Insertion Sequence Elements in *Ralstonia solanacearum*; Roles in Genomic Heterogeneity

by

Eun-Lee Jeong, M.Sc.

A thesis submitted for the degree of

Doctor of Philosophy

in

The University of Adelaide

Discipline of Genetics

Department of Molecular Biosciences

(Faculty of Science)

September, 2000

Bacterial Wilt on

Potato

Tomato

Eggplant

Table of contents

ABSTRACT

Ralstonia solanacearum is the causative agent of bacterial wilt disease in a taxonomically diverse range of plant hosts. The species exhibits enormous genetic and biochemical complexity which is reflected in its highly heterogeneous genome. This thesis examines how insertion sequence (IS) elements, which are the simplest form of mobile DNA, contribute to genomic heterogeneity in many bacterial species. A characteristic of these mobile elements is that they occur in multiple copies within bacterial genomes and their number and organisation in one isolate may differ significantly from that in another. These features make them attractive probes in the study of genomic heterogeneity in *R. solanacearum*, particularly as IS elements have been shown to be associated with changes in virulence in several plant and animal pathogenic bacteria.

In order to isolate IS elements that show various degrees of heterogeneity between genomes of different strains of *R. solanacearum*, a strategy of differential hybridisation of genomic DNA derived from several strains was utilised. Four novel IS elements, designated IS*Rso2*, IS*Rso3*, IS*Rso4* and IS*61*, were isolated from strains ACH0158 (biovar 2) and ACH0171 (biovar 3). Chapter 3 describes the nucleotide and deduced amino acid sequence analysis of these IS elements and their classification based on unique structural features. IS*Rso3* and IS*Rso4* are IS*5* family members while IS*61* belongs to the IS*3* family and IS*Rso2* apparently belongs to an unknown IS family. The four IS elements displayed particular structural characteristics which were shared with other IS elements identified in distantly related bacterial species suggesting DNA transfer occurs across wide taxonomic distance.

In chapter 4 the transposition of one of the IS elements was investigated as a possible cause of phenotype conversion (PC). Three spontaneous PC mutants were investigated to determine the involvement of IS*Rso4* elements. PC mutants do not produce the main virulence determinant of the pathogen, exopolysaccharide (EPS). One of the three, ACH0158-M81C, was found to be caused by an IS*Rso4* insertion into *phcA*, a global transcriptional regulatory gene controlling coordinated expression of a series of genes controlling virulence and *in planta* growth. The other two PC mutants, ACH0158-M8 and ACH0158-M3, showed sequence lesions in the *phcA* gene but these were not related to IS element movement. The results showed that movement of IS*Rso4* may be involved in virulence gene regulation in *R. solanacearum* and confirmed the importance

of *phcA* gene in phenotype conversion. No revertants that produce large amounts of EPS could be recovered from ACH0158-M81C in either *in vitro* or *in vivo* experiments.

In chapter 5 Southern analysis using the four IS elements as probes was used to reveal major differences in hybridisation profile between different groups of isolates of *R. solanacearum*. IS*Rso4* and IS*Rso2* were predominantly present in biovar 1 and biovar 2 (division II) groups of isolates while IS*61* was confined to biovar 3, 4 and 5 (division I) groups of isolates. The results indicate that the two different groups (division I and division II) of *R. solanacearum* are separate lineages such that specific genomic tools may be developed to partition these genetically diverse strains into particular groups. Transfer of IS elements between groups is also a possibility as Southern analysis revealed significant frequency of occurrence of IS*Rso3* in both the major divisions of the species.

Chapter 6 describes the isolation and characterisation of genes, directly flanking IS*Rso4* and IS*Rso3* insertions, that encode proteins that are potentially associated with virulence and pathogenicity in *R. solanacearum*. A gene encoding the protein porin was isolated and sequence analysis revealed that it is in high G+C content in the coding region and employs unusual codon usage compared with three other genes of *R. solanacearum*. A second IS-flanking gene encoding an enzyme involved in peptidoglycan layer biosynthesis was more conserved than porin in comparisons with homologous in the data bases.

The development of a practical method for detection and quantification of *R. solanacearum* in plant and soil samples is described in chapter 7. Two standard DNA molecules were constructed from clones p092-03B5 and pBs2.10 and tested in PCR experiments for the ability to quantify *R. solanacearum* DNA. It was possible to detect 1000 genome copies of *R. solanacearum* but additional modifications are necessary to improve the sensitivity and practicability of the method. PCR primers derived from p092-03B5 amplify 282 bp fragment that is specifically diagnostic for *R. solanacearum*. The genomic region responsible and its flanking regions were fully characterised in ACH0158. The 282 bp sequence is within the promoter of *lpxC*, a gene encoding UDP-3-*O*-acyl-GlcNAc deacetylase which is an enzyme involved in lipid A biosynthesis.

DECLARATION

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed:

Eun-Lee Jeong

September, 2000

ACKNOWLEDGEMENTS

Above all, I am indebted to my supervisor Dr Jeremy Timmis for his generosity and support. I also thank all the academic staffs in Genetics for their encouragement throughout the course.

