Spring 1993

Insertion sequence elements in Yersinia: Nucleotide sequence of IS100 of Yersinia pestis

Stephen Dale Torosian
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

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.

Insertion sequence elements in *Yersinia*: Nucleotide sequence of IS100 of *Yersina pestis*

Torosian, Stephen Dale, Ph.D.

University of New Hampshire, 1993

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

INSERTION SEQUENCE ELEMENTS IN *YERSINIA* :

NUCLEOTIDE SEQUENCE OF IS100 OF *YERSINA PESTIS*

BY

Stephen D. Torosian

BS, University of Massachusetts at Amherst, 1977

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of
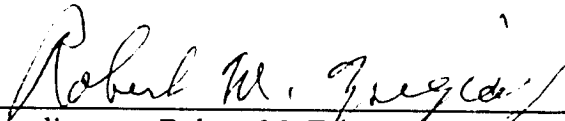
Doctor of Philosophy

in
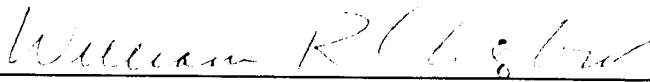
Microbiology

May, 1993
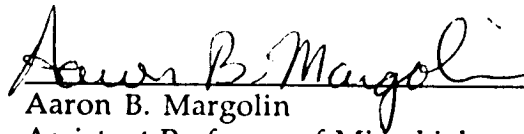
This dissertation has been examined and approved.

_____
Dissertation director, Robert M. Zsigray
Professor of Microbiology and Genetics

_____
William R. Chesbro
Professor of Microbiology

_____
Aaron B. Margolin
Assistant Professor of Microbiology

_____
John J. Collins
Assistant Professor of Biochemistry and Genetics

_____
Donald M. Green
Professor of Biochemistry and Genetics

5/5/93
_____
Date

DEDICATION

This dissertation is dedicated to Laura Ackerman

iii

ACKNOWLEDGEMENTS

I would like to thank the Central University Research Fund, the Department of Microbiology, and the Graduate School for providing research funds, teaching assistantships, research assistantships, and teaching assistant fellowships.

I would like to thank the Microbiology Staff Linda, Carol, Alberta and Bob for their assistance and friendship.

I would like to thank the members of my committee; Dr. Chesbro, Dr. Margolin, Dr. Collins and Dr. Green for their guidance.

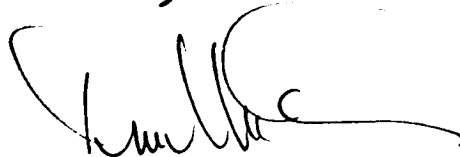I would like to thank Dr. Zsigray for his assistance, support and friendship.

I would like to thank all the graduate students in the Department of Microbiology; with special thanks to John D'Elia, Tod Gavron and Robin Ross for technical assistance and discussions.

I would like to thank Elizabeth for her assistance in typing and editing this manuscript.

I would especially like to thank my children, Ellyn, Stephen John, and Patrick for patiently enduring my tenure as a graduate student.

Finally, I would like to thank Karen for the daily support, positive attitude and smiles on the worst and best of days.

iv

TABLE OF CONTENTS

vi

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

x

ABSTRACT

INSERTION SEQUENCE ELEMENTS IN *YERSINIA* :

NUCLEOTIDE SEQUENCE OF IS100 OF *YERSINIA PESTIS*

by

Stephen D. Torosian

University of New Hampshire, May, 1993

The World Health Organization classified (Williams, 1983) *Y. pestis* as *Y. pseudotuberculosis* subsp. *pestis* on the basis of DNA homology, yet the two organisms cause markedly different disease. Portnoy and Falkow (1981) reported IS100 to influence the virulence of *Y. pestis*. IS100 was shown to be found in *Y. pestis* but not *Y. pseudotuberculosis*.

IS100 from *Y. pestis* was sequenced and shown by sequence analysis to fulfill the requirements of being an IS element. pIS1C, an 821 bp fragment of IS100 was transferred to *Y. pseudotuberculosis* Trp-Ca-, resulting in the ability of the cells to ferment rhamnose and delayed production of urease at 37^0C but not at 26^0C. These traits are normally associated with *Y. pestis*. Maxicell analysis revealed more products and larger products than could be expected from the coding capacity of the clone. Two-dimensional polyacrylamide gel

electrophoresis analysis demonstrated a complex rearrangement of protein

profile based on the presence of pIS1C. These data clearly indicate the

thermo-regulatory involvement of IS100 at the molecular level.

INTRODUCTION

The genus *Yersinia* is comprised of seven species of gram negative rods to coccobacilli ranging from 0.5-0.8 μm in diameter to 1-3 μm in length (Bercouvier and Mollaret, 1984) They are non-motile at 37^0 C, and all but *Yersinia pestis* are motile at 30^0 C. The size of the *Y. pestis* genome has been determined by macrorestriction pattern polymorphism and two dimensional pulsed field gel electrophoresis to be 4,397.9 $\pm$ 134.6 kbp (Lucier and Brubaker,1992). The genus *Yersinia* contains seven species of which three are medically important to humans, *Y. pestis*, *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. *Y. pestis* was shown to be the etiological agent of plague in 1894 by Alexandre Yersin. *Bacterium enterocoliticum* (presently *Y. enterocolitica*) was first proven to cause infection in humans in 1939 by Schliefstein and Coleman (Butler 1983). Daniels and Goudzwaard (Butler 1983) renamed it *Pasturella x* in 1963 and the bacterium was placed in the genus *Yersinia* in 1964 by Frederickson (Butler 1983). *Bacillus pseudotuberculosis* (presently *Y. pseudotuberculosis*) was discovered in 1889 by Pfeiffer (Butler 1983) and renamed *Pasturella pseudotuberculosis*, also by Pfeiffer. It was first reported to cause infection in humans in 1909 (Saisawa, 1913). It was called *Shigella pseudotuberculosis* in 1935 and finally placed in

1

its current genus, *Yersinia* (Smith and Thal, 1965). *Y. pestis* has a

guanine/cytosine content of 46%, *Y. pseudotuberculosis* has a %GC of 46.5%

and *Y. enterocolitica* has a %GC of 48.5% (Bercouvier and Mollaret, 1984) as

determined by thermal melt. Percent GC is an accepted indicator of

relatedness between species but even more important is % DNA homology.

Y. pestis shares approximately 50% DNA homology with *Y. enterocolitica*

and these organisms represent distinct species (Brenner et al, 1976; Bercovier

et al, 1980). On the other hand , *Y. pestis* and *Y. pseudotuberculosis* are

essentially identical with respect to their DNA homology; however, the type

of disease that each produces is dramatically different. Because of their close

relatedness, the World Health Organization changed the names of the

etiologic agents of plague and pseudotuberculosis to *Y. pseudotuberculosis*

subsp. *pestis* and *Y. pseudotuberculosis* subsp. *pseudotuberculosis,*

respectively. The scientific community has rejected this nomenclature based

not only upon confusion that may result during the cultivation of this

bacterium but also upon Devignat's hypothesis that *Y. pseudotuberculosis*

evolved from *Y. pestis,* but not **vice versa** (Mollaret, 1987).

For over 55 years, a number of investigators have been convinced that

culture of *Y. pestis* would yield variants that are essentially indistinguishable

from *Y. pseudotuberculosis.* As examples, Harvey (1933) speculated that *Y.*

pseudotuberculosis might be a plague bacillus living in symbiosis with a

bacteriophage. Likewise Soviet scientists have reported the spontaneous

"transformation" of *Y. pestis* to *Y. pseudotuberculosis* (Bessonova et al, 1936) and the conversion of *Y. pestis* to *Y. pseudotuberculosis* by either the action of a bacteriophage (Zhukov-Verezhnikov et al, 1955) or by prolonged storage (Korobdova, 1963). In contrast, Burrows and Bacon (1960) hypothesized that *Y. pseudotuberculosis* "may, in some way, be concerned in the endemicity of plague. The sporadic derivation of *P. pestis* from *P. pseudotuberculosis*, an organism having wide geographical and host range, would nicely explain the otherwise puzzling alteration of plague epidemic and quiescence throughout history." To my knowledge, only two additional reports have appeared in the literature that suggested that *Y. pestis* may originate from *Y. pseudotuberculosis* (Surgalla, 1965, Sodeinde et al, 1992).   Because of the extreme degree of relatedness of *Y. pestis* and *Y. pseudotuberculosis*, the possibility that the difference between the two bacteria may be due to insertion sequence (IS) elements was considered and the concept formed the basis of the studies to be described in this dissertation.

3

Literature Review

Some of the applicable phenotypic traits of the medically important *Yersinia* are listed in Table 1. It should be noted that *Y. pestis* and *Y. pseudotuberculosis* are most closely related by trait. Indeed some phenotypic revertants characteristic of one sub-species can readily be isolated from the other sub-species, a phenomenon called meiotrophic reversion (Table 2).

Virulence in *Y. pestis* is due to five established virulence factors (Brubaker, 1972), as well as a number of putative virulence factors (Brubaker 1991). These virulence factors may either be chromosomal or plasmid mediated. The five established virulence factors are production of capsular antigen or Fraction 1 (Fra+), pigmentation (*pgm+*), ability to synthesize purines (*pur+*), production of V and W antigens, and the production of pesticin. Fraction 1 is a surface glycoprotein produced at 37^0C but not at 26^0C (Fox and Higuchi, 1958). Fraction 1 functions by interfering with complement components C_2 and C_4 (Williams et al., 1972), thus preventing phagocytosis by monocytes and neutrophils. Fraction 1 is encoded by the 65 MegaDalton (MD) plasmid. The *pgm+* trait is due to peptide F which is an outer membrane protein involved in storage of iron, but not in its transport (Mulder et al., 1989, Pendrak and Perry, 1991). Iron transport appears to be regulated by Fur and *fur* sequences in *Yersinia* (Staggs and Perry, 1990)

4

TABLE 1

Some of the Phenotypic Traits of the Three Medically Important Yersinia [a]

	Y. pestis	Y. pseudotuberculosis	Y. enterocolitica
Motility at 25°C	-	+	+
Ornithine Decarboxylase	-	-	+
Urease	-	+	+
β xylosidase	+	+	-
Citrate (Simmons) 25°C	-	-b	-
Voges Proskauer	-	-	+
Indole Production	-	-	W
γ Glutamyl Transferase	-	W	+
Acid Production From			
Rhamnose	-	+	-
Sucrose	-	-	+
Cellibiose	-	-	+
Melibiose	W	+	-
Sorbose	-	-	+
Sorbitol	-	-	+
Raffinose	-	Wb	-
Malonate Utilization	-	-	-
Citrate(Christensen)	-	-	W
Nitrate reduced to Nitrite	W	+	+
D-Glucose-gas production	-	-	W
Tetrathionate Reductase	-	W	W
Lipase	-	-	W
Deoxyribonuclease	+	W	W
Glycerol	W	+	+
i Inositol	-	-	+
D-Xylose	+	+	W
Esculin	+	+	W
Amygdalin	-	-	V
Salicin Dextrin	W	W	V
Lactose	-	-	W
Fraction 1	+	-	-
Pesticin	+	-	-
Fibrinolysin	+	-	-
Coagulase	+	-	-
Pigmentation	+	-	-
Murine Toxin	+	-	-

V- Variable reaction
W- Weak reaction
b- Serogroup IV strains are positive

5

Table 2. Virulence Determinants and Meiotrophic Reversions of the Three Medically Important *Yersinia*[b]

Property	Virulence Determinant	Y. pestis	Y. pseudotuberculosis	Y.enterocolitica
6 MD plasmid		+	-	-
Plasminogen Activator	+	+	-	-
Posttranslational degradation of Yops	+	+	-	-
45 MDplasmid		+	+	+
Yops	+	+	+	+
YadA	+	-	+	+
V,W antigens	+	+	+	+
65 MD plasmid		+	-	-
Fraction 1	+	+	-	-
Murine exotoxin	+	+	-	-
Peptide F	+	+	-	-
inv product	+	a	+	+
ail product	+	-	-	+
pH 6 antigen	+	+	+	-
Catalase	+	+	+	-
Assimilation of low levels of NH_4^+	-	a	+	+
Urease	-	a	+	+
Rhamnose	-	a	+	-
Melibiose	-	a	+	-
Biosynthesis of Methionine	-	a	+	+
Phenylalanine	-	a	+	+
Threonine-glycine	-	a	+	+
Isoleucine-valine	-	a	+	+

a-capable of meiotrophic reversion
b- (Brubaker, 1991)

6

and is chromosomally determined. Ability to synthesize purines (Pur+) is also chromosomally determined. Mutants of *Y. pestis* auxotrophic for purine synthesis show reduced virulence (Burrows and Bacon, 1958). Production of V and W antigens are secreted proteins encoded by the 47MD plasmid. The V antigen is a 38 KiloDalton (KD) monomer as evidenced by autoproteolysis of purified preparations (Brubaker et al., 1987). The W antigen is a larger (140 KD) and may be a GroEL-V antigen complex (Brubaker, 1991). GroEL is a chaperone protein involved in selective secretion (Lecker et al 1989). Pesticin, encoded by the 6 MD plasmid, has bactericidal activity due to hydrolysis of peptidoglycan by N-acetylglucosaminidase (Ferber and Brubaker, 1979). It has recently been demonstrated that pesticin specific mutations on the 6 MD plasmid are not responsible for reduction in virulence (Brubaker, 1992). Loss of expression of plasminogen activator, also encoded by the 6 MD plasmid, inhibits invasion of tissues (Brubaker, 1991). This putative virulence determinant was probably the actual target of mutation that resulted in postulation that pesticin was a virulence determinant.

All pathogenic *Yersinia* possess an approximately 47 MD plasmid which governs the production of V and W antigens and outer membrane proteins (Yop's). Loss or mutation of this plasmid results in loss or reduction of virulence (Price et al., 1991). These effects are different between species and are more pronounced in the enteropathogenic *Yersinia*. When *Yersinia* grown at 37^0C are deprived of Ca++ (conc. below 2.5 mM), RNA synthesis and

7

cell division stop and there is a reduction in the adenylate energy charge (Charnetsky and Brubaker, 1982). This is termed the low calcium response (Lcr), and the cells are said to be restricted. During restriction, virulence factors encoded by the Lcr plasmid are produced. Some of these virulence factors are Yop's and V antigen (Brubaker, 1991). There are at least 11 Yop's in *Y. pestis*, 11 in *Y. pseudotuberculosis* (not necessarily the same), and 7 in *Y. enterocolitica*. In *Y. pestis* at least 5 Yop's are required for virulence. 1) Yop E is a 25 KD protein involved in resistance to phagocytosis (Forsberg and Wolf-Watz, 1990). 2) YopH, which is a protein tyrosine phosphatase, causes dephosphorylation of host proteins (Bliska et al., 1991) as well as showing involvement with antiphagocytic activity (Rosqvist et al., 1988). 3+4) Yop's K and L are required for rapid growth in tissues and have been implicated in preventing cell mediated immunity (Straley and Cibull, 1989). 5) Yop M by virtue of its structure is similar to a human platelet surface protein and inhibits platelet aggregation (Leung and Straley, 1989). Posttranslational degradation of these Yop's as a result of proteolysis by plasminogen activator (Sodeinde et al., 1988) is also closely tied to virulence.

Murine toxin is very lethal in rats and mice but not in humans. It is a polymer composed of 15 KD subunits and is an andrenergic blocker (Brown and Montie, 1977). Classical antigens 4 and 5 are both implicated in virulence. Antigen 4 or pH 6 antigen is a polymer of 15 KD subunits expressed at pH 6, but not pH 8, and at 37°C but not at 260C. Because of this, *Y. pestis* may be

8

suited to growth in the phagolysosome and the extracellular environment such as buboes and lesions. Loss of this antigen results in a 200 fold increase in LD_{50} (Lindler et al., 1990). Antigen 5 is probably a novel catalase essential for, though ill defined, in virulence (Brubaker, 1991).

Although YadA and *inv* are not necessary for virulence of *Y. pestis*, these interesting loci are present in *Y. pestis* but are normally inactive. Yad A is inactive due to a frameshift mutation and *inv* is an established meiotrophic factor. Both of these loci are involved in virulence of *Y. pseudotuberculosis* and/or *Y. enterocolitica*. The *inv* product allows for penetration of phagocytes (Finlay and Falkow,1989). Binding to receptors alone mediates entry into the phagocyte (Rankin et al., 1992). *Y. enterocolitica* and *Y. pseudotuberculosis* both possess the active *inv* locus; in *Y. pestis*, it is meiotrophic. In *Y. pseudotuberculosis, inv* mediates penetration of mammalian cells by binding beta 1 chain integrin receptors (Leong et al., 1991).

The *ail* gene is important for the virulence of *Y. enterocolitica* only. It is smaller than the product of the *inv* gene, 17 KD, and mediates attachment and invasion of cultured epithelial cells as well as promoting resistance to complement killing (Bliska and Falkow, 1992). Attachment and serum resistance require that *Y. enterocolitica* be grown at 37^0C, indicating the need for a thermally induced co-factor for *ail*.

The YadA gene is essential to virulence in both *Y. enterocolitica* and *Y.*

9

pseudotuberculosis. YadA promotes binding to collagen and mediates agglutination of erythrocytes and autoagglutination (Emody et al., 1989). YadA has also been found to mediate binding of bacteria to fibronectin (Tertti et al., 1992), while their plasmid cured counterparts bound fibronectin much less fastidiously.

The other major virulence factors in *Y. enterocolitica* and *Y. pseudotuberculosis* are shared with *Y. pestis.* These include Yop's and production of V and W antigens. Although all pathogenic *Yersinia* possess an approximately 47MD plasmid, the products (Yop's) are not always the same. All pathogenic *Yersinia* produce Yop's BCDE and H. Only Yop's E and H are established as virulence determinants in all three pathogenic species of *Yersinia.* Only *Y. pestis* produces Yop's K, L and M which are essential for full virulence. The precise interrelationships of mechanisms for the role of virulence determinant Yop's in *Y. enterocolitica* and *Y. pseudotuberculosis* remain to be elucidated.

Insertion Sequence Elements (IS) in *Yersinia*

There are only a few reports of IS elements in *Yersinia.* A transposon, Tn951 conferring the lactose+ phenotype has been identified (Cornelius et al., 1978) on the pGC1 conjugative plasmid of *Y. enterocolitica.* The transposon is 16.6 kbp possessing short terminal inverted repeats and *lac* genes *i, z* and *y.* This transposon presents identification problems of normally *lac-* pathogenic bacteria.

10

A 1.2 kbp IS-like element has been characterized in *Y. enterocolitica* 8081. This unnamed IS was found 5 bp from the *yop*E gene. Its DNA sequence was shown to have imperfect terminal inverted repeats, as well as duplicated target sequences. At least three copies of this IS were located on the chromosome of *Y. enterocolitica*, one region of homology on the *Y. pestis* chromosome and no homology with the chromosome of *Y. pseudotuberculosis* or any of the virulence plasmids of either *Y. pestis* or *Y. pseudotuberculosis* (Forsberg and Wolf-Watz, 1990).

In *Y. pestis* strain EV76, Ca++ independent mutants are easily obtained. Some of these mutations are the result of deletions, some the result of plasmid loss, and some due to the insertion of a 2.2 kbp IS (termed IS100) into one of two adjacent sites of the 47 MD virulence plasmid. IS100 was found to insert into either the Bam H1-5 or Bam H1-8 fragment of the plasmid. Hybridization studies, using IS 100 as probe, indicated that 10-13 copies of this element existed in the *Y. pestis* chromosome, but no homology was detected with *Y. enterocolitica* 8081 (Portnoy and Falkow, 1981). In a Russian study (Fillippov et al., 1990) in which Ca++ independent mutations of the same type were found, it was discovered that IS 100 transposed into three different sites as well as caused extensive deletions. Slightly smaller IS 101 was also located but found to possess a Hind III site absent in IS 100 and demonstrating a high specificity for integration.

To date, no IS elements have been reported in *Y. pseudotuberculosis*.

11

Forms IS Elements Can Take

By generally accepted definition, insertion sequence elements meet four criteria. 1) They are bordered by DNA sequences varying from 8-35 bp and are normally inverted but may be direct repeats. Frequently these are perfect repeat sequences, but this is not always the case. IS 801 from *Pseudomonas syringae* has no repeats at all (Romantschuk et al., 1990). 2) There is a DNA sequence in the region between the repeats which contain at least 1 open reading frame presumably capable of encoding a transposase. However, even this most basic rule has been waived in the case of an IS in *Leptospira interrogans* which has no identifiable ORF (Woodward and Sullivan, 1991). 3) The element must demonstrate the ability to transpose. 4) Upon transposition, the "target site" of the insertion sequence may be duplicated. As with the other "rules", this is not always the case. None of the IS found in *Coxiella burnetii* duplicate the target site upon transposition (Hoover et al., 1992). IS elements may differ greatly on the number of bases duplicated during transposition and one study classified IS elements on this basis (Ishiguro and Sato, 1988).

IS elements can transpose on their own or in conjunction with other IS elements or parts of IS elements so that occasionally the term transposons will be mentioned. Transposons are generally accepted to be defined as genes (commonly drug resistant genes) flanked by at least one complete IS element, and all or part of another IS element. As in the case of Tn2921 in *Serratia*

12

marcesans, an IS on one end of the transposon and a part of IS 10 on the other end still promotes transposition (Navas et al., 1985). A situation similar to this is the demonstration that resistance plasmid pRR30 which encodes resistance to neomycin and kanamycin is flanked by IS1 on one end and 840 bp of the chloramphenicol acetyl transferase (cam) on the other. This plasmid could be shown to transpose (Peterson and Rownd, 1985). In a related vein, Tn2680 contains IS903 in the middle of the transposon. IS903 can mediate cointegration of a novel transposon Tn2681 which undergoes reciprocal recombination between directly repeated IS26's (Mollet et al., 1985). The complete IS element donates the transposition capabilities of the transposon. Effects caused by IS elements in concert are as frequent as effects caused by a single IS. The highest degree of stability is when the flanking IS elements are in inverted orientation (Iida et al., 1981). It is quite common, however, for transposable elements to insert in a direct orientation (Schmitt et al., 1981). It is not uncommon to get site specific deletion from one terminus of an inverted repeat, although this may be due to the insertion of an IS element within another IS element. It is possible for insertion sequences to undergo intramolecular recombination which leads to different consequences depending on IS structure. For IS elements in direct repeat orientation intramolecular recombination can lead to loss of intervening sequences. For inverted repeat IS elements, intramolecular recombination can lead to reversal of orientation of intervening sequence (Berg et al., 1981).

13

This force would be of more severe consequence to IS elements as deletion of intervening sequences could lead to loss of the IS, but in a less drastic condition this force could be evolutionarily more significant.

No IS element is known to exist autonomously; however, it has been shown that introducing an IS into the central part of another transposable element allows a self replicating unit to form (IS plasmid) but only when the IS element is flanked by sequences that carry a promoter as initiation site for primer RNA synthesis (Lusky et al., 1981).

In a relevant study on the effects of IS50 of *Pseudomonas aeruginosa*, plasmids were constructed containing various pieces of IS50. It was found that a plasmid containing the IS50 terminal 485 bp could excise precisely but when the whole IS50 was present it could not excise precisely (Goldberg et al., 1990).

Inverted and direct repeats at the ends of an IS very seldom are a perfect match usually differing by only one or two bases but sometimes as much as 50% of bases. As an example, IS50 has non-identical ends termed inside end and outside end and this configuration has a profound effect on translation (Makris and Reznikoff, 1989).

IS1 is found in large numbers in the chromosome of some *Enterobacteriaceae* species. In fact, in *Shigella dysenteriae* there are up to 200 copies per chromosome that vary in composition from the original IS1 sequence by as little as two bases up to almost 50% of sequence. These are

14

called iso-IS1 elements (Ohtsubo et al., 1981).

As shown in Table 3, IS elements have been identified in *Thiobacillus, Proteus, Clostridium, Enterobacter, Acetobacter, Bradyrhizobium, Haloferax, Porphyromonas, Leuconostoc, Coxiella, Thermus, Lactococcus, Leptospira, Borrelia,Vibrio, Klebsiella, Neisseria, Rhodopseudomonas, Haemophilus, Camplyobacter and Acinetobacter* as well as *Trypanosoma, Sacchromyces, Nicotiana, Aspergillus, Bombyx-mori*, maize, wheat, *Drosophila*, algae, mice, herpes viridae, human genes, soybeans and *Spodoptera*. This is by no means a complete list and as will be seen many of the IS elements reported for one organism may be found in others. This is especially true of the *Enterobacteriaciae*, probably because they are among the most studied organisms to date.

