 (Fig. 17 and Table 4). HaeII cleaves pN2 into only four fragments. In p14 and p28 subclones (except for p14JC5), bands corresponding to fragments between bp positions 6 and 1150 were missing as compared to control p0 pN2/HaeII.

Gel assays of p14 and p28 subclones cleaved with DdeI showed one large band missing that suggested flanking regions ranging from bp positions 2626-to-3055 and 995-to-1939. A deletion was thought to occur between bp positions 3055-to-995.

Screening Additional Subclones for Deletions

It was not known whether the deletion process arose out of an accumulation of small deletions or via a single event. Thus, further subcloning was done to try to find additional small (50 to 200 bp) and/or intermediate-sized deletion mutants (200 to 600 bp). This screening was accomplished by transforming DH5αF’IQ cells with the plasmid DNA prepared from each of the initial 28 passages. Each transformation culture was plated onto a two-layer LB-agar plate containing amp (Fleischmann and others 1995). The two-layer plate allowed freshly transformed cells to be plated immediately,
Restriction Maps of pN2 cut with Alul, Ddel, HaeII and MspI

Figure 16. Restriction maps of pN2 cut with AluI, DdeI, HaeII or MspI. The cleavage positions in pN2 plasmids are shown. Each of these enzymes has four or more cut sites in pN2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 17. A 5% PAGE analysis of selected p14JC3 and p28JC4 subclone DNAs digested with AluI, DdeI, HaeII or MspI. p0 pN2 and p0 pTZ19U were digested with the same enzymes and used as gel standards. Bands missing from p14JC3 and p28JC4 restriction digests represent specific fragments that have been deleted. Bands that differ from those seen in the pN2 control digests represent junctions of deletions.
TABLE 4. ALUI AND MSPI RESTRICTION FRAGMENTS OF P14JC3 AND P28JC4 SUBCLONES CORRESPONDING TO FIGURE 17

<table>
<thead>
<tr>
<th>AluI Restriction Fragment Number</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (bp): 136 226 118 64 95 22 27 635 96 215 257 621 63 100 521 257 226</td>
<td></td>
</tr>
<tr>
<td>+ + + NR NR NR NR - NR - - + NR NR + -</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mspl Restriction Fragment Number</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (bp): 26 147 489 138 895 711 404 110 67 34 404 190 320</td>
<td></td>
</tr>
<tr>
<td>+ + - + - + + + + + +</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Haell Restriction Fragment Number</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (bp): 370 1136 8 1939 ~1200-1500</td>
<td></td>
</tr>
<tr>
<td>+ - NR +</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ddel</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (bp): 2338 549 166 409 ~1600</td>
<td></td>
</tr>
<tr>
<td>- + NR +</td>
<td></td>
</tr>
</tbody>
</table>

- = Band is absent
+ = Band is present
NR = The band is too small to be resolved on a 5% acrylamide gel
New = New band is present that is different from the expected bands and its size is estimated relative to control bands
without the usual hour incubation and allowed all transformed cells to have the same starting growth advantage to form a colony. Twenty-four to 48 colonies derived from each plate were selected, cultured, and plasmid DNA isolated by a Qbomp method. Electrophoretic analysis showed that deletion mutants were found as early as p2 and that this early deletion included the insert and part of the vector (denoted as a "large" deletion). No small or intermediate-sized deletions were discovered out of 632 clones examined. Thus, it appeared either that the deletion process was not a series of small deletion events accumulating over time or that small deletions were selected against because large deletion mutants (smaller plasmids) had replicative advantage over small deletions (larger plasmids). With increasing passage number, the percentage of subclones containing large deletions increased; by p28, 94% of subclones had large deletions (Table 5).

Comparison of pN2, pN6, and pRh Plasmids

Another primate repetitive element (pRh) was ligated into plasmid pTZ19U and used as a control recombinant in the JC7623 cell strain. This element is a tandemly repetitive rhesus monkey alphoid DNA with higher GC content than the L2Hs element (55% vs 28%), is slightly larger (0.68 kb vs 0.60 kb) and represents four tandem repeats of 170 bp (Pike...
TABLE 5. RESULTS OF GEL ASSAYS OF PLASMID DNA ISOLATED FROM SERIAL-PASSAGED JC7623 CELLS CONTAINING PN2

This table is based on the 1% agarose gel assays of pN2 plasmids isolated from 24 to 48 single colonies from passages p1, p3 and all even passages p2 through p28. Subcloned plasmid DNA was isolated from single colonies grown for 12 hours at 37°C in microtiter plates containing one ml of TB plus amp (150 μg/ml) per colony by the Liu and Mishra QBmp method (1995). The DNA was electrophoresed on 1% agarose gels. Possible dimer plasmids and plasmids with large deletions (deletions of the insert plus flanking vector sequence) were observed in early passages. No small-to-intermediate sized deletions were discovered out of 632 Qbmp samples examined.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Inconclusive¹</td>
<td>2/48</td>
<td>1/48</td>
<td>1/24</td>
<td>1/48</td>
<td>0/48</td>
<td>0/48</td>
<td>0/48</td>
<td>0/48</td>
<td>0/48</td>
<td>12/34</td>
<td>4/34</td>
<td>3/34</td>
<td>6/34</td>
<td>14/34</td>
<td>24/34</td>
<td>10/34</td>
</tr>
<tr>
<td>% inconclusive</td>
<td>4%</td>
<td>2%</td>
<td>4%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>35%</td>
<td>12%</td>
<td>9%</td>
<td>24%</td>
<td>41%</td>
<td>71%</td>
<td>26%</td>
</tr>
<tr>
<td>Possible</td>
<td>1/48</td>
<td>3/48</td>
<td>1/24</td>
<td>1/48</td>
<td>4/48</td>
<td>4/48</td>
<td>1/48</td>
<td>5/48</td>
<td>5/48</td>
<td>1/34</td>
<td>1/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
</tr>
<tr>
<td>plasmid dimers²</td>
<td>2%</td>
<td>6%</td>
<td>4%</td>
<td>2%</td>
<td>65%</td>
<td>65.4%</td>
<td>2%</td>
<td>10.4%</td>
<td>29%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>% possible</td>
<td>1/48</td>
<td>1/48</td>
<td>1/24</td>
<td>1/48</td>
<td>1/48</td>
<td>1/48</td>
<td>1/48</td>
<td>5/48</td>
<td>5/48</td>
<td>1/34</td>
<td>1/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
</tr>
<tr>
<td>plasmid dimers</td>
<td>2%</td>
<td>6%</td>
<td>4%</td>
<td>2%</td>
<td>65%</td>
<td>65.4%</td>
<td>2%</td>
<td>10.4%</td>
<td>29%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>deletions³</td>
<td>0%</td>
<td>76%</td>
<td>8%</td>
<td>6%</td>
<td>1.25%</td>
<td>42%</td>
<td>48%</td>
<td>56%</td>
<td>85%</td>
<td>74%</td>
<td>65%</td>
<td>29%</td>
<td>29%</td>
<td>71%</td>
<td>94%</td>
<td></td>
</tr>
<tr>
<td>% large deletions</td>
<td>0/48</td>
<td>39/48</td>
<td>2/24</td>
<td>2/48</td>
<td>4.1%</td>
<td>12.5%</td>
<td>42%</td>
<td>48%</td>
<td>56%</td>
<td>85%</td>
<td>74%</td>
<td>65%</td>
<td>29%</td>
<td>29%</td>
<td>71%</td>
<td>94%</td>
</tr>
<tr>
<td>Runs with std.</td>
<td>45/48</td>
<td>6/48</td>
<td>20/24</td>
<td>44/48</td>
<td>46/48</td>
<td>41/48</td>
<td>27/48</td>
<td>20/48</td>
<td>0/34</td>
<td>1/34</td>
<td>0/34</td>
<td>4/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0%</td>
</tr>
<tr>
<td>(No deletion)</td>
<td>94%</td>
<td>12.5</td>
<td>83%</td>
<td>91%</td>
<td>65%</td>
<td>65.4%</td>
<td>56%</td>
<td>42%</td>
<td>0%</td>
<td>3%</td>
<td>0%</td>
<td>12%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>% runs with std.</td>
<td>0/48</td>
<td>0/48</td>
<td>0/24</td>
<td>0/48</td>
<td>0/48</td>
<td>1/48</td>
<td>0/48</td>
<td>0/48</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
</tr>
<tr>
<td>Exceptions⁵</td>
<td>0/48</td>
<td>0/48</td>
<td>0/24</td>
<td>0/48</td>
<td>0/48</td>
<td>1/48</td>
<td>0/48</td>
<td>0/48</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
<td>0/34</td>
</tr>
<tr>
<td>% exceptions</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

¹DNA yield was insufficient to make a conclusion
²DNA is larger than the marker and is similar in pattern to the dimer standard
³Major band is smaller than the pN2 marker, suggesting deletion(s) around 0.5 kb or more
⁴Major band runs with the standard pN2 major band
⁵DNA bands larger than pN2 standard but bands do not resemble dimer bands
and others 1986). To compare the stability of pN2, pN6, pTZ19U, and pRh plasmids in *E. coli*, JC7623 cells transformed with these plasmids were serially passaged in quadruplicate every 12 hours for 14 passages in TB. Plasmid DNA was isolated using a Qbmp procedure and linearized by *SacI* cleavage to size the plasmids and sort out mixtures of different topological forms. The pTZ19U and pRh plasmids were stable throughout all passages. Plasmid pN6 displayed a deletion species by p6 while a pN2 deletion was observed by p8.

**Screening JC7623 Cells For Reduced pN6 Plasmid Copy Number**

Some later-passaged JC7623 cells had reduced plasmid yields while amp<sup>3</sup> remained high. To test whether part of the β-lactamase (*amp<sup>3</sup>*) gene was incorporated into the bacterial chromosome, colonies with low plasmid yield were isolated. First, several low-yield pN6 cultures were subcloned on LB-agar plates containing 150 μg/ml amp and colonies were selected. The colonies were cultured for 12 hours in TB containing amp (150 μg/ml). The number of cells per ml of culture for each sample p11A1, p11B3, p11C2, and p11D1 in Figure 18 was determined by colony counts of 12 hour cultures plated on LB-agar plates. The p11A1 12 hour culture had 7.67 x 10<sup>5</sup> cells/ml, p11B3 had 1.33 x 10<sup>6</sup> cells/ml, p11C2 culture had 7.10 x 10<sup>5</sup> cells/ml and p11D1
had $1.24 \times 10^6$ cells/ml. Plasmid DNA was isolated from 1 ml of each 12 hour culture. Subclone p11C2 showed lower plasmid yield (Fig. 18) and was studied further.