I wish to thank Professor Bruce Holloway, Dr Viji Krishnapillai in Monash University and Dr Chris Hayward in the University of Queensland for helping me to obtain bacterial DNA samples. I also thank Dr Jaw-Fen Wang at the Asian Vegetable Research and Development Center (AVRDC) in Taiwan for providing pictures of bacterial wilt disease.

I thank to all the members of the discipline of genetics for their help, in particular Dr Andreas Houben, Dr Robin Lockington and Ms Daniela Marshall. I thank to everyone in the JNT lab, both past and present, specially Dr Tamzin Donald and Ms Dawn Verlin for helping me to solve so many problems. I am also grateful to many PhD colleagues in Genetics for their friendship.

I thank to my parents and my brothers from my heart for their trust and support. Nothing would be possible without their support. Thank you.

CHAPTER 1

Introduction

1.1 Introduction

Ralstonia solanacearum E. F. Smith (Yabuuchi *et al.*, 1995) is a plant pathogenic bacterium causing a bacterial wilt disease in the tropics, subtropics and warm temperate regions (Hayward, 1991). It has a wide host range (Hayward, 1994a), a high level of biochemical complexity (Hayward, 1964, 1991) and enormous genetic variation among isolates (Cook *et al.*, 1989; Cook and Sequeira, 1994; Poussier *et al.*, 1999). Genomic

heterogeneity in many bacterial species is often caused by transposable DNA, which is common among bacterial genomes (Healy *et al.*, 1999). Insertion sequence (IS) elements are the simplest form of transposable DNA in bacteria (Galas and Chandler, 1989) and they have been identified as a major cause of genomic changes such as insertion, deletion and inversion of DNA sequences. In pathogenic bacteria, IS elements modify gene expression and involved in the transposition of genes controlling the interaction of the pathogen with the host organism (Galán and Collmer, 1999; Hacker *et al.*, 1997; Kallastu *et al.*, 1998), suggesting an important role for DNA transposition and rearrangement in the evolution of the host/pathogen interaction.

1.2 Bacterial wilt disease

1.2.1 Symptoms

Infection of tomato plants by *R. solanacearum* initially occurs during secondary root emergence (Kelman and Sequeira, 1965; Schmit, 1978). The bacteria colonise intercellular spaces in the parenchyma layer of the primary root, digest cell walls and spread into the plant (Schmit, 1978; Wallis and Truter, 1978). Tyloses, which are filled with bacteria, develop from the xylem parenchyma cells and partially block the xylem elements (Wallis and Truter, 1978). As tyloses break, the bacteria are released into the vascular system where they spread throughout the conducting vessels of the stem, multiply to more than 10^{10} cells per plant and cause wilt symptoms (Wallis and Truter, 1978).

In the initial stage of development typical symptoms of the wilt disease are enfolding leaves and drooping petioles (van Alfen, 1989). The wilt symptoms progress to become more severe, with the leaves yellowing and the stem of the plant wilting (van Alfen, 1989). Severely infected plants display decayed roots which are characteristically slimy due to *R. solanacearum* oozing out of contaminated roots (Akiew *et al.*, 1993; Husain and Kelman, 1958).

1.2.2 Dispersal mechanisms

Dispersal of the wilt disease is possible *via* insect, seed and seedling as carriers of *R. solanacearum* (Kelman *et al.*, 1994). Transmission of the pathogen by abiotic means (eg. mechanical, aerial and rain-splash) and root-knot nematode is also responsible for disease dispersal (Kelman *et al.*, 1994; Napiere, 1980).

In banana, ginger and potato plants, latently infected planting materials, such as corms (Buddenhagen, 1986), rhizomes (Lum, 1973) and tubers (Ciampi and Sequeira, 1980; Ciampi *et al.*, 1980; Nyangeri *et al.*, 1984; Sunaina *et al.*, 1989) have been a major means of dispersal (Kelman *et al.*, 1994; Hayward, 1991). Distribution of latently infected plants and tubers is responsible for spreading the wilt disease globally (Kelman *et al.*, 1994).

Recently, outbreaks have occurred on potato in some European countries, where *R. solanacearum* was not detected in seed tubers (Caruso *et al.*, 1997). In Europe, seed potato samples are screened by indirect immunofluorescence (IIF) microscopy, which is a method approved by the European Plant Protection Organisation (Anonymous, 1992). Further identification of bacteria in IIF positive potato sample extracts is made after plating on a selective medium. The presence of the pathogen is confirmed by species specific fatty acid analysis (Janse, 1991), IIF staining and an *in vivo* pathogenicity test on tomato seedlings. These techniques are laborious and time-consuming. Wullings *et al.* (1998) developed an alternative method of fluorescent *in situ* hybridisation (FISH) with an *R. solanacearum* specific probe developed from 23S rRNA sequence analysis. However, a considerable amount of time (minimum 2 days) and sophisticated equipment are required for this analysis.