As will be deduced from this section, some of the hard and fast rules of what make up an insertion sequence may need to be redefined in the future. Some generally accepted IS elements such as IS 200 and 492 in Pseudomonas and an undesignated IS from *Cyanobacterium anabena* have all the accepted characteristics of IS elements but do not contain inverted repeats. In some cases, IS mediated effects do not result in duplication of the IS. Sawyer et al., (1987) studied the ECOR reference collection of 71 natural isolates of diverse geographical and host origins of *E. coli* for the presence of six unrelated IS elements. They were IS elements 1, 2, 3, 4, 5 and 30. This study was undertaken to test nine models differing in strength of regulation of

15

Table 3 Distribution of IS in Nature

Bacterium	IS designation	size (bp)	#bp target	IR or DR	size DR	# of ORF's
Agrobacterium tumefaciens	427	1.1 kbp	2	IR	39	2
Bacillus	240					
Bacteroides fragilis	4351	1553	3	IR	25	2
Bordetella pertussis	481v1					
Bordetella pertussis	481v2					
Caulobacter crecentus	298					
Corynebacterium diptheria		1500	9	DR		
Citrobacter freundii	IS 1-like	.7-.8 kbp				
Cyanobacterium anabena		1351	dnd*	no IR's		1
Cyanobacteria calothrix	701	1.4 kbp				
E. coli	1	800		IR		
E. coli	2	1322		IR	4	1
E. coli	3	1258	3	IR	39	2
E. coli	4	1426	11	IR	18	
E. coli	5	1195				3
E. coli	10	1.3 kbp				
E. coli	30					
E. coli	46	810				
E. coli	50	**				2
E. coli	91					
E. coli	101	**				
E. coli	102	**				
E. coli	103	1441	6	IR	23	
E. coli	186					
E. coli	903	**				
gram-ves	150	1443	3	IR	32	3
Group B strep	861	1442	3	IR		2
Halobacterium halobium	H1	***				
	H1.8	***				
Halobacterium halobium	H2	***				
	H23	***				
	H24	***				
	H25	***				

16

Organism	Strain	Size		IR		
	H26	***				
	H27		5	IR	16	
	H28	***				
	H50	***				
Lactobacillus caseii	ISL1					
Mycobacterium pseudotuberculosis	900	1.45 kbp				
Mycoplasma pneumonaie		1550		IR	28	2
Pseudomonas atlantica	492	1202	5		0	1
Pseudomonas aeruginosa	21	2131	4	IR	11	2
Pseudomonas aeruginosa	22	7.3 kbp				
Pseudomonas aeruginosa	222	1350		IR	40	
Pseudomonas syringae	51	1.3 kbp				
Pseudomonas cepacia	931	1477				
Pseudomonas cepacia	932	3400				
Rhibozium lupini	ISR1	1.15 kbp		IR		
Rhizobium melioti	Rm2	2.7 kbp	8	IR	22	
Salmonella	160					
Salmonella	200	700	0			
Shigella sonnei	600	1264	3	IR	29	2
	629	1310		IR	25	
	630	1159		IR	32	1
	640	1092		IR	15	1
Staphlococcus aureus	257					
Strep coelicolor	110					
Streptococcus lactis	1 W	808	8	IR	18	1
	1N	809	8	IR	18	1
Streptomyces lividans	493	1.6 kbp		IR		2
Yersinia enterocolitica	unnamed	1.2 kbp				
Yersinia pestis	100	2.2 kbp				
Xanthomonas	511					
Xanthomonas campestris	476					

Key:
* does not duplicate
** characterization either incomplete or done previously
*** all as yet uncharacterized

17

transposition and effect of copy number on reducing fitness.

A number of interesting observations were made, one of which was that the number of copies of IS increased rapidly in a relatively short time after introduction. They also ascertained that the number of IS elements and copies of IS elements per strain varied greatly. Of the seventy-one strains, only two had none of the five IS elements tested and one strain had only one copy of one IS element. The rest of the strains had at least two known IS elements present and in multiple copies. Some strains had as many as sixty-two copies of IS elements present at an average of 1.2 kbp. Thus, a total of 74 kbp of sequences were essentially devoted to the ability to rearrange the chromosome. Since the average *E. coli* contains 4×10^3 kbp, some *E. coli* cells contain IS elements that constitute 2% of the genome, most of which is devoted to chromosome rearrangement. In no manner is this an isolated case. I mentioned above that in *Shigella dysenteriae* as well as other *Enterobactericiae* there are up to 200 copies of IS 1 on the chromosome (Ohtsubo et al., 1981). This would add up to an even higher percentage of the genome dedicated to rearrangement.

Another interesting aspect of this study was the finding that the majority of the IS elements were found on the chromosome. In some cases, *E. coli* strains harbored as many as twenty-seven copies of one IS on its chromosome, in addition to multiple copies of other IS elements as well. Although some strains had as many as fifty total copies, they only contained

18

two copies on their resident plasmids and, in fact, many strains with as many as forty-one chromosomal copies of IS elements had no IS elements on their resident plasmids. Distributions like this contradicts the contentions of some researchers that the distribution of IS elements is evolutionarily recent and due to dissemination by plasmids. The conclusion of the researchers in this study was that transposition is mildly regulated and that harmful effects on fitness increase mildly if at all with an increase in copy number (Sawyer et al., 1987).

Another study (Hall et al., 1989) examined the distribution of a new IS, IS 103, on the same ECOR collection of seventy-one natural isolates of *E. coli*. Thirty-six of seventy-one strains contained IS 103 and using the same models as the previous study, there was a strong indication of regulation and a weak to non-existent effect on fitness. It was found that IS 103 had a strong preference for plasmids in that sixty percent of the copies could be found on these replicons. On the other hand, IS 1-5 and IS 30 could be found on plasmids only ten percent of the time. The suggestion was that IS 103 was evolutionarily "newer" and was disseminated by plasmids which have not yet been lost.

Much research has been done to place individual IS elements phylogenetically. The result is a feeling for ascent from a small number of IS elements to the large number that is seen today. A study was undertaken (Rouch and Skurry, 1989) on the pivotal role thought to be played by IS

19

elements in the emergence of multi-drug resistance in pathogenic organisms like *Streptococci* and *Staphlococcus aureus*. IS 257 of *S. aureus* is found on many of the pSK1 multi-resistance plasmids, aminoglycoside resistance plasmids, mercury and heavy metal resistance plasmids as well as chromosomal determinants for the same resistances. Very often IS 257 is present in multicopy with these resistance genes, and as such forms composite transposons. In fact, sequence analysis is complete on six transposons of the IS 257 family and significant homology exists with the IS 15 family which is found widely among R plasmids in enterobacteria, *Acinetobacter* and *Camplyobacter*. Included in the IS 15 family are IS 15, IS delta15, IS 26, IS 46 and IS 140. Homology was also detected to ISS 1 from *S. lactis*. These three families of IS have from forty to sixty-four percent sequence homology in nucleotides as well as polypeptide sequences. Twelve out of fourteen bases are conserved in their terminal inverted repeats.

Another interesting aspect of this study is the idea that base substitutions by IS elements are allowable and necessary to bring %GC content into agreement with the host's %GC. The conclusion is that these three families span a proposed IS 257 superfamily of IS elements and that evolutionarily, this family goes back in time before the split of gram positive and gram negative bacteria (Rouch and Skurry, 1989).

The distribution of IS elements in the ECOR collection of *E. coli* natural isolates was used to do a phylogenetic analysis by IS fingerprinting.

20

The conclusion of these researchers is that the pattern of distribution of IS elements in the bacterial genome is sufficiently stable for epidemiological studies (Lawrence et al., 1989). Epidemiological studies have been carried out using IS distribution in *Salmonella* (Stanley et al., 1991), *Mycobacterium tuberculosis* (Otal et al., 1991) and *Enterobacter cloacae* (Matsutani, 1992).

In most organisms in which some DNA sequencing had been performed, IS elements have surfaced. It would not be surprising if many more IS elements remain to be found. To date, approximately ten per cent of the *E. coli* genome has been sequenced and analyzed. It will be interesting to see how many more IS elements will be found in the remaining ninety percent.

Structure and Function

In some cases, IS elements cause insertions which in themselves cause a variety of effects, such as deletions and rearrangements. The rate of transposition is typically at least an order of magnitude greater than the rate of deletion (Sawyer et al., 1987). It has been assumed that IS elements that transpose by the non-replicative mechanism, i.e. those that do not leave a copy of the IS behind, cannot give rise to deletions. Adjacent deletions have been demonstrated, apparently by intramolecular recombination, at a low frequency with IS 10 (Roberts et al., 1991).

Tn10 is 9300 bp and has inverted repeat ends of IS 10 and genes for tetracycline resistance in the non-repeated central portion of the transposon.

21

The IS 10's are denoted IS 10$_R$ and IS 10$_L$. IS 10$_R$ is 1329 bp containing one long open reading frame encoding transposase. It also encodes three promoters (Simons, 1983), pIN which is oriented inward (toward transposase), pOUT which is a strong promoter, is oriented away from the transposon. Another promoter pIII is a weak promoter found near the 5' end of the transposase gene oriented toward the inside (toward Tc^r genes of Tn10). IS 10-left is defective and relies on the transposition functions of IS 10-right (Simons and Kleckner, 1983).

Tn10 and IS 10 transpose at a low frequency. About once in 10^3 generations for IS 10, once in 10^7 generations for Tn10 and once in 10^5 generations for rearrangements. In every instance, active transposase is required in addition to a host protein either IHF or HU which probably alter DNA structure. When they do transpose, they do not leave a copy behind and only a little repair synthesis is done at the very ends (Raleigh and Kleckner, 1986). This is in contrast to IS 21 which is a 2131 bp IS with 11 bp IR's at either end. Unlike IS 10, IS21 has two adjacent reading frames, istA and istB. IS 21 is normally a resident on plasmids and produces replicon fusions at high frequency in *E. coli*. This is contingent on the presence of two tandem copies of IS 21 on the plasmid. In fact, IS 21 does not express *ist*A and *ist*B if only one copy is present and results of promoter analysis indicated that. IS 21 acts as a mobile promoter which causes the expression of genes downstream from the insertion site (Schurter and Holloway, 1986). This is

22

why tandem IS 21's are needed to express the *ist*AB operon, resulting in replicon fusions. IS 21 in single copy is known to transfer poorly, probably due to the need for a promoter. Once the transposase is produced in the tandem arrangement, it opens the plasmid at the IS 21-IS 21 junction and insertion is by the "cut and paste" method into the target replicon. The high transpositional activity of plasmids carrying tandem IS 21's is due to transposase specificity for IS 21 junctions rather than outer ends (Riemmann et al., 1989).

This situation is similar to IS 50 in that a 19 bp segment of the inside end is needed for efficient transposition (Dodson and Berg, 1989). This requirement for certain sequences is a result of interaction with not only the transposase but associated host proteins as we saw with IS 10 and will see with other IS's as well. However, it is not the case for IS 10 as it is for IS 21 to require an external promoter to produce transposase. A different situation exists for IS 1 which, like IS 21, has two coding frames essential for its transposition. IS 1 is further like IS 21 in needing an alteration to bring about the production of transposase. In the case of IS 1, a -1 frame shift from the *ins* A frame is required to read the *ins* B frame (Sekine and Ohtsubo, 1989). This frame shift occurs in the sequence AAAAAC where the two coding frames overlap. It is interesting to note that this signal to shift reading frame is also present as a frameshift signal in retroviruses transframe production of poly protein. Further diversity of the structure of IS elements is provided by IS 5

23

which contains three open reading frames. One long ORF (Rasmussen et al., 1987) oriented in one direction has a weak promoter and a ribosome binding site while the other two shorter ORF's (5a + 5c) have their corresponding promoters and ribosome binding sites oriented in the opposite direction. The native production of 5B protein appears to be very weak but when joined to a tac promoter is quite strong (Chernak and Smith, 1989). The *in vitro* orientation may be analogous to the situation with IS 21 requiring a tandem to activate a promoter region and this may also be the case *in vivo* in that 5B protein is produced when an external promoter is present. However, this situation is not always the case. It has been demonstrated (Riemmann et al., 1989) that external promoters have a negative effect on transposition frequency and synthesis of transposase from IS 1, IS 10 and IS 50.

One of the most apparent features of IS structure is their differences in the strength of promoters. The direct consequence of this is the frequency of transcription initiation that takes place and the amount of transposase that is produced varies from IS to IS depending on promoter strength. In the case of IS 10, the amount of transposase that is produced per cell per generation is known and sheds some light on regulation of IS's. IS 10 transposase genes were fused to lac Z (Raleigh and Kleckner, 1986), and allowed to produce protein. Analysis of the products reveals an average of 0.58 of translational product per transcript. It has been shown that approximately twelve units of transposase are required for Tn10 transposition (Huisman et al., 1990),

24

indicating that average transcription is insufficient for the observed frequencies of transposition, especially in the light of the fact that transposase's tend to be degraded easily *in vivo*. This was determined using the transposase of IS 903 (Derbyshire et al., 1990). The transposase produced in a transposase-β-gal fusion was more stable than native transposase and could act equally well in cis or trans. This is contradictory to native transposases which act almost exclusively in cis. It also should be noted that the La protease, a product of the *lon* gene, was found to be an important determinant of transposase stability. Stability of native transposase in *lon-* strains increased 10-100 fold.

There are several factors required to regulate initiation of transcription as well as regulation at the translational level. The same numbers hold for the relative amounts of transposase produced and frequency of initiation of translation in IS 10.

Two factors have been investigated as to why transcription from the transposase gene promoter is infrequent. One of the most important factors is probably the degree of *dam* methylation (Roberts et al., 1985). Tn10, IS 10, Tn5, IS 50, Tn903, IS 903 have all been shown to undergo increased transposition in *dam-* strains in which DNA adenine methylase is inactivated. The loss of methylation of two *dam* methylation sites (GATC) within IS 10, one of which overlaps the -10 region of the transposase promoter and the other which occurs within the transposase binding site at

25

the other end of IS 10, not only reduces the level of transcription but ensures that it occurs only at the correct time. DNA in w.t. *E. coli* is never unmethylated and only briefly hemimethylated during passage of the replication fork. This hemimethylation results in the activation of IS 10 transiently when an IS 10 enters a new host by a mechanism that involves transfer of single stranded DNA followed by synthesis of the complementary strand in the recipient, because IS 10 transposes by a non-replicative mechanism (Kleckner, 1990). These findings are further complicated by the orientation of the IS ends, relative to the transposase gene *tnp*. In studies with IS 50 (Dodson and Berg, 1989) the presence of an inside end and an outside end is the preferred arrangement for most efficient transposition even in a *dam*- background. This clarifies somewhat why inverted repeats are the preferred arrangement for IS elements since a direct repeat configuration results in inside end-inside end or outside end-outside end structure resulting in significantly reduced transposition. This preference for orientation of IR's of the IS is due to the mechanism of transposition by transposase. Sequences on the ends of the IS are recognized and bound by the transposase prior to transposition. It must be remembered here that the IR's are frequently imperfect so IR's on each end of the IS will bind the transposase ends with differing affinity.

There are other controls at work to reduce or limit the level of transposase in the cell. Translation is inefficient due to a lack of obvious

26

Shine Delgarno ribosome binding sequences, at least in IS 10, 50 and 903. Translation is confined to initiation in the first few seconds after transcription is initiated. This is the result of two factors. Fold back inhibition reduces translation several fold by burying the ribosome binding site at the beginning of the message. Failure to initiate translation at the AUG start codon reduces the stability of the 5' end of the mRNA (Kleckner, 1990). Possibly of more importance is the situation revealed, again in IS 10, where the mRNA encoded by pIN that codes for transposase (RNA IN) is destabilized by the message encoded by opposing and stronger promoter pOUT. In this situation, RNA OUT and RNA IN pair in a 35 bp region of homology that includes the ribosome binding site of *tnp* (Case et al., 1990). The pairing is subject to ribonuclease III cleavage which limits the life of the RNA. The important part in regulating translation is the demonstration in vivo that the pairing of antisense RNA OUT to sense RNA IN has been shown to block ribosome binding (Ma and Simons, 1990). Inhibition by RNA OUT has been shown to function in trans which has wide implications for reducing transposition when an element is present in more than one copy in the cell. Couple this with the idea that transposase is cis-acting because its action is linked to its transcription by *dam* regulation and you see a regulatory process to reduce transposition when an element is present in more than one copy (Kleckner, 1990).

Defense mechanisms do exist for protecting IS 10, and presumably

27

other IS's, from activation by external promoters. This is important because read-through transcription would disrupt both *dam* methylation which depends on specific activation of pIN, and inhibition by RNA OUT which requires a precise 5' end of RNA IN. Externally activated transposase genes yield less than 1% of the transcripts than from pIN activated transposase. This is due to a number of factors. RNA-IN which is externally activated lacks the structure to pair with RNA OUT. Transcription from an external promoter inhibits transposition even when transposase is provided in trans, as the result of formation of secondary structure. *dam* methylation provides protection at this point also since the transposase made by external promoters will only be active in the small fraction of the cell cycle when the inside end is activated. Lastly, transposition into actively transcribed regions is inhibited so external promoters are less likely to be active (Kleckner, 1990).

Host factors are also known to influence transposition. IS 1 and IS 5 frequencies of transposition are reduced in Rho mutants (Datta and Rosner, 1987) of *E. coli.* In most IS examined so far Integration Host Factor (IHF) plays a major role in target specific action of transposase. IHF has been implicated in binding to one or both ends of the IS or transposons and directing the IS to the target site. The requirement for IHF is not rigid as transposition can be demonstrated in *ihf-* cells.

It is quite likely that there are more mechanisms of regulation to be discovered both internally and externally to IS elements especially since

28

analysis of gross structure alone shows that many IS do not have the capacities for regulation that IS 10, 50, 903, and others do.

Effects caused by IS

The discovery of controlling elements by Barbara McClintock (1941) was favored by the highly visible mutations they cause in maize. The discovery that IS elements also existed in procaryotes was again facilitated by highly visible mutations within the galactose operon of *E. coli*. Independent analysis on the causes of severe polar effects on the *gal* operon lead Jordan and Shapiro to the conclusion that insertion sequence elements function in procaryotes. Since those early beginnings IS elements were located mostly by the very visible effects that they caused. In the last five to six years, the development of relatively simple procedures for DNA sequencing and molecular cloning has dramatically increased our knowledge on how IS elements affect gene expression in many microorganisms and several examples are provided.

Effects on Plasmids. When the incP plasmid RP4 is transferred to *Rhizobium lupini* it is quite unstable. This is in contrast to the relative stability of RP4 in other bacteria. In *Rhizobium*, plasmid loss occurred 60% of the time and 10% of the time plasmids displaying mutations arose. An analysis of the plasmids indicated that they contained ISR1, a new *Rhizobium* IS element (Preifer et al., 1981). Similar effects were seen in R plasmids contained in *E. coli*. In this case, however, rearrangement was not as frequent

29

and was due to IS1 (Chandler et al., 1981). IS27 has been found in the *bop* region of plasmids associated with *Halobacterium halobium*. In fact, there are at least eight IS elements found because of frequent deletions occurring on all *Halobacterium* plasmids (Pfeifer and Blaseio, 1989). In addition, the plasmids resident in *Streptococcus lactis* are non-conjugal but become conjugal upon insertion of an IS26-like sequence called IS S1S (Haandrikman et al., 1990). Repression of conjugal transfer of R plasmids is due to two genes *fin* O and *fin* P. Constitutive transfer of F is believed to be due to lack of *finO*. In fact, F does contain the *finO* sequences but are inactivated by the insertion of IS 3 (Yoshioka et al., 1987).

Effects on Chromosomal Genes. In an attempt to clone the *ilv*A gene from *Pseudomonas cepacia* into *E. coli*, the *ilv*A locus was poorly expressed. However, isolates that grew better were selected and found to contain IS2 (Barsomian and Lessie, 1987). IS2 also activated the *trp* gene in *Saccharomyces cerevisiae* (Broisus and Walz, 1983), the *arg* genes from *Methanococcus voltae* (Wood and Konisky, 1985) and the *arg* genes from *Aspergillus nidulans* (Dmochllowska et al., 1986). *E. coli* does not grow normally on salicin but can be induced to grow on prolonged incubation in the presence of salicin. Growth on salicin requires two mutations and the excision of IS103 from the *bgl* operon (Hall, 1988).

The *erm*F locus codes for an rRNA methylase in a number of gram positive organisms. The *erm* operon codes for resistance to macrolide-

30

lincosamide-streptogramin B resistance. IS 4351 provides transcriptional start signal for *ermF* and possibly *ermC* and *ermA* as well (Rasmussen et al., 1986). The transposon Tn3411 confers on the cell the ability to utilize citrate. When inserted into pBR322, Tn3411 caused deletion of citrate genes by intramolecular recombination between two copies of IS3411 (Ishiguro and Sato, 1985). An insertion sequence IS 5B is present in the 3' end of the *ompC* gene which is a porin and doubles as receptor for lambda. The IS 5B is also found within a defective gsr' phage (Highton et al., 1985).

The exopolysaccharide xanthan is produced by *Xanthomonas campestris* pv. *campestris*. Insertion of an uncharacterized IS produced surface mutants deficient in exopolysaccharide production and this deficiency could be restored in trans (Hotte et al., 1990). In *Xanthomonas campestris* pv *vesicatoria*, the acquisition of virulence was found to be caused by insertion of IS 476 into the *aur*Bs1 locus (Kearney and Staskawicz, 1990). Another example of insertional mutagenesis is in the *cya* locus of *E. coli*. *cya* is the structural gene for adenylate cyclase. Mutation of *cya* appeared as possible regulatory control because mutants failed to synthesize cAMP or grow on maltose, arabinose, lactose or glycerol. *cya* R1 was found to be the result of the insertion of IS 1 (Glaser and Danchin, 1989). IS 481v1 and IS 482v2 in *Bordetella pertussis* cause a loss of expression of adenylate cyclase and agglutinin genes (Mendiola and De la Cruz, 1989). As the result of the presence of IS 931 and IS 932 *Pseudomonas cepacia* (Haugland et al., 1990) has

31

the ability to metabolize 2,4,5 trichlorophenoxyacetic acid. The same strain in the absence of IS 931 and IS 932 cannot degrade 2,4,5 -T (Barsomian and Lessie, 1987). IS 1, 2 and 3 have been shown to interrupt *rho* causing the production of a truncated *rho* protein, resulting in the lack of termination of transcription at *rho* sensitive sites (Datta and Rosner, 1987).

Current Usages of IS Elements

With the elucidation of the structure and control of IS elements and the mechanisms by which they transpose, it is only logical that some innovative uses would be developed for the IS elements. One of the earliest that I am familiar with was the analysis of structure of DNA by electron microscopy to identify the stem loop configurations of IS elements. This obviated the need for other methods to alleviate the tedium involved with isolating IS elements. One method that I used was the DNA denaturation-brief renaturation procedure (Ohtsubo and Ohtsubo, 1976) that takes advantage of the presence in a DNA sample of a fairly large number of homologous sequences that are more likely than surrounding DNA to renature. But still more specific procedures were required for IS elements that are of low copy number. Gay et al., (1985) developed a positive selection vector pUCD800 which is designed to entrap IS elements from gram-bacteria. The heart of the vector is the *sac* B gene from *Bacillus subtilis* which does not allow for growth on media containing 5% sucrose unless *sac* B is inactivated (Demeirsman et al., 1989). A similar system was developed for

32

Streptomyces utilizing the C1 857 gene on a plasmid. Upon disruption by transposition the host becomes apramycin resistant (Solenberg and Burgett, 1989). Both have the disadvantage of requiring active transposition. In a great many cases IS need to be induced to transpose.

IS elements have been used in other systems to effect transposon mutagenesis by artificial transposons that facilitate subsequent cloning of the mutated gene. One such system based on IS 1 generates insertional mutagenesis in a wide range of gram negatives at frequencies ranging from 10^{-5} to 10^{-9} (Joseph-Liauzum et al., 1989). The same strategy has been exploited in *Rhizobium* via plasmid containing Tn5 which was mobilizable due to cointegrate formation by IS 150 (Donald et al., 1985).

Amplification of drug resistance genes is possible in transposons flanked by inverted repeats of IS 1. Recombination between the IS 1 sequences resulting in the first tandem duplication appears to be the rate limiting step for further amplification (Iida et al., 1987). Another valuable usage for IS is for differentiation of *Mycobacterium psedotuberculosis* from *M. avium*. *M. PTB* causes severe disease in cattle whereas *M. avium* does not. IS 900 was used as a probe for differentiating the two species (Vary et al., 1990).

Mutations Caused by IS That Affect Virulence

The role of IS's in virulence is implicated strongly by the close association of IS elements with antibiotic resistance genes. These usually

33

result in the creation of transposons of which there are numerous examples. However, there is a myriad of effects caused by IS elements that directly effect virulence. In the case of *Yersinia*, the insertion of IS 100 into the region of Ca^{++} dependency results in a reduction of virulence. In *Klebsiella pneumoniae*, IS 26 has been implicated directly in the development of a new antibiotic resistance operon. In this case, the IS element provides a much stronger promoter region by replacing the weak promoter formerly controlling it (Lee et al., 1990).