Cells were plated with amp ranging from 0.25 µg/ml to 150 µg/ml to screen for differences in amp. Host cells without plasmids were also plated as controls and were inviable on plates containing as low as 0.1 µg/ml amp. By comparison, colonies containing p11C2 subclones maintained amp with amp concentration as high as 150 µg/ml. However, with increasing levels of amp there was a reduction in the number and size of colonies from p11C2 subclone cells. It was thought that cells with little-to-no plasmid that could be maintained on low amp but not on high amp either could have incorporated the β-lactamase gene into the bacterial chromosome or just have very few plasmids. Therefore, individual colonies that grew on 0.25 µg/ml amp plates were transferred to plates containing 0.25 µg/ml amp and also replicated on plates containing 150 µg/ml. This was done to screen for cells with little-to-no amp. Out of 288 clones, 117 colonies grew on low amp (0.25 µg/ml) and on high amp (150 µg/ml) but 171 colonies grew only on low amp. All colonies from low and high amp plates were cultured in 1 ml of TB with amp (0.25 µg/ml) and DNA was isolated by the Qbmp procedure. Yield of plasmid from colonies grown on
Figure 18. 1% agarose gel of four different p11 subclones: A1, B3, C2, D1. Markers are control pTZ19U and pN2. There are four samples from each subclone set represented: p11A1, p11B3, p11C2 and p11D1. All four p11C2 subclones have a similar band intensity pattern: bands are much less intense than other p11 subclones. This suggests that p11C2 subclones have reduced plasmid yield compared to the other subclones, yet retain ampicillin resistance. Also, it is observed that the lower band in p11A1, p11B3, and p11D1 subclones is not present in the p11C2 subclones. The p11C2 bands run similarly to the pN2 marker.
both low and high amp plates was at least 5 ng/μl. Little-to-no plasmid was obtained from those colonies that grew on low amp plates only.

Cells reserved from 22 samples that yielded no observable plasmid and from five samples containing plasmid were cultured and molded into gel plugs for intact total cellular DNA preparation. Plug slices were loaded onto a 1% agarose gel and the host chromosomal DNA separated from the plasmid DNA by FIGE. The DNA was transferred to a nylon membrane and hybridized with a total pN2 probe. DNA in the wells and some high molecular weight DNA bands were positive for pN2, even though no plasmid bands were present for some of the samples (data not shown). Although non-specific binding of the probe to non-pN2 DNA was observed (i.e. hybridization of the probe to DH5αF'IQ and JC7623 containing no plasmid) some homology might exist between pN2 and the bacterial chromosome through the lacZ gene, engineered into the pTZ19U plasmid and its recombinants and occurring in the E. coli chromosome of JC7623 and DH5αF'IQ strains (Table 2, p. 30). Thus, homologous recombination may be possible between plasmid and bacterial chromosome through the lacZ locus.

To refine what areas of pN2 might be homologous to or recombined into the chromosome, a second FIGE analysis was done using pN6 pII subclones along with linearized and uncut
Figure 19. FIGE analysis of the distribution of ORI probe sequences. (A) A 1% agarose FIGE gel of total bacterial cellular DNA was run in 45 mM TBE. Positive controls were uncut and ScaI-linearized pN2, dimer pN6 DNA, uncut and ScaI-linearized pTZ19U, pN6 plasmid DNA and four pl1C2 pN6 subclones (1-4) known to contain pN6. Size markers were the λ ladder (M1) and a marker made from pN2 restricted DNAs (M2). Genomic and pN6 plasmid DNA derived from DH5αF’IQ cells and genomic DNAs from DH5αF’IQ and JC7623 cells are other marker DNAs. (B) The hybridization pattern of the ORI probe is shown. The FIGE gel was blotted and the membrane hybridized with the ORI probe. The positive control DNAs in lanes one through six had strong hybridization with the ORI probe, as expected. However, unexpected positive hybridization to the ORI probe was found in some wells (i.e. wells 10 and 12 for pl1 subclones 4 and 6, respectively). This suggested that the RNA I, II locus might have been incorporated into the bacterial genome or that the pN6 plasmid DNA might have been trapped in the well.
pN2 and pTZ19U as controls. A bidirectional transfer of the DNA to two membranes was accomplished using an adaptation (Smith and Summers 1980) of the Southern transfer method (Southern 1975). Meanwhile, PCR was used to synthesize two biotinylated probes: one probe for the β-lactamase gene region (βLac) and one for the pTZ19U origin of replication including the RNA I and II replication sites (ORI). Positive hybridization to chromosomal DNA was observed in several samples as shown in Figure 19B for the ORI probe and for the βLac probe (data not shown).

To test whether plasmid DNA might have been trapped in the wells of the FIGE gels, the gel plugs were digested with HindIII. Any trapped plasmid DNA would be cleaved, leave the well, and migrate through the gel as higher molecular weight bands that could be clearly distinguished from bands of plasmid DNA. HindIII was chosen because it cleaves chromosomal DNA fewer times than HaeIII or other four-base cutters. In addition, HindIII leaves a more distinctive chromosomal banding pattern. Any plasmid DNA entangled in the chromosomal DNA also is free to migrate. A FIGE analysis of the restricted DNAs was run, blotted, and hybridized with the various probes. To test whether chromosomal DNA had been properly digested, the blot was probed for E. coli ribosomal RNA (rRNA) genes which are present in multiple copies. A similar banding pattern was
found for every DNA, indicating that the ribosomal DNA had undergone cleavage with HindIII, yielding a singular pattern of five bands for the five rRNA genes. This finding suggested that the chromosomal DNA had undergone cleavage. This blotted DNA also was hybridized with βLac and ORI probes. Hybridization was detected faintly in two of the wells and was most likely background since the rRNA gene cleavage pattern indicated complete cleavage of the DNA (data not shown).

**DNA Sequencing Data**

DNA sequencing was done to map the deletion boundaries at nucleotide resolution. It was unknown whether a contiguous stretch of DNA was deleted or whether several interspersed deletions occurred. Standard and specifically-designed primers were employed for sequencing deletions within the L2Hs element and those encompassing insert plus vector sequences.

Using the Universal Reverse primer, the site of the small, single 82-base deletion in pN2 p5JC10 was resolved. Boundaries of the deletion were confirmed by sequencing from the opposite direction using the Universal Forward primer. Boundaries of the deletions extended from positions 434 through 515 (Table 6).

Development of primers close to the origin of
### TABLE 6. PN2 SUBCLONE DELETION BOUNDARIES AND SIZES OF DELETIONS

<table>
<thead>
<tr>
<th>Subclones:</th>
<th>Primers used in sequencing:</th>
<th>5' Deletion Boundary (bp position)</th>
<th>3' Deletion Boundary (bp position)</th>
<th>Deletion Size (bp)</th>
<th>Features of the 3' Flanking Deletion Boundary</th>
<th>Direct Repeat Motif and Positions</th>
</tr>
</thead>
<tbody>
<tr>
<td>p5JC10**</td>
<td>UR¹ UF²</td>
<td>433</td>
<td>516</td>
<td>82</td>
<td>bp 516; 20% GC 5'-CTATAAT-3'</td>
<td>5'-GTACCCAAA-3' bp 507 bp 426</td>
</tr>
<tr>
<td>p14JC2,3,4</td>
<td>pNAMP1678³ pN1343R⁴ OriReverse⁵</td>
<td>3276</td>
<td>1134</td>
<td>1311</td>
<td>bp 1134; 90% GC-rich 5'-GCCAGCCGC-3'</td>
<td>5'-GCCAGC-3' bp 1134</td>
</tr>
<tr>
<td>p14JC6</td>
<td>pNAMP1678³ pN1343R⁴ OriReverse⁵</td>
<td>3276</td>
<td>1111</td>
<td>1289</td>
<td>bp 1111; 90% GC-rich 5'-CGCCAGCCT-3'</td>
<td>5'-GCCAGC-3' bp 1134</td>
</tr>
<tr>
<td>p14JC7</td>
<td>pNAMP1678³ pN1343R⁴ OriReverse⁵</td>
<td>3276</td>
<td>1193</td>
<td>1371</td>
<td>bp 1193; 70% GC-rich 5'-CGGTATCCC-3'</td>
<td>5'-GGCT-3' bp 1194</td>
</tr>
<tr>
<td>p28JC1**</td>
<td>pNAMP1678³ OriReverse⁵</td>
<td>3371</td>
<td>1510</td>
<td>1592</td>
<td>bp 1510; 27% GC 5'-CTATAAT-3'</td>
<td>5'-TGCG-3' bp 3341 bp 1546</td>
</tr>
<tr>
<td>p28JC2,3,4,6,7</td>
<td>pNAMP1678³ pN1343R⁴ OriReverse⁵</td>
<td>3308</td>
<td>1192</td>
<td>1377</td>
<td>bp 1182; 73% GC-rich 5'-CGGTATCCC-3'</td>
<td>5'-AGGAA-3' bp 3304 bp 1176</td>
</tr>
<tr>
<td>p28JC8</td>
<td>pNAMP1678³ pN1343R⁴ OriReverse⁵</td>
<td>3276</td>
<td>1192</td>
<td>1369</td>
<td>bp 1192; 73% GC-rich 5'-CGGTATCCC-3'</td>
<td>5'-GGCT-3' bp 3264 bp 1194</td>
</tr>
</tbody>
</table>

* Deletion boundaries and sizes inconclusive

** The flanking sequence is AT-rich in this subclone, rather than GC-rich

¹Priming positions (5'→3') of Universal Forward (UF) in pN2 are bp 951-928
²Priming positions (5'→3') of Universal Reverse (UR) in pN2 are bp 200-221
³Priming positions (5'→3') of pNAMP1678 in pN2 are bp 1678-1659
⁴Priming positions (5'→3') of pN1343R in pN2 are bp 1343-1324
⁵Priming positions (5'→3') of OriReverse in pN2 are bp 3039-3062

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
replication (OriReverse primer), the β-lactamase gene (pNAMP1678 primer), and a primer located between the β-lactamase gene primer site and the L2Hs insert (pN1343R primer) enabled the large deletion mutants to be sequenced. Any deletions extending upstream of the ori site would not be seen because replicative ability would be destroyed and these plasmids would be lost. Likewise, cells containing plasmids with deletions that extended into the β-lactamase gene (amp sensitive plasmids) would be selected against by amp in the growth media.

Using the OriReverse primer, pl4 and p28 pN2 subclones were sequenced. A single large deletion was found for these subclones, except for pl4JC5 which was like intact pO pN2. Also discovered were direct repeat motifs within close proximity of the deletion boundaries.

All the pl4 subclones, except for pl4JC1, pl4JC5 and pl4JC8, had a 5' (left) deletion boundary at bp position 3276. Subclone pl4JC5 was intact. Sequencing of pl4JC1 and pl4JC8 was incomplete due to unreadable gel patterns. The 3' (right) deletion boundaries occurred at bp positions 1111 (pl4JC6), 1134 (pl4JC2-4) or 1193 (pl4JC7), in sequences that are GC-rich (Table 6). The deletion sizes range from 1.1 kb to nearly 1.6 kb. Position 3276, the left flanking deletion boundary of pl4JC2-4, 6, 7 and p28JC8, is at the origin of replication (bp positions 3273 to 3279). A six bp
direct repeat (5'-GCCAGC-3') is found at the boundaries 1134 and 3276 in p14JC2-4 plasmids.