Epidemiological and ecological studies of geographical distribution of *R. solanacearum* in soil and water are hampered by the lack of reliable contamination tests (Wullings *et al.*, 1998). Therefore, there is a urgent need for sensitive and specific detection methods, such as serological or DNA-based methods, to facilitate routine monitoring of potentially infected planting materials (Caruso *et al.*, 1997), soil and irrigation water samples.

1.2.3 Epidemiology and disease management

The epidemiology of bacterial wilt disease is highly complicated and several interacting factors, such as temperature and soil conditions, can influence the disease progression (Hayward, 1991; Kelman *et al.*, 1994). Air temperature and soil conditions are crucial factors affecting *R. solanacearum*-host plant interaction, survival of the pathogen, disease development and consequent disease management (Hayward, 1991). Generally, the disease is more prevalent and resistant plants can become susceptible when the temperature rises to between 30 °C and 35 °C (Hayward, 1991). Additionally, the disease is more wide-spread in some soil conditions, as suggested by He (1990),

however, systematic investigations into the relationship between soil type and incidence of the disease are still required (Hayward, 1991).

R. solanacearum is capable of surviving for long periods in plant debris, in cryptic weed hosts and in some soil types even in the absence of major host plants (Hayward, 1991), resulting in the need for a high level of complexity in disease management. In particular, weed hosts are likely to increase the bacterial population in the soil and to be alternative shelters until susceptible crops are planted. As a result, efforts in disease management, such as crop rotation and fallow, to reduce the number of bacteria in soil, are often ineffective (Hayward, 1991).

Despite these problems, the use of uninfected seed material, crop rotation and fallow, to reduce the population of *R. solanacearum* in soil, have been common strategies for control of bacterial wilt (French, 1986, 1988; Shekhawat *et al.*, 1988; Buddenhagen, 1986). However, to be effective, the clean seed approach requires fast and reliable diagnostic techniques in latently infected material, particularly seed potato tubers (Buddenhagen, 1986). Likewise, effective crop rotation and fallow systems need more practical examination of soil and climatic factors influencing the survival and maintenance of the pathogen (Buddenhagen, 1986). Selection and development of resistant varieties and cultivars have not been sufficiently successful (Abdullah, 1982; Schmiediche, 1986, 1988; Thurston, 1976; Buddenhagen, 1986), but breeding for resistance to bacterial wilt offers the only long-term prospect for controlling this disease.

1.3 *Ralstonia solanacearum*

R. solanacearum [previously named *Pseudomonas solanacearum* and *Burkholderia solanacearum* (Yabuuchi *et al.*, 1992, 1995; Anonymous, 1993)] was first described by E. F. Smith (1896) as a causal agent of the wilt disease of solanaceous plants.
R. solanacearum is a rod-shaped, gram-negative, soil-borne bacterium. The genus *Ralstonia* is a member of the β-subdivision of the *proteobacteria* within eubacteria (Busse and Auling, 1988; Stackebrandt *et al.*, 1988; Kersters *et al.*, 1996; Ludwig *et al.*, 1995) and is classified into rRNA homology group II of the taxon *Pseudomonas*, as determined by rRNA gene sequence analysis (Palleroni *et al.*, 1973). Recently, some *Pseudomonas* species were transferred into *Burkholderia* genus (Yabuuchi *et al.*, 1992) but of these, two [*B. pickettii* (Ralston *et al.*, 1973) and *B. solanacearum* (Smith, 1896)] and *Alcaligenes eutropha* (Davis, 1969) have been transferred into the *Ralstonia* genus (Yabuuchi *et al.*, 1995).