A report in the Russian literature implicates IS elements in rearrangements of penicillinase plasmids in *N. gonorrhea*. Another report implicates IS elements in *N. gonorrhea* in the variability of surface antigen, possibly affecting virulence (Roy et al., 1988). IS's have also been implicated in recombination between silent and expression pilin loci in *N. gonorrhea* (Swanson et al., 1985). *Xanthomonas campestris* pv. *vesicatoria* can become virulent to tomatoes and peppers by the insertion of IS 476 into the *avr*Bs1 locus for copper resistance. The transposase of IS 476 is homologous to the transposase of IS 3 and the inverted repeats of IS 476 show significant homology to the IS 51 of another plant pathovar *Pseudomonas syringae* (Kearney and Staskawicz, 1990). In the *erm* BC locus, which codes for resistance to macrolides, IS 15 and IS 1 are implicated in transposition of this locus to gram positives. The IS 15 transposes into sequences of gram positive origin followed by IS 1 inserting into IS 15 and mediating rearrangement such

34

as transposition and integration (Brisson et al., 1988). In a situation analogous to the fundamentals of my research, *Mycobacterium pseudotuberculosis* causes severe disease in cattle whereas the very closely related *Mycobacterium avium* does not. It was discovered that *Mycobacterium pseudotuberculosis* contains a 1.45kbp IS (IS 900) whereas *M. avium* does not (Vary et al., 1990, Kunze et al., 1992). In the marine bacterium *Pseudomonas atlantica*, IS 492 causes reversible inactivation of extracellular polysaccharide. This is caused by variable insertion and precise excision from the *eps* locus which could be an antigenic variability strategy (Bartlett and Silverman, 1989). In *E. coli*, a copy of IS 2 is contained within the *iss* gene which encodes for increased serum survival and surface exclusion (Chuba et al., 1989). In *Shigella flexneri*, IS are involved in a couple of known virulence functions. Most enteric bacteria including *E. coli* and *Shigella flexneri* utilize high affinity transport systems in iron deficient environments to acquire iron for growth. These transport systems consist of low molecular weight Fe^{3+} chelating compounds called siderophores and various membrane proteins required to transport the iron bound siderophone through the cell envelope. Enterobactin and aerobactin are examples of siderophores. In the *ent-* locus of *Shigella flexneri* as well as *E. coli*, IS 1 has been found. Such insertions lead to loss of virulence because of reduced ability to compete for iron (Schmitt and Payne, 1989). IS elements are also implicated in phase variation in *Salmonella typhimurium* resulting in virulence changes (Bartlett and

35

Silverman, 1989) and IS 1 is also found present in the aerobactin genes of *Salmonella* and *Aerobacter* as well (Colonna et al., 1985). Also, in *Salmonella*, frequent rearrangements of the drug resistant R plasmids occur and these are due to multiple copies of IS 160 (Nies et al., 1985). Similar to the phase variation of *Salmonella* is the reversible expression of capsular antigen Vi in *Citrobacter freundii*. The loci *via*A and *via*B are the components of the Vi operon. Analysis of Vi strains reveal the presence of a .7kbp IS1-like element in the *via*B gene. Excision resulted in Vi+ (Ou et al., 1986). In *Shigella flexneri*, a spontaneous insertion of an IS element into the *vir*F gene resulted in the inactivation of this plasmid encoded positive regulator of *ipa* gene expression, resulting in the loss of virulence (Mills et al., 1992). The capsulation gene cluster of *Haemophilus influenzae* has been mobilized in the chromosome as a compound transposon, as a result of the insertion of IS 1016. This has resulted in amplification of the capsulation product resulting in biological advantage (Kroll et al., 1991).

Unusual Properties of IS or Tn's

IS 4 is found in the *E. coli* chromosome but no matter what the strain, it is found in the same region. There are actually three insertion sites within this region and one is the *gal* T gene. The surrounding DNA region is different and when IS 4 deletes, it also deletes variable lengths of surrounding DNA (Klaer et al., 1981). An unusual occurence among IS is the IS 511 of *Caulobacter crecentus* which does not cause a duplication of the target site, felt

36

to be due to the creation of blunt end cuts prior to transposition rather than staggered cuts (Ohta et al., 1990). An example of an unusual structure of an IS is IS 492 of the marine bacterium *Pseudomonas atlantica* which meets all the criteria of an insertion sequence but there are no inverted or direct repeats at the termini (Bartlett and Silverman, 1989).

IS elements in plasmids containing Tn5 can be mobilized into *Rhizobium* and genomic integrates stably formed. The unusual thing is that IS 50 transposes independently to give IS 50 triplicate (IS 50 is at both ends of Tn5). It is felt that this triplicate is a result of cointegrate formation due to IS 50 (Donald et al., 1985). A similar situation has been shown to occur with IS 30 in which cointegrate formation of replicons carrying IS 30 did not leave intervening IS elements at the junction of cointergration. A mechanism for these types of transposition is proposed (Dalrymple, 1987). Although at present it seems strange, it may turn out that IS 103's strong affinity for plasmids over chromosomes may be the definition of another class of IS. It may also indicate the presence of some other mechanism(s) in target selection by IS elements (Hall et al, 1989). IS 630 of *Shigella sonnei* is unusual in that it cannot mediate cointegration between two replicons in *E. coli* which is unusual but in this respect alike to IS 10 and IS 50 (Tenzen et al., 1990). IS 150 of *E. coli* is highly selective of IS 1 as a target for integration (Schwartz et al., 1988). In *Bacteroides*, Tn4351 can mediate integration of the incP plasmid R751 of *E. coli* into the *Bacteroides* chromosome. R751 is then unable to

37

mobilize itself (Shoemaker et al., 1986).

IS 1 has been shown to contain 2 coding frames *ins*A and *ins*B which are essential for transposition. An event which results in a -1 frameshift near the 3' region of *ins*A toward an open reading frame, B' frame, is involved in the production of *ins*A-B'*ins*B transposition protein. This frameshift occurs identically to one of two sequences identified in retroviruses as frameshift signals for production of transframe proteins from overlapping genes (Sekine and Ohtsubo, 1989). It has also been shown that IS 3 and 5 cause mutations in lambda, and IS H 1.8 of *Halobacteria* is contained in a prophage. IS 121 is found on Mu as well as IS on the herpes virus. Another IS element IS222 (Gertman, et al., 1986) was found to transpose to or from its lysogenized phage D3 which is a member of *Siphoviridae*.

The 4.2kbp cryptic plasmid of *N. gonorrhoeae* is described as a variable plasmid. It has within it two 44 bp repeats that have been reported to be involved in site specific recombination. A novel insertion of 156 bp. causes an alteration in the B protein which is a variable surface antigen of gonococcus (Roy et al., 1988).

Materials and Methods

Bacterial Strains and Growth Conditions

The strains used in this study are listed in Table 4. *E. coli* for plasmid isolations was grown in Luria-Bertanni (LB) medium (Maniatis et al., 1982A) at 37°C with appropriate antibiotics. Unless designated otherwise, *Y. enterocolitica* was grown in Heart Infusion Broth (HIB, Difco Laboratories) at 37°C while *Y. pestis* and *Y. pseudotuberculosis* were grown in HIB at 26°C. Selective agents for this study were ampicillin (amp) at 50 or 100 μg/ml and tetracycline (tet) at 25 μg/ml. Selection for recombinant plasmids was on Violet Red Bile agar (VRB, Difco laboratories). Cells capable of metabolizing lactose produced purple colonies while those unable to metabolize lactose produced white colonies. Cultures were maintained in LB or HIB broth with appropriate antibiotics and 10% Dimethyl Sulfoxide (DMSO) at -70°C.

Isolation of Genomic DNA

Prokaryotic genomic DNA was isolated by a modified method (Marmur, 1961). Bacterial cells were grown to log phase in 1 L HIB (Difco Laboratories) at 37° C and collected by centrifugation at 10,000 x g for 10 min. The pellet was washed three times in 100 ml saline-EDTA (0.15 M NaCl, 0.1 M EDTA[pH 8.0]). The pellet was resuspended in 40 ml saline EDTA and 400 μg/ml lysozyme. The mixture was held at 37°C for 30 minutes; then 1.0 ml of 25% SLS (sodium lauryl sulfate) was added, the mixture shaken and held at 60°C for 15 min. The solution was cooled to 26°C

39

TABLE 4. BACTERIAL STRAINS

Strain	Plasmid Content	Revelant Genotype	Source
Escherichia coli			
DH5α	-	*end*A1, *hsd*R17(r-k,m+k), *sup*44, *thi*-1, *rec*A1,*gyr*A96, *rel*A1, Δ(*arg*F-*lac*ZYA)U169, φ80d*lac*ZΔM15, F-, λ-	BRL, Gaithersburg MD
JM83	-	φ80d*lac*ZΔM15, *ara*, Δ(*lac*-*proAB*), *rps*L, *thi*, F-	John Hopper
JM83	pUC19	same as JM83 without pUC 19, *amp*+	John Hopper
JM83	pISC	*amp*-	this study
LCD44	-	*recA*-*srl*, *lac*U169, *ara*D139, *thi*, *non*, *nal*A, *rps*L, *met*E	John D'Elia
SURE	F′ *proA*B *lac*19ZΔ15 Tn10	Δ(*hsd*RMS), *mcr*B,*mrr*, *end*A1, *sup*E44, *thi*-1, *gyr*A96, *rel*A1, *lac*, *rec*B, *rec*J, *sbc*C, *umu*C, *uvr*C, λ-, *e*14-, *pin*, *lit*, *sfi*C	Stratagene La Jolla CA
SURE ISC	same as SURE and pISC	same as SURE and *amp*+, *lac*-	this study
SURE IS9	same as SURE and pIS9	same as SURE and *amp*+, *lac*-	this study
SURE PUC	same as SURE and pUC19	same as SURE and *amp*+, *lac*+	this study
TB1	pUC19	*amp*+	John Hopper
Yersinia enterocolitica			
3973-76 0:4	47MD		David Yu
Yersinia pestis			
EV7651F	61,47,6MD	*lac*, *pgm*	D. Cavanau
EV7651F Ca-2	61,49,6MD	*lac*, *pgm*	R. Zsigray
EV7651F E6	6MD	*lac*, *pgm*	R. Zsigray
Yersinia pseudotuberculosis			
Ca- trp-	-	*trp*-, Ca-	R.Zsigray
Ca-trp-ISC	pISC	*trp*-, Ca-, *amp*+,*lac*-	this study
Ca-trp-IS9	pIS9	*trp*-, Ca-, *amp*+,*lac*-	this study
Ca-trp-PUC	pUC	*trp*-, Ca-, *amp*+,*lac*+	this study
w.t.	47MD		R. Brubaker

40

and 10 ml of 5 M $NaClO_4$ and 50 ml of a 24:1 mixture of chloroform isoamyl alcohol was added. The mixture was shaken for 30 min on a wrist action shaker and centrifuged at 10,000 x g for 10 min. The supernate was treated with RNase A (final conc. 10 μg/ml) and incubated at 37°C for 15 min. An equal volume of chloroform isoamyl alcohol (24:1) was added and was again shaken for 15 min. These extractions were repeated until precipitate free. The supernate was overlayed with 2 volumes of 95% ethanol and the DNA collected by spinning onto a glass rod. The DNA was air dried and resuspended in TE buffer (0.01 M Tris, 0.001 M EDTA [pH 8.0]), then DNA analysis (Burton, 1956) was performed to determine concentration.

Isolation of Insertion Sequence (IS) Elements

IS isolation was carried out by the rapid renaturation procedure (Ohtsubo and Ohtsubo, 1976) as follows: 5 mg of DNA was resuspended in 0.7 ml sterile dH_2O and denatured in a boiling water bath for 5 min. The sample was immediately immersed in ice water and the salt concentration adjusted to 30 mM by the addition of 0.3 ml of 0.1 M NaCl. The denatured DNA was then exposed to 68°C for exactly 1 min and immersed in ice water. The DNA was precipitated with 2 volumes of 95% ethanol and pelleted at 10,000 x g for 10 min. The pellet was air dried and resuspended in 1 ml S 1 reaction buffer (30 mM Sodium acetate, 50 mM NaCl, 1 mM zinc acetate [pH 4.6]) and one unit of S 1 nuclease (Gibco/BRL, Gaithersburg, MD) was added per μg of DNA. The mixture was incubated at 37° C for 90 min, until the viscosity dropped to that of water. An equal volume of chloroform isoamyl alcohol (24:1) was added and the mixture shaken for 10 min, then centrifuged at 10,000 x g for 10 min. The extraction was repeated until precipitate free. The supernate was precipitated with 2 volumes of 95% ethanol and pelleted by

41

centrifugation at 10,000 x g for ten minutes. The pellet was air dried and saved for DNA gel electrophoretic analysis and cloning.

Plasmid Isolations

Plasmids used in this study are listed in Table 5. Plasmid DNA was isolated by the alkaline lysis method (Birnbiom and Doly, 1979) for use in agarose gel electrophoresis. Five ml of an overnight culture of cells were collected by centrifugation (10,000 x g for 5 min). The cell pellet was resuspended in 90 µl wash buffer (50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA, [pH 8.0]) and transferred to a microfuge tube. Lysozyme was added to a final concentration of 4 mg/ml. The mixture was stored on ice for 30 min at which time 200 µl of lysis solution (1% sodium lauryl sulfate, 0.2 N sodium hydroxide [pH 12.5]) was added. The tube was inverted several times and kept on ice for 5 min. Chromosomal DNA was precipitated by addition of 150 µl of 5M potassium acetate (pH 4.8). The mixture was kept on ice for 60 min, then centrifuged at 10,000 x g for 5 min. The supernate containing the plasmid was either precipitated with 2 volumes of 95%EtOH or used immediately for gel electrophoresis. For electrophoresis the pellet was resuspended in TE buffer (1 mM EDTA, 10 mM Tris-HCl [pH 8.0]) (Crosa and Falkow, 1981) and mixed in a 3:1 ratio with tracking dye (33% glycerol, 7% sodium lauryl sulfate, 0.07% bromphenol blue). Samples were loaded onto either horizontal (10 cm x 15 cm or 5 cm x 15 cm) or verticial (12 cm x 14 cm x 3 mm) 0.8% agarose gel. Electrophoresis was at 90V for 25 h for vertical gels and 50V for 2 h for horizontal gels. After electrophoresis, gels were stained for 30 min with ethidium bromide (1µg/ml) and viewed on a transilluminator.

Plasmid DNA used in DNA sequencing studies was isolated by a modification of the rapid boiling method (Wang et al., 1988). One and a half

42

TABLE 5. PLASMIDS

PLASMID	BACTERIAL HOST	ORIGIN OF INSERT	SIZE OF PLASMID	SOURCE
E6	EV7651F		9.6 kbp	R. Zsigray
p6E2	JM83	EcoR1 cut 6 MD	1.2 kbp	this study
p6E6	JM83	EcoR1 cut 6 MD	1.2 kbp	this study
p6E16	JM83	EcoR1 cut 6 MD	7.2 kbp	this study
p6PS4	JM83	Pst 1 and Sma 1 cut 6 MD	1.2 kbp	this study
p6PS5	JM83	Pst 1 and Sma 1 cut 6 MD	1.4 kbp	this study
p6PS6	JM83	Pst 1 and Sma 1 cut 6 MD	2.2 kbp	this study
pCl13	SURE	IS 100 cut w Pst 1	1.4 kbp	this study
pIS2	SURE	Y. pestis IS	1.2 kbp	this study
pIS9	SURE	Y. pestis IS	9 kbp	this study
pIS11	SURE	Y. pestis IS	9 kbp	this study
pIS13	SURE	Y. pestis IS	2.2 kbp	this study
pIS39	JM83	Y. pestis IS	1.2 kbp	this study
pIS42	JM83	Y. pestis IS	0.6 kbp	this study
pISC	JM83	IS 100 cut w Pst 1	0.8 kbp	this study
pISC	LCD44	IS 100 cut w Pst 1	0.8 kbp	this study
pISC	Y. P.T.B.	IS 100 cut w Pst 1	0.8 kbp	this study
pISC	SURE	IS 100 cut w Pst 1	0.8 kbp	this study
pISC	DH5α	IS 100 cut w Pst 1	0.8 kbp	this study
pUC19	JM83		2.7 kbp	J. Hopper
pUC19	LCD44		2.7 kbp	J. Hopper
pUC19	TB1		2.7 kbp	J. Hopper
pUC19	Y.P.T.B.		2.7 kbp	J. Hopper
VWa	EV7651F		75 kbp	D. Cavanaugh
VWa Ca-	Y. P.T.B.		77 kbp	R.Zsigray

43

ml of an overnight culture were pelleted by centrifugation (10,000x g for 1 min). The pellet was resuspended in 210 μl STET Buffer (8% sucrose, 0.5% Triton x-100, 50 mM EDTA, 50 mM Tris-HCl [pH 8]), 1 μl of RNase (40 mg/ml heat inactivated for 5 min at 100°C) was added as well as 15 μl of fresh lysozyne (10 mg/ml in TE buffer). The solution was boiled for 50 sec and transferred immediately to a microfuge and centrifuged for 10 min at 10,000 x g. The supernate was collected and placed on ice. One tenth volume of sodium acetate (3 M) and one volume of isopropyl alcohol was added, the tube inverted and placed on ice for at least 10 min. The preparation was centrifuged and the pellet rinsed in 70% ethanol, then dried. One quarter to one half of the DNA isolated by this method was used for sequencing.

Photography

Gels were photographed after staining with ethidium bromide while illuminated on a model C63 transilluminator (Ultra violet products Inc., San Gabriel, CA). Polaroid type 55 positive/negative film was used by exposure through a Kodak #16 gelatin filter. Lens distance varied by gel size and exposures were for 2-3 min depending on fluorescence levels. The f-stop was set at 4.7 and shutter speed at B.

Maxicell Procedure

Five ml of M9 (Sancar et al., 1979) medium (0.04 M Na_2PO_4, 0.015 M KH_2PO_4, 0.02 M NH_4Cl, 10 μM $MgSO_4$, 1 μM $CaCl_2$, 0.5% casamino acids, 0.005% thiamine-HCl, 0.4% glucose) were inoculated with 0.1 ml of an overnight culture of *E. Coli* LCD 44 with or without plasmid. Incubation was at 37°C until an OD_{420} reached 0.1. Cells were then irradiated in sterile Petri dishes at 360 μW/cm^2 for 16 sec, then transferred to foil covered tubes to prevent photo-reactivation. The cultures were incubated for 60 min at 37°C

44

and 0.1 ml was removed and plated on L-agar (1% tryptone, 0.5% yeast extract,1% sodium chloride, 1.5% agar) to determine post-UV survival rate. Twenty µl of freshly prepared cycloserine (50 mg/ml) were added to 5 ml of cells and incubated in foil covered tubes overnight. Cells were washed twice in 1x M9 salts (0.04 M Na_2PO_4, 0.015 M KH_2PO_4, 0.02 M NH_4Cl) resuspended in 1 ml of labeling media (1x M9 salts, 1 µM $CaCl_2$, 0.005% thiamine), and incubated at 37^0C. After 60 min, 300 µCi of ^{35}S-methionine (DuPont) was added and incubated for an additional 60 min. Cells were washed twice in M9 medium by centrifuging at 10,000 x g for 1 min and the pellet resuspended in 25 µl TE buffer. An equal volume of TE buffer plus 2% SLS was added, mixed and boiled for 3 min. Tubes were centrifuged at 10,000 x g for 10 min at room temperature, 5 µl were removed for scintillation counting and the rest stored at -20^{0}C for analysis by SDS-PAGE.

DNA Quantitation and Digestion

Chromosomal or plasmid DNA was quantified either by Burton analysis or by comparing the intensity of DNA on gels containing known standards. Appropriate amounts were digested by restriction endonucleases (Gibco/BRL Gaithersburg, M.D.) according to manufacturer's instructions.

DNA Purification and Ligation

Two methods were used to purify DNA for ligation. One method, Magic Mini Preps©, (Promega, Madison, WI) used a purification resin to bind DNA and to allow impurities to be washed away before the DNA was released from the matrix prior to use or storage. Manufacturer's procedures were followed.

Electroelution into membrane traps was employed to collect DNA from agarose gel slices. After visualizing the gels on a transilluminator,

45

bands were excised from the gels with a sterile scalpel and transferred to an Elutrap™ (Schleicher and Schuell, Keene, NH) electrophoresis system. Electroelution was in TE buffer (10mM Tris, 1 mM EDTA [pH 7.5]) according to manufacturer's specifications for the size fragment being eluted.

Ligations (Struhl, K., 1985) were performed dependent on the nature of the ends of the molecules to be ligated. In all ligations, a maximum of 1.0 μg of DNA was used. Vector to insert ratios were varied to preferentially allow the desired substrate molecules to join. Nine μl of DNA were combined with 10 μl of 2X ligase buffer and 1 μl 10 mM ATP. For blunt end ligations, 10U (units) of T_4 ligase (Gibco/BRL, Gaithersburg, MD.) were added and for cohesive ends 1 U of T_4 ligase was used. The mixture was held at 15°C for 12-24 hr. A 2 μl sample was removed to verify ligation and the rest was used to transform or electroporate host cells.

Bacterial Transformation

Transformation of *E. coli* strains was achieved by a slight variation of the calcium chloride method used to transform competent cells (Mandel and Higa, 1970). An overnight culture was diluted 1:100 into 40 ml of Lurian Bertani (LB) (1% tryptone, 0.5% yeast extract, 0.5% sodium chloride) medium, and incubated with shaking at 37°C for approximately 3 h until A_{600} = 0.6. Cells were collected by centrifugation for 5 min at 5000 x g at 4°C and washed in 20 ml 500 mM $CaCl_2$ for 30 min at 0°C. Cells were again collected by centrifugation for 5 min at 4°C. The pellet was resuspended in 2 ml of 500 mM $CaCl_2$ and held for 30 min on ice. Competent cells were used immediately by combining 0.1 ml cells 10 μl ligation mix, and 90 μl transformation buffer (0.09 M Tris-Cl, 0.01 M Tris [pH 7.2]) and incubating on ice for 10 min. The mixture was transferred to 37°C for 2 min and 1 ml

46

prewarmed LB broth added and incubated for 20 min. This allows time for expression of the transformed plasmids. One-tenth ml cells was plated on VRB Agar containing 50 µg/ml amp and the plates were incubated at 37°C for 24 h. White colonies were presumed to contain inserts because of an interrupted ß gal gene as well as their ability to grow on amp. The colorless colonies were isolated on VRB to verify their inability to utilize lactose. Single colonies were picked and propagated on LB plates containing 50 µg/ml amp. The plasmids were isolated and the plasmids analysed by restriction endonuclease digestion to verify insertion of foreign DNA.

Electroporation

Electroporation was performed using the BioRad method (Dower et al., 1988). Bacteria were grown to 0.5 at A_{600} and collected by centrifugation at 5000 x g for 10 min. The pellet was washed twice in cold buffered sucrose medium (270 mM sucrose, 1 mM $MgCl_2$, 7 mM Na_2 HPO_4 [pH 7.4]). Cells were resuspended in electroporation medium until used. Purified plasmid was resuspended in TE buffer to a final concentration of 2 µg/ml in 0.2 ml of cell suspension. The mixture was placed in a Gene Pulser cuvette (0.4 cm gap) and kept on ice for 10 min prior to pulsing. The mixture was exposed to a single pulse (peak voltage 2.5 kv; capacitance 25 µF) which generated a field strength of 6.25 kV/cm. Time constants ranged from 2.2 to 6.0 m sec. After the high voltage pulse, cells were placed on ice for 30 min. One hundred ml of the cell suspension were spread on VRB plates containing 50 µg/ml amp. Analysis of clones was the same as for $CaCl_2$ transformation.

Denaturation of Plasmid and Primer Annealing for DNA Sequencing

Two different methods were used to denature plasmid DNA prior to sequencing. Plasmids using M 13 primers were alkaline denatured (Wang et

47

al.,1988). Custom primers and difficult sequencing templates were heat denatured (Andersen et al.,1992)

For alkaline denaturation 2 µg of fresh plasmid DNA were resuspended in 18 µl sterile H_2O and 6 µl of 1M NaOH were added. After 5 min of incubation at room temperature, 3 µl of cold 3M Ammonium Acetate [pH 4.5] and 70 µl of 95% EtOH were added, inverted and held on ice for 2 h or placed at -70^0C for 15 min. Denatured plasmid DNA was centrifuged for 5 min at 10,000 xg. The pellet was rinsed in 70% ethanol and air dried. The DNA was then ready for primer annealing according to manufacturer's reaction protocols for Sequenase® version 2.0 (United States Biochemical, Cleveland,OH). Denatured DNA was resuspended in 7 µl sterile distilled water and then 2 µl 5X reaction buffer (250 mM NaCl, 100mM $MgCl_2$, 200 mM Tris-Cl [pH 7.5]) and 1 µl of appropriate primer (0.5 pmol/µl) were added. The mixture was warmed to 65^0C for 2 min and allowed to cool slowly below 300 C and the sample placed on ice until used in sequencing reactions.

For heat denaturation and primer annealing, 2 µg of freshly prepared plasmid DNA were dissolved in 7 ul sterile H_2O, 1 ul primer (0.5 pmol/µl) was added and the mixture incubated at 97^0C for 5 min. The DNA/primer mixture was then centrifuged for 30 sec at room temperature. Two µl 5X sequenase buffer were added and mixed, followed by incubation at 37^0C for 15 min. The DNA and annealed primer were ready for sequencing reaction.

