12-2008

Site Directed Mutagensis of Bacteriophage HK639 and Identification of Its Integration Site

Madhuri Jonnalagadda
Western Kentucky University

SITE DIRECTED MUTAGENESIS OF BACTERIOPHAGE HK639 AND IDENTIFICATION OF ITS INTEGRATION SITE

A Thesis presented to

The faculty of the Department of Biology

Western Kentucky University

Bowling Green, Kentucky.

In Partial Fulfillment

Of the Requirement for the Degree

Master of Science

By

Madhuri Jonnalagadda

December, 2008

Site Directed Mutagenesis of Bacteriophage HK639 and Identification of its Integration Site

Date Recommended 12/4/08

Dr.Rodney A. King

Director of Thesis

Dr.Jeffery Marcus

Dr. Sigrid Jacobshagen

_____

Dean, Graduate Studies and Research Date

DEDICATION

I would like to dedicate my thesis to my dear father Dr. J. Padmanabha Rao, my mother, Dr.V. Lakshmi, my sister Dr. J. Sushma for their support and encouragement.

ACKNOWLEDGEMENTS

I wish to express my heartfelt gratitude to my graduate committee in the Department of Biology, Western Kentucky University.

Primarily, I am thankful to my advisor, Dr. Rodney A. King for his encouragement and support which kept me going through the past year and who through his tireless efforts provided me with academic assistance and funding for my research. Thank you very much for everything you have done for me, without which I could never have completed this project. I am thankful to my committee members Dr. Jacobshagen and Dr. Jeffery Marcus for their help. I would also like to thank the Dean of Graduate Studies, Dr. Richard Bowker and all the departmental staff for their support. I also want to thank Bridget A. Sutton, Donna M. Kridelbaugh and Ali Wright for their help.

A special thanks to my fiancé Sandeep for his constant support and encouragement.

TABLE OF CONTENTS

Table of contents **iii**

List of Figures **v**

List of Tables **vii**

Abstract **viii**

Introduction 3

Materials and Methods

Bacterial Growth Media, Biochemicals, Oligonucleotides, Restriction enzymes

And Antibiotics 20

Isolation of genomic DNA 20

Plasmid DNA Isolation 21

Restriction Digestion 21

PCR Reaction Conditions 21

In vitro transposition of EZ-Tn5 <R6K*γori*/KAN-2> transposon into HK639 22

In vitro transposition of EZ-Tn5 <R6K*γori*/KAN-2> transposon into pUC19 23

Incorporation of restriction endonuclease sites that flank the Kanamycin Sequence 23

Cloning of Kanamycin gene into pRK729 23

Cloning of Kanamycin gene with flanking HK639 homology into plasmid

pML042 24

Inverse PCR 24

Recombineering the Kanamycin gene into RK1145 25

Restriction Digest of pMLO42 25

Gel Purification of EcoRI Fragment of pMLO42 25

Preparation of Electrocompetant Cells 26

Preparation of Electrocompetent Cells for Recombineering 26

Generation of RK1157 lysogen 27

Identification of Integration Site of HK639 lysogen by plasmid rescue 28

Dephosphorylation of DNA 28

Purification of DNA By Microfiltration 29

Sequencing reactions 29

Purification of DNA Sequencing Reactions 29

Purification of PCR Products 29

Results

Isolation of Genomic DNA from an HK639 lysogen 33

Identification of the HK639 integration site by Inverse PCR 36

Mutagenesis of HK639 38

Cloning of Kanamycin gene into pRK729 44

Site Directed Mutagenesis of HK639 prophage 48

Identification of Integration Site of HK639 lysogen by plasmid rescue 52

Discussion **56**

References **58**

LIST OF FIGURES

Figure 1. The Lysogenic Cycle 4

Figure 2. The Lytic Cycle 5

Figure 3. Lambda transcriptional unit 9

Figure 4. Similarity between the right and left put sites of HK022 13

Figure 5. Sequence conservation of the rpoC zinc binding domain 14

Figure 6. Comparison of the genetic organization of HK022 and Lambda 15

Figure 7. A cartoon depiction of the HK639 genome 16

Figure 8. Genomic DNA isolated from strain RK1033 34

Figure 9. Gel electrophoresis of PCR products amplified from an HK639 lysogen 35

Figure 10. Schematic representation of Inverse PCR 37

Figure 11. Plasmid vector pUC19 40

Figure 12. Cartoon of R6Kγ*ori*/KAN-2 transposon 41

Figure 13. Sequence junction between pUC19 & the R6Kγ*ori*/KAN-2 transposon 42

Figure 14. A cartoon depiction of pRK1132 43

Figure 15. A cartoon depiction of the plasmid pRK729 45

Figure 16. Confirmation of cloning of Kanamycin gene into pRK729 46

Figure 17. A diagrammatic representation of the plasmid pRK1142 47

Figure 18 A. Gel confirming the presence of the R6KoriKan Transposon sequences on the HK639 prophage 50

Figure 18 B. Gel confirming the presence of the R6KoriKan Transposon sequences on the HK639 prophage 51

Figure.19. Sequence showing the location of the junction 53

Figure 20. Representation of the insertion site in *E. coli* genome 54

Figure 21. Gel confirming the integration sites of HK639 in *E.coli* chromosome 55

LIST OF TABLES

Table 1. Bacteria, Phages and Plasmids 18

Table 2. Primers 31

SITE DIRECTED MUTAGENESIS OF BACTERIOPHAGE HK639 AND IDENTIFICATION OF ITS INTEGRATION SITE

Name: Madhuri Jonnalagadda Date: December, 2008. Pages 60

Directed by: Rodney King, Sigrid Jacobshagen, and Jeffery Marcus

Department of Biology Western Kentucky University

ABSTRACT

Bacteriophages affect bacterial evolution, pathogenesis and global nutrient cycling. They are also the most numerous and diverse group of biological entities on the planet [1, 2, 3, 4, 5, 6]. Members of the Lambda phage family share a similar genetic organization and control early gene expression by suppressing transcription, a process known as antitermination. Transcription antitermination in Lambda is mediated by a phage-encoded protein whereas in lambdoid phage HK022, antitermination is mediated by a phage-encoded RNA molecules. Recent results suggest that another bacteriophage called HK639 also appears to use RNA-mediated antitermination. To characterize this newly identified phage we generated site directed mutations and identified where the phage integrates into the chromosome of its bacterial host.

Introduction

Bacteriophages are viruses that infect bacteria. They were independently discovered more than 80 years ago by Frederick Twort (1915) and Felix d'Herelle (1917) [7]. There are two basic types of bacteriophages: temperate phages and lytic phages. Temperate phages establish a long-term genetic association with their host, typically by inserting their genome directly into the host chromosome by a process known as integration. An integrated phage is known as prophage. The process by which the phage integrates its entire sequence into the host chromosome is known as lysogeny (Figure 1). Temperate phages usually do not cause the bacterial cell to lyse [8, 9] because most of the prophage genes are repressed. Repression is mediated by a phage encoded DNA binding protein. This trans-acting protein confers immunity to the lysogenic cell, because it can repress gene expression from a super-infecting phage of the same type.

Lytic phages do not integrate into host chromosomes. Infection by a lytic phage results in the production of additional virus particles that are released when the host bacterium is lysed. The process by which a phage infects a bacterium and propagates by lysis of the host bacterial cell is known as lytic infection (Figure 2).

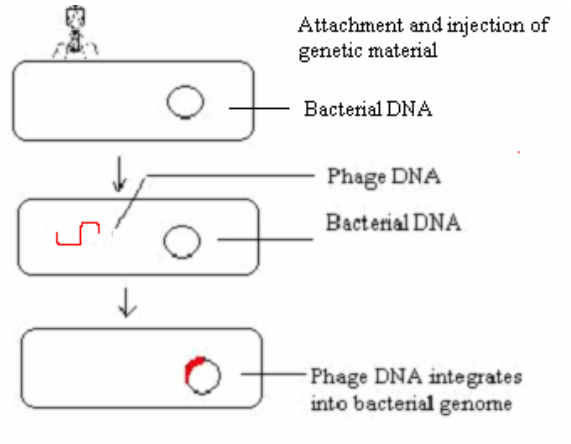

Figure.1. The lysogenic cycle. The figure represents the infection cycle of a lysogenic phage. The phage attaches to the bacterial cell and injects its DNA into the host. Recombination occurs between a specific site on the phage DNA and a specific site on the bacterial chromosome. This process is catalyzed by phage- and host-encoded proteins and is called site-specific recombination. The phage DNA is replicated along with the host chromosome and is passed on to the daughter cells at the end of each replicative cycle.