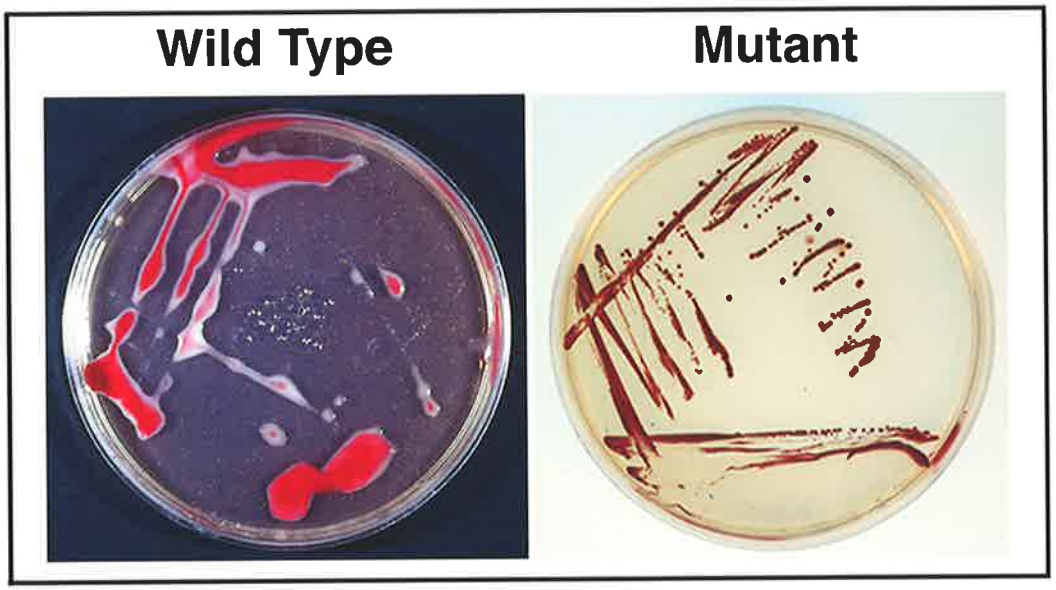


Figure 1.1 *R. solanacearum* on TZC-amended plates

1.3.1 Subspecies classification

R. solanacearum isolates differ in host range, geographical distribution, pathogenicity and physiological properties (Buddenhagen and Kelman, 1964). *R. solanacearum* strains were divided into five different races according to host range (Buddenhagen *et al.*, 1962; He *et al.*, 1983; Pegg and Moffett, 1971), into six biovars based on the utilisation of three disaccharides and three hexose alcohols (Hayward, 1964; Harris, 1972; Hayward *et al.*, 1990; He *et al.*, 1983; French *et al.*, 1995) and into two divisions based on restriction fragment length polymorphism (RFLP) patterns (Cook *et al.*, 1989; Cook and Sequeira, 1994). Fatty acid analysis (Janse, 1991; Stead, 1993) and protein profiling (Dianese and Dristig, 1994) were also utilised to differentiate the genetically diverse isolates of the species.

Kelman (1953) reported that numerous plant species in 35 plant families with most genera in the *Solanaceae* family were hosts of the pathogen. Recently, several hundred plant species in more than 50 plant families were identified as susceptible hosts of the pathogen and the number of new and unexpected crop hosts continues to increase (Kelman *et al.*, 1994; Hayward, 1994a; Buddenhagen, 1986). The host range of race 1 strains is very wide, as many *Solanaceae*, some diploid bananas, numerous other crops and weeds are affected (Buddenhagen *et al.*, 1962). Race 1 strains are genetically variable and widely distributed throughout the lowlands of the tropics and subtropics (French, 1986). However, race 2 strains affect mainly triploid banana and other *Muscaceae* (Buddenhagen *et al.*, 1962; Buddenhagen, 1986; French, 1986). Race 3 strains have a very narrow host range, affecting potato and rarely tomato (Buddenhagen *et al.*, 1962; Buddenhagen, 1986). Race 4 and race 5 strains predominantly affect ginger and mulberry, respectively (Pegg and Moffett, 1971; He *et al.*, 1983; Buddenhagen, 1986).

In initial biovar groupings, *R. solanacearum* isolates were classified into four different biovars on the basis of the ability to oxidise three disaccharides (lactose, maltose and cellobiose) and three hexose alcohols (mannitol, sorbitol and dulcitol) (Hayward, 1964, 1976). Biovar 5, whose strains can utilise all three disaccharides and also mannitol, but not dulcitol and sorbitol, was added later (He *et al.*, 1983). Recently, biovar 2 isolates were differentiated into two subgroups, designated as biovar 2-A (A for Andean) and biovar 2-T (T for tropical) (French *et al.*, 1993, 1995). Biovar 2-A is metabolically less active and originated from the highlands of the Andean region, while biovar 2-T (French *et al.*, 1993, 1995) [or also designated as biovar N2 (new biovar 2) (Hayward

et al., 1991; Hayward, 1994b; Gillings *et al.*, 1993; Seal *et al.*, 1993)] is metabolically more active and was isolated from the tropical lowlands (French *et al.*, 1995).

Race	Host range	Biovar	RFLP Division
1	*Solanaceae*, diploid bananas	1, 3, 4	I, II
2	*Muscaceae*, triploid banana	1	II
3	potato	2	II
4	ginger	3, 4	I
5	mulberry	5	I

Table 1.1
Characteristics of races and their relationship to other subdivisions of *R. solanacearum*.

The biovar scheme shows no clear relationship with the race classification (Table 1.1). Biovar 2, however, is almost equivalent to race 3 as it affects a limited range of hosts, especially potato (Hayward, 1964; Fegan *et al.*, 1998). These strains were thought to have originated in South America, moving to most potato-growing regions of the world *via* infected potato tubers (Buddenhagen, 1986; Cook *et al.*, 1989).

Cook *et al.* (1989) analysed restriction fragment length polymorphism (RFLP) patterns of 62 isolates of *R. solanacearum* which were distinguished into 28 unique groups. More RFLP groups were added after more strains were analysed (Cook and Sequeira, 1994). Similarity coefficients of RFLP patterns separated the isolates tested into two major divisions. Division I contains all members of race 1-biovar 3 and biovar 4 (Cook *et al.*, 1989), confirming that there are no major genomic differences between biovar 3 and biovar 4 strains (Harris, 1972; Poussier *et al.*, 1999). In addition, Cook *et al.* (1989) identified division II that contained all race 1-biovar 1 and race 3-biovar 2 strains. Cook *et al.* (1989) concluded that race 2 and race 3 strains are restricted to narrow ranges of muscaceous and potato hosts, respectively, and have easily distinguishable RFLP patterns, while race 1 strains, having a wide host range, have more complex RFLP patterns.