Plasmid Preparations for Cloning

A discontinuous cesium chloride (CsCl) ethidium bromide (EtBr) gradient method (Garger et al., 1984) was used for purification of plasmid for cloning and restriction endonuclease analysis. Beckman quik seal polyallomar ultracentrifuge tubes (Beckman, Palo Alto, CA) 5/8 X 3 were

48

partially filled with a less dense CsCl solution (14.9 g CsCl in 20 ml TE). The dried plasmid prep was resuspended in 2.4 ml TE and 4.2 g of CsCl were added and dissolved. After the addition of 0.4 ml EtBr (10 mg/ml) to the plasmid-CsCl mixture, the more dense CsCl preparation was slowly added to the bottom of the quik seal tube with a syringe and cannula. Any remaining air space at the top of the tube was filled with less dense CsCl solution by means of syringe and needle. The tubes were balanced and sealed. Centrifugation was in a Beckman L80 ultracentrifuge for 18 h at 45,000 rpm in a Beckman type 80 Ti rotor. After centrifugation, the plastic tube was viewed with a (Model UVG 54) hand held short wave (254 nm) UV light (UV Products Inc., San Gabriel, CA). The band containing the plasmid of interest was removed by puncturing the tube with an 18 gauge needle fitted to a 3 ml syringe and drawing off liquid until the band disappeared. Ethidium bromide was extracted by succesive washings with equal volumes of water saturated isoamyl alcohol until the sample was clear. The CsCl was removed by dilution with three volumes of water and the DNA precipitated by addition of two volumes of cold 95% ethenol and used immediately or pelleted at 10,000 x g for 1 min, air dried and stored for later use.

Polymerase Chain Reaction(PCR)

The urease structural gene was selectively amplified by PCR using synthetic oligodeoxy nucleoside primers (Fig. 10) designed to anneal to urease genes. The forward primer corresponds to bases 411-426 (5'-ATGGAATTAACACCA-3') of the *Proteus vulgaris* structural genes for urease (Mörsdorf and Kaltwassen, 1990). The reverse primer corresponds to the complementary bases 2744-2728 (5'-AATAGAAAATAACGC-3') of the *Proteus vulgaris* structural gene for urease. PCR reaction mixtures were compiled according to

49

the manufacturer's protocol (Perkin Elmer, Norwalk, CT). The reaction mixture (50 mM KCl, 10 mM Tris-Cl[pH8.3], 4 mM $MgCl_2$, 200 µM of each NTP{dATP, dTTP, dGTP, dCTP}, 0.2 µg of each oligonucleotide primer, 100 ng template DNA, ddH_2O to 100 µl, and 2 units Taq DNA polymerase) were overlayed with 100 µl mineral oil and subjected to 40 cycles in a Perkin Elmer thermal cycler (Perkin Elmer, Norwalk, CT). Each cycle consisted of denaturation at 92^0C for 1.5 min followed by primer annealing for 1 min at 37^0C then nucleotide chain extension at 72^0C for 2 min. Amplified product (15 µl) was electrophoresed on a 0.8% agarose gel.

DNA Sequencing and Sequence Analysis

All DNA sequencing for this study was performed on double stranded plasmid by the dideoxy chain termination method (Sanger et al., 1977). Reagents used were supplied in the Sequenase Version® 2.0 sequencing kit and their protocols were followed. The only reagents not supplied in the kit were ^{35}S labeled dATP (NEN, Boston, MA) and synthetic oligonucleotides for sequence extension (National Biosciences, Plymouth, MN). DNA sequencing gels were cast using a premixed novel crosslinking acrylamide solution HydroLink Long Ranger (AT Biochem, Malvera, PA) using the supplied protocol. Electrophoresis was on a model S2 sequencing system (BRL, Gaitherburg, MD) in TBE buffer (1M Tris, 0.9 M Boric acid, 0.01 M EDTA) at 55 W for times varying from 2 h to 14 h. After electrophoresis was complete, gels were transferred to Whatman 3 MM Chromotography paper (Whatman Ltd., Mardstone, England) and dried at 75^0C under vacuum on a Hoefer Model 540 slab gel drier (Hoefer Scientific Instruments, San Fransisco, CA) with a Flexi dry (FTS Systems, Inc, Stove Ridge, NY) cold finger in the vacuum line to remove moisture. The dried gel was exposed to Hyperfilm

50

MP (Amersham Life Sciences, Arlington Heights, IL) for 24-48 h. Autoradiographs were developed by immersion in D-11 (Kodak, Rochester, NY) developer for 5 min followed by a 30 sec rinse in H_2O and a 5 min soak in Fixer (Kodak, Rochester,NY). The autoradiographs were then washed in water for 15 min and air dried. Sequences were read directly from film and analysed by Genetics Computer Group (GCG Software, Madison, WI) and compared against sequence database (Altschul et al., 1990) and filed with GenBank (San Francisco, CA).

Labelling of DNA for use as Probe

DNA from one of the described isolation methods was quantified and randomly labelled (Maniatis et al., 1975) with ^{32}P-α labelled CTP (NEN). Unicorporated label was removed (Maniatis et al., 1982 B) by elution through a Sephadex G50 column equilibrated with STE buffer (10 mM Tris-Cl, 1 mM EDTA, 0.1 M NaCl [pH 8.0]). The labelled DNA was passed through the column and was collected and used immediately or stored at -20°C until used.

Southern Transfer and Hybridization

Agarose gels to be used in Southern blot analysis (Southern, E.M., 1975) were first partially hydrolysed (Wahl et al., 1979) by twice soaking in 0.25 M HCl for 15 min each. The gel was then denatured (1.5 m NaCl, 0.5 M NaOH [pH 8.0]) for 1 h with shaking at room temperature, followed by neutralization (1M Tris HCl, 1.5 M NaCl [pH 8.0]) for 1 h. Transfer to either S+S Nytran (Schleicher and Schuell, Keene, NH) or Transblot transfer membrane (BioRad, Richmond, CA) was accomplished by using 3MM chromotography paper as a wick on a supported glass plate. The DNA was transferred via capillary action due to transfer of 10 X SSC buffer (88.7 g NaCl 38.7 g sodium citrate [pH 7.0]/liter) through the gel then onto a piece of transfer membrane

51

cut to the size of the gel to a stack of paper towels topped by a 500 g weight. Transfer was for 24 h. When transfer was complete, the membrane was soaked in 6X SSC for 5 min, then baked in a vacuum oven for 2 h at 80°C.

Hybridization (Maniatis, T., 1982 B) was accomplished by first floating the dried nitrocellulose filter on 6 X SSC then immersing it for 2 min. The filter was then placed in a heat sealable bag and 0.2 ml/cm² of prehybridization fluid (6 X SSC, 0.5% SDS, 5 X Denhardts solution[Ficoll 1 g, polyvinyl pyrolidine 1 g, 1 g BSA in 100 ml] 100 µg/ml denatured salmon sperm DNA) that had been prewarmed to 68°C was added. The bag was sealed and incubated in a 68°C water bath with occasional agitation for 2-4 h. At the end of the incubation period, the prehybridization fluid was removed and 50 µl/cm² of hybridization fluid (6 X SSc, 0.01 M EDTA, ^{32}P labelled probe, 5 X Denhardts, 0.5% SDS, 100 µg/ml denatured salmon sperm DNA) was placed in the bag, sealed and incubated at 68°C overnight. The filter was removed and soaked (2 X SSC, 0.5% SDS) at room temperature for 5 min then soaked for 15 min (2 X SSC, 0.1% SDS). The filter was transferred to a solution of 0.1 X SSC and 0.5% SDS and incubated at 68° for 2 h with gentle agitation. The solution was changed and the filter washed for another 30 min. The filter was dried and exposed to Hyperfilm MP (Amersham Life Sciences, Arlington Heights, IL) for 12-36 h. Autoradiographs were developed by the same procedures as sequencing gels.

Electrophoresis of Maxicell Proteins

Maxicell preparations were electrophoresed in one dimension (Hanes and Rickwood, 1981) on a 11.5% acrylamide SDS horizontal slab. The dimensions of the gel were 12 cm X 14 cm X 1.5 mm. The composition of the gel was 11.5 ml of 30% acrylamide stock (292 g acrylamide, 8 g bis acrylamide

in 1 L of H_2O), 7.5 ml lower buffer (188.4 g Trizma, 4.0 g SDS [pH 8.9] in 1 liter), 11 ml H_2O, 99 µl 10% ammonium persulfate (APS), and 15 µl TEMED. After the gel solution was poured between the glass plates, 0.4 ml H_2O was overlayed to allow the slab to polymerize. This slab gel was topped by a 4.5% acrylamide stacking gel composed of 1.87 ml 30% acrylamide stock, 3.12 ml upper buffer (60.5 g Tris base, 4.0 g SDS/liter) 7.5 ml H_2O, 7 µl 10% APS, 4.75 µl TEMED carefully overlayed with 0.4 ml H_2O. Samples were loaded and electrophoresis through the stacking gel was for approximately 1/2 h at 200 V until the ion front reached the slab gel. Electrophoresis through the separating gel was for 3-4 h at 50 W until the ion front was approximately 1/4" above the bottom of the glass plates.

2-D Electrophoresis

The protocol used in this study for two dimensional polyacrylamide gel electrophoresis of proteins was supplied by Dr. F.C. Niedherdt's Lab (University of Michigan Medical School, Ann Arbor, MI) and is a compilation of methods (O'Farrell, R.H. 1975., O'Farrell et al., 1977., Celis and Bravo, 1984., Young et al., 1983, Garrels, J.I., 1983).

Sample preparation for 2-D gels was as follows. Ten ml cultures were grown at the appropriate temperature either with or without plasmid to mid log phase. Two hundred µCi of TransLabel (NEN, Boston MA) were added and shaken continuously for 30 min. Cells were pelleted at 5,000 x g for 5 min and the pellet resuspended in 40 µl SDS/ß-ME solution (0.3 g SDS, 5 ml ß-mercaptoethanol (BME), 0.444 g Tris-HCl, 0.265 g Trizma base adjusted to 100 ml with milli Q [Millipore Systems, Bedford, MA]). The suspension was boiled for 2 min, cooled on ice and 4 µl RNase/DNase solution (2 mg/ml both RNase [2.7.7.16] and DNase [3.1.4.5] in Tris-MgCl buffer [1.2 g Tris base, 1.02 g

53

MgCl; 6H$_2$O] to 100 ml with MilliQ]) were added and incubated at room temperature for 20 min.

After incubation, 160 µl lysis buffer (5.7 g Urea, 2.0 ml Triton X-100, 0.5 ml ßME, 0.4 ml pH 5-7 ampholines and 0.1 ml pH 3-10 ampholines brought to 10 ml with MilliQ) were added and the samples placed immediately at -70°C. Isoelectric focusing gels were constructed by first scrupulously cleaning and rinsing 14 cm X 3 mm glass tubes. The tubes were dried and marked 12 cm from the bottom. The bottoms were tightly wrapped in parafilm, aligned vertically, and filled with isoelectric focusing (IEF) gel mix (1.15 g urea, 0.266 ml IEF acrylamide [2.838 g acrylamide, 0.162 piperazine diacrylamide, 7.1 ml H$_2$O] 0.4 ml 10% NP40 or Triton X-100, 0.4 ml H$_2$O, 0.1 ml 4-8 ampholines [0.08 ml 5-7 ampholines, 0.02ml 3.5-10 ampholines may be used in place of 4-8 ampholines] 2 µl 10% APS, 1.4 µl TEMED). This amount of mix was sufficient for 2-14 cm X 3 mm tubes. A syringe filled with a cannula or long needle was used to fill the tubes from the bottom up to the mark at 12cm. The gel mix was carefully overlaid with 0.2 ml H$_2$O and 30-60 min was allowed for polymerization. Once polymerized, the parafilm was removed and the water removed from the top. The tubes were carefully checked to ensure that there were no bubbles and then inserted into the 1-D apparatus (Hoefer model GT3, Hoefer Scientific Instruments, San Francisco, CA). The lower chamber buffer (0.01 M H$_3$PO$_4$) was added followed by the tube holder, making sure no air bubbles appeared at the bottoms of the tubes. Protein samples were placed on top of the gel (this study used autoradiographic analysis and samples contained 500,000 cpm^{35}S; however, 10-20 µg of protein is optimal for this system but may range from 1-50 g) and 30 µl of sample overlay (5.5 g Urea, 5.0 ml H$_2$O, 0.25 ml 4-8 ampholines) were carefully added.

54

Any additional space in the tube was overlayed with upper chamber buffer (0.02 NaOH-degassed) prior to filling the chamber with buffer. Electrophoresis was at 400V for 15-20 h. The voltage was increased to 1000 V for 1 h to sharpen the bands. The gels were gently extruded from the tubes using a syringe fitted with tubing the same diameter as the tube. The gels were placed into SDS sample buffer (125 ml upper gel buffer, 100 g Glycerol, 22.5 g SDS, 50 ml ß-ME, 1 ml 1% bromophenol blue, in 1 L), equilibrated for 20-30 min, then loaded onto a 2-D gel or frozen at -70°C until used.

The 2-D gels were constructed by the same protocol as for maxicell gels except that the stacking gel was poured to within 1/4" of the top of the plates, overlayed with 0.4 ml H_2O and allowed to polymerize for 1 h. The water layer was removed and the tube gel from 1-D IEF was carefully placed atop the glass plates. The 2-D and IEF gels were sealed together by pipetting melted agarose (1 g agarose, 25 ml upper gel buffer, 75 ml H_2O) over the tube gel, making sure to remove all air bubbles. The electrophoresis unit (Hoefer) was loaded with running buffer (1 g Tris base, 14.4 g glycine, 1 g SDS, in 1 L) and electrophoresed at 200V (constant voltage) until the dye was 1/2" above lower gel. The gel was then run at 50W (constant power) until the dye front reached the bottom of the gel. The gel was removed from the apparatus and soaked in fixative (45% methanol, 10% Acetic Acid) for a minimum of 45 min and no longer than 12 h. The gel was then placed on 3 MM paper and vacuum dried (as sequencing gels) for 2 h at 60°C. The gels were exposed to X-ray film and developed by the same procedures as sequencing gels.

RESULTS

IS Isolation

This study was undertaken to sequence IS 100 (Portnoy and Falkow, 1981) of *Y. pestis* in order to confirm its standing as an IS element. A rapid renaturation technique (Ohtsubo, 1976) was employed in which double stranded DNA was heat denatured and allowed to briefly renature and thus allowing the inverted repeats of IS elements to form duplexed molecules while most DNA did not (Fig. 1). The single stranded DNA was then digested with S_1 nuclease leaving double stranded IS elements. The S_1 treated DNA was subjected to electrophoresis on a 0.8% agarose gel. Strains tested by this method were *Y. pestis* EV7651F, *Y. pseudotuberculosis* w.t. and *Y. enterocolitica* 3973-76 0:4. The result of one such separation is shown in Fig. 2. Measurements for molecular sizing were done using a Pst 1 digest of phage λ. *Y. pestis* IS were compared to *Y. pseudotuberculosis* and *Y. enterocolitica* IS (Fig. 3) and their sizes are listed in Table 5. Gels containing IS elements of *Y. pseudotuberculosis* and *Y. enterocolitica* required soaking in EtBr overnight for the bands to become visible.

IS Cloning

Initial cloning studies involved cutting pUC 19 (Fig. 4) with Sma1 because such digests resulted in a blunt end which was compatible with the blunt ends formed during the isolation of the IS elements. None of the initial cloning experiments were successful so a new approach was needed. As can be seen in Figure 3, a significant background of nucleotides (smearing) was evident which I felt was due to single and double stranded DNA interfering with the ligation mix. Consequently, the 2.2 kbp band from *Y.*
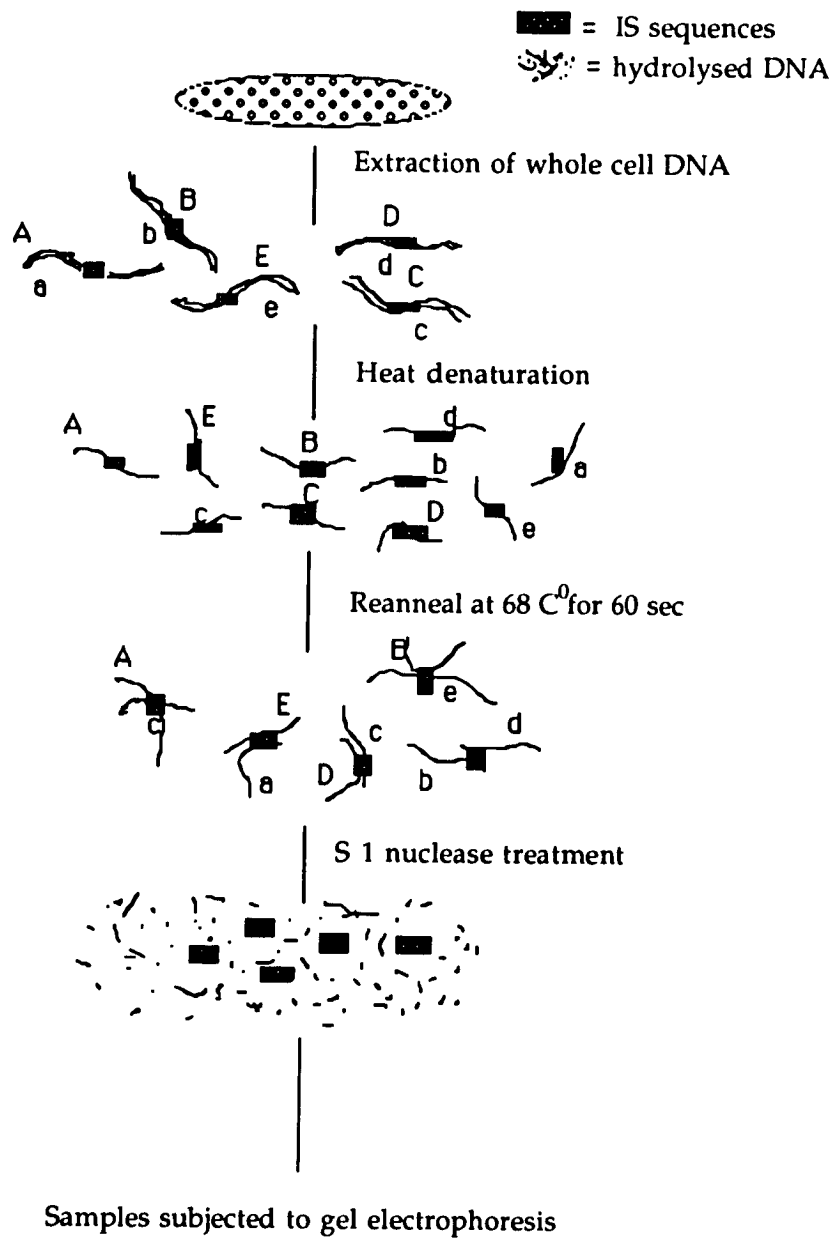
56
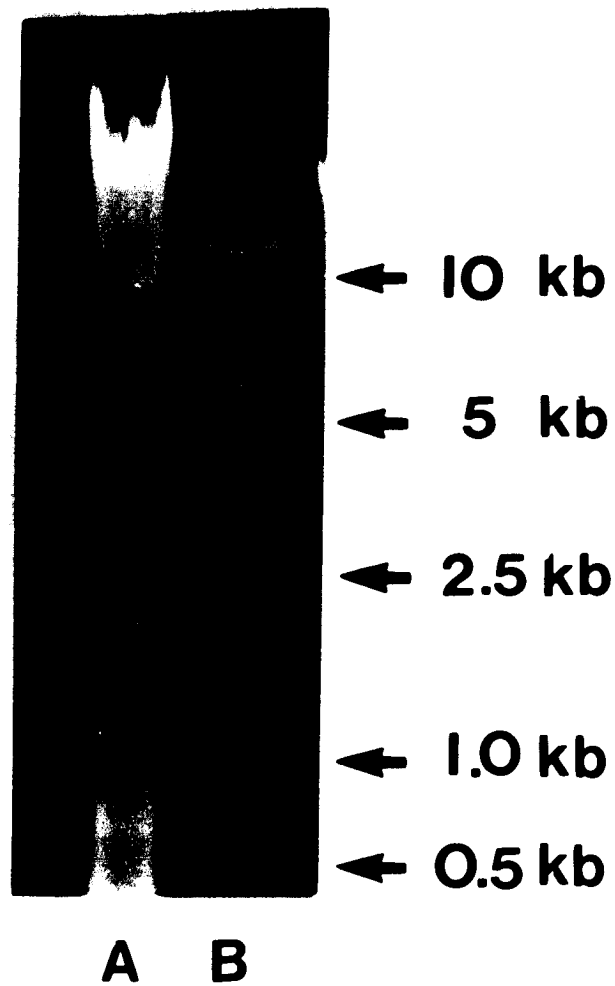
Fig. 1 Schematic protocol for IS isolation

57

Fig. 2 Sizing of IS elements from rapid renaturation protocol. Lane A is a rapid renaturation product using *Y. pestis* DNA. Lane B is λ DNA cut with Pst1.

58
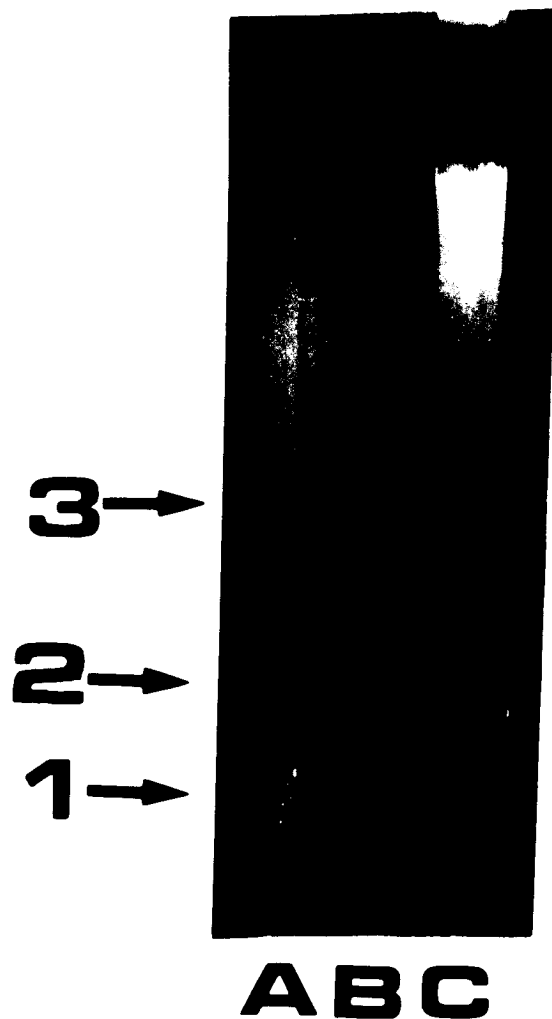
Fig. 3 Putative IS elements from rapid renaturation technique of the three medically important *Yersinia*. Lane A; *Y. enterocolitica*, with bands appearing at 1, 2 and 3; lane B; *Y. pseudotuberculosis*, with weaker bands at 1 and 2; lane C; *Y. pestis*, bands are strongest and appear at 1 and 2 as well as between 2 and 3. Another DNA band is found just below the chromosomal band.

59

Fig. 4 pUC19 (Yanisch-Perron et al, 1985) vector used in all cloning experiments in this study.

60

Table 6 Number and sizes of putative IS elements isolated from the three medically important species of *Yersinia*.