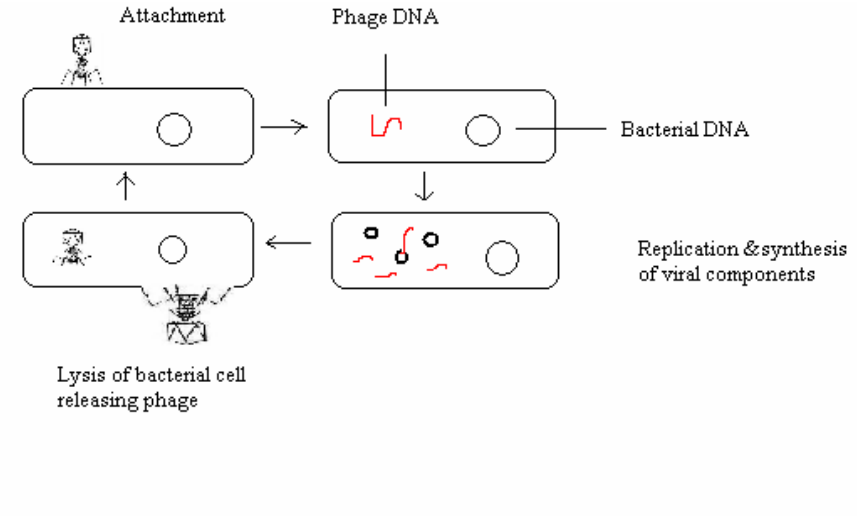

Figure.2. The Lytic Cycle. Figure represents the lytic cycle of phage infection. The phage attaches to the bacterial cell and injects its DNA into the cell. This is followed by replication and synthesis of viral components and finally the lysis of the cell to release the progeny phage.

Gene Expression

The central enzyme in gene expression is RNA polymerase. The *Echerichia coli* core enzyme consists of four distinct subunits: $\alpha_2\beta\beta'\omega$. The holoenzyme is generated when the core enzyme associates with a fifth subunit called sigma (σ)**.** The sigma factor is necessary for promoter recognition. There are different types of σ factors in *E.coli* such as σ^{70}, σ^{32}, σ^{28}, and σ^{54}. The σ^{70} subunit is the "house keeping" sigma factor and directs the transcription of most of the genes. The β and β' subunits constitute the catalytic center of the enzyme and the alpha subunits are responsible for promoting the assembly of the enzyme. The alpha subunits are also involved in the interaction between RNAP and some regulatory factors [10]. Transcription can be divided into three main stages: Initiation, Elongation and Termination [11]. Transcription initiates when RNA polymerase binds to the promoter which is located at the start of a gene. A promoter is a DNA sequence that is recognized and bound by RNA polymerase. No two promoter sequences are the same but most that are recognized by σ^{70} have the following organization: A DNA sequence (TATAAT) centered 10bp upstream of the transcription start site and another sequence (TTGACA) centered 35bp upstream of the transcription start site. The efficiency or strength of a promoter is affected by mutations in these sequences and their spacing.

After initiation, the sigma factor is released and the enzyme enters the elongation phase of transcription. During elongation, the RNA chain is extended by the rapid addition of individual nucleotides. The nucleotides are added to the 3' end of the growing RNA strand in an unwound region called the transcription bubble. Elongation

involves the processive movement of the transcription bubble along the length of the DNA template until RNA polymerase is signaled to stop [10].

Termination is the final stage of transcription where no further bases are added to the RNA chain. In *E.coli,* terminators are of two types: Intrinsic terminators and Rho dependent terminators. Intrinsic terminators are characterized by secondary structure in the nascent transcript (a hairpin and loop) followed by a run of Uracil residues. It is thought that the hairpin facilitates termination by slowing down RNA polymerase. In contrast to intrinsic terminators, Rho dependent terminators depend on the activity of the bacterial Rho protein. Rho is a helicase that binds to the *rut* site (Rho utilization site) of nascent RNA. A *rut* site is rich in Cytosine and poor in Guanine residues. Rho tracks along the RNA until it reaches RNA polymerase. It then releases the RNA transcript from the polymerase and the DNA template [10].

Termination is sometimes prevented by allowing the RNA polymerase to read through the termination site. This mechanism is called antitermination. Transcription antitermination was originally discovered in the bacteriophage lambda [12].

The Bacteriophage Lambda

Lambda is a temperate phage that infects *Escherichia coli* and it is the best-studied temperate phage. Lambda was discovered accidentally in 1951 by Esther Lederberg who crossed the lambda sensitive strain of *E.coli* K-12 [12] with its parent. The lambda genome is a linear double stranded DNA of 48,514 base pairs with unique 12 bp single stranded ends. When lambda infects a host, its DNA is circularized by joining of its cohesive ends also known as *cos* sites. These sites are essential for the phage to form infective particles. Members of the Lambda phage family share a similar genetic

organization and control early and late gene expression by suppressing transcription termination [10].

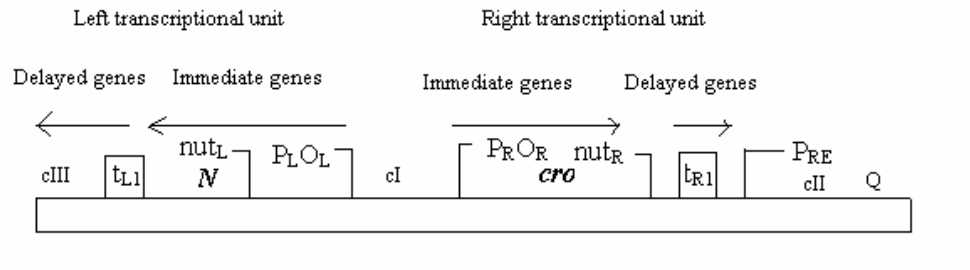

Figure.3. Lambda Transcriptional unit**.** The above figure shows the early transcriptional units of bacteriophage Lambda. The early genes are organized into two operons that are expressed from the major leftward promoter P_L and the major rightward promoter P_R. The right operon consists of the operator O_R, the promoter P_R, a rho-independent terminator t_{R1}, an N utilization site (nut_R), and the *cro* gene. The left operon consists of an operator O_L, the promoter P_L, a rho-independent terminator t_{L1}, an N utilization site (nut_L) and the N gene. Genes *N* and *cro* are immediate early genes separated from the delayed genes by the terminators. Synthesis of N protein allows RNA polymerase to read through the terminators.

Lambda uses immediate early genes and delayed early genes for both the lysogenic and lytic cycle. When a phage infects a cell, the expression of immediate early and delayed early genes is necessary regardless of the pathway. The lytic pathway occurs only if the late genes are expressed. Lambda has two immediate early genes: N and Cro (Figure 3). The N gene codes for an antitermination factor that promotes transcription of delayed genes. N binds to RNA transcript of the N utilization site or *nut* site [13]. There is a copy of the *nut* site in each of the major early operons, *nut*L and *nut*R. The RNA transcribed from the P_R or P_L promoters adopts a secondary structure that is recognized by the N protein. This RNA structure is specified by the box B sequence of the phage *nut* site. The association of N and several host-encoded proteins with the phage *nut* site converts RNA polymerase into a termination resistant form [13]. The Cro protein has two major functions. One is to prevent the synthesis of a repressor. This is a necessary action for the lytic cycle to proceed. The other major function of Cro is to turn off the expression of early genes. The delayed early genes include replication genes, recombination genes and regulator genes. Most importantly, Cro represses the P_{RM} promoter of the cI repressor. The Cro protein as well as the phage repressor protein recognize regulatory DNA sequences called operators. The operators overlap the major early promoter sequences. The late genes are expressed from the $P_{R'}$ promoter. Full expression of the late genes requires the product of the Q gene. Unlike N, Q is a DNA binding protein. It recognizes a DNA site near the $P_{R'}$ promoter. Similar to N, it associates with the ternary transcription complex and converts it to a termination resistant form.

In a lysogen, the phage encoded repressor protein binds to the phage operator sequences that overlap the major early promoters. This turns off most of the phage genes and thus the phage remains in a quiescent state. However, the prophage can be induced to enter the lytic cycle. Exposure to UV rays causes damage to a bacterial DNA and this damage activates the bacterial RecA protein. Activated RecA cleaves between the carboxyl and amino domains of the repressor and consequently the cleaved repressor dissociates from the operator sequences [7]. This allows polymerase to initiate transcription at the P_L and P_R promoters. The first Cro protein that is synthesized binds to the O_{R3} site. This prevents the polymerase from binding to P_{RM} and prevents further synthesis of the repressor.