In the p28 subclones, the same large 1377 bp deletion was found in subclones p28JC2, p28JC3, p28JC5, p28JC6, and p28JC7 between position 3308 (5') and 1192 (3'). A four bp direct repeat at positions 3264 and 1194 (5'-GGCT-3') and a five bp direct repeat at positions 3304 and 1176 (5'-AGGAA-3') are near the deletion boundaries 3308 and 1192. Although these subclones originated from individual colonies, it is possible that some subclones were derived from the same cell because the deletions are identical. One exception is subclone p28JC1 which had a deletion of 1.59 kb between positions 3371 (5') and 1510 (3'). A four bp direct repeat at positions 3341 and 1546 (5'-GTCG-3') is in a range of ~30 bases from the deletion boundaries. Another example is p28JC8 which had a 1.37 kb deletion between positions 3276 (5') and 1192 (3'). A four bp direct repeat is found at positions 3264 and 1194 (5'-GGCT-3'), within close proximity of the deletion boundaries.

Many potential stem-loop structures have been predicted in regions near the deletion boundaries. A very clear example is for the p14 subclones which have a 5' deletion boundary at 3276 and a 3' deletion boundaries at 1111, 1134 or 1193. Near the 3' deletion boundary is an 18 bp inverted repeat at positions 1096 to 1113 and its complement at 1117.
to 1134 with a loop of 3 bases. At the 5' deletion boundary there is an 8 bp inverted repeat extending from 3276 to 3283 and its complement from 3211 to 3218 with an intervening region of 27 bases. These examples would be flanked by the direct repeats already mentioned.

To confirm that there were no small deletions flanking the β-lactamase gene, the counterclockwise pNAMP1678 primer (5'-to-3' positions 1678 to 1659) within the β-lactamase gene was used to sequence p14 and p28 subclones. Sequencing results showed that these plasmids were intact between positions 1646 to 1426 in pN2.

Primer pN1343R was used to sequence p14 and p28 pN2 subclones (except p14JC5 and p28JC1) in the counterclockwise direction from the 3' deletion boundary toward the 5' boundary. The region of RNA I and II, which is upstream of the ori site, was sequenced. These sequencing results demonstrated that the RNA I and II coding regions were intact between positions 2600 to 3200. Restriction mapping data confirms the intactness of this region (Fig. 13, Table 3, Fig. 15, Fig. 17, Table 4,).

Finally, it was observed that most of the 3' deletion boundaries of the p14 and p28 pN2 subclones consisted of GC-rich motifs except for p14JC5 and p28JC1 (Table 6, final column). Perhaps the GC-richness has provided transient stability at this flanking region.
A pl4 pN6 subclone was sequenced by an outside facility using the OriReverse primer. Results indicated a deletion of 1.6 kb from position 1 to 1624.
Repetitive DNA, ubiquitous to eukaryotic genomes, seems to show accelerated evolution via length and copy number polymorphisms in comparison to single-copy sequences. This rapid evolution is dependent on the sequence content, length and surrounding genomic environment (Epplen and others 1996). Simple repetitive DNA, referring to di-, tri-, and tetranucleotide repeats (i.e. microsatellites), is most often associated with this phenomena of genome instability; however, interspersed moderately repetitive sequences, like the L2Hs family, also fit into this category. The multiplicity of the L2Hs family has been shown by genomic fingerprinting and slot blotting: elements can vary from one individual to another and between tissues of the same individual (Musich and Dykes 1986, Musich 1996). Moreover, L2Hs elements differ in size and sequence, and are interspersed throughout the human genome. Thus, flanking sequences also differ among members of the L2Hs family, making it difficult to study specific interactions among the L2Hs elements, and between them and their adjacent sequences.

The essence of the genomic instability featured in the
L2Hs family has been ascertained in a bacterial model system in which a single L2Hs element was inserted into the plasmid pTZ19U. Subsequently, the recombinants were transformed into *E. coli* strains CES200, CES201, C600, DH5αF′IQ, and JC7623, and serially propagated for 28 passages. The control plasmid pTZ19U was stable throughout each passage while the L2Hs recombinants showed instability by passage 12 (p12) in strain JC7623, indicative of plasmids with significant deletions. The other host strains also appeared to generate a heterogeneous molecular mixture of pN2 plasmid. In order to distinguish whether the heterogeneity was due to deletions or just different topoisomeric forms, isolated plasmid DNA was linearized and assayed by gel electrophoresis. In all strains except for JC7623, the plasmid DNAs appeared intact. Some of the heterogeneity observed in these other host strains possibly was due to a mixture of different topoisomeric forms of plasmids or multimeric species such as dimers. Higher molecular weight DNA may have been due to linear multimers (Cohen and Clark 1986). Only the JC7623 strain, however, actually showed size reduction in the linearized DNAs.

The observation that the L2Hs recombinants became unstable in one host strain led to the hypothesis that the JC7623 cell strain was causing, influencing, or allowing the
instability. However, pTZ19U plasmid alone was stable in the JC7623 cell strain. Also, a pTZ19U recombinant containing a 680 bp rhesus monkey alphoid DNA (pRh) was stable when similarly passaged. This rhesus element contained four tandem repeats of 170 bp (Pike and others 1986). Thus, it seemed that features of the L2Hs element itself influenced plasmid stability in the JC7623 cell strain.

An important difference of the JC7623 host strain with regard to the other hosts is its recombination genotype: recBrecCsbcBsbcCsbcD (often reported as a quadruple mutant: recBrecCsbcBsbcC). SbcCD, the gene product of wild type sbcCsbcD, is a multimeric protein possessing ATP-dependent double-strand exonuclease and ATP-independent single-stand DNA endonuclease activities (Connelly and Leach 1996). This gene product aids in palindrome-mediated inviability associated with an arrest of DNA replication (Shurvinton and others 1987, Connelly and Leach 1996). SbcCD may attack hairpins formed through replication slippage at palindromes, forcing collapse of the replication fork (Connelly and Leach 1996). The hairpin formation is thought to occur on only one of the DNA strands, most likely the lagging-strand (Trinh and Sinden 1991, Rosche and others 1995). Repair of the fork is possible through the most common recombination
system encountered in the bacterial chromosome: the RecBCD complex and associated Chi site (χ-site). When a replication fork collapses, the RecBCD complex continues to degrade the fork until a special sequence motif, known as the χ-site, is encountered. This χ sequence converts RecBCD into a recombinase which continues unwinding the DNA. The single strands bind RecA, a strand-exchange protein in homologous recombination, and are enabled to invade the intact sister strand. Leading- and lagging-strand synthesis is re-initiated (Asai and Kogoma 1994, Ryder and others 1996). Leach refers to this type of repair as "arm-directed secondary structure repair" (Fig. 20).

Thus, interaction of the RecBCD-χ complex with the collapsed replication fork initiated by SbcCD allows the sister strand to resume synthesis and replication can begin again at the original start site.

In the JC7623 cell strain, a different recombination and repair system is utilized because recBrecCrecD and sbcCsbcD genes are aberrant. Hairpins created by single-stranded palindromic sequences would not be removed by the nonfunctional protein product of mutant sbcCD. Also, repair would not be made through the activities of mutant recBrecCrecD gene products. Instead, four scenarios could occur in JC7623 cells. First, if no secondary DNA structure
Figure 20. Collapse of a replication fork by SbcCD and the fork's repair via the RecABCD-χ-mediated recombination/repair proposed by Leach in the arm-directed repair theory. During replication of a palindromic DNA sequence, intrastrand base-pairing may occur on the single-stranded portion of the lagging strand, downstream of the replication fork. Pausing of DNA replication occurs at the secondary structure. The SbcCD protein removes the secondary structure, generating a double-strand break. The double-strand break recruits RecBCD and RecA (homologous recombinatory enzymes) to allow the reconstitution of a replication fork by recombination with the intact sister arm. The broken arm is digested by RecBCD until a χ-site is encountered. A Holliday junction is formed by RecA protein and is resolved. This regenerates the replication fork. No recombination occurs with the intact copy of the palindromic sequence. Thus, the sequence is replicated after replication is reinitiated. Source: Leach, D.R.F., Okely, E.A. and Pinder, D.J. (1997). Repair by recombination of DNA containing a palindromic sequence. Mol. Microbiol. 26, 597-606.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
were made, replication proceeds unhindered. If non-B-form secondary structures were formed during replication, stalling of DNA polymerase and its loss of processivity at the secondary structure might lead to truncation of replication and inviability of the progeny. A third possibility is that the DNA polymerase might skip over the secondary structure and continue replication (i.e. strand extrusion) on the other side of this structure, leading to deletion of the structural motif (i.e. palindromic sequence). Finally, stalling at the secondary structure could allow time for DNA "breathing" and intrastrand realignment of other repeated sequences that might effect slippage events. Thus, the result might be deletion or expansion of the genome. Recombinants pN2 or pN6, which have many palindromes/inverted repeats in the L2Hs element, have the potential to form secondary structures at replication forks. When hosted in JC7623 cells, genomic instability may occur because of defective "arm-directed secondary structure repair" due to a mutant SbcCD protein and lack of the RecABCD-\(\chi\)-mediated recombination/repair pathway.

Replication of plasmids containing palindromes 22 bp or greater also may be aberrant due to linear plasmid multimer formation (Biek and Cohen 1986, Cohen and Clark 1986, Leach
1996). Because each multimer segregates into the daughter cells as a single plasmid, the effective copy number is lowered in each cell during cell division. Thus, the chance of plasmid loss is increased with each cell division. In plasmids such as ColEl, multimers can be resolved before cell division at the cer locus by the Xer proteins, site-specific recombinases that act on the cer site to resolve multimers after replication (Snyder and Champness 1997). Plasmid pBR322 and derivatives such as pTZ19U lack the cer site, resulting in the increased stability of plasmid multimers (Pouwels 1991). Furthermore, these plasmids lack a partition function needed for proper segregation of plasmid vectors to daughter cells during cell division. Thus, these multimers are maintained and are not resolved by Xer. Moreover, uneven segregation into daughter cells may result, reducing copy number even more in some of the daughter cells. Additionally, the stability of L2Hs recombinants in some cell strains and not in others may be influenced by the levels of topoisomerases which act as multimer resolvases. Topoisomerase I and DNA gyrase both can resolve dimers generated in pBR322 replication (Kornberg and Baker 1992). Finally, it still is unclear if other E. coli enzymes exist for this particular function of multimer resolution and, if so, under what conditions their action is
While L2Hs recombinants are unstable in strains like JC7623, they are stable in recA mutant strains like DH5αF’IQ. The recA wild-type gene is essential in most homologous recombination events (Kowalczykowski and others 1994). This recA mutant is recombination deficient but also has the wild-type sbcCsbcD and recBrecC genotype. recA mutants also are capable of maintaining stable direct repeats (usually greater than 50 bp) (Hanahan 1983). Ergo, L2Hs plasmids are stable in strain DH5αF’IQ.