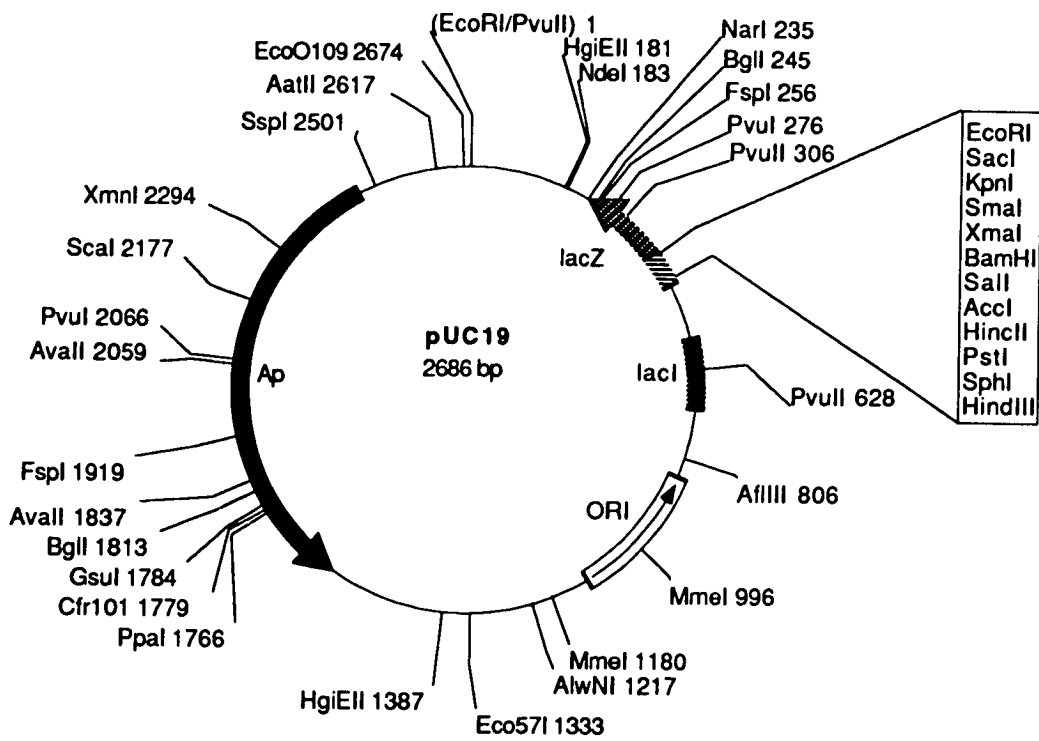| Species	number of bands in gels	size (kbp)
Y. pestis	4	9, 2.2, 1.2 and 0.6
Y. pseudotuberculosis	2	1.2 and 0.6
Y. enterocolitica	3	4, 1.2 and 0.6

pestis, which most likely represented IS100, was isolated from gels. A number of gels containing the IS elements were run and the 2.2 kbp band excised and electroeluted as described in "Materials and Methods". The band was then subjected to several restriction endonucleases to determine which would be most suitable for cloning. Sau3A, HaeIII, Ava II, AccI cut IS 100 into many small pieces (data not shown) Pst 1 and EcoR1 each cut IS 100 into two fragments, 1.4 and 0.8 kbp and 1.6 and 0.6 kbp respectively. A Pst1 digest was performed on the electroeluted 2.2 kbp band and the products run on a 0.8% agarose gel. The interfering DNA fragments in the gel were still visible at the 2.2 kbp range while the Pst1 digested IS 100 DNA was located in two bands at 0.8 and 1.4 kbp, away from the interfering single stranded DNA (Fig. 5). The 1.4 kbp and 0.8 kbp fragments were excised, electroeluted and then subjected to blunt end ligation with Sma1 cut pUC19. The ligation product was transformed (see Methods and Materials) into *E. coli* JM83, resulting in the formation of 201 colonies. Of these, 181 contained no insert. Twenty white colonies were screened by double restriction endonuclease digestion with EcoR1 and Xba1 in order to detect the size of the insert (Fig. 6). Only one sized

insert appeared and it was calculated to be approximately 0.8 kbp. Southern analysis was carried out on a selected clone and the construct designated as pISlC. The probe was the 2.2 kbp fragment that had been digested with EcoR1 and electroeluted. The fragments were combined and labelled with ^{32}P for use as probe. Analysis showed homology between pISlC and the IS 100 probe. Portnoy and Falkow (1981) described the presence of IS100 in the BamH1-2 fragment of a digest of the virulence plasmid of *Y. pestis* 019. Additionally Portnoy and Falkow (1981) located IS100 on the BamH1-5 fragment of the

62
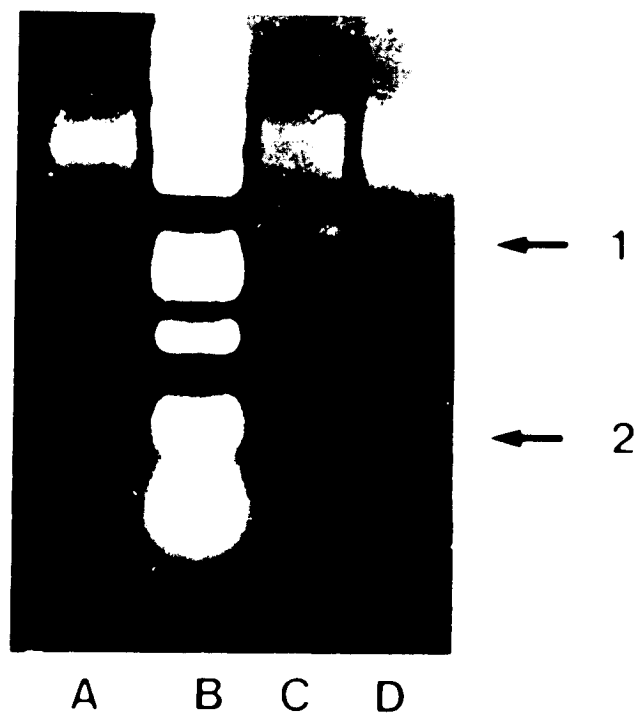
Fig. 5 Preparation of 2.2 kbp fragment for cloning. Lanes A and D are electroeluted 2.2 kbp fragments from the rapid renaturation technique. Lane B is λ DNA cut with Pst1. Lane C is the 2.2 kbp fragment cut with Pst1. Note the 1.4 kbp fragment denoted by arrow 1 and the 0.8 kbp fragment denoted by arrow 2.

Fig. 6 Sizing of clones. Lanes B-G are plasmid DNA samples from clones containing inserts of the 2.2 kbp fragment digested with Pst1. The plasmid DNA was first restricted with EcoR1 and HindIII before loading onto the gel. Lane A is λ DNA digested with HindIII.

64

virulence plasmid of Y. pestis EV76. Homology between the probe (the 2.2 kbp fragment) and a BamH1 digest of the plasmid of each strain was observed (Figs. 7 and 8). Also depicted in Fig. 8 is the homology of the 2.2 kbp fragment used as probe with plasmids p6E16, pIS1C and pE6.

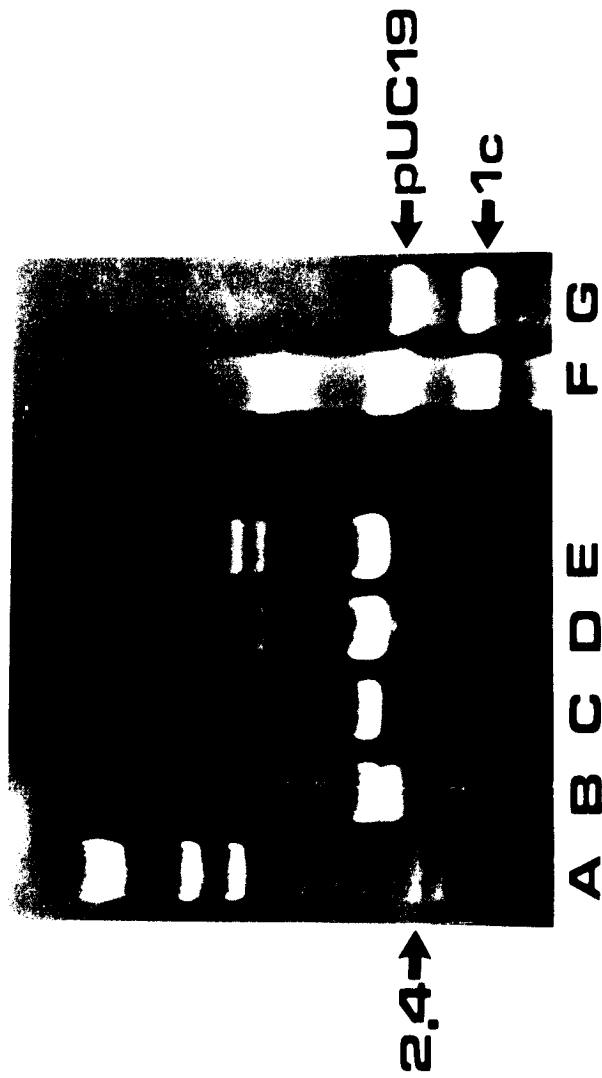Additional cloning studies were performed to clone the remaining putative IS elements from Y. pestis by the shotgun method. The 9, 1.2 and 0.6 kbp fragments were successfully cloned.

Since the 6 MD plasmid of Y. pestis was known to contain IS100 (Portnoy and Falkow, 1981), a map (Fig. 9) was compiled from the data of Sodeinde et al. (1988) and plasmid data obtained in our lab. This map provided the basis for my cloning strategy. Because treatment of the 2.2 kbp band resulted in fragments of 1.4 and 0.8 kbp, double digestions using Pst1 and Sma1 were carried out on the 6 MD plasmid as well as pUC19. The products of the digestions were used in shotgun experiments and colonies suspected to have insertions were picked for screening either by double digestions or by Southern analyses using the 2.2 kbp fragment as probe. In addition, the 6 MD plasmid was digested with EcoR1 and the fragment ligated in pUC19. Clones from this experiment were treated in the same manner as those from the Pst1-Sma1 cut clones.

DNA Sequencing of the Putative IS Elements

Sequence analysis was performed initially on the pIS1C clone. The sequencing strategy (Fig. 10) was to use M13/pUC19 forward and reverse primers to obtain initial sequence information. Appropriate priming sites were selected by choosing regions with the highest AT% and fewest possible
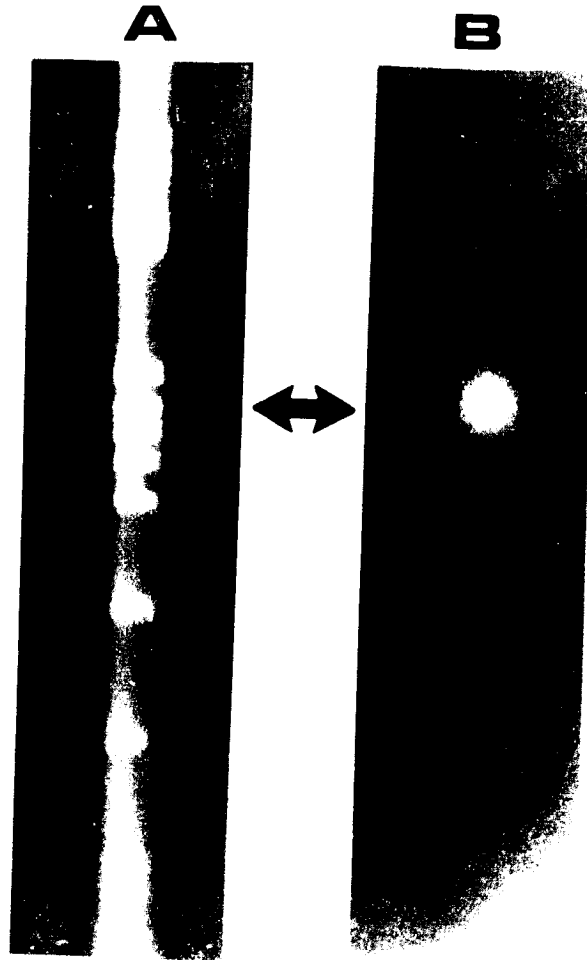
Fig. 7 Restriction profile of the 47 MD plasmid of *Y. pestis* EV7651F and corresponding autoradiograph. Lane A is a BamH1 digest of the 47 MD. The arrow denotes the BamH1-5 fragment. Lane B is the corresponding autoradiograph of Lane A when probed with the 2.2 kbp fragment from IS isolation.

66

Fig. 8 Homology with the 2.2 kbp fragment from IS isolation protocol. 1 is the agarose gel and 2 is the corresponding autoradiograph. Lane A is p6E16, lane B is bad plasmid prep of pIS1C, lane C is a BamH1 digest of the virulence plasmid of *Y. pestis* 019, lane D is a BamH1 digest of the virulence plasmid of *Y. pestis* 019 Ca-4. *Y. pestis* 019 Ca-4 is a calcium independent mutant that has suffered an insertion in the BamH1-2 fragment of the virulence plasmid. Lane E is the pE6.

67

Fig. 9 Map of 6 MD plasmid of *Y. pestis* EV7651F. Area containing IS100 is darkened.

68

Fig. 10 Sequencing strategy for pIS1C. Arrows denote direction of primer extension. Solid arrows are forward sequencing reactions. Unfilled arrows are reverse sequencing reactions.

69

repeats. Synthetic oligonucleotides (Table 7) were purchased (NBI, Plymouth, MN) to extend sequence information until the entire insert was sequenced over both the plus and minus strands. pISIC was shown to consist of 821 bp (Table 8).
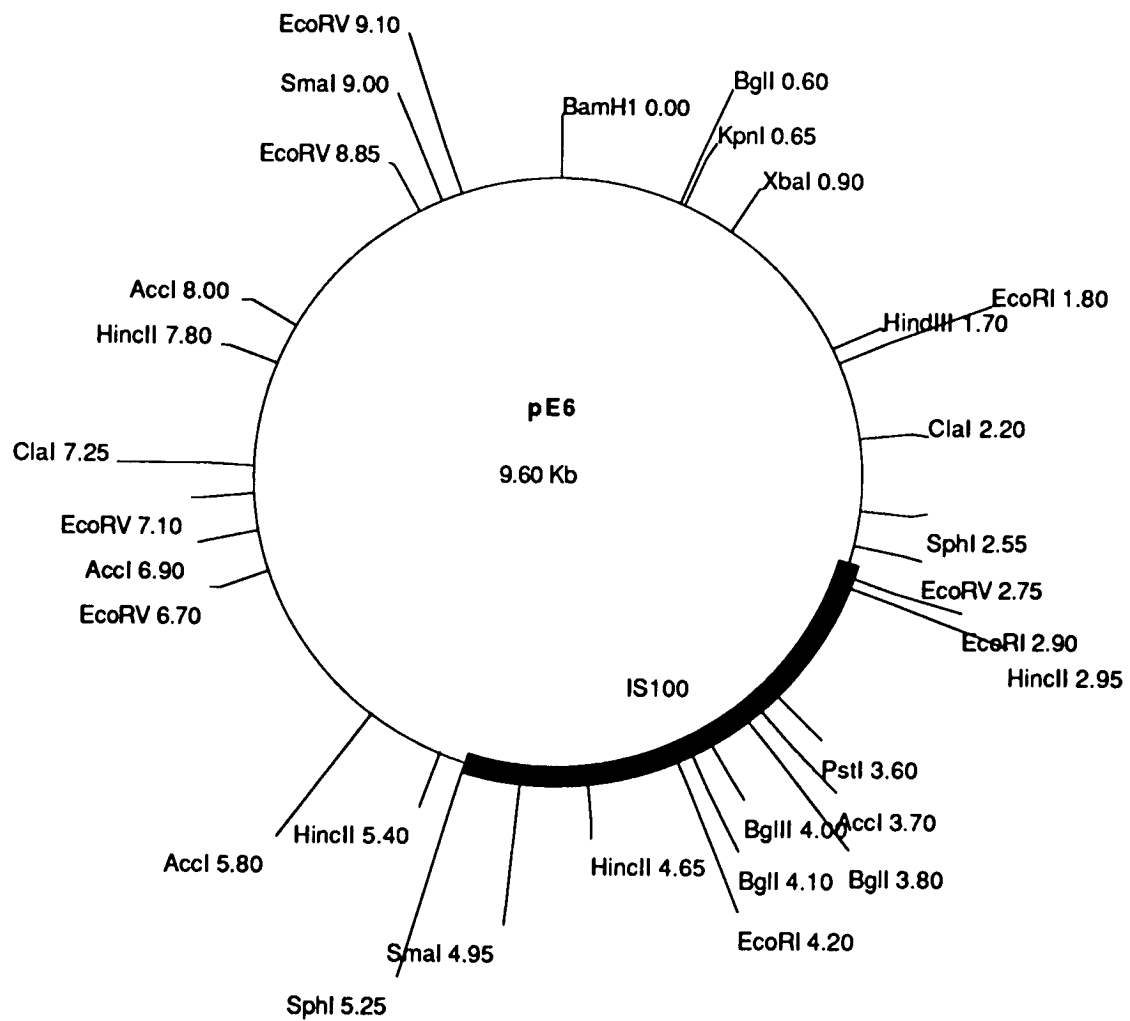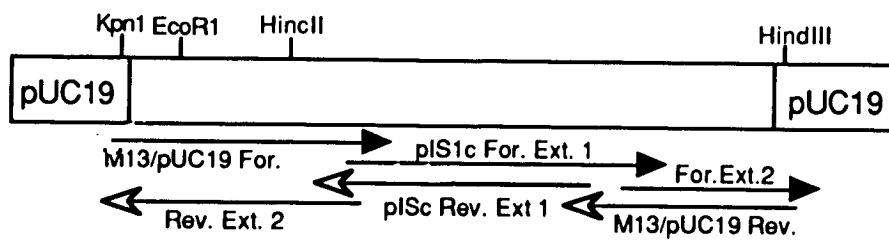
Sequence analysis was preformed using the GCG software package and all sequences were compared against the GenBank database via blast analysis (Altschul et al., 1990) at NCBI.

All attempts to clone the larger piece of IS100 were unsuccessful. Consequently I undertook a multipronged strategy to obtain sequence information for the 1.4 kbp fragment of IS 100. As indicated above, IS 100 was shown to exist on the 6 MD plasmid (Portnoy and Falkow, 1981) as well as the 47 MD plasmid. An Oligonucleotide primer that was homologous to the end of pISIC and which directed inward of IS 100 was obtained and this primer was used to sequence the 6 MD plasmid (pE6) directly. Using the same approach to sequencing pE6 as for pISIC, I was able to extend the sequences associated with IS 100 using the 6 MD plasmid as template. The 6 MD plasmid of *Y. pestis* is of lower copy number than pUC19 however and sequencing proved to be difficult. Ongoing efforts to clone DNA fragments from restriction digests of the 6 MD were successful; therefore, the strategy to obtain sequence data directly from the 6 MD was abandoned.

Six clones containing inserts from the 6 MD plasmid (pPS4, pPS5, p6E6, p6E16, pISIC and p6E6) were used for additional sequencing studies as they were of appropriate size. A summation of my sequence strategy is shown

70

Table 7. Sequencing and PCR primers used in this study

M13/pUC19 Forward	5′-GTAAAACGACGGCCAGT-3′
M13/pUC19 Reverse	5′-CAGGAAACAGCTATGAC-3′
pIS1c Forward Ext. 1	5′-AAGCGAACTTTGATA-3′
pIS1c Reverse Ext. 1	5′-AAGTTCGCTTCACAA-3′
pE6	5′-TTTCTCGCGTTACTA-3′
pE6 Forward	5′-TTATAAGGCTTTCCA-3′
p6E16 Forward Ext.	5′-TAACAGTTTTTACA-3′
pIS1c Reverse Ext. 2	5′-GTATCTTACACACT-3′
pPS4 Forward Ext. 1	5′-TAGCTATTCACATGT-3′
pPS4 For. Rev. Ext. 1	5′-AGTGGGATCACTTTT-3′
Urease Forward	5′-AAATGGAATTAACAC-3′
Urease Reverse	5′-AATAGAAAATAACGC-3′

71

Table 8. DNA Sequence of IS100.

```
   1 GTATCAACGA CGGATGAAAG TGATCCAACT TATATCTCCA CCAACGGCCC

  51 AATATTGATC CACCGTTTTA CTCAGGAATT CGTTCTGCTA TAACCCCGGC

 101 CTTTCGTTTC TGTCTGAGTC GATGTTGACT CTCCTTTGAT TTGAACGACA

 151 TGTGAGTGGT GTAAGATACG GTCCAGCATC GCTGAGGTCA GTGCTGCATC

 201 ACCGGCGAAC GTTTGATCCC ACTGCTCGAA CGGCAGATTG GATGTCAGGA

 251 TCATTGCGCT CTTTTCGTAA CGTTTAGCGA TGACCTGGAA GAACAGCTTT

 301 GCTTCTTCCT GACTGAACGG CAGATAGCCT ATTTCATCAA TGATGAGCAG

 351 GCGGGGGGCC ATTACTCCAC GCTGAAGCTC GTTTTATAAC GGCCCTGACG

 401 TTGTGCCGTA GATAACTGAA GTAACATATC TGCTGCTGTT GTGAAGCGAA

 451 CTTTGATACC TGCACGGACT GCTTCATAGC CCATCGCTAT TGCCAGATCG

 501 GTTTTCCCCA CACCTGATGG CCCCAGTAAT ACGATATTTT CATTACGTTC

 551 TATGAAGCTG AGTGAGCGTA ACGACTGGAG TTGCTTCTGC GGTGCTCCGG

 601 CGGCGAATGT GAAGTCATAC GCTTCGAACG TTTTCACCGC CGGGAAGCGC

 651 CATTCGGCTA TACATCGCCT GTTTACGTTG ATGACGGCCA GTTTTTCTGC

 701 AGAAGCAGAG GCCAGGAAGT CCATATAACT CCATTCCTGC TCTACTGCCT

 751 GTTGTATACC GCAGCCGCGC TTATAAGGCT TTCCAGTTGC AACTGCCCGC

 801 GAGCGCCATC ACGGCTCGTT GATGTTGCAG TTGCATCATC ACGCCACTCC

 851 TCAGTGTTGT CCTCATCTGG TCCTTACCTC AATATACCTG AAGACCTCGT

 901 AGAGCAAGTA CTTCTTTTTG ACCGTCATAG TTGCATTTGT CCGCTACATA

 951 TGGGCTTACG TCGGAGGGCC GGACTTTTGC AGCTTCTCAT ACTGAAGTGT

1001 AAGTGTCTCG TGCTCTTCGT GAGTCAGCAA TCGATGATCG AAGTATCTTG

1051 CATTACTTTT TAGCATAATG ACCCCGTACT CACACCCCCT TTTGCTACAT

1101 CTCGTTATCG CATCCCGATA CTCTGAGCAC GTCCTAGTGC GAATGTGTCG
```

72

1151 <u>TCGTCTGACA TGAATCATCA TCGTTCACTG ACAATATCCG TTATGACACG</u>

1201 <u>CTGGAGACCT GCCATCGTGG CAATGCCGTTC CGCTTCTTTG GTGGTGTGCC</u>

1251 <u>GCGCGAAGTG TTGTATGACA ATATGAAAAC TGTGGAATTC TGCAACGTGA</u>

1301 <u>CGCATATCAG ACCGGTCAGC ACCGTTCCAT CCTTCGTCGT GGCAGTTCGG</u>

1351 <u>CAAGGAGATG GGCTTCTCTC CCCGACTGTG TGCCCCTTCA GGCACAGACT</u>

1401 <u>AAAGGTAAGG TGGAGCGGAT GGTGCAGTAC ACCCGTAACA GTTTTTACAT</u>

1451 <u>CCCCACTAAT GACTCGCCTG CGCCCGATGG GATCACTGTC GATGTTGAAA</u>

1501 <u>CAGCCAACCG CCACGTCTGC GCTGGTGCAC GATGTCGCTA ACCAACGAAG</u>

1551 <u>CATGAACATC CAGGCCGTCC GCGATCGCTG GCTCGAACAC AGTCCATCGT</u>

1601 <u>ACATTTGCTG CCTACTTTTC ACTAGGGTGA TATAGAGGTG GTTGCCGGGT</u>

1651 TATAACTAGG TGGCAAAATG AGTCCTAATC GAAGACGATA TTGGGGCCGG

1701 AAAGCAAAGA CAGACTCAGC TATCGAAAGA GGAAACTAAA CTTGCTGCAC

1751 ACTCACCACA TTCTATGCCA GTCGTAGCGA CTCCAGTCAC GACATGGTGG

1801 CCGCTGCACT AGGTTGACGG GCTGCGTCTA ACTACAGTCC TAGTAACGCG

1851 AGAAAGCATC GAATCGCTAC TGACTCTGTT CAACAGAGAC TGACTGCGTC

1901 TATCGATAAG TGTACACCGC GCCCGTATAG TCTCATCAGC AAAGGAGTGT

1951 AATTAAATAA TGATATAATG ATATTAATAA GTCGTAATTT TATTAATATC

2001 ATTATAGTTA TGCAATTACA CCCGGGCGTA CCTCTCTATG AACAGCTTTA

2051 GTAACCGTTT GAGTAACGCG TTGTCACCCA TTAGCGACCT TACGTTAAGC

2101 CACTCTACCT AAAACAGCAG AAGTGCATTA TTTACTGCAC GCTAATTGCC

2151 GCTCGGTCCC ATCATCGCAT GCAAAAAGCG GTGCCAATTT TATCAACTAA

2201 CGCGAGTATA TGCCCCGATT AGACGCCCAT T**CAGTAGCGA CGACGG**

Sequence of IS 100. Bold letters are imperfect direct repeats. The first 821 base pairs are the pIS1c clone. The underlined sequences are the significant open reading frames (ORF) in the forward direction. Significant ORF's are on the complementary strand are bases 40-360 and 1190-1600.
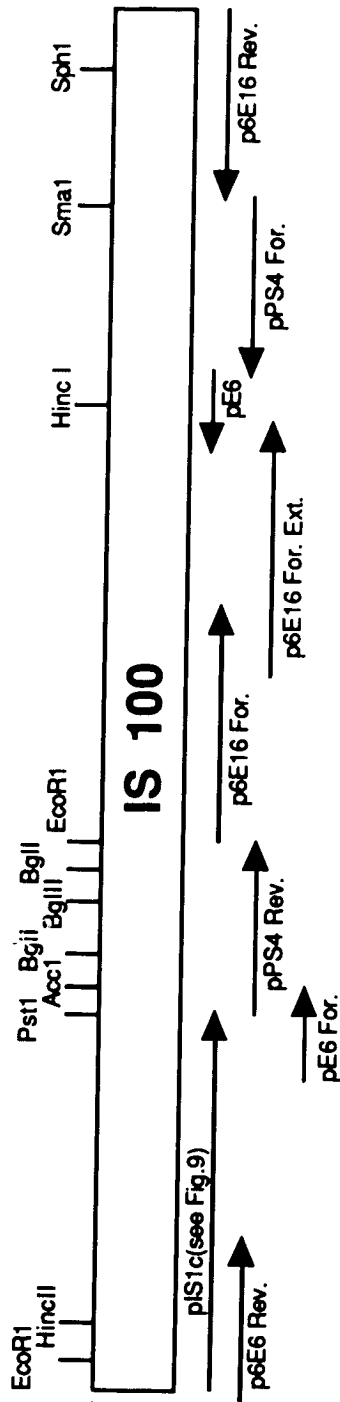
73

Fig. 11 Restriction map and sequencing strategy for IS100.