Bacteriophage HK022

Bacteriophage HK022 is related to Lambda because it shares a similar genetic organization (Figure 6) and crosses between HK022 and Lambda form viable hybrids. HK022 also forms lysogens by inserting its genome into the host chromosome.
As outlined above, antitermination in most lambdoid phages is mediated by a phage-encoded protein. However, antitermination in the lambdoid phage HK022 is mediated by RNA molecules. These antiterminator RNAs are located in the early operons and their association with RNA polymerase converts the enzyme into a termination resistant form and thus facilitates transcription of downstream genes [14]. The antiterminator RNAs are encoded by the *put*L (polymerase utilization site) and *put*R sites. These sites are 70bp long. The *put*L site is located at the beginning of the P_L transcript and *put*R is located 280 bp downstream of P_R. Each of the put transcripts forms two stem loops separated by an unpaired base (Figure 4). Mutations that prevent the formation of the stem loops

prevent antitermination. Second site mutations, which restore base pairing but not the original sequence, were shown to restore antitermination [14]. This proves that the secondary structure is important for the antitermination mechanism to occur in HK022. Host mutants that prevent HK022 growth (antitermination) were identified by Clerget et al [15]. These single amino acid substitutions occur in a highly conserved domain of *E. coli* RNA polymerase (Figure 5).

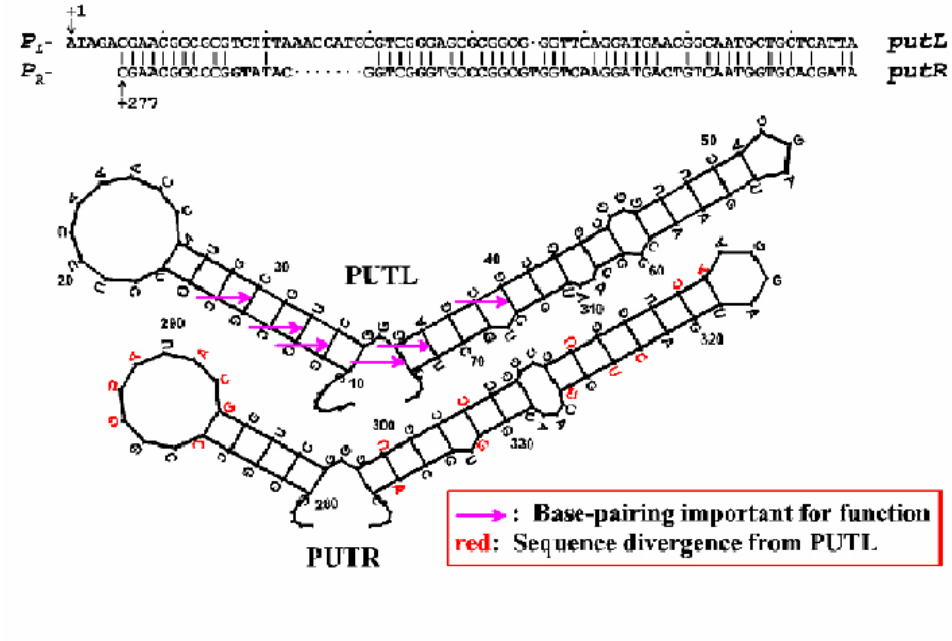

Figure.4. Similarity between the right and left *put* sites of HK022. *Put* sites encode antiterminator RNAs in the genome of the lambdoid bacteriophage HK022.

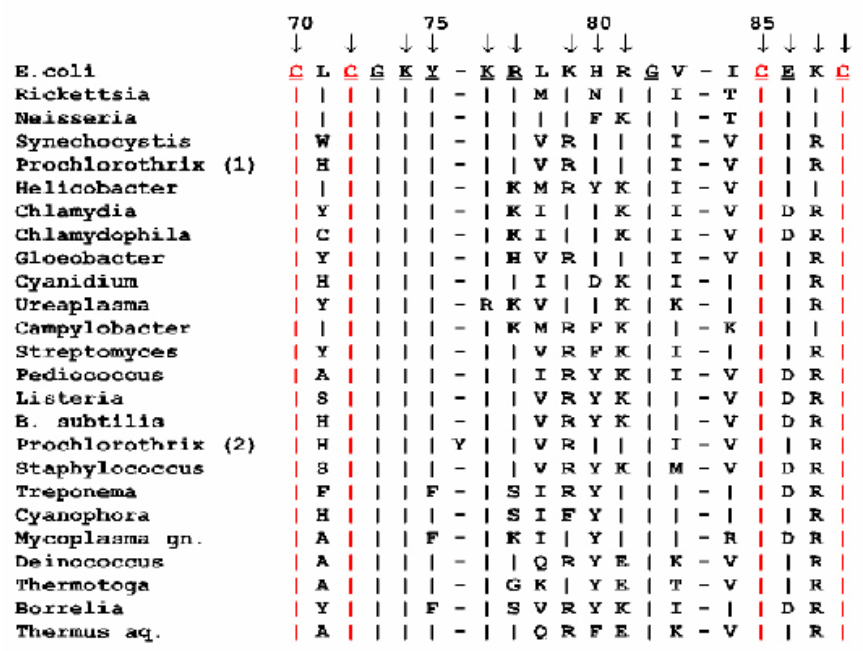

Figure.5. Sequence conservation of the *rpo*C zinc binding domain. The substitution of Tyrosine at position 75 with Asparagine blocks growth of HK022 and HK639.

King et al recently performed a screen to identify additional phages that use RNA antiterminators (Rodney King, personal communication). The screen was based on the differential growth of phages on a wild type *E .coli* host versus a host that carries the *rpo*CY75N mutation. A new bacteriophage, HK639, was identified that was incapable of growing on the *rpo*CY75N mutant strain.

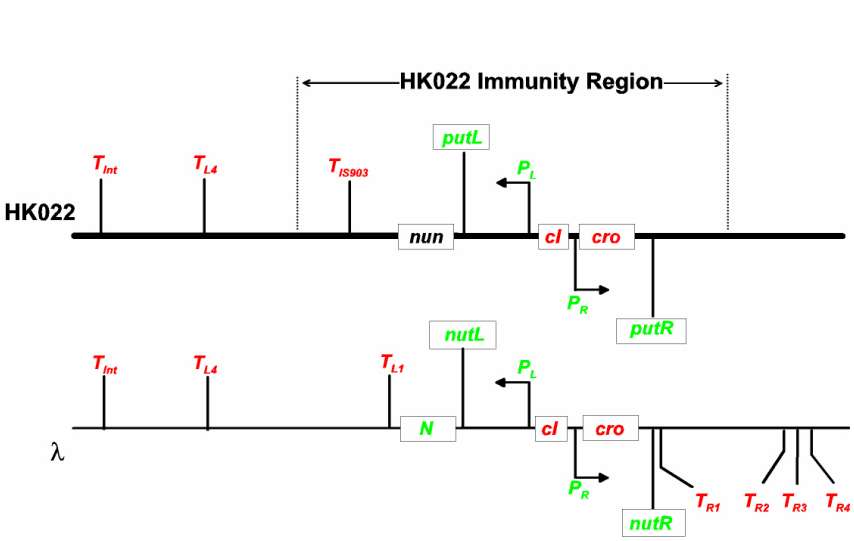

Figure.6. A comparison of the genetic organization of bacteriophages HK022 and Lambda. Polymerase utilization sites (*put* sites) of HK022 and the *nut* (N utilization) sites of Lambda are shown. The phage repressor genes, *cro* and *cI*, the promoters (P_L and P_R) and known terminators (red) are shown. The immunity region indicates all the genetic information necessary to confer immunity on a lysogenic cell.