Other factors inherent in the L2Hs element may contribute to plasmid instability. For example, the small deletion found in subclone p5JC10 may have resulted from replication slippage on the lagging strand in DNA synthesis. The L2Hs insert, which is AT-rich and thermodynamically less stable than the vector, may become unwound. With the contribution of delayed synthesis on the lagging strand, the opportunity for slippage from one direct repeat to another on the lagging strand is likely. In pN2, direct repeats (5’-GTACCCAAA-3’) are found at positions 426 to 434 (DR2) and 507 to 515 (DR1) (see Fig. 21 and Appendix B). These positions encompass positions 433 and 516, the boundaries of the small deletion of subclone p5JC10 that have been verified by DNA sequencing. Also several interspersed
inverted repeats are located between the positions 433 and 516. One example of inverted repeats is that occurring at positions 435 to 439 (IR2 is 5'-TATAT-3') and positions 443 to 447 (IR1 is 5'-ATATA-3'). When helicase-mediated strand separation occurs during replication, interspersed inverted repeats such as IR1 and IR2 may form a hairpin secondary structure. DR1' could be replicated. However, when the replication machinery reaches the hairpin structure, pausing might occur with dissolution of the synthetic complex. DR1 may denature and anneal with the DR2' template beyond the hairpin. Thus, DR1' and the hairpin sequence would not be replicated and would be deleted. Figure 21 depicts the hypothesized p5JC10 deletion event.

Some observations regarding the large deletion mutants of pN2 should be noted. The p14 and p28 mutants of pN2 plasmids have left deletion boundaries positions 3276, 3308, and 3371 (Table 6) that are close to or within the origin of replication. The right deletion boundaries for the p14 and p28 subclones extend to positions 1111, 1134, 1192, 1193, and 1510. The first 10-11 bases of the right deletion boundaries are GC-rich (73-90% as shown in Table 6) in all except subclones p14JC5 (not shown in Table 6) and p28JC1. Deletion boundaries for p14JC2, p14JC3, and p14JC4 mapped to positions 3276 and 1134. Also, direct repeats (GCCAGC) were found at these positions. These repeats may be involved in
Figure 21. The deletion mechanism of mutant p5JC10. A step-by-step account of a possible deletion process converting pN2 to p5JC10. (1) The lagging strand replicates through DR1'. (2) Simultaneously, the bases between DR1' and DR2' have looped out. This could occur because of the inverted repeats and quasi-inverted repeats found in this locus. The replication machinery halts at the secondary structure. (3) The newly synthesized DR1 disassociates from the lagging strand and realigns with template DR2', downstream of the loop. Replication continues from this point. (4) The result of a second round of replication of the upper strand of (3) is depicted. Bases 434 to 515 have been lost and there is only one DR motif left. DR1' in (4) means the complement of DR1 has been replicated. This figure is an adaptation of Sinden’s mechanism (Fig. 5).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
slipped mispairing during replication. For example, in the p14JC2, -JC3, and -JC4 subclones, the region between base positions 3276 and 1134 may form a series of unstable, dynamic secondary structures involving the L2Hs element on the single-stranded lagging strand during replication. To further support this notion, inverted repeats of 18 bases (positions 1096 to 1113 and 1117 to 1134) with a loop of 3 bases (positions 1114 to 1116) are found at the 3' end deletion boundary. A possible stem-loop structure with inverted repeats of 8 bases (positions 3276 to 3283 and 3211 to 3218) and a loop of 27 bases, located at positions 3284 to 3310, is found at the 5' deletion boundary. Another stem-loop structure with inverted repeats of 15 bases (positions 65 to 79 and 1038 to 1052) and a loop size of 958 (positions 80 to 1037) is another possibility. The direct repeats discovered at the deletion boundaries would flank these inverted repeats. There are several hundred predicted stem-loop structures, similar to the examples given, that are internal to the deletion boundaries. Thus, a hypothetical multiplex structure, depicted in Figure 22, utilizes the potential for internal stem-loop structures flanked by direct repeats at the deletion boundaries. This potential structure may stabilize the pairing of the direct repeat at 3276 and the complementary distant repeat at 1134. When replication occurs, the intervening sequences are then

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
skipped over, causing a large deletion. Figure 23 depicts this deletion process using p14JC2 as an example.

When the ori site unwinds, the entrance of DnaBC helicase and DnaG primase needed to lay down the initial RNA primer for DNA replication is stimulated. Thus, a replication fork is initiated. As leading strand replication proceeds, a large segment of plasmid at least 1 to 2 kb becomes the single-stranded lagging strand template (Kornberg and Baker 1992). The delay of replication on the lagging strand may allow intrastrand base pairing to occur, especially between palindromes or inverted repeats. For the L2Hs elements, intrastrand base pairing as indicated in Fig. 22 may involve nearly 40% of the element.

A comparison of pN2 and pN6 subclones was done to see if insert orientation influenced the types of instability observed. In the initial pN2 study, a mixture of intact plasmids and possible deletion species was seen by p12. Subcloning passages 2, 5, 14, and 28 revealed a small deletion in one of the p5 subclones and deletions ranging from 1.29 kb to 1.59 kb for p14 and all but one of the p28 subclones. A more extensive subcloning of pN2 was done looking at 24 to 48 subclones from each passage. Large deletions were found as early as p2, but it was not until p12 that approximately 50% of the subclones had large deletions. No other small deletions (less than the insert
Figure 22. A possible multiplex stem-loop structure in the L2Hs element. (A) The multiplex structure showing the size of the loops in bp (and the position). The bp positions within the loops and stems correspond with the underlined stems and loops shown in the Table in (B). The Stem Energy depicts the stability of the cruciform in the Kcal/mole in 1 M Na+ at pH 7.25 as calculated by the Hitachi DNAsis program (Hitachi America). The bp coordinates are expressed relative to plasmid pN6. Base 1 indicate 48 bases upstream of the Kpnl site in plasmid pN6 (bp 301 in pN6) and would correspond to base 253 in pN6. Source: Musich, P.R. (1996). Human genome instability - features of a dynamic repetitive element. Einstein Quart. J. Biol. Med. 12, 74-84.
Figure 23. Schematic of proposed deletion mechanism generating p14 and p28 subclones involving replication slippage and unusual non-B-form DNA structures. (A) The pN2 plasmid has two direct repeats DR1 (3276 to 3281) and DR2 (1134 to 1140) flanking the L2Hs element (305 to 895). (B) The plasmid strand separates when unidirectional replication begins. As leading strand replication continues, denaturation includes the region encompassing the L2Hs element. (C) Unusual non-B-form structures may form on the lagging strand, such as hairpins, cruciform or the multiplex structure shown in Fig. 24A. Pausing of the replication machinery on the lagging strand may occur at such a secondary structure. At this point, replication of DR2' on the lagging strand is assumed to have occurred. With slippage of the lagging strand, DR1' slips back on the newly synthesized DR2'(as illustrated in Fig. 22). (D) The final outcome shows the completed replication of the nascent strand synthesized from the lagging strand after a second round of replication. A large deletion is the result, which includes the region between bp 3276 and 1134. This may explain how the mutant pN2 p14JC2 was generated.
leading strand
lagging strand
replication of leading strand
Second Round of Replication

pN2
OR1
pN2

p14JC2
DR2-1
DR2-2nd
DR1-2nd

size) were discovered. This suggests that the deletion process was not an accumulation of small deletions but a single, large deletion event. An extensive subcloning of pN6 subclones revealed large deletions (>1 kb) by p6. No small deletions were observed. Restriction cleavage studies (HaeIII, HinfI) show that observable left flanking deletion boundaries stop at the ori site, just as with the pN2 large deletion mutants. Size of the pN6 deletion mutants based on ScaI and SspI restriction mapping indicated deletions ranging from 1.2 kb to 2.0 kb. When pN2 and pN6 DNAs were serially propagated simultaneously, pN6 DNAs were observed to have deletions by p6, while deletions were not apparent in pN2 DNAs until p12. In another serial propagation study of pN2 to screen for intermediate-sized deletions, large deletion mutants were observed as early as p2. The one small deletion mutant p5JCl0 in pN2, discovered out of more than 1000 subclones of pN2 and pN6 analyzed, suggests a replication error as the source of the deletion process. Because the other mutants had large deletions, events producing large deletions may have occurred early in the serial passage. Due to the replicative advantage of smaller plasmids, the deletion species became dominant in the plasmid population. This process may be exacerbated by the fact that pTZ19U and its recombinants are high copy-
number plasmids.

The sequencing of a p14 pN6 subclone with a 1.6 kb deletion reveals differences between the deletion boundaries of pN2 and pN6 mutants. Analysis of the sequencing data indicated that the large deletion occurred between positions 1 and 1624, downstream from rather than at the ori site. Because no intermediate-sized deletions were found in the pN6 mutants, the deletion process was probably a single event rather than an accumulation of many smaller deletions.

There is an obvious limit on the size of observable deletions. Plasmids with deletions that interrupted the RNA I and II loci or β-lactamase gene locus would be inviable. Such deletions may occur but would not be detectable in our assay.

The regulation of plasmid replication is important to plasmid viability. Also involved is a mechanism inhibiting the initiation of plasmid replication when the number of plasmids in the cell reaches a certain level (Snyder and Champness 1997). This regulation of replication and control of plasmid copy number occurs by plasmid-encoded RNA I and RNA II. Of crucial significance is the fact that any mutations or deletions in the RNA I- and RNA II-encoding regions could inhibit plasmid replication and/or the control of copy number. Without plasmid replication, the host becomes increasingly sensitive to ampicillin.
Cells with high copy-number plasmids carrying the β-lactamase gene for amp\(^3\) could survive higher concentrations of amp treatment than cells with fewer plasmids. With loss of plasmids from cells, amp\(^3\) would be maintained but lower concentrations of amp would have to be used. The plasmid copy-number can also be manipulated by the amount of antibiotic used to grow cells that carry plasmids that confer resistance to that antibiotic. A threshold level of antibiotic selects for cells carrying plasmids. Plasmid copy-number is lowered when little-to-no antibiotic is used. There is no selective advantage for a cell to expend the metabolic resources to maintain plasmids if there is no antibiotic selection. Eventually, plasmids might be lost from the dividing cells.