74

diagramatically in Fig. 11. As a result of overlapping by use of GCG programs, the sequence of IS 100 was assembled and analysed.

Sequence Analysis of the 9, 2.2, 1.2, and 0.6 kb Fragments of Y. pestis

Analysis of IS100 using GCG software (Altschul et al., 1990) revealed two large open reading frames (ORF's) in the plus strand and two in the minus strand. The ends of IS 100 were bounded by a 13 bp imperfect direct repeat (Table 8). The total length of the sequence was 2246 bp. Comparison of the IS 100 sequence against the GenBank sequence database using Blast analysis (Altschul et al., 1990) revealed 87% homology over 33 bp (bases 1398-1430) of the plus strand with E. coli transposable element IS 21 istA and istB genes and 62% homology over 150 bp (bases 206-355) of the minus strand. Likewise, 71% sequence homology existed over 56 bp (bases 1228-1283) of the plus strand with Bacillus stearothermophilus IS5376, and 70% homology over 55 bp (bases 1398-1452) of the plus strand of Shigella sonnei IS640 .

In addition, 70% sequence homology was noted over 54 bp (bases 534-481) with E. coli ATP dependent protease binding subunit clpB gene. Unexpectedly, 95% homology over 40 bp (bases 468-507) of the plus strand existed with the virF gene of Shigella dysenteriae and 95% homology over forty bases (468-507) with the plus strand of the virF gene of Shigella flexneri which is contained on a 140 MD plasmid.

9kb Insert. The clone containing the 9 kbp putative IS was partially sequenced and the partial sequence shared homology with the capsule formation regulator caf1 of Y. pestis as well as the attachment invasion locus(ail) of Y. enterocolitica and the plasmid virulence protein yopH of Y. pseudotuberculosis. In addition, this clone had homology with the Streptococcus agalactiae IS 861.

75

1.2kb Fragment The clone of the 1.2 kbp putative IS element was also partially sequenced and the sequence analysed. This clone showed extensive homology with common 16s ribosomal RNA sequences with literally hundreds of organisms.

0.6kb Fragment Likewise, the clone containing the 0.6kbp putative IS element was partially sequenced. This clone shared sequence homology with the *Y. pestis caf*1 gene as well as the *clp*B gene of *B. nodosus* that encodes a regulatory subunit of ATP dependent protease.

Comparison of the amino acid sequence translated from the nucleotide sequence of IS 100 was also run against the protein database of GenBank using the blast X program (Altschul et al., 1990). The *ist*B protein of *E. coli* IS 21 in the -1 reading frame showed 69% positive amino acid homology corresponding to nucleotides 389-469 of IS100. In the -2 reading frame, there was 66% positive homology with the *ist*B protein. In addition, 84% positive amino acid homology was observed with *ist*B in the +3 reading frame corresponding to nucleotides 1377-1451 and 58% positive homology with *ist*B corresponding to nucleotides 1225-1296 in the +1 reading frame of IS100.

Relative to *Bacillus stearothermophilus* IS5376, the amino acid sequence of IS100 had 63% positive homology in the -2 reading frame over bases 649-368 of and 68% positive homology in the -1 reading frame for amino acid's corresponding to nucleotides 371-126. In the +3 reading frame, there is 74% positive homology to the amino acids corresponding to nucleotides 1377-1457 and 100% positive homology in the +1 reading frame corresponding to nucleotides 1225-1284 of IS100.

The *ist*B gene protein of *B. thurengensis* IS232 was 55% homologous to IS100 in the -1 reading frame for amino acids corresponding

76

to nucleotides 353-207 and 76% for bases 200-126. In the -2 reading frame, there was 55% homology between amino acids corresponding to nucleotides 646-380. The amino acid homology was 80% in the +1 reading frame corresponding to bases 1225-1284 of IS100.

There was 84% positive amino acid homology of IS100 with IS 640 of *Shigella sonnei* corresponding to bases 1377-1451 in the +3 reading frame and 58% positive homology in the +1 reading frame corresponding to nucleotides 1225-1296.

There was 73% positive amino acid homology of IS100 with *Pseudomonas syringae* in the +1 coding frame and 62% positive homology to the amino acids of transposon Tx1 of *Xenopus laevis* corresponding to nucleotides 1821-1916.

There was 44% positive amino acid homology in the -2 reading frame of the *clp*A protein ATP binding subunit of *E. coli* corresponding to nucleotides 589-374 of IS100. There was 100% positive homology in the -2 reading frame of *Brevibacterium* ATP dependent preotease and corresponding nucleotides 538-494 of IS100. There was 46% positive amino acid homology in the -2 reading frame to the ATP dependent *clp* protease ATP binding subunit of *Rhodopseudomonas blastica* and IS 100 corresponding to nucleotides 547-374.

There was 72% positive amino acid homology in the -2 reading frame between *E. coli* 5 methylcytosine-specific restriction enzyme B and IS 100, corresponding to nucleotides 541-476. There was also 57% positive amino acid homology in the -3 reading frame between *Azotobacter vinelandii* and IS 100.

Expression of pIS1C in Maxicells

77

The largest cloned fragment (pISC) of the IS100 and pUC19 was electroporated separately into two maxicell (*E. coli* LCD44) populations. Three strains of maxicells, one containing pIS1C, one containing pUC19 and the other harboring no plasmid were subjected to the protocol for maxicells as stated in "Materials and Methods". The proteins were extracted and subjected to polyacrylamide gel electrophoresis. The resulting autoradiogram is shown in Fig. 12. A minimum of 13 additional protein bands were observed in the protein profile of pIS1C than in either the pUC19 or plasmidless profiles.

Two-dimensional(2-D) gel electrophoresis

pIS1C and pUC19 were electroporated into an attenuated strain of *Y. pseudotuberculosis* (Trp⁻ Ca⁻). Duplicate cultures (26^0C and 37^0C) of each of the following, *Y. pseudotuberculosis* Trp⁻ Ca⁻, *Y. pseudotuberculosis* Trp⁻ Ca⁻ pUC19 and *Y. pseudotuberculosis* Trp⁻ Ca⁻ pIS1C were prepared for 2-D electrophoretic analysis according to protocols described in "Materials and Methods". The resulting autoradiographs (Figs. 13-18) of the strains grown at 26^0C showed the presence of at least 2 new proteins present in the pIS1C containing strain and 5 proteins present in both the pUC19 and pIS1C containing strains. In addition, there were at least 3 new proteins present only in the plasmidless strain. These proteins disappear in pUC19 and pIS1C containing strains. There are 4 proteins present in the plasmidless and pUC19 containing strains which disappear in the pIS1C containing strain. The coordinates of these proteins are compiled in Table 9. At 37^0C at least 12 proteins were present in the pUC19 and plasmidless strains which disappear in the pIS1C containing strain. There are 2 new proteins present in the pIS1C and pUC19 possessing strains. The plasmidless strain had at least 4 proteins which disappear in the other strains and 1 protein was found in the

78
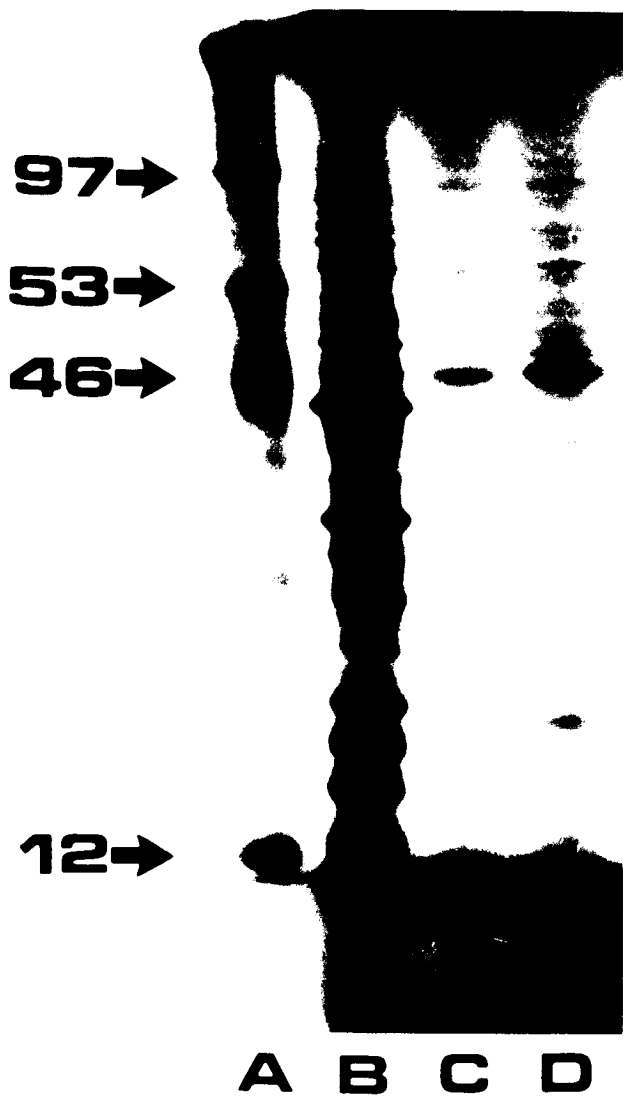
Fig. 12 Protein profiles from *E. coli* LCD44 (maxicells) cultures with and without plasmids. Lane A, protein standard (sizes in margin are in KD); Lane B, *E. coli* LCD44 containing pIS1C; lane C, *E. coli* LCD44 containing pUC19 and lane D is *E. coli* LCD44.

79
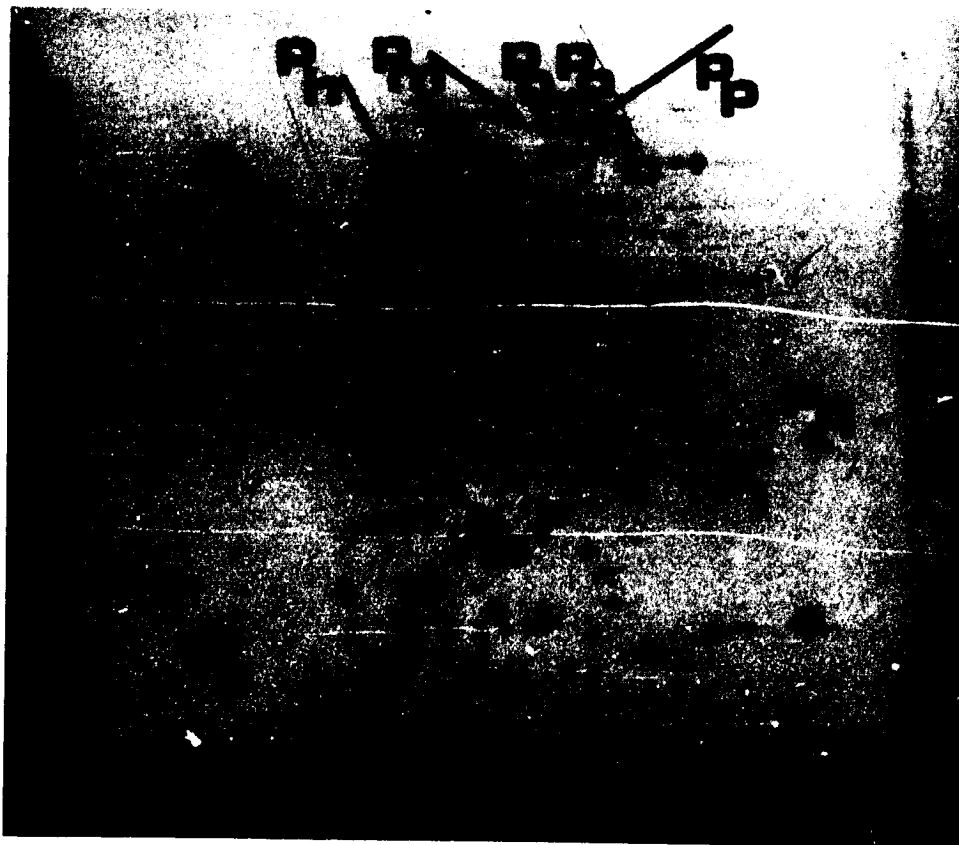
Fig. 13 Autoradiograph of two dimensional gel of proteins from Y.
pseudotuberculosis Trp- Ca- grown at 37⁰C.

Fig. 14 Autoradiograph of two dimensional gel of proteins from *Y. pseudotuberculosis* Trp- Ca- pUC19 grown at 37⁰C.

81

Fig. 15 Autoradiograph of two dimensional gel of proteins from *Y. pseudotuberculosis* Trp⁻ Ca⁻ pIS1C grown at 37⁰C.

82

Fig. 16 Autoradiograph of two dimensional gel of proteins from *Y. pseudotuberculosis* Trp⁻ Ca⁻ grown at 26⁰C.

83
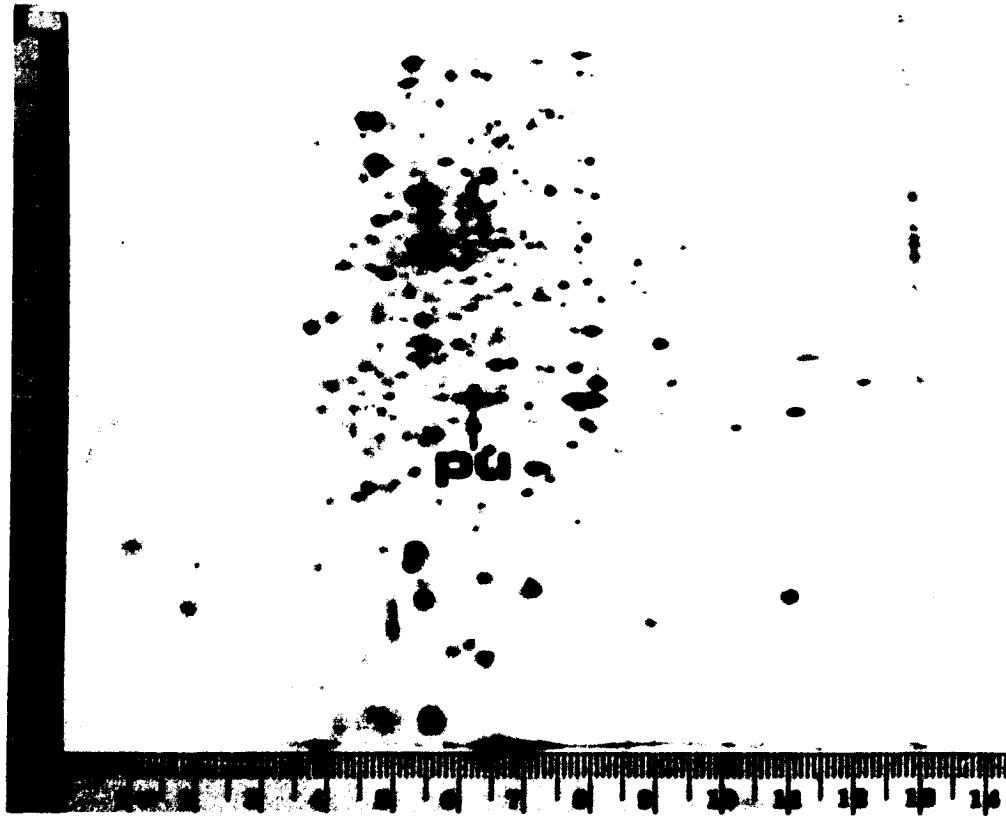
Fig. 17 Autoradiograph of two dimensional gel of proteins from *Y. pseudotuberculosis* Trp⁻ Ca⁻ pUC19 grown at 26⁰C.

84
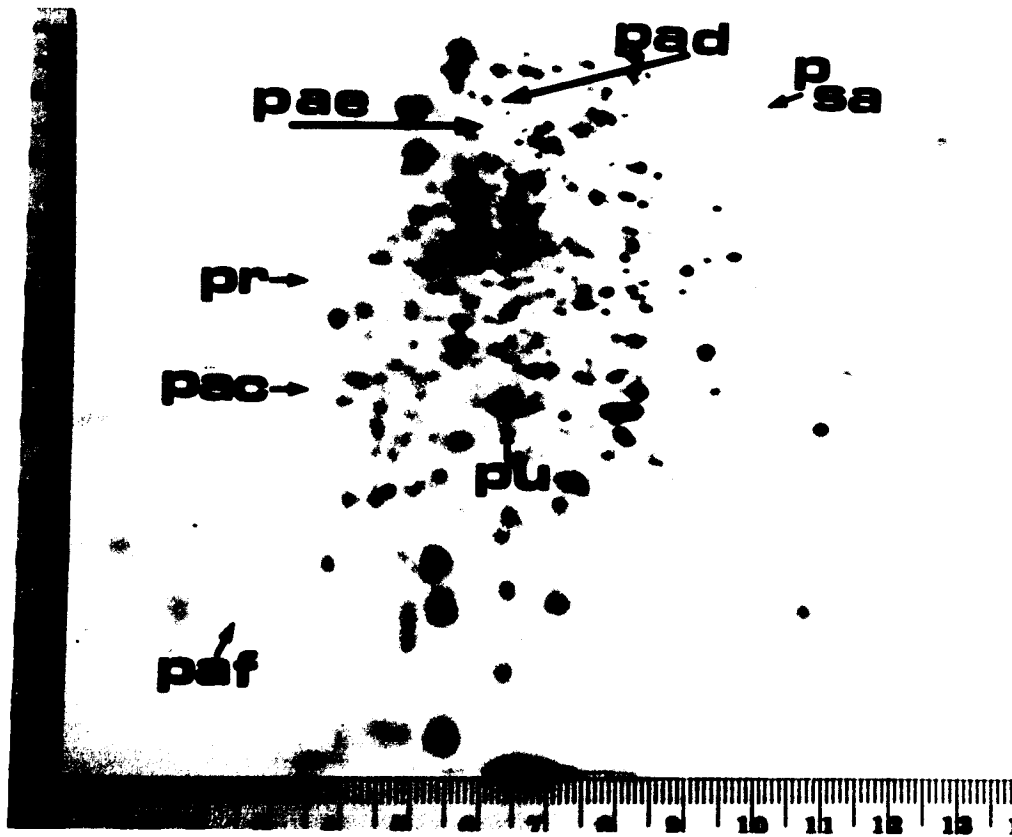
Fig. 18 Autoradiograph of two dimensional gel of proteins from *Y. pseudotuberculosis* Trp- Ca- pIS1C grown at 26⁰C.

85

| Protein Designation | Protein present in clone with | | | Coordinates | |
	no plasmid	pUC19	insert	vertical	horizontal
Pac	-	+	+	5.6	3.6
Pad	-	+	+	0.85	6.15
Pae	-	+	+	1.1	6.0
Paf	-	+	+	8.5	2.5
Pu	-	+	+	5.5	6.3
Pv	+	-	-	6.25	11.0
Paa	+	-	-	7.9	7.4
Pab	+	-	-	3.2	4.1
Pr	-	-	+	3.7	3.55
Psa	-	-	+	1.0	10.05
Px	+	+	-	5.6	14.0
Py	+	+	-	10.1	5.5
Pz	+	+	-	10.7	6.45
Pw	+	+	-	9.5	9.6

Table 9. Coordinates of proteins of interest in 2-D gels of *Y. pseudotuberculosis* Trp⁻ Ca⁻ cells grown at 26⁰C.

86

plasmidless and pIS1C strains which disappear in the pUC19 containing strain. A minimum of 3 proteins were present only in the pIS1C strain. The coordinates of these proteins are compiled in Table 10.

Utilization of Rhamnose, Melibiose and Urease Production by Y. pseudotuberculosis Containing pIS1C

The strains constructed for 2-D polyacrylamide gel analysis were examined for their ability to utilize rhamnose and melibiose, and for production of urease at 26°C and 37°C. These traits are commonly used to distinguish Y. pestis from Y. pseudotuberculosis. None of the traits was affected at 26°C but at 37°C, detection of urease was delayed by 24-48 hr and the utilization of rhamnose was inhibited completely. These results are summarized in Table 11.

Detection of Urease Gene Sequences by the Polymerase Chain Reaction (PCR)

The polymerase chain reaction DNA amplification (Mullis and Faloona, 1987) techniques were employed to investigate whether or not the sequences associated with urease genes were present in Y. pestis. The results obtained from agarose gels of the products of PCR amplification using chromosomal from agarose gels of the products of PCR amplification using chromosomal DNA from Y. pestis, Y. pseudotuberculosis, and Y. enterocolitica as template and synthetic oligodeoxynucleoside primers derived from published sequence of Proteus vulgaris (Mörsdorf and Kaltwassen, 1990) are shown in Fig 19. Both Y. pestis and Y. pseudotuberculosis amplification resulted in an approximately 1 kbp product, whereas no amplification was evident with Y. enterocolitica.

87

Protein Designation	Protein present in clone with			Coordinates	
	no plasmid	pUC19	insert	vertical	horizontal
Pa	+	+	-	1.8	6.7
Pb	+	+	-	1.6	8.1
Pc	+	+	-	8.7	2.7
Pd	+	+	-	7.9	1.9
Pe	+	+	-	6.0	6.1
Pf	+	+	-	4.4	4.7
Pg	+	+	-	3.3	5.7
Ph	+	+	-	2.55	5.9
Pi	+	+	-	4.0	5.3
Pj	+	+	-	3.7	5.8
Pk	+	+	-	2.75	3.9
Pl	+	+	-	5.8	11.0
Bla	-	+	+	5.7	8.0
Blb	-	+	+	6.2	7.05
Pm	+	-	-	0.45	7.35
Pn	+	-	-	0.45	7.6
Po	+	-	-	0.45	8.0
Pp	+	-	-	5.0	7.95
Pq	+	-	+	5.5	11.9
Pr	-	-	+	10.7	6.4
Ps	-	-	+	10.8	6.3
Pt	-	-	+	2.0	4.3

Table 10. Coordinates of proteins of interest in 2-D gels of Y. *pseudotuberculosis* Trp- Ca- cells grown at 37ºC.

88

Strain or Isolate	26°C after			37°C after		
	24hr	48hr	72hr	24hr	48hr	72hr
	RMU	RMU	RMU	RMU	RMU	RMU
Y. pseudotuberculosis Trp⁻ Ca⁻	+++	+++	+++	+-+	+-+	+-+
Y. pseudotuberculosis Trp⁻ Ca⁻ pUC19	+++	+++	+++	+-+	+-+	+-+
Y. pseudotuberculosis Trp⁻ Ca⁻ pIS1C #1	+++	+++	+++	---	---	--+
#2	+++	+++	+++	---	---	--+
#3	+++	+++	+++	---	---	--+
#4	+++	+++	+++	---	---	--+
#5	+++	+++	+++	---	---	--+
#6	+++	+++	+++	---	---	--+

Table 11 Repression of rhamnose (R) utilization and the delay in urease (U) detection in six isolates of *Y. pseudotuberculosis* harboring a cloned fragment from *Y. pestis*.

89

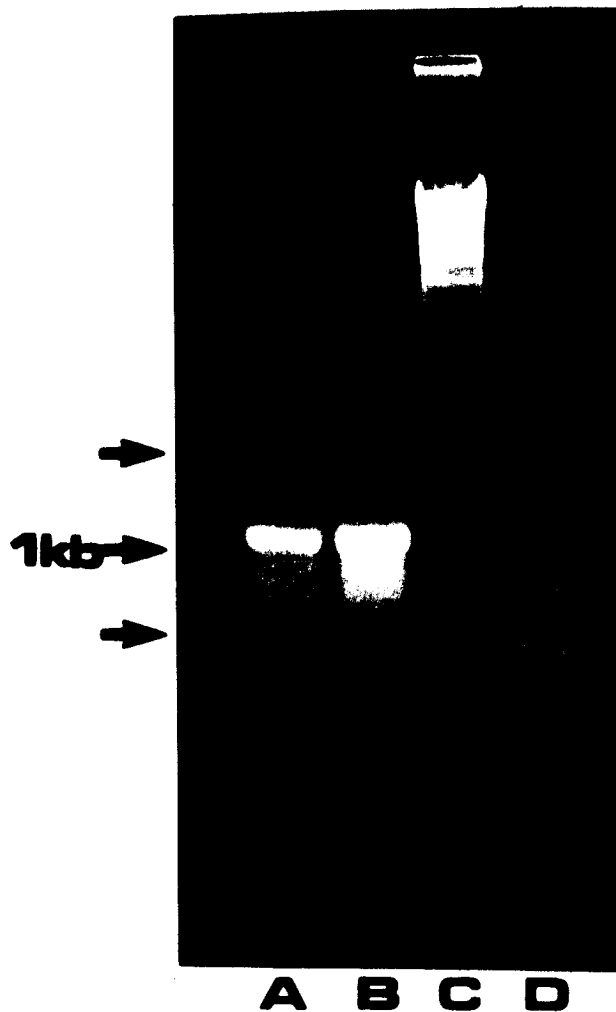**Fig. 19** Amplification of a urease gene in *Yersinia* by PCR. Chromosomal DNA from the three medically important strains of *Yersinia* was used as template for amplification by PCR. Primers were derived published sequence (Mörsdorf and Kaltwassen, 1990) of urease genes from *Proteus vulgaris*. Lane A,*Y. pestis*; lane B, *Y. pseudotuberculosis*; lane C, DNA digested with Pst1; lane 3 *Y. enterocolitica* .