1.3.2 Genetic heterogeneity of *R. solanacearum*

R. solanacearum is a genetically heterogeneous species (Cook *et al.*, 1989; Hayward, 1991; Opina *et al.*, 1997). A wide range of host plants, broad geographical distribution and diverse physiological characteristics suggest that significant genomic variations are present among *R. solanacearum* strains (Hayward, 1964, 1991). Genomic variation of the species was confirmed by several recent genetic studies (Cook *et al.*, 1989; Cook and Sequeira, 1994; Opina *et al.*, 1997; Poussier *et al.*, 1999). However, mechanisms contributing the genomic variation in *R. solanacearum* are poorly understood.

Cook and Sequeira (1991a) isolated a 2 kb genomic fragment from a race 3 strain (UW23) by subtractive hybridisation. The 2 kb fragment was present in all race 3 strains tested, a result which is commensurate with the similar RFLP profiles between race 3 strains (Cook and Sequeira, 1991a). However, the presence of the 2 kb fragment was observed in a few non-race 3 strains (5 out of 90 strains tested) (Cook and Sequeira, 1991a). To explain this result, horizontal transfer of genomic DNA among distantly related strains within *R. solanacearum* might have occurred (Cook and Sequeira, 1991a, 1994).

Bertolla *et al.* (1997) discovered that *R. solanacearum* can develop competence for transformation *in vitro* during its exponential growth phase after a cultivation under nutrient limitation. More recently, Bertolla *et al.* (1999) showed that the species can take up DNA by transformation during infection within a host plant. In addition, it has been proven that horizontal gene transfer occurs *in planta* between two strains by transformation, not by conjugation (Bertolla *et al.*, 1999). It was therefore hypothesised that some cells of *R. solanacearum* are competent to be transformed naturally *in planta* to take up donor DNA horizontally from other cells which may be releasing genomic DNA after the cells lysed (Bertolla *et al.*, 1999). This is an important discovery, because the ability of *R. solanacearum* to participate in lateral gene transfer can assist the species to incorporate foreign DNAs to increase genomic variation, to improve virulence and pathogenicity and to modify bacterial antigenic molecules. Acquisition of DNAs containing genes to strengthen bacterial genetic variability, adaptability and compatibility would be significantly advantageous to *R. solanacearum* during co-evolution with host plants.

1.3.3 Pathogenicity factors

Pathogenicity determinants in *R. solanacearum* are classified into three groups: *hrp* (hypersensitive response and pathogenicity) genes, *dsp* (disease specific) genes and *avr* (avirulence) genes (Boucher *et al.*, 1987). *hrp* genes are required to induce a hypersensitivity response (HR) in non-host plants and to cause disease in host plants, which is a phenomenon confined to plant pathogenic bacteria. *dsp* genes are not involved in HR induction but can interfere with disease development in host plants. *avr* genes are involved in the control of race/cultivar-specific interactions so that their presence in pathogenic bacteria restricts the potential host range to a limited number of cultivars which do not carry the complementary resistance genes. Similar to *dsp* genes, race-specific *avr* genes are not involved in the induction of the HR in non-host plant species.

The major pathogenicity factor of *R. solanacearum* is encoded by the *hrp* gene cluster (Brito *et al.*, 1999). In *R. solanacearum* strain GMI1000, which is pathogenic in many solanaceous plants including tomato and induces the HR in tobacco, transposon mutagenesis was utilised to clone the *hrp* gene cluster localised on a megaplasmid in the strain (Boucher *et al.*, 1985, 1986, 1987). The gene cluster was shown to span more than 23 kb and to be comprised of a minimum of six transcriptional units (Arlat *et al.*, 1992; van Gijsegem *et al.*, 1995). Expression of the *hrp* genes was shown to be regulated by the nutritional status of *R. solanacearum* (Arlat *et al.*, 1992). Genetic studies on one of the *hrp* genes, *hrpB*, proved that the gene encodes a transcriptional activator for the expression of transcriptional units 2, 3 and 4 of the *hrp* gene cluster during bacterial growth in a minimal medium (Genin *et al.*, 1992). The function of most of the *hrp* genes are still to be determined though it has been speculated that fully functional *hrp* gene products are required in the early process of plant-pathogen interaction, but not in the infection steps. DNA sequencing of the *hrp* gene cluster identified a number of open reading frames that potentially encoded seven proteins all containing putative transmembrane α-helix domains (Arlat *et al.*, 1993). Additionally, the amino acid sequences of five other *hrp* proteins were significantly similar to those of proteins which consist of secretory signals specifically involved in the transmembrane transport of extracellular proteins (primary determinants of pathogenicity in host cells) of *Yersinia pestis*, *Y. enterocolitica* and *Shigella flexneri* (Gough *et al.*, 1992). Therefore, the *hrp* genes in *R. solanacearum* may be involved in the secretion of extracellular bacterial proteins including PopA1, which was shown to be a HR elicitor that was directly active in plant cells (Arlat *et al.*, 1994). More recently,