In some of the pN2 and pN6 middle-to-late passages, it was observed that high amp\(^3\) remained while the yield of plasmid seemed reduced. It was hypothesized that perhaps the β-lactamase gene may have been incorporated into the bacterial chromosome, still conferring amp\(^3\) to the cell. To test this, an experiment was done to observe the effects of various concentrations of amp on the host cell and on the plasmid copy number. Colonies derived from individual cells plated on LB-agar were more abundant and grew larger on plates with 0.25-0.5 µg/ml amp. Colonies were smaller and fewer on plates containing 150 µg/ml amp. However, plasmid
yield from cultures (estimated $1.5 \times 10^3$-2.5 x $10^8$ cells/ml in an overnight culture) derived from colonies grown on 0.25-0.5 μg/ml plates appeared much reduced relative to cells from colonies picked from 150 μg/ml amp. In addition, some cell cultures from low amp plates failed to yield plasmid DNA, yet the cells were ampR (0.25-2 μg/ml amp). These cells were then tested to see if very low levels of plasmid were present or, alternatively, if the β-lactamase gene portion of plasmid DNA had been incorporated into the bacterial chromosome. Intact cellular DNA was prepared in agarose plugs for FIGE analysis. With the FIGE conditions employed (Figs. 19), chromosomal DNA should remain in the wells while the smaller plasmid DNA migrates through the gel. Hybridization analyses with probes for the pN2 plasmid, the β-lactamase gene and ori region were inconclusive, perhaps due to possible plasmid trapping in the well. For instance, in the analysis using the β-lactamase gene probe, positive hybridization signals were obtained from both plasmid DNA bands and chromosomal DNA in some of the wells. A weak signal might be expected in the region of the well if the β-lactamase gene had incorporated into the bacterial chromosome since there are only a few copies (~2-4) of the replicating chromosome per rapidly growing cell. Alternatively, these weak bands could be the result of plasmid trapping. Restriction cleavage was used
to distinguish whether these weak signals were due to covalent linkage of the β-lactamase gene to chromosomal DNA or if they were the result of trapping of plasmid DNA in the DNA plug. DNA was cleaved into smaller fragments to assure that all DNA would leave the wells. The plasmid and chromosomal fragments could be distinguished from each other by predicted plasmid banding patterns. There was no indication of incorporation of the plasmid β-lactamase gene or ori region into the host chromosomal DNA. This finding suggests that it takes only a few plasmids to maintain high resistance against antibiotics.

Several explanations address the cause of low plasmid yield from some of the passaged cells: that of unequal partitioning of plasmids during replication, plasmid inviability through deletion of the RNA I and II loci or cell inviability through the deletion of the β-lac gene. Those few remaining plasmids retained ampicillin resistance. Another explanation is that differences in the number of cells to prepare the DNA may have biased the outcome.

The AT-richness, sequence motifs and potential to form secondary non-B-form DNA structures appear to contribute to the instability of the L2Hs element in the E. coli strain JC7623. Several models of replication slippage and error suggest how deletions might occur, given particular
conditions of supercoiling and torsional tension (Zheng and others 1991, Trinh and Sinden 1993, Rosche and others 1995, Wierdl and others 1996, Kramer and Sinden 1997). In a mutant recB recC sbcB sbcC sbcD genotype, the SbcCD nuclease is nonfunctional. Therefore, stem-loop structures, formed from palindromes during replication and normally cleaved by SbcCD, would remain. However, pausing of the replication machinery at the stem-loop is likely and would result in truncation of replication, strand slippage or the replication machinery skipping over the entire stem structure leading to deletion. Leach and others suggest that palindromes cannot be stably maintained in plasmids with this mutant recB recC sbcB sbcC sbcD genotype (Leach 1996, Akgün and others 1997). The JC7623 genotype may affect the repair mechanisms involved at collapsed replication forks induced by stem-loop structures. Hence, the inherent dynamics of the L2Hs element and the succeeding instability, are demonstrated directly in the JC7623 strain.

Retrospection and Additional Concerns

The L2Hs element exhibits instability under very specific conditions. The L2Hs element is unstable in plasmids in E.coli host strain JC7623 in which the nuclease SbcCD is nonfunctional. This nonfunctional protein implies that replication errors that may occur at unusual non-B-form
structures are unable to be resolved and repaired. In turn, this suggests that the actual deletion process occurs during replication. Would instability occur only during replication, while exponential growth phase of the host was taking place? This question could be answered by designing an experiment in which plasmid DNA was isolated from a cell culture at various time points. An essential requirement is that cell counts would have to be done at every time point (i.e. cell doubling time), along with very careful quantitative isolation of plasmid DNA.

Another area of retrospection concerns the model explaining the large deletions in the prokaryotic system. This model relies on a hypothesis that the L2Hs element may undergo various multiplex conformations when the DNA becomes single-stranded. The multiplex structures were adapted from single-stranded RNA pseudoknot designs. Although substantial differences between RNA and DNA exist, including base composition and structural forms, it seems likely that L2Hs DNA would have the potential to fold into the hypothesized multiplex structures because of its repetitive motifs, especially the inverted repeats. The documentation of single-stranded regions of the L2Hs element under different supercoiling environments (Ratnasinghe 1993) strengthens the argument for the existence of multiplex
A limitation in understanding both the deletion processes and the resolution of non-B-form DNA structures is an area of concern. The L2Hs element's AT-richness and palindromic and repeat sequence features appear to influence the stability of flanking sequences within the JC7623 host strain. A replication slippage model seems best to explain how the instability could occur. However, it is conceivable that there may be contributing factors such as specific, although still undefined, nucleases and other host proteins that resolve these postulated non-B-form structures and aid the slippage process during DNA replication.

The Relationship of the Bacterial Model to the Human System

Those features of the L2Hs element that affect change in the bacterial model system may also influence the stability of eukaryotic systems. Certain sequence motifs present in the L2Hs element may form unusual secondary non-B-form DNA structures in a eukaryotic system, too. Rather than instability through deletion as observed in the bacterial model, there may be amplification of the L2Hs element. Evidence for expansion of repetitive elements in the human genome comes from studies regarding trinucleotide repeat arrays and their correlation to human disease. Gene
amplification also is found to be a part of large palindromes. Butler and others (1996) report a model involved in generation of a large palindrome in which a double-strand break is introduced next to a pair of short inverted repeats. This model has been observed in *Tetrahymena* and *Saccharomyces*. Many small inverted sequence motifs are found in the L2Hs element which may undergo such an event. Thus, in a eukaryotic system, amplification rather than deletion of L2Hs elements may occur.

The L2Hs element’s instability in the JC7623 host strain was related to replication and repair errors. Similar instability of the L2Hs element and flanking regions may occur in the human system. Although the basis for the L2Hs element’s polymorphisms has not been fully appraised in the human system, reasons for instability in several other comparable repetitive DNA sequences have been determined. Often the instability relates to replication slippage and repair errors. Several examples of eukaryotic genome instability have been elicited via human disease. For instance, a mutator phenotype that leads to intrinsic genetic instability has been postulated to play a major role in the development of carcinogenesis (Parsons and others 1993). Characteristic size instability (deletions or expansions) in di- or trinucleotide repeats, which also
occurs in the L2Hs element, have been found in hereditary nonpolyposis colorectal cancer (Radman and Wagner 1993). Most tumors developing in hereditary nonpolyposis colorectal cancer patients have mutations in tandem dinucleotide repeat arrays (Parsons and others 1993). Slipped mispairing of these repetitive arrays has been postulated, leading to insertion or deletion mutagenesis during replication. Parsons and others (1993) have shown that vectors containing dinucleotide repeat arrays which were transfected into cells derived from colorectal tumor lines characterized with replication error phenotypes, were more mutable than those vectors transfected into colorectal cancer cells derived from a non-replication error phenotype. It was discovered that the tumor cells with the replication error phenotype also exhibited defective mismatch repair. This finding is consistent with the known role of mismatch repair stabilizing (CA)$_n$ repeat sequences in E. coli and yeast (Levinson and Gutman, 1987, Parsons and others 1993, Strand and others 1993). Thus, it may be conjectured that mismatch repair mechanisms in humans may stabilize potentially dynamic sequences like those in the L2Hs family. Otherwise, elements like the L2Hs might be completely absent from the human genome or unduly amplified. These phenomena have been discovered with some repetitive sequences observed in
certain instances of cancer or other disease states, such as fragile X syndrome, myotonic dystrophy, and Huntington’s disease in which a deficient repair mechanism is present.

Another consideration pertains to protein binding in the eukaryotic system. In the bacterial model used, a nonfunctional *sbcCD* gene product contributed to a replication repair error that increased the likelihood of deletions of palindromic sequences. A comparable SbcCD protein in eukaryotes may be discovered in the future, although undisclosed at this time. However, are there other specific proteins that bind to L2Hs elements and moderate the formation of potential non-B-form DNA structures? Possible candidates include topoisomerases, helicases, resolvases or transcription factors known to recognize unusual DNA motifs (Giffin and Haché 1995, Steinmetzer and others 1995, Bennett and West 1996, Pearson and others 1996, Tomonaga and Levens 1997).

The L2Hs element’s stability in the bacterial model may have been influenced by the element’s juxtaposition with the origin of replication and also by the affects of lagging strand synthesis. How would stability of the L2Hs element be effected by multiple origins of replication in a eukaryotic system and what would be the results of lagging strand synthesis? Because origins of replication are AT-
rich, it might be predicted that L2Hs elements would be close to or even part of the origin. Also, lagging strand synthesis might facilitate non-B-form DNA structural formation. Misalignment or slippage might follow, as hypothesized in the bacterial model. Comparable mechanisms of repair are a possibility, yet not clearly appreciated.

Finally, one last prediction about the L2Hs element may be surmised. In the bacterial model, the L2Hs DNA was observed only in the context of a plasmid system, with the host chromosome supplying the necessary enzymes for repair and recombination and the L2Hs element serving no function except via its potential for secondary structure. In contrast, the L2Hs element in the human system may function as a nuclear matrix attachment region (MAR). Features of the L2Hs element, such as AT-richness, inverted and direct repeat motifs, are shared by putative MARs (Bode and others 1992, Boulikas 1993, Vogt 1992, Kay and Bode 1994, Nadir and others 1996, Kramer and Sinden 1997). MARs serve as anchorage sites of chromatin loops to the nuclear matrix and are thought to be involved in replication, acting as a "torsional sink" by absorbing negative supercoiling generated by transcribing RNA polymerase (Sinden 1994). MARs also may be involved in transcription and its control. Thus, the L2Hs element may play a significant role in human
genome activity.

Now that characterization of the L2Hs element in a bacterial model answers questions about the intrinsic potential of the L2Hs element to develop non-B-form DNA structures and affect flanking regions, additional issues may be approached. The intrinsic instability of the L2Hs element is enhanced by the host strain in the bacterial model. Thus, it may be postulated that in vivo processes in human cells probably affect L2Hs elements. The significance of repetitive sequences like the L2Hs elements with regard to organization of the genome as well as possible regulatory roles concerning chromosome structure, replication and transcription may be elucidated one day. The relationship of repetitive sequences to mutation and disease also might be uncovered. These challenges multiply with each new discovery and piece of data.
BIBLIOGRAPHY
BIBLIOGRAPHY


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX A

Buffers and Solutions

Alkaline lysis solution
0.2 N NaOH, 1% SDS (made fresh)

Bacterial lysozyme solution for FIGE protocol
10mM Tris (pH 7.5), 50 mM NaCl, 100 mM EDTA, 0.2% Na deoxycholate, 0.5% sarcosyl, Na+ salt, 1 mg/ml lysozyme solution (10 mg/ml stock)

Blocking solution for Phototope™ chemiluminescent detection
5% SDS, 25 mM phosphate buffer (pH 7.2)

Denhardt's reagent (50X stock solution)
5 g Ficoll, 5 g polyvinylpyrrolidone, 5 g bovine serum albumin, add 450 ml dH₂O, mix to dissolve, bring to a final volume of 500 ml with dH₂O, filter to sterilize, store at -20°C

Digestion Buffer for FIGE DNA preparation protocol
0.5M EDTA (pH 8.0), 1% lauroyl sarcosine, sodium salt, 0.1mg/ml Proteinase K.