90

DISCUSSION

The aim of this study was to sequence IS100 of *Y. pestis*, with the ultimate objective of using this information to determine if IS elements were the cause, at least in part, of speciation in the *Yersinia*. One of my goals, therefore, was to clone IS 100 in its entirety for use in additional studies. If the initial premises were accurate, one would expect to see an organism more characteristic of *Y. pestis* when the IS elements are introduced into *Y. pseudotuberculosis*. Unfortunately, I was unable to clone IS100 in its entirety, largely due to the interference by factors inherent in the IS isolation protocol. These factors included single stranded DNA of variable length as well as random sized double stranded DNA fragments remaining in the reaction mixture. These factors most likely interfered with or prevented ligation. It was clear that a method was needed to separate the desired DNA products (IS) from the undesirable products (single stranded, random double stranded DNA). Glass beads, sephadex, organic solvents all proved to be detrimental to successful cloning. The problem was solved by restricting the desired band to give smaller fragments which migrated to different positions on agarose gels and which were recoverable.

The only clone that was recovered directly from the 2.2 kbp fragment was pIS1C and this clone was completely sequenced in both directions. I was able to obtain the sequence of IS100 by primer extension, using sequence information from pIS1C. Sequence extension was carried out on clones of the 6 MD plasmid. These clones were not chosen at random, rather by agreement of 6 MD restriction data and restriction data of IS100 compiled by Portnoy and Falkow (1981) and our lab. This combined data determined the position of

91

IS100 on the map of the 6 MD plasmid and allowed me to select particular parts of the 6 MD plasmid as targets for cloning. Sequence analysis of these clones resulted in overlaps that were detected using computer analysis. Some of the overlapping sequence data were generated by direct sequence analysis of the 6 MD plasmid (pE6). To my knowledge this represents the first report of sequence information derived directly from a Yersinia plasmid. Using pE6 did present difficulties, however, one of which I felt was inferior plasmid preparations due to lower copy number than that of pUC19. This resulted in the presence of more contaminants and thus more interference with the sequencing reaction.

Blast analysis of pIS1C showed homology with IS21 of E. coli, virF of Shigella and with the E. coli ATP dependent protease ClpB. This finding was of particular interest because the structure of IS100 was more consistent with that of a transposon rather than a simple IS element. Further evidence supporting this assumption was the presence of two large ORF's in both the positive and negative strands. A simple IS element usually has only one or two large ORF's. In addition, IS100 was bounded by an imperfect 13 bp direct repeat rather than inverted repeats as noted with IS elements.

Also connected with these findings was the report of Portnoy and Falkow (1981). These authors indicated that IS100 was approximately 2.2 kbp, and was located on the BamH1-2 fragment of the VWa plasmid of Y. pestis 019 and on the BamH1-5 fragment of Y. pestis EV76. One of the products of the IS isolation was approximately 2.2 kbp, and when used as a probe, significant homology with the BamH1-5 fragment of the VWa plasmid of Y. pestis EV76 and with the BamH1-2 fragment of Y. pestis 019 existed. Likewise, sequence analysis of part of the 2.2 kbp fragment (pIS1C) showed

92

homology with other IS elements. From these data I concluded that the 2.2 kbp fragment from the IS isolation procedure of Ohtsubo (1981) was IS100.

A total of 2246 bp of sequence were obtained for IS100. This sequence was obtained by overlapping sequence analysis of 6 MD clones, pPS4, p6E6, pE6, and p6E16. As indicated above, IS100 was found to have 13bp imperfect direct repeats with no known homology to other known repeats of IS elements. These repeats are consistent with many other IS elements (Table 3) by size and the fact that they are in direct orientation. Direct repeat IS elements lend themselves to intramolecular recombination more readily than inverted repeat sequences (Roberts et al., 1991) and this tends to result in loss of intervening sequences. Direct repeat IS elements also are less likely to leave a copy of themselves behind. IS100 was not duplicated on transposition as shown by Portnoy and Falkow (1981). Thus the structure of IS100 may have direct consequences for speciation between Y. pestis and Y. pseudotuberculosis. If Y. pseudotuberculosis received a copy of IS100 in a given gene and the gene was disrupted becoming "pestis-like", the gene could be induced to "turn on" by complementation in trans by transposase or by acquisition of another IS100 resulting in the loss of both IS elements by intramolecular recombination and a return to normal gene expression.

IS100 has two major ORF's of approximately 400 and 500 bp in the plus strand capable of encoding proteins of approximately 15 and 19 KD respectively. These proteins would be smaller than the majority of transposases but as in the case of IS3 (Timmerman and Tu, 1985) and many other IS elements, a short sequence encodes their transposases. IS100 may possibly encode a dysfunctional transposase or require another IS to act as promoter as in the case of IS21 (Riemann et al., 1989). This would in part

93

explain the low frequency of transposition seen by Portnoy and Falkow for IS100. Low levels of transposition can also be explained by relative promoter strength (Raleigh and Kleckner, 1986). An incoming IS100 would provide the extra promoters to enhance the production of transposase.

Computer analysis of IS100 sequences revealed a number of possible promoters in different reading frames oriented outward as well as inward of the molecule. This is analogous to many IS elements such as IS10. These promoters can influence the effects of IS element in many ways. IS10 has been shown to have reduced transposition capabilities when more than one copy resides in the cell due to reduced ribosome binding caused by pairing of products of promoters directed inward (pIN) and promoters directed outward (pOUT) (Ma and Simons, 1990). Another consequence of multiple promoters in multiple reading frames is the possibility of turning on and off genes. This phenomenon is observed in many IS elements and is responsible for the discovery of controlling elements (McClintock, 1941).

The blast program used to analyse IS100 gives a "best fit" scenario to homologies. This program is designed to reduce the background and leave the researcher with meaningful homologies. This approach omits the feeling for the relatedness of two sequences over the entire course of the sequence. I was interested in seeing if IS100 as a whole was closely related to other IS elements. Consequently, I obtained the entire DNA sequences of some of the close homologs from blast analysis. I compared these sequences against these homologs one on one using the Gap algorithm in the GCG software package. The results are summarized in Table 11. There are two important findings from this analysis. One is the extent of homology associated with a broad range of IS elements. The other is that every significant homology of IS100 is

94

Table 12 Results of Gap analysis of IS100.

IS100 homology with	peptide/ DNA	#bp/aa	region of match	%similarity	%identity
clpB of E. coli	DNA	3503	701-3089		39.6
IS640 of S. sonnei	DNA	1097	1-1097		40.2
IS21 of E. coli	DNA	2131	1-2131		38.9
IS5376 of B. stearothermophilus	DNA	2107	1-2107		35.9
virF of S. flexneri and S. dysenteriae	DNA	1018	1-1018		35.6
IS5376	aa	702	1-400	46.5	19.6
IS640	aa	365	1-315	44.1	20.5
nif regulatory protein	aa	522	1-522	43.4	20.1
virF	aa	339	1-262	42.7	15.8
transposable element Tx1c protein B	aa	1308	351-1308	41.9	19.5
transposable element Tx1c protein A	aa	571	1-377	41.1	19.1
5-methylcytosine specific enzyme B	aa	465	1-465	39.8	17.0
IS5376 ORF B	aa	702	1-251	43.4	20.3
clpB	aa	1167	1-390	43.4	22.7
IS21	aa	710	1-390	43.4	22.7
clpA	aa	594	1-594	40.2	16.9
istB of IS21	aa	250	1-250	43.7	21.4

95

with a gene product or DNA sequence that is either involved in regulation (has affinity for DNA or protein) or potentially has the ability to cleave DNA or protein. These homologies are particularly interesting at the amino acid level where they do not share as much identity as they do similarity. This means that although the sequence may not be exactly the same, the corresponding protein has identical or similar shape and possibly function.

In experiments to duplicate the work of Portnoy and Falkow (1981), a number of Ca⁻ mutants were obtained from *Y. pestis* 019. Some of these mutants had no detectable plasmid, while some mutants harbored plasmids of smaller or larger sizes. One mutant, *Y. pestis* 019 Ca-4, had an approximately 2.2 kbp insert into the BamH1-2 fragment. When these data and the data from Fillipov et al. (1990) were examined together, a plausible explanation for every type of Ca⁻ mutant is evident. When subjected to the stress due to lack of calcium, IS100 may be incited to transpose. If one transposition takes place, the new IS100 may insert into any number of sites on the VWa plasmid. This can result in cointegrate formation resulting in larger plasmids, smaller plasmids, plasmid loss or possibly integration by reciprocal recombination with IS100 located on the chromosome, depending on the site of integration of the incoming IS. If the transposition does not result in cointergate formation, it may be due to activation in trans, resulting in a stable insert as in Ca-4.

During the time it took to develop and implement a sequencing strategy for the remainder of IS100, I attempted to determine the effects that pIS1C might have on the "variable" traits of *Y. pestis*. I felt that the introduction of the IS elements native to *Y. pestis* into *Y. pseudotuberculosis* may cause *Y. pseudotuberculosis* to become "pestis-like". I tested this

96

hypothesis by electroporating pIS1C into an attenuated strain of *Y. pseudotuberculosis* (*Y. pseudotuberculosis*Trp⁻Ca⁻) and determining the production of urease and utilization of rhamnose and melibiose. At 26⁰C, all of the traits remained unchanged as in the wild type, but at 37⁰C detection of urease was delayed by 24 to 48 h and rhamnose utilization was completely repressed. Fermentation of melibiose was the same as in the wild type. These results clearly suggested that there was a temperature dependent product from pIS1C that was being expressed in *Y. pseudotuberculosis* Trp-Ca⁻ at 37⁰C. This conclusion was reinforced by results of analyses of both maxicells and two-dimensional electrophoresis. The results of maxicells clearly showed a host of proteins not present in either plasmidless maxicells or maxicells containing pUC19. pIS1C itself has only 1 small open reading frame of approximately 200 bp. This would encode for 66 amino acids resulting in a protein of about 7 KD. Analysis of maxicell plus pIS1C protein production showed at least 10 proteins ranging in size from 92 KD to 10 KD that did not appear in plasmidless maxicells or maxicells containing pUC19. Three plausible explanations may account for the large number of proteins produced. The product of the pIS1C may be a protease, or subunit thereof, capable of digesting the products of *amp*R and β-gal genes present on the pUC19 vector. The other obvious alternative is that the approximately 7 KD protein is actually a subunit that is capable of forming a variety of aggregates despite the presence of SDS, resulting in the formation of the numerous bands seen in Fig.12. A third possible explanation is that the pIS1C provided promoters directed outward that resulted in a variety of truncated products. This phenomenom has been demonstrated in many IS elements including IS21 which has sequence homology with pIS1C.

97

The result of delayed urease production by *Y. pseudotuberculosis* harboring pIS1C raised the possibility that IS100 could be involved in repressing the urease operon in *Y. pestis*. Urease+ isolates of *Y. pestis* can be isolated readily indicating that the urease operon must be present in this bacterium but is inactive. Thus PCR amplification of a portion of the urease operon from *Y. pseudotuberculosis*, *Y. pestis* and *Y. enterocolitica* was attempted. The nucleotide sequence for urease production is known for *Helicobacter pylori* (Labigne et al., 1991) and *Proteus vulgaris* (Mörsdorf and Kaltwassen, 1990). Because the urease operon of *Proteus vulgaris* is simpler and because I felt the organism was more closely related to the *Yersiniae*, primers to the urease D gene of *P. vulgaris* were constructed. Products of identical size resulted from amplifications using *Y. pestis* and *Y. pseudotuberculosis* DNA as template. No detectable product appeared using *Y. enterocolitica* as template. These results are the first to detect urease genes in *Y. pestis*. A search of the entire urease operon would be necessary to confirm the postulated influence of IS100.

The clone of the 9 kbp putative IS from *Y. pestis* has been electroporated into *Y. pseudotuberculosis* Trp-Ca- and screened for urease production and rhamnose utilization. Production of urease is completely inhibited at 26°C and 37°C, while rhamnose fermentation is delayed at 26°C (Jennifer Matteau, personal communication). These results suggest that there are two mechanisms of thermal regulation.

I decided that two-dimensional (2-D) polyacrylamide gel electrophoresis of whole cell protein may provide some clarification as to what effects pIS1C was producing on the cell. The strains of *Y. pseudotuberculosis*Trp Ca- with and without plasmids used in the carbohydrate utilization study

98

were grown at both 26⁰C and 37⁰C, and their protein content analysed. At 26⁰C there were 2 proteins, P_r and P_{sa}, produced in the culture harboring pIS1C that were not produced in either of the other two strains. Protein P_r was of slightly acidic isoelectric point (pI) and approximately 50 KD and P_{sa} was about 72 KD and basic pI. At 37⁰C, three proteins P_r, P_s, and P_t were present in the culture harboring pIS1C but not in the plasmidless or pUC19 containing strains. P_s was slightly acidic to neutral and about 10 KD while P_r was slightly more acidic than P_s and slightly larger at about 10.5 KD. P_t was an acidic protein of about 65 KD. From this evidence it is clear that there are proteins being produced by or as a result of the presence of pIS1C. Determining what these proteins are may only be decided by 2-D analysis of mutants.

There appears to be some type of regulatory function associated with pIS1C. *Y. pseudotuberculosis*Trp-Ca- plus pIS1c at 26⁰C produces five proteins P_{ac}, P_{ad}, P_{ae}, P_{af}, and P_u that are also produced in *Y. pseudotuberculosis*Trp-Ca-plus pUC19 but not in the plasmidless *Y. pseudotuberculosis*Trp-Ca-. In addition four proteins, P_x, P_y, P_z and P_w, are produced in *Y. pseudotuberculosis*Trp-Ca- and *Y. pseudotuberculosis*Trp-Ca- plus pUC19 that are not produced in *Y. pseudotuberculosis*Trp-Ca- plus pIS1C at 26⁰C. Similar but more dramatic results are realized at 37⁰C with 12 proteins produced in *Y. pseudotuberculosis*Trp-Ca- and *Y. pseudotuberculosis*Trp-Ca- plus pUC19 but not *Y. pseudotuberculosis*Trp-Ca- plus pIS1C. Only two proteins were produced in *Y. pseudotuberculosis* Trp-Ca- plus pUC19 and pIS1C that were not produced in the plasmidless strain. The best evidence of regulatory control exerted by pIS1C is the disappearance of 16 proteins at 37⁰C and 7 proteins at 26⁰C when pIS1C is present. These data have even more meaning when looked at in conjunction with data concerning homologs of pIS1C. The 821 bp piece of IS100 shows

homology with *virF*, *clpB*, *istA* of IS21, and with plasmid RP4 *aphA* gene encoding aminoglycoside 3'-phosphotransferase. There is 80% homology in the first 100 bases of the amino terminus.

In conclusion, I have successfully sequenced IS100 of *Y. pestis*. The direct result of this work opens many avenues for future research. This and future studies will make significant progress to our understanding of speciation and control, not only in the *Yersiniae*, but possibly in other pathogens as well.

LITERATURE CITED

Altschul, S. F., Gish, W., Miller, W.,Myers, E. W., Lipman, D. J.. (1990) Basic Local Alignment Search Tool. J. Mol. Biol. 215: 403-410.

Andersen, A. S., Pettersson, A. F., Kjeldsen, T. B.. (1992) A Fast and Simple Technique for Sequencing Plasmid DNA with Sequenase Using Heat Denaturation. BioTechniques 13 (5): 678-680.

Barsomian, G., Lessie,T.G.. (1987) IS2 Activates the *ilvA* Gene of *Pseudomonas cepacia* in *Escherichia coli*. J. Bacteriol. 169(4):1777-1779.

Bartlett, D.H., Silverman, J.. (1989) Nucleotide Sequence of IS492, a Novel Insertion Sequence Causing Variation in Extracellular Polysaccharide Production in the Marine Bacterium *Pseudomonas atlantica*. J. Bacteriol. 171 (3): 1763-1766.

Bercouvier, H., Mollaret, H. H.. (1984) In: Bergey's Manual of Systematic Bacteriology, Holt, J. G., Krieg, N. R. (eds.). (Williams and Wilkins, Baltimore) Vol. 1 pp. 498-506.

Bercouvier, H., Mollaret, H. H., Alonso, J. M., Brault, J., Fanning, G. R., Steigerwalt, A. G., Brenner, D. J.. (1980) Intra- and Interspecies Relatedness of *Yersinia pestis* by DNA Hybridization and its relationship to *Yersinia pseudotuberculosis*. Curr. Microbiol. 4: 225-229.

Berg D. E., Egner, C., Hircshel, B. J., Howard, J., Johnsrud, L., Jorgensen, R. A., Tlsty, T. D.. (1981) Insertion, Excision, and Inversion of Tn5. pp.115-124. In: Movable Genetic Elements, C.S.H. Symp. Quan.Biol. Vol.15.

Bessonova, A., Lenskaya, G., Moldtsova, P., Mosalova, O.. (1936) Reports Concerning Certain Facts of the Spontaneous Transformation of *Bact. pestis* into *Bact. pseudotuberculosis* rod. Vestn. Mikrobiol. Epidemiol. Parazitol. 15: 2-151.

Birnboim, H. C., Doly, J.. (1979) A Rapid Alkaline Extraction Procedure for Screening Recombinant Plasmid DNA. Nucleic Acids Res. 7: 1513-1523.

Bliska, J. B., Falkow, S.. (1992) Bacterial Resistance to Complement Killing Mediated by the Ail Protein of *Yersinia enterocolitica*. PNAS 89 (8): 3561-3565.

Bliska, J. B., Guan, K., Dixon, E., Falkow, S.. (1991) Tyrosine Phosphate Hydrolysis of Host Proteins by an Essential *Yersinia* Virulence Determinant. PNAS 88: 1187-1191.

Brenner, D. J., Steingerwalt, A. G., Falcao, D. P., Weaver, R. E., Fanning, G. R.. (1976) Characterization of *Y. enterocolitica* and *Y. pseudotuberculosis* by Deoxyribonucleic Acid Hybridization and by Biochemical Reactions. Int. J.

Syst. Bacteriol. 26: 180-194.

Brisson-Noel, A., Arthur, M., Courvalin, P.. (1988) Evidence for Natural Gene Transfer from Gram-Positive Cocci to *Escherichia coli*. J. Bacteriol. 170 (4) 1739-1745.

Broisus, J., Walz, A.. (1983) DNA Sequences Flanking an *Escherichis coli* Insertion Element IS2 in a Cloned Yeast (*Saccharomyces cerevisiae*) TRP5 Gene. Gene (AMST) 17 (2): 223-228.

Brown, S. D., Montie, T. C.. (1977) Beta-andrenergic Blocking Activity of *Yersinia pestis* Murine Toxin. Infect. Immun. 18: 85-93.

Brubaker, R. R.. (1972) The Genus *Yersinia*: Biochemistry and Genetics of Virulence. C. T. in Microbiology 57: 112-157.

Brubaker, R. R.. (1991) Factors Promoting Acute and Chronic Diseases Caused by *Yersiniae*. Clin. Micro. Rev. 4 (3): 309-324.

Brubaker, R. R., Sample, A. K., Yu, D.-Z., Zahorchak, R. J., Hu, P. C., Fowler, J. M.. (1987) Proteolysis of V Antigen from *Yersinia pestis*. Microb. Pathog. 2: 49-62.

Burrows, T. W., Bacon, G. A.. (1958) The Effects of Loss of Different Virulence Determinants on the Virulence and Immunogenicity of Strains of *Pasteurella pestis*. Brit. J. Exp. Path. 39: 278-291.

Burrows, T. W., Bacon, G. A.. (1960) V and W Antigens in Strains of *Pasteurella pseudotuberculosis*. Br. J. Exp. Pathol. 41: 38-44.

Burton, D.. (1956) A Study on the Conditions and Mechanisms of the Diphenylamine Reaction for the Colorimetric Estimation of Deoxyribonucleic Acid. Biochem. J. 62: 315-322.

Butler, T.. (1983) Plague and Other *Yersinia* Infections. Plenum Press, N.Y., N.Y.

Case, C. C., Simons, E. L., Simons, R. W.. (1990) The IS10 Transposase mRNA is Destabilized During Antisense RNA Control. EMBO 9 (4): 1259-1266.

Celis, J. E., Bravo, R. (eds.). (1984) Two Dimensional Gel Electrophoresis of Proteins, Methods and Applications. Academic Press, Inc. Orlando, FL.

Chandler, M., Clerget, M., Caro, L.. (1981) IS1-Promoted Events Associated with Drug-resistance Plasmids. pp.157-166. In: Movable Genetic Elements, C.S.H. Symp. Quan. Biol. Vol.15.

Charnetzky, W. T., Brubaker, R. R.. (1982) RNA Synthesis in *Yersinia pestis* During Growth Restriction in Calcium-Deficient Medium. J. Bacteriol. 149: 1089-1095.

102

Chernak, J. M., Smith, H.O.. (1989) Use of Synthetic Ribosome Binding Site for Overproduction of the 5B Protein of Insertion Sequence IS5. Nuc. Acids. Res. 17:1933-1951.

Chuba, P. J., Leon, M.A., Banerjee, A., Palchaudhuri, S.. (1989) Cloning and Sequence of Plasmid Determinant*iss* Coding for Increased Serum Survival and Surface Exclusion, Which has Homology with Lambda DNA. Mol. Gen. Genet. 216:287-292.

Colonna, B., Nicoletti, M., Visca, P., Casalino, M., Valenti, P., Maimone, F.. (1985) Composite IS1 Elements Encoding Hydroxamate-Mediated Iron Uptake in F1*me* Plasmids from Epidemic *Salmonella* spp.. J. Bacteriol. 162 (1) 307-316.

Cornelius, G., Ghosal, D., Saedler, H.. (1978) Tn951: A New Transposon Carrying a Lactose Operon. Molec. Gen. Genet (MGG). 160: 215-224.

Crosa, J. H., Falkow, S.. (1981) Plasmids; in Gerhardt, Manual of Methods for General Bacteriology. (American Society for Microbiology, Washington, D. C.) pp. 266-282.

Dalrymple, B.. (1987) Novel Rearrangements of IS30 Carrying Plasmids Leading to the Reactivation of Gene Expression. Mol. Gen. Genet. 207: 413-420.

Datta, A.R., Rosner, J.L.. (1987) Reduced Transposition in *rho* Mutants of *Escherichia coli* K-12. J. Bacteriol. 169 (2): 888-890.

DeMeirsman, C., Croes, C., Desair, J., Verreth, C., Van Gool, A., Vanderleyden, J.. (1989) Identification of Insertion Sequence Element IS427 in pTiT37 Plasmid DNA of an *Agrobacterium tumefaciens* T37 Isolate. Plasmid 21: 129-137.

Derbyshire, K.M., Kramer, M., Grindley, N. D. F.. (1990) Role of Instability in the *cis* Action of the Insertion Sequence IS903 Transposase. PNAS 87: 4048-4052.

Dmochllowska, A., Berse, B., Kowalska, I., Skrzypek, M., Weglenski, P.. (1986) Expression of *Aspergillus nidulans arg*B Gene in *Escherichia coli*. J. Gen. Micro. 132: 1729-1737.

Dodson, K.W., Berg, D.E.. (1989) Factors Affecting Transposition Activity of IS50 and Tn5 Ends. Gene 76: 207-213.

Dodson, K.W., Berg, D.E.. (1989) Saturation Mutagenesis of the Inside End of Insertion Sequence IS50. Gene 85: 75-81.

Donald, R. G. K., Raymond, C. K., Ludwig, R. A.. (1985) Vector Insertion Mutagenesis of *Rhizobium* sp. Strain ORS571: Direct Cloning of Mutagenized DNA Sequences. J. Bacteriol. 162 (1): 317-323.

103

Dower, W. J., Miller, J. F., Ragsdale, C. W.. (1988) High Efficiency Transformation of *E. coli* by High Voltage Electroporation. Nucleic Acids Res. 16: 6127-6145.

Emody, L., Heeseman, J., Wolf-Watz, H., Skurnik, M., Kapperud, G., O'Toole, P., Wadstrom, L.. (1989) Binding to Collagen by *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*: Evidence for a *yop*A-Mediated and Chromosomally Encoded Mechanisms. J. Bacteriol. 171: 6674-6679.

Ferber, D. M., Brubaker, R. R., (1979) Mode of Action of Pesticin: N-acetylglucosaminadase Activity. J. Bacteriol. 139: 495-501.

Fillippov, A. A., Oleinikov, P. N., Drozdov, A. V., Protsenko, O. A.. (1990) The Role of IS Elements of *Yersinia pestis* in Forming Mutations of Calcium Independence. Genetika 26 (10): 1740-1748.

Finlay, B. B., Falkow, S.. (1989) Common Themes in Microbial Pathogenicity. Microbiol. Rev. 53: 210-230.

Forsberg, A., Wolf-Watz, H.. (1990) Genetic Analysis of the *yop*E region of *Yersinia* spp.: Identification of a Novel Conserved Locus, *yer*A, Regulating *yop*E Expression. J. Bacteriol. 172: 1547-1555.

Fox, E. N., Higuchi, K.. (1958) Synthesis of the Fraction 1 Antigenic Protein by *Pasteurella pestis*. J. Bacteriol. 75: 209-216.

Garger, S. J., Griffith, O. M., Grill, L. K.. (1984) Rapid Purification of Plasmid DNA by a Single Centrifugation in a Two-Step Cesium Chloride-Ethidium Bromide Gradient. Biochem. Biophys. Res. Commun. 117 (3): 835-842.