the transcriptional units 1, 2, 3 and 4 of the GMI1000 *hrp* gene cluster were shown to be induced when bacteria are cultured with plant cell suspensions (Marenda *et al.*, 1998). Novel genes, *prhJ, hrpG* and *prhA*, flanking the *hrp* gene cluster, were identified in a plant signal dependent regulatory pathway required for the induction of the *hrp* genes in *R. solanacearum* (Marenda *et al.*, 1998; Brito *et al.*, 1999).

Boucher *et al.* (1985) isolated by transposon mutagenesis, mutants of the GMI1000 strain, that were able to induce the HR in tobacco (a non-host), but not cause the disease in tomato (a host), and identified four *dsp* loci. Some of these were located adjacent to the *hrp* gene cluster, while others were scattered in the genome (Boucher *et al.*, 1987; Arlat and Boucher, 1991). An approximately 15 kb DNA region was found to contain *dsp* genes that induced virulence in tomato plants and conferred sensitivity to acidic pH (Arlat and Boucher, 1991). The function of *dsp* genes is still unknown although it is possible that the genes encode proteins for regulating the growth of *R. solanacearum* under acidic conditions and for the pathogen to adapt to the plant tissue during the infection process (Arlat and Boucher, 1991).

Carney and Denny (1990) isolated an *avrA* gene from the *R. solanacearum* strain AW1 using an approach similar to that used by Ma *et al.* (1988). A gene library of strain AW1 (showing HR in tobacco) was transferred into the K60 strain (pathogenic in tobacco) and the *avrA* gene was identified by screening transconjugants that were able to induce the HR but unable to induce wilt in tobacco plants. Interestingly, all 27 cultivars of *Nicotiana tabacum* with very different lineages appeared to have resistance genes corresponding to the *avrA* gene in AW1, suggesting that a gene-for-gene specificity was not present at the host cultivar level (Carney and Denny, 1990).

1.3.4 Virulence factors

Production of polysaccharides outside the bacterial cell wall is common in many pathogenic bacteria. These include exopolysaccharides (EPS) and lipopolysaccharides (LPS) (Costerton *et al.*, 1987). The copious, slimy EPS produced by wild-type *R. solanacearum* is a major virulence factor (Denny and Baek, 1991) and EPS negative mutants of *R. solanacearum* are much less virulent (Buddenhagen and Kelman, 1964; Denny *et al.*, 1988; Husain and Kelman, 1958; Staskawicz *et al.*, 1983). EPS is a determinant that enables *R. solanacearum* to colonise the conducting vessels of the host (Costerton *et al.*, 1987), to block pits in the vascular system and to interfere with water

transport in the plant (Husain and Kelman, 1958; van Alfen, 1982, 1989; van Alfen *et al.*, 1983).

Orgambide *et al.* (1991) found that EPS produced by *R. solanacearum* may be separated into four discrete fractions by ion-exchange chromatography. A main part of EPS slime is EPS I, which is a larger than 1,000 kDa. acidic polymer consisting of a repeating trimeric unit of *N*-acetylgalactosamine, 2-*N*-acetlygalactosaminuronic acid and 2-*N*-acetyl-2,4,6-trideoxygalactose (Orgambide *et al.*, 1991; Schell *et al.*, 1993). Denny and Baek (1991) isolated structural gene clusters, named *epsI* and *rgnII* (or *epsII*), for EPS production in *R. solanacearum* strain AW1. Production of EPS I requires the 18 kb *eps* gene cluster, which encodes several membrane-associated soluble proteins involved in EPS I biosynthesis and export (Denny and Baek, 1991; Schell *et al.*, 1993; Huang and Schell, 1995). In addition, the *eps* gene cluster requires portions of an *ops* (outermembrane polysaccharide) gene cluster which encodes a universal sugar precursor for both EPS and LPS (Cook and Sequeira, 1991b; Huang and Schell, 1995; Kao and Sequeira, 1991, 1994).

In addition to EPS and LPS, *R. solanacearum* produces several extracellular proteins (EXPs) that are proven virulence determinants. Generally, extracellular enzymes assist bacteria to invade and to destroy host cell walls, pectic gels around lateral root emergence points and pit membranes separating xylem vessels. Therefore, several EXPs in *R. solanacearum* are plant cell wall degrading enzymes such as endoglucanase (Egl) (Schell, 1987; Huang and Schell, 1990a, 1992; Roberts *et al.*, 1988), polygalacturonases (PglA/B and homologue PehA/B) (Schell *et al.*, 1988, 1994; Huang and Schell, 1990b; Allen *et al.*, 1991; Huang and Allen, 1997) and pectin methylesterase (Pme) (Spök *et al.*, 1991).