LB broth
10 g bactotryptone, 5 g bacto-yeast extract, 10 g NaCl, add 800 ml dH₂O, mix to dissolve, bring to 1 liter with dH₂O, autoclave to sterilize

PEG solution
27% PEG 8000 (w/v) in 3.3 M NaCl

Prehybridization and hybridization buffer
3 X SSPE, 3% SDS, 5X Denhardt's reagent plus 100 μg/ml single-stranded herring sperm DNA

RNase A
20,000 U/ml in 10 mM Tris-HCl (pH 7.5), 15 mM NaCl, boiled for 15 minutes

Sequencing stop buffer
80% deionized formamide, 10mM NaOH, 1 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue
SET solution
20% sucrose, 50 mM Tris-HCl (pH 8.0), 50 mM EDTA

SBQB
60% sucrose/0.1% bromophenol blue

SSC (1X)
0.15 M NaCl, 15 mM sodium citrate, (pH 7.0)

SSPE (1X)
0.18 M NaCl, 10 mM NaPO4, 1 mM EDTA, (pH 7.7)

STT
8% sucrose, 5 mM Tris-HCl, (pH 8.0), 5% Triton X-100

T-E-
10 mM Tris-HCl (pH 8.4), 1 mM EDTA

T-E-
1 mM Tris-HCl (pH 8.4), 0.1 mM EDTA

TAE
50 mM Tris base, 40 mM acetic acid, 1 mM EDTA, final pH ~8.4

TB
Add 100 ml of a sterile solution of 0.17 M KH2PO4 and 0.72 M K2HPO4 to 900 ml of base broth. Base broth contains 12 g tryptone, 24 g yeast extract and 4 ml glycerol brought to 900 ml with deionized water and autoclaved

TBE
50 mM Tris base, 50 mM boric acid, 1.25 mM EDTA, final pH ~8.3

Wash solution I for Phototope™ chemiluminescent detection
0.5% SDS, 2.5 mM phosphate buffer (pH 7.2)

Wash solution II for Phototope™ chemiluminescent detection
100 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl2 (pH 9.2)
APPENDIX B

Recombinant plasmid pN2  3453 bp

BASE COUNT:  965 A  806 C  762 G  920 T

5'-'AGCGCCCAAT AGCAGAACCG CTCTCTCCCG CCGCTTGCCC GATTCATTAA TGCGAGCTG-

61  71  81  91 101 111
-AGCACAGGTT TCCCCACTGG AAGAGGGCAG TGAGGCGCAA CGCAATTAAT GTGAGTGGC-

121  131  141  151  161  171
-TCACTCATTA GGCACCCCAG GCTTTGTTCT TTAATGCTCG GTGTGTTGGA-

181  191  201  211  221  231
-CTGTGAGCGG ATAAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAGCTCT-

241  251  261  271  281  291
-AATACGACTC ACTATAGGGA AAGCTTGCACT GTCTGCAGCT CGACTCTTAGA GGATCCCCCG-

301  311  321  331  341  351
-GTACCCAAAAC TATACTATTAT ATACTGTCAG TAGATAAGGA AATTAACATCA ATATATTATA-

361  371  381  391  401  411
-TATAGGTAC ATAAAAATATG AAATGTACATC AAATATAGAT ATATACTGT ACATAAAATA-

421  431  441  451  461  471
-TCAAGTACCAAATATATA TCAATATGTT GTATGAAATT CAGTTCACAA ACTATATTAT-

481  491  501  511  521  531
-ATATACGTGCT ACATABAATA TCTAAAGTAC CCAAAACTATA CATTATATAC GTGATACATA-

541  551  561  571  581  591
-AATATGAAAT ACATCAAATAT TGTGAATTTAT ATTAGGTAGT TAAATATGAA GTACATCAA-

601  611  621  631  641  651
-ATATAGGTAT ATATAAGGTG TACATCAAAA ACCAAGTAC CCCGAAATAT ATATATTATA-

661  671  681  691  701  711
-CTGTACATGA AATATCAAAG TCTCTCAAAC ATATATTATA TACTGACAT AAAATATCAA-

721  731  741  751  761  771
-ATGACCCCGATG ATAATATATAC TAAATATGAA TCTAAAGTAC CCAAAACTATA CATTATATAC-

781  791  801  811  821  831
-TATAATCCGACA ATACAAATAT TGGAAAGTAC TCAAATATAT ATTTATTTTT GTACATAAAA-

841  851  861  871  881  891
-TATACAAATGC ACAAATAG TATATATTAT ACTGTACATA AAATATAGAT ATGACGACCT-

901  911  921  931  941  951
-CGAATTGACAT GGCCTGTGTT TTACAACGCT GTGACTGGGA AAACCATGGC GTTACCACAC-

961  971  981  991  1001 1011
-ATTAATCCGATA TACAATGAAAT TCTAAAGTAC CCAAAACTATA CATTATATAC GTGATACATA-

1021 1031 1041 1051 1061 1071
-CCGATGCCAC TCCCCTCTCG CCCGCTGGCC TAAATAGCAA RTGGACGCG CCGTGAGTGC-

1081 1091 1101 1111 1121 1131
-CGCGACAGGG CGCGCCGGGT GTGGTGTTTA CGCGCAAGCT GACCCGTCACA TGGCGACCC-

1141 1151 1161 1171 1181 1191
-CCCTAGGGCG CGCTCTCTCT CCTTCTTTCT CTCTCCTTCT GCCACAGTGC GCCGGCTTTC-

1201 1211 1221 1231 1241 1251
-CTGGCTACAC GCTAAATGCG GCCCTCTTCC TTAGGTGTCAT TTGAGTGCTT TTACGCCACC-

1261 1271 1281 1291 1301 1311
-TCGACCCCAA AAAACTTGAT TAGGCTGATG GTCACGAGTG TGGCCCTACG CCGTGATAGA-
-CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA-
-CTGGAACAC ACTCAACCT ATCTCGTCT ATTTTTGGA TTTATAAAGG ATTTTGCCGA-
-TTCGGCCCTA TTGGTTAAA AATGAGCTGA TTAAACAAAA ATTTAAGCGG ATTTTTAACA-
-AAAATATACG TTTTACAATT TCAGGTGCGA CTTTTCGCGG AATGTGCGGG GGAACCCTAA-
-TTGTGTATT TTTCTAAATA CATTCAAATA TGATCAGCGT CATCAGACAA TAAACCCTGAT-
-AAATTCTTCA ATAATATTGR AAAAAAGAAG ATATGGATAT TCAACATTTC CGTGTGCAGCC-
-TTATTTCCCTTT TTTTGCCGCA TTTTGCCTTC TGATTTTTGC TGACCCAGAA ACCCTTGAGA-
-TAAGTAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA-
-AACAGCCGTTAA GATCCTTGGG ATTTTTCCGCC CGAAGAAGC TTTTCAATAG ATGAGCACTT-
-TTTAAGTCTT GCATATGCGC GCGTTATTAT CCGTTATGGA CGCCGGCGCA GAGCAATCAG-
-GTGGCCGGAT AACTATATTCT CAGAAGCTGA TTTAACAAAA ATTTAACGCGA AATTTTACAA-
-ACAGCGGTAA GATCCTTGAG AGTTTTTCCGCC CGAAGAAGC TTTTCAATAG ATGAGCACTT-
-TGCAACACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG GGAACCGGAG CTGAATGAAG-
-CAATACAAA TGAGGAGCGT GACACCACGA TGCCCTGTAGC AAAGCAAACA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
-TGCAACACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG GGAACCGGAG CTGAATGAAG-
-CCATACAAA TGAGGAGCGT GACACCACGA TGCCCTGTAGC AAAGCAAACA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
-AAGGGATGAA AGTTGCGAGA CCACCTTCCG GCTCGCGCCT TCCGCTTGCG TTGGTTATTG-
-CCTGATAAATC TGGCACCCGG GACGTCGTTG CTCCGCGCTAT CATTGCAGCA GTGGGGCCAG-
-ATGTTAGCC CTCCCGTATCT GTAGTTATCT ACACGAGGG GAGTCGCGCA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
-AACGGAAATAG ACAGATGCCT GAGATAGGCG CCTACGTGAT TRAGCATTG GAACCTGAGG-
-TCTGATAAATC TGGCACCCGG GACGTCGTTG CTCCGCGCTAT CATTGCAGCA GTGGGGCCAG-
-ATGTTAGCC CTCCCGTATCT GTAGTTATCT ACACGAGGG GAGTCGCGCA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
-ACGGAAATAG ACAGATGCCT GAGATAGGCG CCTACGTGAT TRAGCATTG GAACCTGAGG-
-TCTGATAAATC TGGCACCCGG GACGTCGTTG CTCCGCGCTAT CATTGCAGCA GTGGGGCCAG-
-ATGTTAGCC CTCCCGTATCT GTAGTTATCT ACACGAGGG GAGTCGCGCA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
-ACGGAAATAG ACAGATGCCT GAGATAGGCG CCTACGTGAT TRAGCATTG GAACCTGAGG-
-TCTGATAAATC TGGCACCCGG GACGTCGTTG CTCCGCGCTAT CATTGCAGCA GTGGGGCCAG-
-ATGTTAGCC CTCCCGTATCT GTAGTTATCT ACACGAGGG GAGTCGCGCA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
-ACGGAAATAG ACAGATGCCT GAGATAGGCG CCTACGTGAT TRAGCATTG GAACCTGAGG-
-TCTGATAAATC TGGCACCCGG GACGTCGTTG CTCCGCGCTAT CATTGCAGCA GTGGGGCCAG-
-ATGTTAGCC CTCCCGTATCT GTAGTTATCT ACACGAGGG GAGTCGCGCA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
-ACGGAAATAG ACAGATGCCT GAGATAGGCG CCTACGTGAT TRAGCATTG GAACCTGAGG-
-TCTGATAAATC TGGCACCCGG GACGTCGTTG CTCCGCGCTAT CATTGCAGCA GTGGGGCCAG-
-ATGTTAGCC CTCCCGTATCT GTAGTTATCT ACACGAGGG GAGTCGCGCA ACTATGGATG-
-ACACTCGCCG CAACTTACTT CTGACAAGCA TGAGGAGGCGA GAGGAGCTA ACAGCTTTTT-
Note: The single underlined region is the L2Hs element.