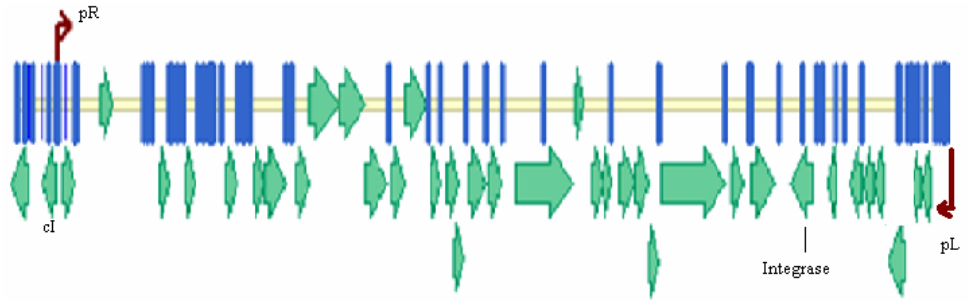

Figure.7. A cartoon depiction of the HK639 genome. Green arrows represent the transcripts and their predicted direction. Blue lines represent the smaller genes and predicted regulatory sites. pR represents the right promoter of HK639 and pL the left promoter. The location of the integrase gene and the repressor protein (cI) gene are shown.

Experimental Goals

Lambda has served as a paradigm for understanding gene regulation for over 50 years. In Lambda, the phage-encoded N proteins together with several host encoded factors promote full expression of the phage genes by suppressing transcription termination. In contrast, antitermination in HK022 is mediated by specifically structured RNA molecules. This unique mechanism of genetic control is the focus of the research in our laboratory. Recently, a newly discovered phage, bacteriophage HK639, was found to use RNA-mediated antitermination. The genomic sequence of this phage has been determined [16] (Figure 7). To further characterize this phage and to get a better understanding of its biology we have generated mutations in the phage to see how they would affect phage growth and we have identified the location of the HK639 prophage in the host chromosome.

TABLE.1. Bacteria, Phages and Plasmids

Strain	Relevant genotype
Bacteria	
RK898	MG1655 *rpo*C+ wild type [17]
RK1033	Strain RK638 MG1655 *rpo*C btuB::Tn10KanR (this work)
RK1145	MC1000+HK639 prophage with plasmid with pRK1006 (from this study)
RK1157	RK1145 recombineered with PCR products amplified from RK1142 (from this study).
LE392	Amber suppressing strain used for the propagation of lambda vector and its recombinants. It is a derivative of ED8654. *gln*V44 *sup*F58 (*lac*Y1 or Δ*lac*ZY) *gal*K2 *gal*T22 *met*B1 *trp*R55 dR514(r_K^- m_K^+) [18].
XL1	Recombination deficient strain that supports the growth of vectors carrying some amber mutations but not the Sam100 mutation. *end*A1 *gyr*A96(nal^R) *thi*-1 *rec*A1 *rel*A1 *lac gln*V44 F'[::Tn10 *pro*AB^+ $lacI^q$ Δ(lacZ)M15] *hsd*R17(r_K^- m_K^+)[Stratagene]
Plasmids	
pUC19	Multicopy plasmid of *E.coli* [19]
pRK729	Strain XL-1 with pUC18 clone of phage HK639.
pML042	pGB2ts plasmid. Plasmid used for shuffling.

Table.1. Bacteria, phages and plasmids

pRK1132	Strain RK898 (MG1655) + pUC 19 with R6Kγ*ori*/KAN-2 Transposon (this work).
pRK1142	Strain XL-1 that carries pRK729 with R6Kγ*ori*/KAN-2 Transposon (from this study).
Phage	
HK639	Bacteriophage discovered in sewers of Hong Kong (T.S.Dhillon).

In some applications, the Qiagen PCR Kit was used to purify PCR products in the 100bp-10kbp size range. This kit was also used for purifying DNA fragments (about 10µg) from enzymatic reactions and agarose gels. The system uses the spin-column technology with selective binding properties of a uniquely designed silica membrane. Special buffers are used for the efficient recovery of DNA and removal of contaminants. The DNA is absorbed onto the silica membrane and the contaminants pass through the column. The DNA is eluted from the membrane by Tris buffer or water.

TABLE. **2.** PCR Primers

Primers For Lysogen confirmation
RK179 5’ CATCGAATTCCTGTTTTTGTGAGCATAGCAC 3’
Primes upstream of the putative pL promoter of HK639.
RK180 5’ GATCTCTAGAAAAGAGCTGCTGGTAAATCG 3’
Primes downstream of the putative pL promoter of HK639.
RK199 5’ATGTGCCCTTCGATATTGTC 3’

Primes upstream cI gene of HK639.
RK200 5'CATAACCTCGTTTATATATTG 3'
Primes downstream of cI gene of HK639.
Primers used to find the integration site of HK639.
RK254 5' CCGAACCACTTAAGCGAAC 3'
Primes at the carboxy terminus of the HK639 integrase gene.
RK255 5'CATGAAAACGGAAGTCTTC 3'
Primes at the carboxy terminus of the HK639 integrase gene (reverse primer).
RK299 5' CACGCTTCAGCCTTCAGCTTGT 3'
Primes between nucleotides 41,817-41,839 of HK639.
RK300 5' CACATCTGTGACAC 3'
Primes between nucleotides 40,610-40,598 of HK639.

TABLE. .2. PCR primers

Mutagenic primers for incorporating NcoI restriction sites
RK314 5' ATAGCCATGGCCACGGTTGATGAGAGC 3'
RK315 5' ACGTCCATGGCTCAACCATCATCGATGAA 3'
The NcoI site is underlined.

Mutagenic primers for incorporating EcoRI restriction sites

RK324 5' CTAGGAATTCCACTGAATAATTGCCGTC 3'

RK325 5' CTAGGAATTCCCATGACCTGGCTCATC 3'

The EcoRI site is underlined.

Primers used for the amplification of transposon sequence

R6KAN-2 FP-1 Forward Primer
ACCTACAACAAAGCTCTCATCAACC

R6KAN-2 RP-1 Reverse Primer
5' – CTACCCTGTGGAACACCTACATCT
These primers came along with the epicenter kit.

Primers for confirmation of integration site

RK326 5'GCCGTCATCCCATCGC. Primes upstream of the site near the NcoI s near the HK639 Integrase gene.

RK327 5'GCTCATCGCCGCACTTA. Primes downstream of NcoI site near the HK639 Integrase gene.

RK222 5'CTGATTGCCGACATTATC 3'. Amplify custom transposon.

RK223 5' AAGGATCAGATCACGCATCTTC 3'. Amplify custom transposon

Results

Isolation of Genomic DNA from an HK639 lysogen

Genomic DNA from strain RK1033, an HK639 lysogen made in our lab, was isolated by standard procedures. The integrity of the DNA was analyzed on a 1% agarose

was analyzed on a 1% agarose gel (Figure 8). We also confirmed that the prophage was present in the lysogen by PCR (Figure 9). This was accomplished by amplification with primers RK179 X RK180 and RK199 X RK200 which yielded the expected sized PCR products. Primer pair RK179 and RK180 amplifies the pL promoter region of HK639 and primers RK199 X RK200 amplify the repressor (cI) gene. As a control, pure HK639 DNA was used as a template in PCR reactions with the same sets of primer pairs.

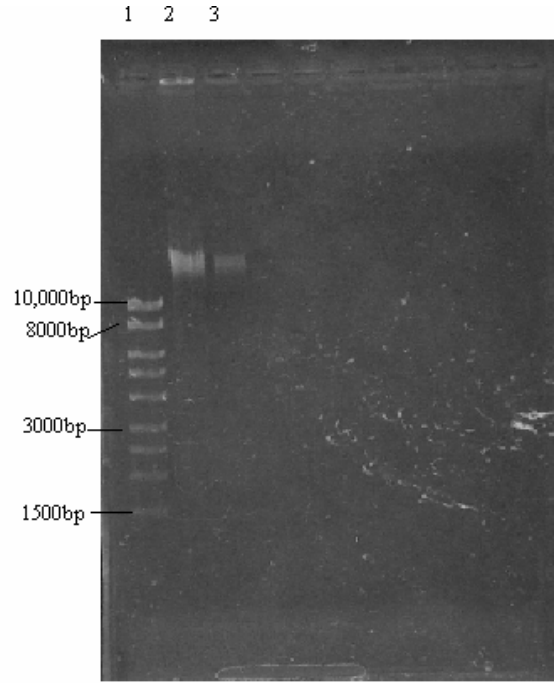

Figure.8. Genomic DNA isolated from the strain RK1033 and RK 1014. Lane 1: DNA high range ladder, Lane 2: RK1033, Lane 3: RK1014 (HK639 lysogen).