Analysis of spontaneously avirulent *R. solanacearum* mutants producing much less endoglucanase suggested that 43 kDa. Egl encoded by the *egl* gene might be necessary for pathogenicity by degrading plant cell wall components such as glucans (Roberts *et al.*, 1988). Polygalacturonases (PglA/B and PehA/B) are hydrolytic enzymes that degrade pectin polymers, the main component of plant cell walls and middle lamellae (Allen *et al.*, 1993; Huang and Allen, 1997). Polygalacturonase-deficient mutants were less virulent but still caused wilt symptoms, suggesting that the enzyme is in highly virulent strains of *R. solanacearum* (Huang and Allen, 1997). Pectin methylesterase (Pme) is an enzyme that removes methyl groups from pectin prior to polygalacturonase

activity. The gene *pme* encoding Pme was characterised in *R. solanacearum* strains K60 (Tans-Kersten *et al.*, 1998) and DSM 59005 (Spök *et al.*, 1991). However, Pme is unlikely to be important for the virulence of *R. solanacearum* because Pme deficient mutants were as virulent as the wild-type K60 (Tans-Kersten *et al.*, 1998). Taken together, the extracellular enzymes produced by *R. solanacearum* clearly have functions in root invasion, acceleration of disease development and increasing aggressiveness (Schell, 1996), though many of the precise mechanisms remain to be determined.

1.3.5 Regulation of virulence and pathogenicity factors

R. solanacearum has evolved complex regulatory mechanisms for virulence and pathogenicity genes including *epsI*. So far, three signal transduction pathways and unique transcriptional regulators have been shown to be involved in the virulence gene regulation of the species (Schell, 1996).

Phenotype conversion (PC) (Brumley and Denny, 1990) is the phenomenon in which wild-type *R. solanacearum* spontaneously loses the ability to produce EPS and extracellular enzymes, resulting in a much less virulent mutant (Fig. 1.1). The *phcA* gene has been shown to be of importance in PC (Brumbley and Denny, 1990). PhcA, a *phcA* gene product, contains a helix-turn-helix (HTH) DNA binding motif at the *N*-terminus and is a member of the large LysR transcriptional regulator family (Brumbley and Denny, 1990, 1993). PhcA has been identified as a global regulator playing a crucial role in the positive control of several genes encoding virulence factors (Schell, 1996) including *epsI, egl, pme* and negatively controls *pglA* and bacterial motility (Schell, 1996). PhcA can bind directly and activate promoters of some virulence genes and the protein also indirectly controls expression of other genes *via* additional regulators such as XpsR (Schell, 1996). Transcription of several PhcA-regulated virulence genes is bacterial cell-density dependent (quorum-sensing regulation) (Kaiser, 1996; Fuqua *et al.*, 1996) such that the amount of *eps* expression is 50-fold higher at 10^9 cells/ml than at 10^7 cells/ml (Clough *et al.*, 1997). In bacterial quorum-sensing gene regulation, the experimental determination of the signalling or inducing molecule of transcriptional regulators is required. Clough *et al.* (1994, 1995) showed that the level of active PhcA was affected by a volatile extracellular factor 3-OH palmitic acid methyl ester (3-OH PAME).

In addition to PhcA, VsrA/VsrD and VsrB/VsrC have been identified as vital two-component signal transduction systems that regulate *eps* and *in planta* growth gene

expression (Schell, 1996). DNA sequencing and maxicell analysis showed that VsrA is a membrane histidine kinase sensor (Schell *et al.*, 1994b) and VsrD is a response regulator for VsrA (Huang *et al.*, 1995). Schell (1996) proposed that the VsrA/VsrD system probably activates expression of the virulence genes, assisting bacterial *in planta* growth, as well as increasing *eps* expression *via* XpsR, an additional transcriptional activator. In the VsrB/VsrC system, VsrC binds and regulates the *eps* promoter, probably in sensing a signal transferred by VsrB (Schell, 1996). XpsR was shown to bind directly at the *eps* promoter and to mediate indirect regulation of the *eps* promoter by VsrA/VsrD and PhcA/PhcSR signal transduction cascades (Schell, 1996). Expression of *pglA* (and a homologue *pehA*), encoding endo-polygalacturonase (endo-Pgl), is regulated by a two-component system PehS/PehR (Allen and Gay, 1995; Allen *et al.*, 1997) and PhcA is probably a negative regulator for PehS/PehR that decreases the level of PglA (Allen *et al.*, 1997; Schell, 1996).

In summary, *R. solanacearum* produces several proteins and polysaccharides that determine pathogenicity and virulence and it has evolved a complex regulatory network containing several two-component signal transduction systems and transcriptional regulators that mediate expression of the genes that control these determinants. *R. solanacearum* shows great genomic heterogeneity among isolates, but genetic mechanisms that generate this variation are not clearly understood.