The double underlined regions are direct repeats found at the deletion boundaries of p5JC10 and of p14JC2.
Recombinant plasmid pN6 3453 bp

BASE COUNT:  965 A  806 C  762 G  920 T

1  11  21  31  41  51
5'-AGCGCCCAAT ACGCAACCG CCTCTCCCC GCGCTTGGCC GATTCATTA TGAGCGTTCG-
61  71  81  91  101  111
-ACGACAGTT TCCCCACTG AAAGCGGGCA GTGAGCGCAG CGCAATTAAT GTGAGTTAGC-
121  131  141  151  161  171
-TCACTCATTAA GGCACCACAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTCGAA-
181  191  201  211  221  231
-ATGGACGGG ATAAACATTAT CACACAGGAA ACAGCTATGA CCATGATTAC GCAAGCTTCT-
241  251  261  271  281  291
-AATAGGACCT AATAGGGA AAAGTTGCGT GCCCTGAGGG CTGCTCTAGA GGATCCCAGG-
301  311  321  331  341  351
-GTACCTTGGAT ATTTATATAT CGGATATAGAA TATATTATTT GCTGAACATTT GATATTATAT-
361  371  381  391  401  411
-GTACAGAATA AAATATATAT TTGATGACT TTATAATTTT ATGACGCGA TATATATTAT-
421  431  441  451  461  471
-GCTTTGGTATTTTTTGTATTTAT TTGATGACG ATGAAATATT ATACCTTGGG TACTTTGATA-
481  491  501  511  521  531
-ATTTTAGGATC ATGATATAT ATGATATTTTG AGAAGCTTGG ATTTGATTTT ATCAGATAT-
541  551  561  571  581  591
-ATATATATTTC GGGGTACTTT TGATATTTTG ATGAGACAGT ATATATTATTTGAT-
601  611  621  631  641  651
-TACTTATATA TATATATAC TATATAAAT CAGATTATAG ATGATATTTT ATGACGCTAT-
661  671  681  691  701  711
-CATCGTATAT ATGTTATT ATGTTATTGT TATATATTTT ATGACGCTAT ATATATTAT-
721  731  741  751  761  771
-TAGTTTAGGATC ATGAAATTTT ATGACATATG ATGATATAA TTTGATATTATTTATTTT-
781  791  801  811  821  831
-ATGACGAT ATGATTAT ATTTGATGATG ATGATATTAT ATGACGCTAT ATATATTATTT-
841  851  861  871  881  891
-ATATGATTG ATATTTACTTT TATGATATAAT ATGATTGTTTG GTAGGGAGCT-
901  911  921  931  941  951
-CGAATTCA TGGCGTTGGTT TTTAAACGCTG TGACTTGGAA AAACCTTGGC GTTACCACA-
961  971  981  991  1001  1011
-CTAACGCTG TATGTTTAT CCCCCTTTTTCC CACGCTCAGGTA TATAGGGAA GAGGGCCACG-
1021  1031  1041  1051  1061  1071
-CCGATCGGGT TTTCCAAACG TTTGCGCGCC TGATCGGCGA ATGGAGCCCG CCTGCTAGCG-
1081  1091  1101  1111  1121  1131
-GCCGATTTAG TGCAGGGAGG GGTGGTGTAGA CGCGCAGC ATGGCGCTGCA CTTGCCAGC-
1141  1151  1161  1171  1181  1191
-CCCTAGCCG CGCTCCTTCT TTTTCTTTT TTCCCCTTCT CGCCAGTGCT GCGGGCCTCT-
1201  1211  1221  1231  1241  1251
-CCCGTCAAGC TCTAAATCGG GGGCTCTCTT TTAGGTTCCG ATTAGTGTG TTACGCACC-
1261  1271  1281  1291  1301  1311
-TGGACCCAAA AAAACCTTGGTATG GATCAGCTAG TGAGCCTACG CCGGACTAGA-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
-CGGTTTTTCCG CCCCCTGAGC TGGGAGTCCCA CTTTTTCTTTATAGTGACTC TTGTTCCAAAG-
-1381
1391 1401 1411 1421 1431
-CTGGAAACAC ACTCAAACCCT ATCTCGGTCT ATTTTTGTTA TTATAAGGG ATTTTGCGCAAG-
1441
1451 1461 1471 1481 1491
-TTTCGGCCTA TTTGTTAAAA AATGAGCTGA TTAAACAAAA ATTTAAGGCC AATTTTAAAC-
1501
1511 1521 1531 1541 1551
-AAATATTAC TTTTACATT TCAGGTTGCA TTCTTTGGGG AAATGTGCGG GAACACCCCTA-
1561
1571 1581 1591 1601 1611
-TTTGTATT TTTTCACATA CTATTACACGT CATGAGACAA TAACCCCTGAT-
1621
1631 1641 1651 1661 1671
-AAATGCTTCA ATAAATATTGAA AAAAAAGAAGA GTATAGATGTC TCAACATTTC CTGTGCCC-
1681
1691 1701 1711 1721 1731
-TTATTTCTTT TTCTGAGCGCA TTCTTTGCCGT GCTTTTTTTGCA TACCCAGAAA ACAGCGGTGAG-
1741
1751 1761 1771 1781 1791
-AGTAGAAGAGA TGGAGAGAGAT CAGTTGCGGGA CAGAGGGCCG TTAATACGGAT-
1801
1811 1821 1831 1841 1851
-ACAGCGTAA AATCTTTGCCG AGTTTTCCGCC CGAAGAAGC TTTTCAATGGTGAGACTCTT-
1861
1871 1881 1891 1901 1911
-TTAGATCTCT GTAATGTGCC GGGCGTATTAG CCCGTATTTGA CGCCGGGCAA GAGCAACTCG-
1921
1931 1941 1951 1961 1971
-GTGGCGCGCAT ACACTATCTCT CAGAATGACT CTGGTTGAGTA CTCACCAGTC ACAGAAAAGC-
1981
1991 2001 2011 2021 2031
-ATCTTTACGGA TGGCATGAGA TATAGAGATGTC TGGGAAAGAG ATGAGACTTGAGA-
2041
2051 2061 2071 2081 2091
-ACACTGCGGCC CAACACTTATT CTGACAACGGA TCGGAGGACC GAAAGGACTA ACCGCTTTTT-
2101
2111 2121 2131 2141 2151
-TGCACACAT GGGGATCAT GTAAGAGTGCA TTGATCGGTTTA GGAACCGGAG CTGAATGAAG-
2161
2171 2181 2191 2201 2211
-CCATAACAAA CGACGGAGGT GACACCAAGA TGCCTGTGCC AACAGCGCAAC ACGTTGCGCA-
2221
2231 2241 2251 2261 2271
-AACTATTAC TGGCGGACTCTG CTTACTTTAG CTTTCCCGGCA ACAAATTTATAGCAGGTGGAG-
2291
2291 2301 2311 2321 2331
-AGGCGGATAA AGTTCGCGGA CCAACCTTGTG GCTGGCGCCTT TCGGGCTGCG TGGTTTATTG-
2341
2351 2361 2371 2381 2391
-CTGATAAACAT TGGAGGCTGG GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG-
2401
2411 2421 2431 2441 2451
-ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACAGACGGGG GAGTCAGGCA ACTATGGATG-
2461
2471 2481 2491 2501 2511
-AACGAAATTAG ACAGATCGCTT GAGATAGGCT CCTACTTATG TACAGATTTG TAACTCGAG-
2521
2531 2541 2551 2561 2571
-ACCAAGTTC ATCTATATATA CTAGGAGATG TTAAACACCT TCAATTATTA TTTAAAGAAAAG-
2581
2591 2601 2611 2621 2631
-TCTAGTGAAA GATGCCCTTTT GTAATTTCAT AGACCAGGG TCTGAAGCAAT CACACAGGTTCG-
2641
2651 2661 2671 2681 2691
-CTGACCGACGT CAGAGGCCTAGA AAGAGGATGTC TCAAGAGGATC TCTCTTTGAGA CTTTTTTTC-
2701
2711 2721 2731 2741 2751
-TGCGCGTAAAT CTTGCTTTGG CAAACAAAAGA ACCACCGGCT ACCACCGGAG CTTTGGTTTCG-
2761
2771 2781 2791 2801 2811
-CGGATCAAGA GCTTACACACT CTTTTCCGGG AGGTAACCTGT CTCAGCGAGA GCGCAGTATC-
Note: The underlined region is the L2Hs element.
APPENDIX D