Garrels, J. I.. (1983) Quantitative Two-Dimensional Gel Electrophoresis of Proteins. Methods Enzymol. 100: 411-423.

Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T., Kado, C.I.. (1985) Positive Selection Procedure for Entrapment of Insertion Sequence Elements in Gram-Negative Bacteria. *J. Bacteriol*. 164: 918-921.

Gertman, E., White B. N., Berry, D.,Kropinski, A. M.. (1986) IS222, a New Insertion Element Associated with the Genome of *Pseudomonas aeruginosa*. J. Bacteriol. 166 (3): 1134-1136.

Glaser, P., Danchin, A.. (1989) Molecular Characterization of Two *cya* Mutations, *cya*-854 and *cya*R1. J. Bacteriol. 171 (9): 5176-5178.

Goldberg, J. B., Won, J., Ohman, D. E.. (1990) Precise Excision and Instability of the Transposon Tn5 in *Pseudomonas aeruginosa*. J. Gen. Micro. 136:789-796.

Haandrikman, A. J., Van Leeuwen, C., Kok, J., Vos, P., DeVos, W. M., Venema, G.. (1990) Insertional Elements in Lactococcal Proteinase Plasmids. Appl. Envir. Micro. 56 (6): 1890-1896.

104

Hall, B. G.. (1988) Adaptive Evolution That Requires Multiple Spontaneous Mutations. I. Mutations Involving an Insertion Sequence. Genetics 120 (4): 887-897.

Hall, B. G., Parker, L. L., Betts, P. W., DuBose, R. F., Sawyer, S. A., Hartl, D. L.. (1989) IS103, a New Insertion Element in *Escherichia coli*: Characterization and Distribution in Natural Populations. Genetics 121 (3): 423-431.

Hanes, B. D., Rickwood, D. (eds.). (1981) Gel Electrophoresis of Proteins. IRL Press Washington, D. C..

Harvey, W. F.. (1933) Bacteriophage with Special Reference to Plague and Cholera. Trop. Dis. Bull. 30: 411-422.

Haugland, R. A., Sangodkar, V. M. X., Chakrabarty, A. M.. (1990) Repeated Sequences Including RS1100 from *Pseudomonas cepacia* AC1100 Function as IS Elements. Mol. Gen. Genet. 220 (2): 222-228.

Highton, P.J., Chang, Y., Marcotte, JR., W. R., Schnaitman, C. A.. (1985) Evidence that the Outer Membrane Protein Gene *nmp*C of *Escherichia coli* K-12 Lies within the Defective *qsr'* Prophage. J. Bacteriol. 162 (1): 256-262.

Hoover, T. A., Vodkin, M. H., Williams, J. C.. (1992) A *Coxiella burnetti* Repeated DNA Element Resembling a Bacterial Insertion Sequence. J. Bacteriol. 174(17): 5540-5548.

Hotte, B., Rath-Arnold, I., Puhler, A., Simon, R.. (1990) Cloning and Analysis of a 35.3-Kilobase DNA Region Involved in Exopolysaccharide Production by *Xanthomonas campestris* pv. *campestris*. J. Bacteriol. 172 (5): 2804-2807.

Huisman, O., Errada, P. R., Signon, L., Kleckner, N.. (1990) Mutational Analysis of IS 10's Outside End. EMBO 8 (7): 2101-2109.

Iida, S., Meyer, J., Arber, W.. (1981) Genesis and Natural History of IS-Mediated Transposons. pp.27-37. In: Movable Genetic Elements , C.S.H. Symp. Quan. Biol. Vol.15.

Iida, S., Kulka, I., Meyer, J., Arber, W.. (1987) Amplification of Drug Resistance Genes Flanked by Inversely Repeated IS1 Elements: Involvement of IS1-Promoted DNA Rearrangements before Amplification. J. Bacteriol. 169 (4): 1447-1453.

Ishiguro, N., Sato, G.. (1985) Spontaneous Deletion of Citrate-Utilizing Ability Promoted by Insertion Sequences. J. Bacteriol. 160 (2): 642-650.

Ishiguro, N., Sato, G.. (1988) Nucleotide Sequence of Insertion Sequence IS3411, Which Flanks the Citrate Utilization Determinant of Transposon Tn3411. J. Bacteriol. 170 (4): 1902-1906.

Joseph-Liauzun, E., Fellay, R., Chandler, M.. (1989) Transposable Elements

for Efficient Manipulation of a Wide Range of Gram-Negative Bacteria: Promoter Probes and Vectors for Foreign Genes. Gene 85: 83-89.

Kearney, B., Staskawicz, B.J.. (1990) Characterization of IS476 and Its Role in Bacterial Spot Disease of Tomato and Pepper. J. Bacteriol. 172 (1): 143-148.

Klaer, R., Kuhn, S., Fritz, H.-J., Tillmann, E., Saint-Girons, I., Habermann, P., Pfeifer, D., Starlinger, P.. (1981) Studies on Transposition Mechanisms and Specificity of IS4. In: Movable Genetic Elements C.S.H. Symp. Quan. Biol. Vol. 15 .

Kleckner, R.. (1990) Regulating Tn10 and IS10 Transposition. Genetics 124: 449-454.

Kroll, J. S., Loynds, B. M., Moxon, E. R.. (1991) The *Haemophilus influenzae* Capsulation Gene Cluster: A Compound Transposon. Mol. Micro. 5 (6): 1549-1560.

Kunze, Z. M., Portaels, F., McFadden, J. J.. (1992) Biologically Distinct Subtypes of *Mycobacterium avium* Differ in Possession of Insertion Sequence IS901. J. Clin. Micro. 30 (9): 2366-2372.

Labigne, A., Cussac, V., Courcoux, P.. (1991) Shuttle Cloning and Nucleotide Sequences of *Helicobacter pylori* Genes Responsible for Urease Activity. J. Bacteriol. 173 (6): 1920-1931.

Lawrence, J. G., Dykhuizen, D. E., DuBose, R. F., Hartl, D. L.. (1989) Phylogenetic Analysis Using Insertion Sequence Fingerprinting in *Escherichia coli*. Mol. Biol. Evol. 6 (1): 1-14.

Lecker, S., Lill, R., Ziegelhoffer, T., Georgopoulos, C., Bassford, Jr.,P. J., Kumamoto, C. A., Wickner, W.. (1989) Three Pure Chaperone Proteins of *Escherichia coli*--SecB, Trigger Factor, and GroEL--Form Soluble Complexes with Precursor Proteins *in vitro*. EMBO J. 8: 2703-2709.

Lee, K.-Y., Hopkins, J. D., Syvanen, M.. (1990) Direct Involvement of IS26 in an Antibiotic Operon. J. Bacteriol. 172 (6) : 3229-3236.

Leong, J. M., Fournier, R. S., Isberg, R. R.. (1991) Mapping and Topographic Localization of Epitopes of the *Yersinia pseudotuberculosis* Invasin Protein. Infect. Immun. 59 (10): 3424-3433.

Leung, K. Y., Straley,. S. C..(1989) The *yop*H Gene of *Yersinia pestis* Encodes a Released Protein Having Homology with the Human Platelet Surface Protein GPIba. J. Bacteriol. 171: 4623-4632.

Lindler, L. E., Klempner, M. S., Straley, S. C.. (1990) *Yersinia pestis* pH 6 Antigen: Genetic, Biochemical, and Virulence Characterization of a Protein Involved in the Pathogenesis of Bubonic Plague. Infect. Immun. 58: 2569-2577.

Lucier, T. S., Brubaker, R. R.. (1992) Determination of Genome Size, Macrorestriction Pattern Polymorphism, and Nonpigmentation Specific Deletion in *Yersinia pestis* by Pulsed Field Gel Electrophoresis. J. Bacteriol. 174 (7): 2078-2086.

Lusky, M., Kroger, M., Hobom, G.. (1981) Direction of Replicational Inceptor Signals in IS5. pp.173-176. In: Movable Genetic Elements C.S.H. Symp. Quan. Biol. Vol. 15.

Ma, C., Simons, R. W.. (1990) The IS10 Antisense RNA Blocks Ribosome Binding at the Transposase Translation Initiation Site. EMBO 9 (4): 1267-1274.

Makris, J. C., Reznikoff, W. S.. (1989) Orientation of IS50 Transposase Gene and IS50 Transposition. J. Bacteriol. 171 (9): 5212-5214.

Maniatis, T., Fritsch, E. F., Sambrook, J.. (1982A) Chapter 2 Propagation and Manitenence of Bacterial Strains and Virus; in Maniatis: Molecular Cloning p.68 (Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.).

Maniatis, T., Fritsch, E. F., Sambrook, J.. (1982B) Chapter 11 Analysis of Recombinant DNA Clones; in Maniatis. Molecular Cloning p. 383-389. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.).

Maniatis, T., Jeffery, A., Kleid, D. G.. (1975) Nucleic Acid Sequence of the Rightward Operator of Phage Lambda. Proc. Natl. Acad. Sci. U.S.A.. 72: 1184-1188.

Mandel, M., Higa, A.. (1970) Calcium Dependent Bacteriophage DNA Infection. J. Mol. Biol. 53: 154.

Marmur, J.. (1961) A Procedure for the Isolation of Deoxyribonucleic Acid from Microorganisms. J. Mol. Biol. 3: 208-218.

Matsutani, S.. (1991) Multiple Copies of IS10 in the *Enterobacter cloacae* MD36 Chromosome. J. Bacteriol. 173 (24): 7802-7809.

McClintock, B.. (1941) The Stability of Broken Ends of Chromosomes in *Zea mays*. Genetics 26: 234-282.

Mendiola, M.V., De la Cruz, F.. (1989) Specificity of an Insertion of IS91, an Insertion Sequence Present in a-Hemolysin Plasmids of *Escherichia coli*.. Mol. Micro. 3 (7): 979-984.

Mills, J. A., Venkatesan, Baron, L. S., Buysse, J. M.. (1992) Spontaneous Insertion of an IS1-Like Element Into the *vir* F Gene is Responsible for Avirulence in Opaque Colonial Variants of *Shigella flexneri* 2a. Infect. Immun. 60 (1): 175-182.

Mollaret, H. H.. (1987) A Personal View of the History of the Genus *Yersinia*. Contr. Microbiol. Immunol. 9: 1-13.

Mollet, B., Clerget, M., Meyer, J., Iida, S.. (1985) Organization of the Tn6-Related Kanamycin Resistance Transposon Tn2680 Carrying Two Copies of IS26 and an IS903 Variant, IS903.B J. Bacteriol. 163 (1) : 55-60.

Mörsdorf, G., Kaltwasser, H.. (1990) Cloning of Genes Encoding Urease from *Proteus vulgaris* andf Sequencing of the Structural Genes. FEMS Microbiol. Letters 66: 67-74.

Mulder, B.,Michiels, T., Simonet, M., Cornelius, G.. (1989) Identification of Additional Virulence Determinants on the pYV plasmid of *Yersinia enterocolitica* W227. Infect. Immun. 57: 2534-2541.

Navas, J., Garcia-Lobo, J. M., Leon, J., Ortiz, J. M.. (1985) Structural and Functional Analyses of the Fosfomycin Resistance Transposon Tn2921. J. Bacteriol. 162 (3) : 1061-1067.

Nies, B. A., Meyer, J. F., Kratz, J., Wiedemann, B.. (1985) R1767, An Example of the Evolution of Resistance Plasmids. Plasmid 13 (3) : 163-172.

O'Farrell, P.H.. (1975) High Resolution Two-Dimensional Electrophoresis of Proteins. J. Biol. Chem. 250: 4007-4021.

O'Farrell, P. Z., Goodman, H. M., O'Farrell, P. H.. (1977) High Resolution Two-Dimensional Electrophoresis of Basic as well as Acidic Proteins. Cell 12: 1133-1142.

Ohta, N., Mullin, D. A., Tarleton, J., Ely, B., Newton, A.. (1990) Identification, Distribution, and Sequence Analysis of New Insertion Elements in *Caulobacter crescentus.* J. Bacteriol. 172 (1): 236-242.

Ohtsubo, H., Nyman, K., Doroszkiewicz, W., Ohtsubo, E.. (1981) Multiple Copies of Iso-Insertion Sequences of IS1 in *Shigella dysenteriae* Chromosome. Nature 292:640-643.

Ohtsubo, H., Ohtsubo, E.. (1976) Isolation of Inverted Repeat Sequences, Including IS1, IS2, and IS3, in *Escherichis coli* Plasmids. PNAS 73 (7) : 2316-2320.

Otal, I., Martin, C., Vincent-Levy-Frebault, V., Thierry, D., Gicquel, B.. Restriction Fragment Length Polymorphism Analysis Using IS6110 as an Epidemiological Marker in Tuberculosis. J. Clin. Microbiol. 26 (6): 1252-1254.

Ou, J. T., Baron, L. S., Rubin, F. A., Kopecko, D. J.. (1986) Specific Insertion of Insertion Sequence 1-Like DNA Element Causes the Reversible Expression of the Virulence Capsular Antigen Vi of *Citrobacter freundii* in *Escherichia coli*.. PNAS 85 (12): 4402-4405 .

Pendrak, M. L., Perry, R. D.. (1991) Characterization of a Hemin-storage Locus of *Yersinia pestis.* Biol. Metals 4: 41-47.

Peterson, B. C.,Rownd, R. H.. (1985) Recombination Sites in Plasmid Drug Resistance Gene Amplification. J. Bacteriol. 164 (3): 1359-1361.

Pfeifer, F., Blaseio, V.. (1989) Insertion Elements and Deletion Formation in a Halophilic Archaebacterium. J. Bacteriol. 171 (9): 5135-5140.

Portnoy, D. A., Falkow, S.. (1981) Virulence-Associated Plasmids from *Yersinia enterocolitica* and *Yersinia pestis* . J. Bacteriol. 148 (3): 877-883.

Price, S. B., Cowan, C., Perry, R. D., Straley, S. C.. (1991) The *Yersinia pestis* V Antigen is a Regulatory Protein Necessary for Calcium Ion-Dependent Growth and Maximal Expression of Low Calcium-Ion response Virulence Genes. J. Bacteriol. 173 (8): 2649-2657.

Priefer, V. B., Burkhardt, H. J., Klipp, W., Puhler, A.. (1981) ISR1: An Insertion Element Isolated from the Soil Bacterium *Rhizobium lupini*.. pp.87-92. In: Movable Genetic Elements C.S.H. Symp. Quan. Biol. Vol. 15.

Raleigh, E. A., Kleckner, N.. (1986) Quantitation of Insertion Sequence IS10 Transposase Gene Expression by a Method Generally Applicable to any Rarely Expressed Gene. PNAS 83: 1787-1791.

Rankin, S., Isberg, R. R., Leong, J. M.. (1992) The Integrin-Binding Domain of Invasin is Sufficient to Allow Entry into Mammalian Cells. Infect. Immun. 60 (9): 3909-3912.

Rasmussen, J.L., Odelson, D. A., Macrina, F. L.. (1987) Complete Nucleotide Sequence of Insertion Element IS4351 from *Bacteroides fragilis*. J. Bacteriol. 169 (8): 3573-3580.

Rasmussen, J. L., Odelson, D. A., Macrina, F. L.. (1986) Complete Nucleotide Sequence and Transcription of *ermF*, a Macrolide-Lincosamide-Streptogramin B Resistance Determinant from *Bacteroides fragilis*. J. Bacteriol. 168 (2): 523-533.

Riemmann, C., Moore, R., Little, S. Savioz, A., Willetts, N. S., Haas, D.. (1989) Genetic Structure, Function and Regulation of the Transposable Element IS21. Mol. Gen. Genet. 215: 416-424.

Roberts, D. E., Ascherman, D., Kleckner, N.. (1991) IS10 Promotes Adjacent Deletions at Low Frequency. Genetica 128 (1): 37-44.

Roberts, D., Hoopes, B. C., McClure, W. R., Kleckner, N.. (1985) IS10 Transposition Is Regulated by DNA Adenine Methylation. Cell 43:117-130.

Romantschuk, M., Richter, G. Y., Mukhopadhyay, P., Mills, D.. (1990) Is 801, An Insertion Sequence Element Isolated from *Pseudomonas syringae* pathovar *phaseolicola*. Mol. Micro. 5 (3): 617-622.

Rosqvist, R., Bolin, I., Wolf-Watz, H.. (1988) Inhibition of Phagocytosis in *Yersinia pseudotuberculosis*: A Virulence Plasmid-encoded Ability Involving

the Yop2b Protein. Infect. Immun. 56: 2139-2143.

Rouch, D. A., Skurray, R. A.. (1989) IS257 from *Staphlococcus aureus*: Member of an Insertion Sequence Superfamily Prevalent Among Gram-positive and Gram-negative Bacteria. Gene 76: 195-205.

Roy, R. N., Bigelow, N., Dillon, J.-A. R.. (1988) A Novel Insertion Sequence in the Cryptic Plasmid of *Neisseria gonorrhoeae* May Alter the B Protein at the Translational Level. Plasmid 19 (1): 39-45.

Saisawa, K.. (1913) Veber die Pseudotuberkulose Beim Menschen. Z. Hyg.Infektionskr 73: 353-400.

Sancar, A., Hack, A. M., Rupp, W. D.. (1979) Simple Method for Identification of Plasmid-Coded Proteins. J. Bacteriol. 137 (1): 692-693.

Sanger, F., Nickelsen,S., Coulson, A.R.. (1977) DNA Sequencing with Chain-Terminating Inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74: 5463-5467.

Sawyer, S. A., Dykhuizen, D. E., DuBose, R. F., Green, L., Mutangadura-Mhlanga, t., Wolczyk, D. F., Hartl, D. L.. (1987) Distribution and Abundance of Insertion Sequences Among Natural Isolates of *Escherichis coli*. Genetics 115: 51-63.

Schmitt, M. P., Payne, S. H.. (1989) Genetics and Regulation of Enterobactin Genes in *Shigella flexneri*. J. Bacteriol. 170 (12): 5579-5587.

Schmitt, R., Altenbuchner, J., Wiebauer, K., Arnold, W., Puhler, A., Schoffl, F.. (1981) Basis of Transposition and Gene Amplification by Tn1721 and Related Tetracycline-resistance Transposons. pp. 59-68. In: Movable Genetic Elements, C.S.H. Symp. Quan. Biol. Vol. 15.

Schurter, W., Holloway, B. W.. (1986) Genetic Analysis of Promoters on the Insertion Sequence IS21 of Plasmid R68.45 Plasmid 15: 8-18.

Schwartz, E., Kroeger, M., Rak, B.. (1988) IS50: Distribution, Nucleotide Sequence and Phylogenetic Relationships of a New *Escherichis coli* Insertion Element. Nuc. Acids Res. 16 (14B): 6789-6802.

Sekine, Y., Ohtsubo, E.. (1989) Frameshifting is Required for Production fo the Transposase Encoded by Insertion Sequence 1. PNAS 86: 4609-4613.

Shoemaker, N. B., Getty, c., Gardner, J. F., Salyers, A. A.. (1986) Tn4351 Transposes in *Bacteroides* spp. and Mediates the Integration of Plasmid R751 into the *Bacteroides* Chromosome. J. Bacteriol. 165 (3): 929-936.

Simons, R.W., Kleckner, N.. (1983) Translational Control of IS10 Transposition. Cell 34: 683-691.

Simons, R.W., Hoopes, B. C., McClure, W. R., Kleckner, N.. (1983) Three Promoters Near the Termini of IS10: pIN, p OUT, and pIII. Cell 34: 673-

110

682.

Smith, J. E., Thal, E.. (1965) A Taxonomic Study of the Genus *Pasteurella* Using a Numerical Technique. Acta Path. Micro. Scand. 64: 213-223.

Sodeinde, O. A., Goguen, J. D.. (1989) Nucleotide Sequence of the Plasminogen Activator Gene of *Yersinia pestis*: Relationship to *omp*T of *Escherichia coli* and Gene E of *Salmonella typhimurium*. Infect. Immun 57: 1517-1523.

Sodeinde, O. A., Sample, A. K., Brubaker, R. R., Goguen, J. D.. (1988) Plasminogen Activator/Coagulase Gene of *Yersinia pestis* is Responsible for Degradation of Plasmid-Encoded Outer Membrane Proteins. Infect. Immun. 56: 2749-2752.

Sodeinde, O. A., Subrahmanyam, Y. V. B. K., Stark, K., Quan, T., Bao, Y., Goguen, J. D.. (1992) A Surface Protease and the Invasive Character of Plague. Science 258: 1004-1007.

Solenberg, P. J., Burgett, S. G.. (1989) Method for Selection of Transposable DNA and Characterization of a New Insertion Sequence, IS493, from *Streptomyces lividans*. J. Bacteriol. 171 (9): 4807-4813.

Southern, E. M.. (1975) Detection of Specific Sequences Among DNA Fragments Seperated by Gel Electrophoresis. J. Mol. Biol. 98: 503-517.

Staggs, L.M., Perry, R. D.. (1990) Identification and Cloning of a Fur Regulatory Gene in *Yersinia pestis*. J. Bacteriol. 173: 417-425.

Stanley, J., Jones, C. S., Threlfall, E. J.. (1991) Evolutionary Lines Among *Salmonella enteritidis* Phage Types are Identified by Insertion Sequence IS200 Distribution. FEMS: 82 (1): 83-90.

Straley, S. C., Cibull, M. L.. (1989) Differential Clearance and Host-Pathogen Interactions of YopE- and YopK- *Yersinia pestis* in BALB/c Mice. Infect. Immun. 57: 1200-1210.

Struhl, K.. (1985) A Rapid Method for Creating Recombinant DNA Molecules. BioTechniques 3: 452-453.

Surgalla, M. J.. (1965) *Pastuerella pseudotuberculosis* Information as Background for Understanding Plague. Public Health Reports 80: 825-828.

Swanson, J., Bergstrom, S., Robbins, K., Barrera, O., Corwin, D., Koomey, J. M.. (1985) Gene Conversion Involving the Pilin Structural Gene Correlates with Pilus+ to Pilus- Changes in *Neisseria gonorrhoeae*. Cell 47:267-276.

Tenzen, T., Matsutani, S., Ohtsubo, E.. (1990) Site-Specific Transposition of Insertion Sequence IS630. J. Bacteriol. 172 (7): 3830-3836.

Tertti, R., Skurnik, M., Vartio, T., Kuusela, P.. (1992) Adhesion Protein YadA

111

of *Yersinia* spp. Mediates Binding of Bacteria to Fibronectin. Infect. Immun. 60 (7): 3021-3024.

Timmerman, K.P., Tu, C.-P. D.. (1985) Complete Sequence of IS3.   Nuc. Acids Res.   16 (3): 2127-2139.

Vary, P. H., Andersen, P. R., Green, E., Herman-Taylor, J., McFadden, J. J.. (1990) Use of Highly Specific DNA Probes and Polymerase Chain Reaction to Detect *Mycobacterium paratuberculosis*.   J. Clin. Micro.   28 (5): 933-93.

Wahl, G. M., Stern, M., Stark, G. R.. (1979) Efficient Transfer of Large DNA Fragments from Agarose Gels to Diabenzyloxymethyl-paper and Rapid Hybridization by Using Dextran Sulfate. Proc. Nat. Acad. Sci. U.S.A. 76 (8): 3683-3687.

Wang, L-M., Weber, D. K., Johnson, T., Sakaguchi, A. Y.. (1988) Supercoil Sequencing Using Unpurified Templates Produced by Rapid Boiling. BioTechniques 6 (9): 839-843.

Williams, J. E.. (1983) Warning on a New Potential for Laboratory-Acquired Infections as a Result of the New Nomenclature for the Plague Bacillus. Bull. WHO 61: 545-546.

Williams, R. C., Gewurz, H., Quie, P. G.. (1972) Effects of Fraction 1 From *Yersinia pestis* on Phagocytosis *in vitro*. J. Infect. Dis. 126: 235-241.

Wood, A. G., Konisky, J.. (1985) Activation of Expression of a Cloned Archaebacterial Gene in *Escherichia coli* by IS2, IS5, or Deletions.   Mol. Gen. Genet.   198 (3): 309-314.

Woodward, M. J., Sullivan, G. J.. (1991) Nucleotide Sequence of a Repetitive Element Isolated from *Leptospira interrogans* serovar *hardjo* type *hardjo-bovis*. J. Gen. Micro. 137 (5): 1101-1110.

Yanisch-Perron, C., Vieira, J., Messing, J.. (1985) Improved M13 Phage Cloning Vectors and Host Strains: Nucleotide Sequences of the M13mp18 and pUC19 Vectors.   GENE 33: 103-119.

Yoshioka, Y., Ohtsubo, H., Ohtsubo, E.. (1987) Repressor Gene *finO* in Plasmids R100 and F: Constitutive Transfer of Plasmid F is Caused by Insertion of IS3 into F *finO*.   J. Bacteriol. 169 (2): 619-623.

Young, D. A., Voris, B. P., Maytin, E. V., Colbert, R. A.. (1983) Very High Resolution Two-Dimensional Electrophoresis Separation of Proteins on Giant Gels. Methods Enzymol: 190-214.

Zhukov-Verezhnikov, N.N., Mayski, I. N., Kalinichenko, L. A.. (1955) More in Relation to the Problem of Species and Species Formation in Microbiology. Advances in Contemporary Biology. (Soviet) XXXIX Volume 2: 245-673.