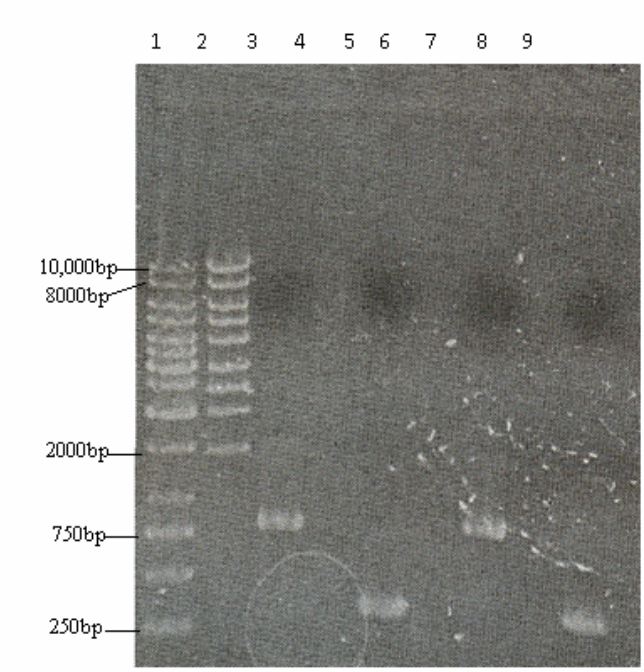

Figure.9. Gel electrophoresis of PCR products amplified from an HK639 lysogen (strain RK1033) and pure HK639 DNA (a positive control). The results confirm the presence of the prophage in strain RK1033. The expected fragment size for the primer pair RK179 X RK180 is 250bp and 750bp for the primer pair RK199 X RK200. Lane 1: 1Kb ladder, Lane 2: DNA high ladder, Lane 3: RK1033 (PCR products using primers RK199 X RK200), Lane 5: RK1033 (PCR products using primers RK179 X RK180), Lane 7: HK639 (PCR products using primers RK199 X RK200), Lane 9: HK639 (PCR products using primers RK179 X RK180)

Identification of the HK639 integration site by Inverse PCR

We attempted to identify the HK639 prophage integration site by a technique called Inverse PCR. Inverse PCR (IPCR) was developed by Ochman et al [20]. It is a method for the rapid *in vitro* amplification of DNA sequences that flank a region of known sequence. A unique feature of this method is that the primers are initially oriented in opposite directions but after circularization of the DNA fragments they prime towards each other. Oligonucleotides were designed such that they are complementary to the opposite strands of a known sequence contained within the prophage–host junction fragment and oriented such that the 3' extension product from each oligonucleotide initially proceeds away from the other. The chromosomal DNA fragments of the *E.coli* lysogen (strain RK1033) generated from a restriction digest were circularized using T4 DNA Ligase. After ligation the flanking host DNA should lie between the 3' ends of the oligonucleotides (Figure 10) [20].

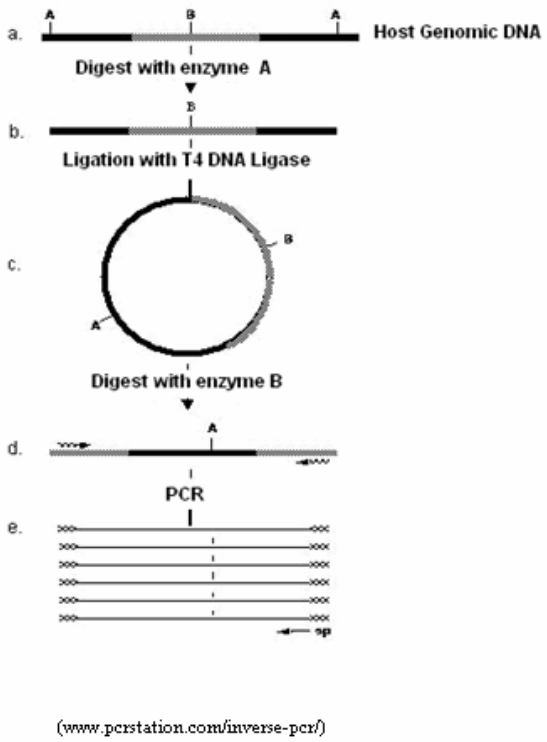

Figure.10. Schematic representation of Inverse PCR. Part (a) represents the genomic DNA of a cell. The genomic DNA is digested with a frequent cutter such as Sau3AI, which does not cut in the prophage (part b). The digested genomic DNA is then ligated using T4 DNA Ligase (part c). The ligated DNA is then digested with AflII to linearize the circular DNA (part d). The linearized DNA molecules are then used as template in PCR reactions where the primers that originally faced away from each other, now prime toward each other. The amplified PCR products were examined by gel electrophoresis.

Purified genomic DNA from the HK639 lysogen-containing *E. coli* strain RK1033 was digested with Sau3AI and treated as described in Figure 10. After ligation of the digestion products, the circular molecules were digested with AflII and then used as template for a PCR reaction with primers RK254 and RK255. The PCR products were gel purified and sequenced using the same primers. The sequences of the IPCR products were then used in a BLAST search for sequences in the GENBANK database. Although good sequence information was recovered from the gel purified PCR products, no *E. coli* sequences were found. We only recovered sequences of HK639 that circularized. We tried additional enzymes such as MspI, EcoRI, EcoRV, FokI, NcoI and AlwI in place of Sau3AI but none of these yielded PCR products. EcoRV did not work with Inverse PCR but worked when the mutant RK1158 was digested with it. It may be because EcoRV has blunt ends and blunt end ligations are not as efficient as sticky end ligations and also because the PCR fragment was either too large or small for the ligation to occur.

Mutagenesis of HK639

Since the inverse PCR experiments were unsuccessful, we decided to genetically tag the phage with a selectable marker. This would provide a useful tool for our characterization of this phage. Transposons encode antibiotic resistance and they are classic tools for generating mutations. A transposon or transposable element is a DNA sequence that is capable of inserting itself or a copy of its sequence into a new location in a genome. Transposons were originally discovered by Barbara McClintock [21]. This approach also has the added benefit of potentially generating a family of mutants whose properties could be characterized. Furthermore, the transposon might insert itself close to

the integration site (*att*P) site of the phage and this would facilitate its identification. The EZ-Tn5™ <R6Kγ*ori*/KAN-2> Insertion Kit from Epicentre was used to make random mutants of the phage. This particular kit was chosen because it encodes Kanamycin resistance and it has an origin of replication that could potentially be used for plasmid rescue (Figure 11). The transposition reaction was set up in vitro according to the manufacturer's instructions. Purified HK639 DNA was incubated with the transposon/transpose complex and the activation of the complex was achieved by the addition of magnesium ions. The DNA was purified and transformed into competent E. coli cells by electroporation. After the expression period, the cells were plated on LB plates supplemented with Kanamycin. No Kanamycin resistant colonies were recovered. To test the efficiency of the kit components a transposition reaction with pure pUC19 plasmid DNA was attempted (Figure 11). This control reaction worked. Kanamycin resistant colonies were recovered and the exact transposon/plasmid junction was determined by DNA sequencing using transposon specific primers (Figure 12). The resulting plasmid (pRK1132) served as a template for amplification of the transposon.

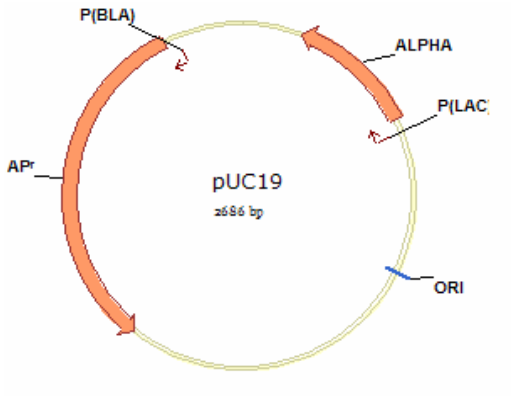

Figure.11. Plasmid vector pUC19. The Origin of replication, the lac promoter (P(LAC)), the beta lactamase promoter (P(BLA)), and the Ampicillin resistance gene (AP) are shown in the above figure.

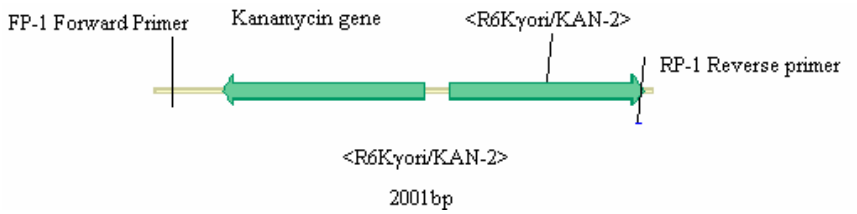

Figure.12. Cartoon of R6K*γori*/KAN-2 transposon. The positions of the forward and reverse primers and the orientation of the Kanamycin resistance gene and origin of replication are shown.