PTZ19U  2863 bp DNA
BASE COUNT:  701 A  722 C  709 G  731 T

1  11  21  31  41  51
5' -AGCGCCCAAT ACGCAAACG CCTCTC CCCG CCGTTG GCC CATTAA TGCAGCTGGC-61
   71  81  91 101 111
   -ACGCAGGTTC TCCCAGCTGAA AAGCGGGCA GTGAGCGCAA CGCAAATAT GTGAGTTAGC-
121  131  141  151  161  171
   -TCACCTATTG GCCACCCCA GCTTTACACT TTAGCTTCCC GGTCTGTATG TTGTTGAGA-
181  191  201  211  221  231
   -TTGTGAGCCG ATAAACAATTT CACACAGGA ACAACTATGA CATGATTAC GCAAGCTCTC-
241  251  261  271  281  291
   -AATACGACTC ACTATAGGGA AAGCTTGCAT GCCCTGAGGT CGACTCTAGA GGATCCCCGG-
301  311  321  331  341  351
   -GTACCGAGCT CGAATTCACACT GCGCGTCGTT TTCAACAGCT GTGACTGGGA AAACCTGGC-
361  371  381  391  401  411
   -GTTACCCACG TTAATCGCCT TCGCACCACT CCCCTTTCCG CGACTGGGCG TAAATAGC-
421  431  441  451  461  471
   -GAGGCCCCGCA CGGAGCCGCC TTTCCCAAGC TTTGCCACCC GGAATGGCG AAGGGACCGG-
481  491  501  511  521  531
   -CCCTGTAGGCC CGCCATAGAA AGCCGCAGGGT GCTGGTGTTA CGGCAGCGT GACCCCTA-
541  551  561  571  581  591
   -CTTGGCCACCG CCGCCAGCCC CGGCTCTTTC GGGTTCTTCC CGGCGACTC CACTCCAG-
601  611  621  631  641  651
   -GCGGTCTTTC CCGGCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT-
661  671  681  691  701  711
   -TTACGCCCACC TCGACCCCAA AAAAACTTGA TAGGGTGATG GTCAGCGTAG TGGCCCCATC-
721  731  741  751  761  771
   -CCCTGTAGAGCC CGCCATAGAA AGCCGCAGGGT GCTGGTGTTA CGGCAGCGT GACCCCTA-
781  791  801  811  821  831
   -TGTGTCTAAA CTGGAAACAC ACTCAACCTT ATCTCGGTCT ATCTTATGGA TTTATAAGGG-
841  851  861  871  881  891
   -ATTGTGCGCA TTGTGGCTTTA GCTGTTAAAA ATAGGACCTA ATTAACAAAA ATTTAAGC-
901  911  921  931  941  951
   -AATTTTAACAA AAATATTACCA TTTACAAATT TCAAGTGCGA CTTTCCCGGG AATTGTTCC-
961  971  981  991 1001 1011
   -GGAAACCTAAT TTTGTATTATT TTCTACAAATA CATTTACAA TGTATCCAGC TGGAGACA-
1021 1031 1041 1051 1061 1071
   -TACCCCATGAA ATATGCCCTA ATATATATTGAA AAAAGGAGA GTATGGAAT CATGAACAA-
1081 1091 1101 1111 1121 1131
   -CGTGTGCCC TTAACCCCTT TTTGCGGGCA TTGTGGCTTC CGGTTTTGGC TCAACCTA-
1141 1151 1161 1171 1181 1191
   -ACGCTCCTGA AAGTTAAAGA TGCTGAAGAT CAGTTGGTTC TACAGTGAGA TACATCGG-
1201 1211 1221 1231 1241 1251
   -CGTGTCTCAAC AGACCGTGTTA ATACCTTGAG ATGTTTTCGCC CGCAAAGA CTTTTCAATT-
1261 1271 1281 1291 1301 1311
   -ATGAGGCTTT TTAAGTCTTC GCTATGGGCC CGCCTATTAT CCCGGTGCAA CAGAGCGG-
1321 1331 1341 1351 1361 1371
   -GAGCAGCTCG GTCGCGCCCAT ACACATTTCT CAGAATGCT CGGTTGAGT AATCCAGATC-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
-ACAGAAAGC ATCTTACGGA TGCCATGACA GTAAAGAAT TATGCAGTGC TGCCATAACC-
1441 1451 1461 1471 1481 1491
-ATGAGTGATA ACRTGCGGG CAACTTACTT CTGCAACAGA TGAGGAGCTA-
1501 1511 1521 1531 1541 1551
-ACCGCTTTTT TGCAACAATC GGGGATCAT GTAACTCGGC TTGATCTGTG GGAACCGGAG-
1561 1571 1581 1591 1601 1611
-CTGAATGAAG CCAATACCAAA CGACGAGCGT GACACCAGA TGCCCTGATG AATGGCAACA-
1621 1631 1641 1651 1661 1671
-ACGTGCCGCA AACTATTAC TGGCAGAATA CTTACTCTAG CTTCGCCGCA ACAATTAATA-
1681 1691 1701 1711 1721 1731
-GACTGGAATGG AGGCAGGATA AGTTGCAAGA CCACCTTCTGC GCTCCGGGCT-
1741 1751 1761 1771 1781 1791
-TGGTTATTTG CTGATAAACAT TGAGCGGCTT GAGCTGATT GCTTGCGAAGCA-
1801 1811 1821 1831 1841 1851
-CTGGGGCGAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGAGGG GAGTCAGGCA-
1861 1871 1881 1891 1901 1911
-ACATATGAGT AAGCAATAGT ACAATCGCCG GATATGGTG CCTCAGTGAT TAAGCATGGA-
1921 1931 1941 1951 1961 1971
-TAACGTGCAAC ACCAGTTATTA CTTATATATA CTTTACAGTT ATTTAAGCCT CTTTATATTA-
-TTTAAAAGAGA CTGTAGGAGA GACTCTTCTT GATAATCTCA TGAAACAAAT CTTTAACTGT-
2041 2051 2061 2071 2081 2091
-GAGTTATTCGT CCTCAGTGACG TCTCGAGACT GATAGAAAAGA TCAAGAAGTG GATAGGGCA-
2101 2111 2121 2131 2141 2151
-CCTTTTTTCT TGCCGCCTAT CTGCTGCTTTG CAAACAAAAA AACCCACGCT ACCAGCGGT-
2161 2171 2181 2191 2201 2211
-GTTTTTTCG CCGATCAAGA GCTACACTCT CTTCCTCGGA AGTAACTCTG CTTCAGCAGA-
2221 2231 2241 2251 2261 2271
-GCGCAGATAC CAATAGACTG CCTCTCTCGT TAGCTGATT GACTGCCACCA CTTCAAGAC-
2281 2291 2301 2311 2321 2331
-TCTTGAGCAC GCCTACATA CTCCGCTCTG CTAAATGTGT TACCCGTTGC TGCTGGCAAT-
2341 2351 2361 2371 2381 2391
-GGCCATAAGGT CTGTCTACGC CCGCGTGAGA TCGAGGCTG AAGAGGCGCA TAAAGGCGCA-
2401 2411 2421 2431 2441 2451
-CGGTGGGGCT CAGAAAACGGG TTGGGTACTC CAGCCGGAGC GACCTACACCCT-
2461 2471 2481 2491 2501 2511
-GAACAGATAC ACCTACGGG TGAGGATGG AAGACGCGCA CGCTTCCCGA AGGGGAAAG-
2521 2531 2541 2551 2561 2571
-CCGGACAGGT ATCCGTTAAG CGGCAGGCAT GAAACAGGAG AGCCGAGGG GAGGCTTCCA-
2581 2591 2601 2611 2621 2631
-GGGGAAACGC CCTGACTATTT TTATATGCCT GTGGCTGAGG AGAACGCTC AAGGCTTCCA-
2641 2651 2661 2671 2681 2691
-CGATTTTGTG CAGCAGCTGTTT GGGGAAAAGG AGCTTATGGA AAACAGCCGCA ACACGCGG-
2701 2711 2721 2731 2741 2751
-TTTTTACGTT TCTGCTGTCT TTGCTCGCT TGGTCTGACG TGCTTTCCGA TGCGTATC-
2761 2771 2781 2791 2801 2811
-CCTGAATCTG TTGATAACCC TATTACCAGT TTGAGTGAG CTGATAACCAC TGCGCAGCGC-
2821 2831 2841 2851 2861
-GAACAGACGG AGCGAGGCGA GTCAGTGGC GAGGAGCGAG AAG-3'
APPENDIX E

L2Hs element 595 bp

\begin{verbatim}
1 11 21 31 41 51
5' -GTTACCTTGA TATTTTATGT ACAGTATAGA ATATACTATT TGCTGAACCT TGATATTTTA
3' -CCATGGAACT ATAAAAATACA TGTCATATCT TATAGTAAT ACCACTTGGA ACTATAAAAAT

61 71 81 91 101 111
- TGTACAGAAT AAAATATATA TTTGTAGTAC TTTTATATT TATATGACG ATATATATA
- ACAATGCTTTA TTTTATATAT AAAACTACATG AAAGTATAAA ATACATGCAG TATATATAT

121 131 141 151 161 171
- TGCTTTGGGT ACTTTGATAT TTTTTATACA GATATGAAATA ATACACTTTG GTATCTTTGAT
- ACGAAACCCCA TGAAACTATA AAAACACGTG CATACTTTAT ATATGGAACC CATGAAACTA

181 191 201 211 221 231
- ATTTATAGTG CAGTATATAA TATAATATTT GAGAACCTTG ATATTTCTAT TACAGATATA
- TAAATAACAC GTCTATATCT ATATATAA ACGTGAACCT TATAAAGTAC ATATCATATT

241 251 261 271 281 291
- AATATATATT TCGGGGTACT TTTGTATTTT AGTGTACAGG TATATATATC TATATTTGAT
- TTAATATATAT AGCCCCATGA AAACCACTAA TGTCATATGC ATATTTGAG ATATAACTA

301 311 321 331 341 351
- GTACTTTTCT ATTTACATAC CTAATTATAA TACATATATT GATGTATTTC ATATATATTG
- CATGAAAGTA TAAATGTATG GATTATATTT ATGTGTAATAT ATACATAAAG TATAAATACA

361 371 381 391 401 411
- ATACGTTTATA TATATGATAG TTTGCTGTTAC TACCATATTG TATATGATCAG TATATAATAT
- GTACGTATAT ATTATGACGG AAAACCATTA GTACATATTT ATATGGAACC CATGAAACTA

421 431 441 451 461 471
- ATCTTTTCTG TICATATATAT TATATAGTTT GAGAATTTTG ATAGTATATAT TATATATATT
- ATATGACGTT GTACATATTT GAATTTATCA TATATATATG TATATATATG

481 491 501 511 521 531
- GTATATATAT ATATATATAT TATATATATAT GCATATATAT ATATATATATG TATATATAT
- TATAGTAGTG TATATATATAT TATATATATAT TATATATATAT TATATATATAT

541 551 561 571 581 591
- TATGGATG TAAATTATCTA TTTATATGAC AGTATATAAT GTATAGTATT GGAGATCC-3' 
- ATATACGTCT ATATATATAT TATATATATAT TATATATATAT TATATATATAT
\end{verbatim}

The \textit{KpnI} restriction site is underlined. This 595 bp sequence includes the uncleaved, entire \textit{KpnI} restriction site.
VITA
VITA

Eugenia Lee Posey

Personal Data: Place of Birth: Atlanta, Georgia

Education:
Public Schools, Fulton and DeKalb County Schools: Atlanta and Decatur, Georgia

Wesleyan College, Macon, Georgia; piano performance, B.M., 1976

University of Hartford, West Hartford, Connecticut; piano and piano teaching, M.M., 1981

Graduate and Post-baccalaureate studies (no degree):
Emory University, Atlanta, Georgia (summer 1976); German and aesthetics in Vienna, Austria

Georgia State University, Atlanta, Georgia (1976-1977): Graduate work in piano performance and music history

Georgia State University, Atlanta, Georgia (1988-1990): Undergraduate work in science and mathematics


Professional Experience:
Studio accompanist for Peter Harrower, Georgia State University, Atlanta, Georgia, 1976-1977.


Substitute teacher, DeKalb County Public Schools, Cardston, Georgia, 1982

Piano teacher, Jim Scott Organ and Piano Studios, Atlanta, Georgia, 1982-1987

Piano teacher, The Gwent School of Music, Lilburn, Georgia, 1983-1989

Piano teacher, The Piano Studio, Atlanta, Georgia, 1983-1988

Organist and choral accompanist at churches in Connecticut, Georgia and Tennessee, 1976-1997

Graduate Assistant, East Tennessee State University, Department of Biochemistry and Molecular Biology, 1992-1996

Tutor in Biochemistry, East Tennessee State University, Department of Biochemistry and Molecular Biology, 1993-1996

Publications:
Abstracts:


Honors and Awards:

Magna cum Laude, Wesleyan College 1976
Dean’s List 1972-1976 Wesleyan College
STUNT Scholarship, Wesleyan College
Susan Martin Catchings Award at Wesleyan Commencement
Freshman Chemistry Award, Georgia State University 1988
Life member of Golden Key National Honor Society, Georgia State University 1989
Member of Gamma Beta Phi Honor Society, East Tennessee State University 1994-present