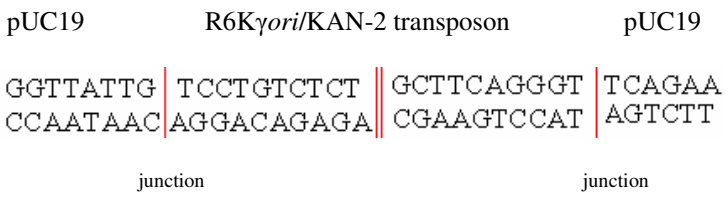

Figure.13. Sequence junction between pUC19 and the R6Kγ*ori*/KAN-2 transposon. The junctions are indicated by the single lines.

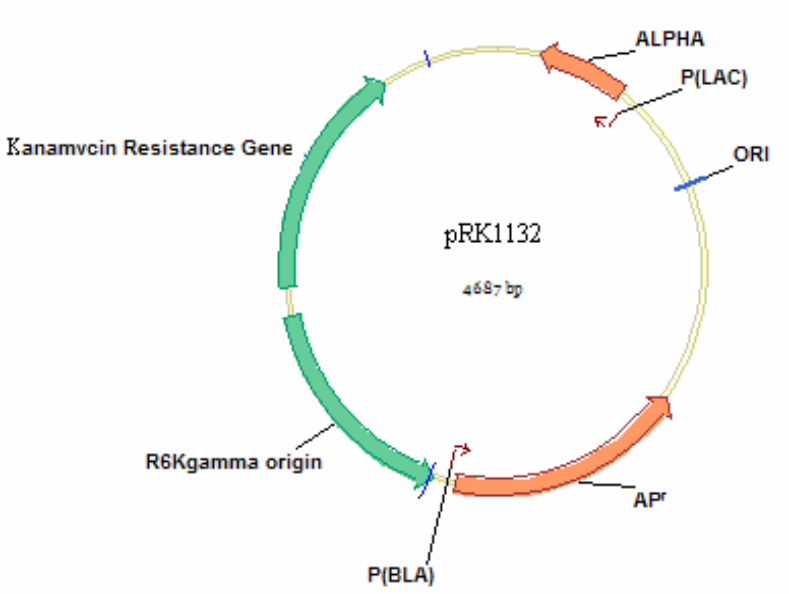

Figure.14. A cartoon depiction of pRK1132 (pUC19 + R6Kγori/KAN-2 transposon) The lac promoter (P(LAC)), the beta lactamase promoter (P(BLA)), the Ampicillin resistance gene (AP) and the R6K*γori*/KAN-2 transposon are shown.

Cloning of Kanamycin gene into pRK729

Our attempts to tag the HK639 with the transposon were unsuccessful. However, it should be possible to insert the Kanamycin resistance gene into the chromosome through homologous recombination, if sufficient flanking homology is provided. In most lambdoid phages, the integration site is located downstream of the integrase gene. We were in possession of a plasmid clone of HK639 that contained the integrase gene and some flanking DNA (pRK729). Close examination of this clone revealed a unique NcoI site immediately upstream of the integrase gene. We decided to clone the R6K-Kan-2 sequence into this site. NcoI sites that flanked the Kan sequence were incorporated into oligonucleotides used for PCR amplification. The amplified DNA was digested with NcoI and cloned into the NcoI digested pRK729. The cloning was successful. The resultant plasmid, pRK1142, conferred kanamycin resistance. Overnight cultures of the transformed strain were grown and plasmid was extracted. The plasmid DNA (approximately 10kb) was digested with NcoI to see if it would release the Kanamycin gene (2001bp) (Figure 16).

Deleted:

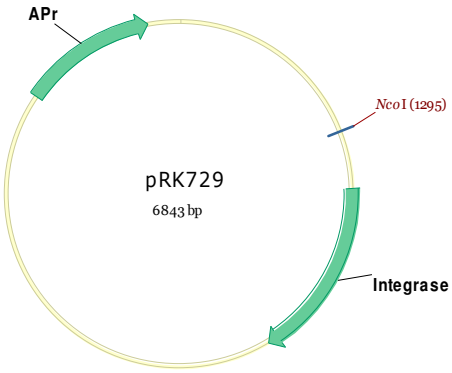

Figure.15. A cartoon depiction of the plasmid pRK729. The integrase gene and NcoI site, the origin of replication are shown. The lac promoter (P(LAC)), the beta lactamase promoter (P(BLA)), and the Ampicillin resistance gene (AP) are shown in the above figure.

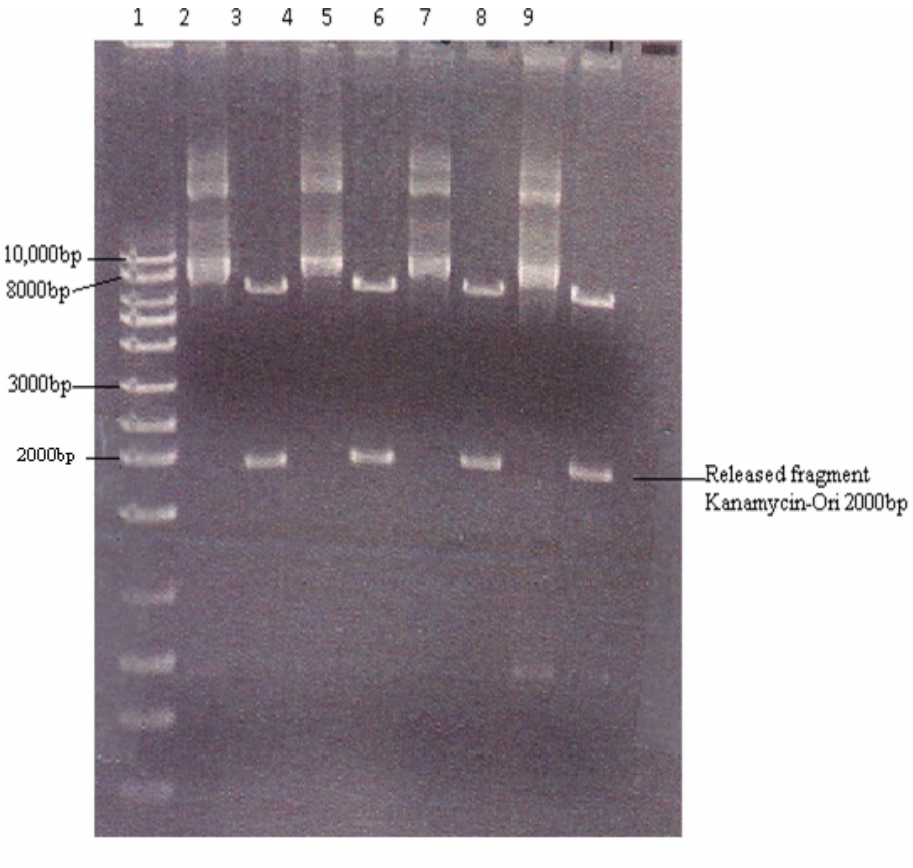

Figure.16. Confirmation of cloning of Kanamycin gene into pRK729. Lane 1: 1Kb ladder, Lane 2,4,6,8: pRK1142 uncut candidates , Lanes 3, 5, 7, 9: pRK1142 NcoI digested candidates.

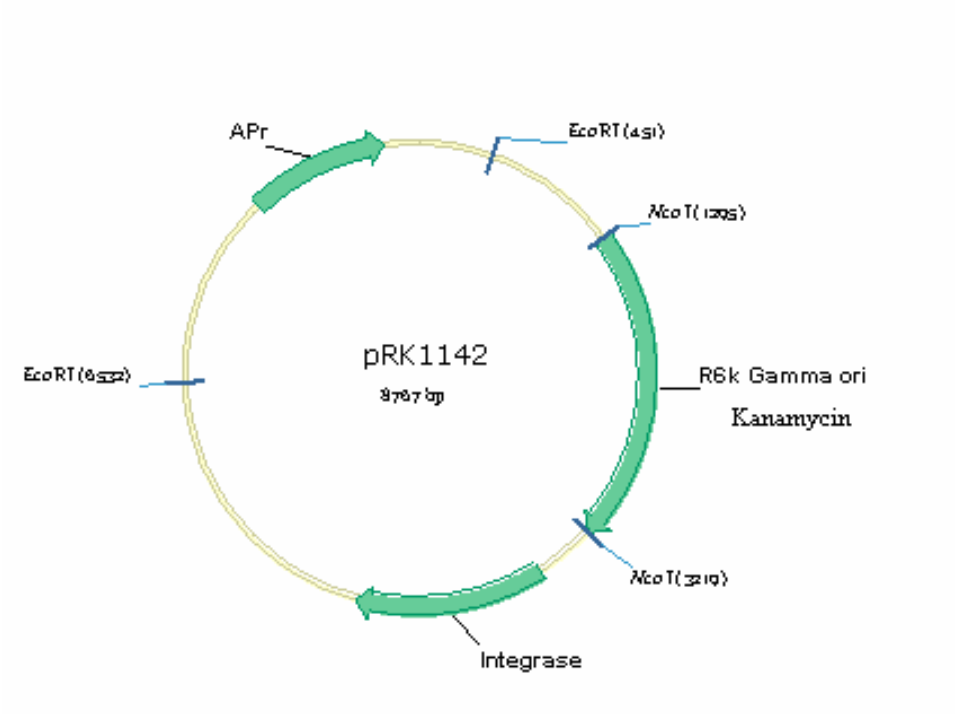

Figure.17.A diagrammatic representation of the plasmid pRK1142. The integrase gene and R6K gamma ori and the Kanamycin gene are shown.

Site Directed Mutagenesis of HK639 Prophage

For unknown reasons, HK639 was incapable of infecting the strain that carries plasmid pRK729 or its derivatives. To accomplish our goal of site directed mutagenesis of the prophage we pursued two parallel approaches. One approach was to subclone the region of interest (Kanamycin gene) into plasmid pML042. pML042 is pGB2ts plasmid, which is used for plasmid shuffling. The second approach was to use a technique called recombineering to introduce the sequences into an HK639 prophage. HK639 was incapable of infecting pRK1142. There may be some HK639 genes on the pRK1142 plasmid that prevent HK639 infection. Therefore, we chose to subclone a small segment of pRK1142 into pMLO42. As detailed below, we successfully made the desired mutant by recombineering. Therefore, we abandoned our attempts to clone the Kanamycin gene into pML042. Recombineering is a recently developed method of genetic engineering which does not rely on restriction sites for cloning. It uses short regions of homology and bacteriophage proteins, which are capable of catalyzing recombination between these homologies to form combinations of genes and other genetic elements in vivo [22]. This technique has been gaining importance over the past few years because of its simplicity and efficiency. With the help of this technique, the exact desired construct can be made [22]. The Kanamycin resistant recombinants were purified and analyzed by PCR using primers RK326 X RK223 and RK327 X RK222. The mutation did not affect the viability of the prophage. Spontaneously released phages were detected by plating culture supernatants on lawns of susceptible host cells. In addition, phage release could also be

induced with mitomycin C [23]. We used this strain (RK1157) to determine the integration site of the phage HK639 in *E. coli.*

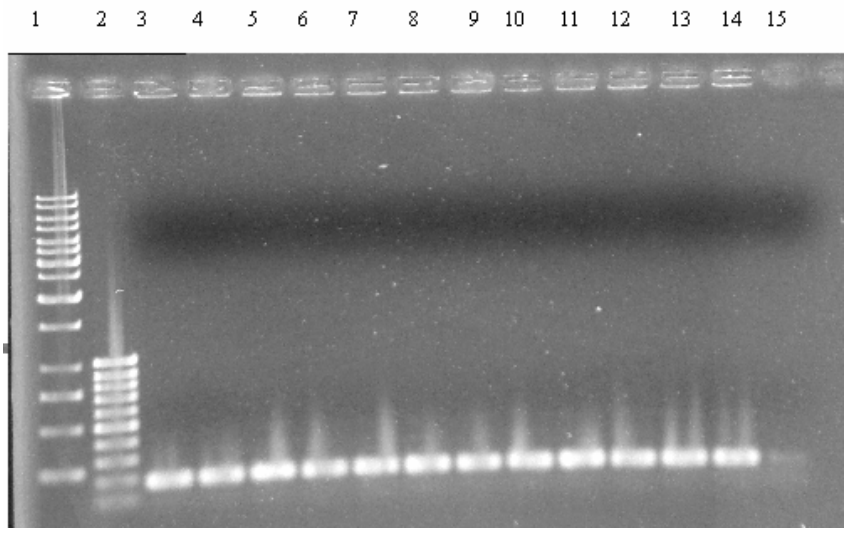

Figure. 18. A Gel confirming the presence of the R6KoriKan Transposon sequences on the HK639 prophage. The site was confirmed by PCR using primer set RK326 X RK223. The expected size is 180bp. The product is running just below the 200bp marker (Low range marker). Lane 1: 1kb ladder, Lane 2: Low MW ladder, Lanes 3-14: Recombineering candidates with ori kan insert at NcoI site upstream of the HK639 Integrase gene. Lane 15: Negative Control.

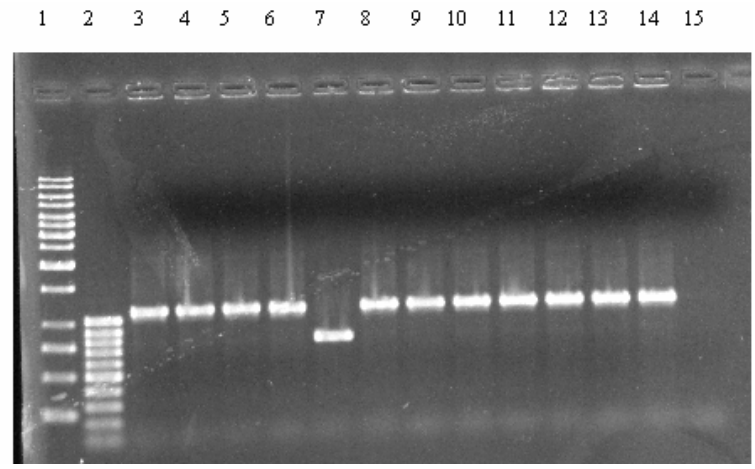

Figure.18. B Gel confirming the presence of the R6KoriKan Transposon sequences on the HK639 prophage. The site was confirmed by PCR using primer set RK222 X RK327. The expected size is 1,131 bp. The product runs just above the 1,031bp marker (Low range marker). Lane 1: 1kb ladder, Lane 2: Low MW ladder, Lanes 3-14: Recombineering candidates with ori Kan insert at NcoI site upstream of the HK639 integrase gene. Lane 15: Negative Control.

Identification of integration site of HK639 lysogen by plasmid rescue

Five microliters of the purified ligation mixture (RK1157 digested with EcoRV and ligated) was added to 50 mL of TransforMax™ EC100D™ *pir*$^{+}$cells from Epicentre. These cells express the π protein (which is the product of the *pir* gene) for the replication of plasmids having the R6Kγ*ori*. They are useful for cloning of unstable DNA sequences and maintain about 15 copies of the R6Kγ*ori* containing plasmids per cell. These cells also have high transformation efficiency and can accept large insert clones. After transformation by electroporation SOC was added and the mixture was incubated at 37^{o}C for 1 hour. The mixture was then spread on LB-Kanamycin (20μg/ml) and incubated overnight at 37^{o}C. The colonies were picked and incubated in LB- Kanamycin media (20 μg/ml). Plasmid DNA was prepared the next day. After confirming the integrity of the DNA sequencing reactions were set up with the primer RK254. We recovered one clone having homology to HK639 and *E. coli*. The junction (Figure 18) is located at 262,136bp in *E. coli* which occurs in the *thrW* tRNA gene (Figure 19). Phage integration in or near tRNA genes is well documented and our result is consistent with these observations [24]. The integration site was confirmed by amplification of the junction using primers RK345 X RK254 and RK347 X RK345 (Figure 20). The correct product sizes were observed.

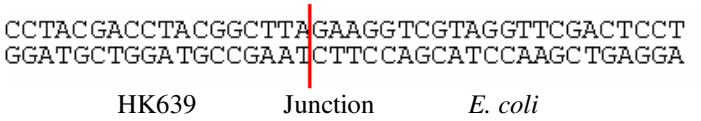

Figure.19. Sequence showing the location of the junction. The junction is indicated by the line.

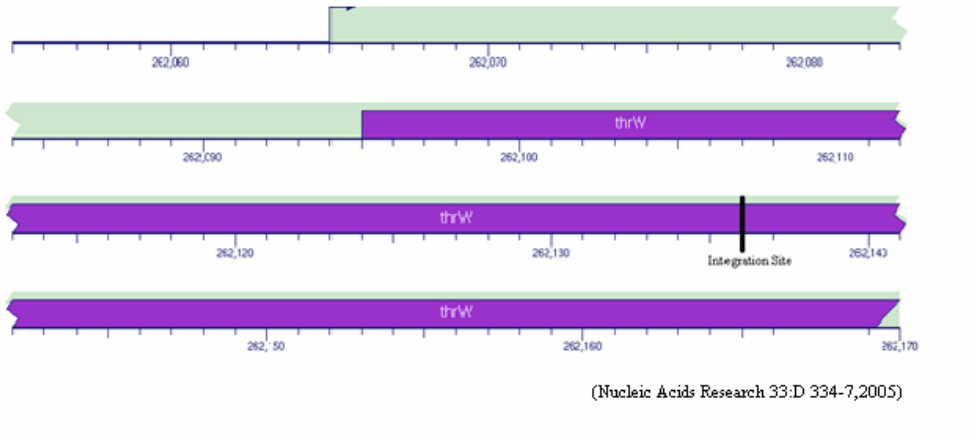

Figure.20. Representation of the insertion site in *E. coli* genome. The location of the HK639 integration site in *E. coli* (Strain K-12 MG1655 genome). The junction is indicated by a straight line.

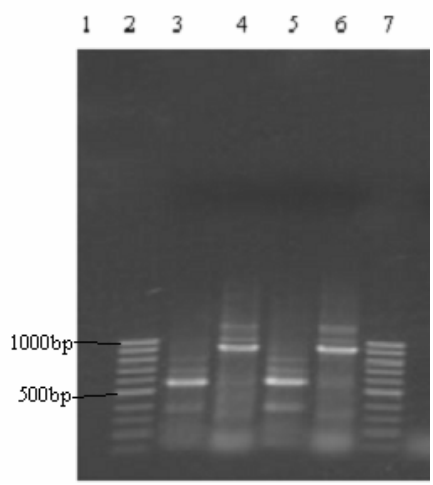

Figure.21. PCR confirmation of the integration site of HK639 in the *E. coli* (Strain K-12 MG1655) chromosome. The junction was amplified using primers mentioned above and products of the expected size 950 bp (for primer pair RK347 XRK345) and 500bp (for primer pair RK346 X RK 254) were observed. Lane 2: 100bp ladder, Lane 3: PCR product (RK 346 X RK254), Lane 4: PCR product (RK347 X RK345), Lane 5: PCR product (RK 346 X RK 254), Lane 6: PCR product (RK347 X RK345) and Lane 7: 100bp ladder.

Discussion

In this study we attempted to characterize a newly discovered phage called HK639. This phage is interesting because it exhibits RNA mediated antitermination unlike Lambda which has antitermination mediated by a phage-encoded protein. This unique mechanism of genetic control is the focus of the research in our laboratory. The focus of this study was to make mutants of the phage and to identify its integration site in *E. coli.*

Our attempt to make random mutants of the bacteriophage HK639 was not successful. There are several possible explanations for this. Low transformation efficiency or low efficiency of electroporation could have prevented the recovery of the desired mutants. Instead of making random mutants we tried to make mutants by cloning the Kanamycin gene near the HK639 integrase gene in pRK729 (=Strain XL-1 with pUC18 clone of phage HK639). Though we successfully cloned the Kanamycin gene, HK639 was not capable of infecting the new plasmid pRK1142 (=Strain XL-1 that carries pRK729 with the R6Kγ*ori*/KAN-2 Transposon). We attempted a couple of different approaches to introduce the desired mutation into HK639 by recombination. Recombineering was highly efficient and yielded the desired mutants. The new strain, RK1157, was Kanamycin resistant. Introducing a mutation into the phage genome did not affect the viability of the prophage or its ability to lysogenize a new host.

We tried to identify the integration site by Inverse PCR but we were not successful. Poor template DNA circularization/ligation and/or low amplification efficiency could explain this result. The mutant made by recombineering, RK1157 was

used to identify the integration site. We found that the integration site is located in the *thrW* gene of *E. coli*. Phage integration in or near tRNA genes is well documented and our results are consistent with these observations [24]. Bacteriophage HK639 is only the second known example of a phage that uses RNA-mediated antitermination. As part of our characterization of this newly identified phage, we have generated a useful mutation near its integrase gene. The location of this mutation allowed us to identify where the phage integrates into the chromosome of its host. The recombineering technique will be used to generate additional mutants so that we can get a better understanding of the biology and evolution of phage that use RNA-mediated antitermination.

References:

1) **Bergh, O., K.Y. Borsheim, G. Bratbak and M.Heldal** (1989). High abundance of viruses found in aquatic environments. *Nature 340; 467-468*.

2) **Brussow, H. and R.W. Hendrix** (2002). Phage genomics; Small is beautiful. *Cell 108; 13-16.*

3) **Fuhrman, J.A.** (1999). Marine viruses; Biogeochemical and ecological effects. *Nature* 399(6736):541-8.

4) **Hendrix, R.W.** (2002). Bacteriophages: Evolution of the Majority. *Theoretical Population of Biology,* Volume 61, Issue 4, Pages 471-480.

5) **Hendrix, R.W., M.C.Smith, R.N.Burns, M.E.Ford and G.F.Hatfull** (1999). Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc .Natl Acad.Sci.USA 96:2192-2197.*

6) **Wommack, K.E. and R.R Colwell** (2000). Viroplankton: Viruses in aquatic ecosystems, *Microbiol .Mol.Biol.Rev.64:69-114.*

7) **Mark Ptashne**. *A Genetic Shift*, Cell Press Blackwell Scientific Publishing, (1986): pg 2-4.

8) **Jonathan M.Budzik**. (2003). *Phage isolation and investigation,Vol 3, No.1.*

9) **Noreen E. Murray and Alexander Gann**. What has phage ever done for us? *Current Biology Vol 17 No 9.* Volume 17, Issue 9, Pages R305 - R312.

10**) Benjamin Lewis**. *Essential genes*, Pearson Education, Inc. (2006): 258-279

11) **Jerome J. Perry, James T. Staley, and Stephen Lory**. *Microbial life,* Sinauer Associates. (2002).

12) **Max E. Gottesman and Robert A. Weisberg** (2004). Little Lambda, who made thee? *Microbiology and Molecular Biology reviews, pg 796-813.*

13) (http://www.mun.ca/biochem/courses/4013/Topics/Lambda/Antitermination/html)

14) **Sarbani Banik-Maiti, Rodney A. King and Robert A.Weisberg** (1997)**.** The antiterminator RNA of Phage HK022. *J. Mol. Biol 272, 677-687.*

15) **Clerget, M., D. J. Jin, and R. A. Weisberg.** (1995). A zinc binding region in the ß' subunit of RNA polymerase is involved in antitermination of early transcription of phage HK022*. J. Mol. Biol. 248:768-780.*

16) **Courtney N. Miles and Rodney A. King**. The immunity region of bacteriophage HK639 encodes an antitermination RNA (unpublished results).

17) **Guyer, M. S., R. E. Reed, T. Steitz, K. B. Low** (1981). Identification of a sex-factor-affinity site in E. coli as gamma delta. *Cold Spr.Harb.Symp.Quant.Biol. 45:135-140.*

18) **Murray, N.** (1977). Lambdoid phages that simplify the recovery of in vitro recombinants *Mol. Gen. Genet.150, 53-61.*

19)**Yanisch-Perron, C., Vieira, J. and Messing, J.,** (1985)**.** Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, *Gene 33, 103-119.*

20**) Ochman, H. et al.** (1998). Genetics 120, 621-623.

21) **Barbara McClintock** (1953). Induction of instability at selected loci in maize. *Genetics* 38:579–99.

22) **Lynn Thomason, Donald L. Court, Mikail Bubunenko, Nina Costantino, Helen Wilson, Simanthi Dutta, and Amos Oppenheim** (2005). Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination. *Current Protocols in Molecular Biology.1.16.1-1.16.21.*

23) **Jeffery W. Roberts and Christine W. Roberts**. (1975). Proteolytic cleavage of bacteriophage Lambda repressor in Induction. *Proc. Nat. Acad. Sci. USA, vol. 72, No.1, pp. 147-151.*

24) **Allan M.Campbell**. (1992). Chromosomal Insertion sites for phages and plasmids. *Journal Of Bacteriology, Dec. 1992, p. 7495-7499 Vol. 174, No. 23.*