1.4 Insertion sequence (IS) elements

1.4.1 Bacterial insertion sequence (IS) elements

DNA sequences that are able to transpose to new genomic locations are often present in multiple genomic copies and are called transposable elements. They are distributed in eukaryotes and prokaryotes. In bacteria, there are two classes of transposable element, which are insertion sequence (IS) elements and transposons (Tn). IS elements are the simplest autonomous transposable DNA and can be found as components of transposons. Transposons are often larger and more complicated often carrying additional genes which are not related to the transposition of the elements (Galas and Chandler, 1989).

IS elements were originally discovered as causes of spontaneous mutations in galactose and lactose operons in *E. coli* and bacteriophage lambda (Jordan *et al.*, 1968; Malamy, 1966; Shapiro, 1969; Brachet *et al.*, 1970). They are the discrete genetic entities carrying gene(s) concerned with transpositional functions and they are involved in a

wide variety of genomic rearrangements, such as bacterial chromosomal DNA deletions, inversions and duplications, resulting in variation of host genomic DNA organisation and consequent alteration of gene expression (Iida *et al.*, 1983).
DNA sequence variations of three homologous IS elements, IS*1*, IS*3* and IS*30*, were investigated in wild strains of *E. coli* and closely related species of enteric bacteria to identify the mechanisms that have influenced their evolution (Lawrence *et al.*, 1992). It was concluded that:

- strong sequence homogeneity and a high rate of transposition of IS copies are common, suggesting that turnover and dissemination of the elements among strains was too rapid to allow substantial genetic divergence.
- the rate of horizontal transfer of IS elements among related bacterial species appears to be much less frequent than the rate of intraspecies transfer.
- intragenic recombination events demonstrated in both IS*1* and IS*3* elements may have played a role in the evolution of the elements.
- functionally active open reading frames are often evolutionary-conserved.

1.4.1a Structural characteristics

IS elements are bordered by inverted repeat (IR) and direct repeat (DR) sequences. The length of different IS elements can vary and the elements usually carry one or more open reading frame(s) that encode a transposase which is absolutely required for active transposition (Ohtsubo and Sekine, 1996) (Fig. 1.2).

Inverted repeats (IRs) of IS elements, which are perfectly or nearly perfectly palindromic, are about 10 bp (up to 40 bp) in length and these are recognised directly by a transposase that is specific for a particular IS element (Ohtsubo and Sekine, 1996; Olasz *et al.*, 1997). Mutational analysis was utilised to localise the functional portion of IRs of IS*30*, where the identified region corresponded to the binding site of IS*30* transposase (Olasz *et al.*, 1997). Some IS element family members, (eg. IS*5*), contain highly conserved DNA sequences in their IR termini (Rezsöhazy *et al.*, 1993), suggesting the presence of functionally important sequences that are specifically recognised by the corresponding transposase during the transpositional process.

Upon transposition, IS elements create small (2 bp up to 13 bp) direct repeats (DRs) of the target DNA at the insertion point (Ohtsubo and Sekine, 1996) (Fig. 1.2), due to the insertional process which involves staggered cutting, ligation and repair of short

stretches of single-stranded DNA. More details about target DNA specificity of some IS elements and the genetic mechanisms that generate DRs as a result of transpositional process will be provided in 1.4.1b Target sequences and 1.4.1c Transpositional process.

Some IS elements such as IS*66*, IS*136* and IS*600* contain more than one ORF (Iida *et al.*, 1983; Lawrence *et al.*, 1992) that encode a transposase and one or more additional proteins. However, many others, such as IS*5* and IS*4* family members, usually possess one long, uninterrupted ORF for a transposase (Iida *et al.*, 1983; Rezsöhazy *et al.*, 1993). In addition, some IS elements, such as IS*1*, IS*3* and their relatives, contain two adjacent ORFs, the smaller 5' ORF being in the -1 reading frame relative to the 3' ORF (Ohtsubo and Sekine, 1996; Chandler and Fayet, 1993). Homologous IS elements normally share highly conserved regions within their transposase coding sequences, as proven in several IS elements in IS*5* family (Rezsöhazy *et al.*, 1993) and IS*3* family (Ohtsubo and Sekine, 1996).

1.4.1b Target sequences

Direct repeat (DR) sequences of IS elements are created as a result of the insertional process, in which the transposase proteins firstly nick the target genomic DNA, staggering the cleavage site and the IS element is inserted by ligation and short stretches of single-stranded nucleotide sequences are repaired (Galas and Chandler, 1989; Lewin, 1994). Transposable elements in bacteria have been classified into three groups by their target specificity patterns (Iida *et al.*, 1983):

- the first group includes elements, such as bacteriophage Mu, that insert into random, short target sequences.
- the second group includes IS*1*, IS*2* and IS*50* elements that have preferences for specific regions and can be inserted into several sites within these regions.
- the third group includes IS*4*, IS*5*, IS*30* and IS*10* elements that have consensus sequences of target DNA of various lengths.

Olasz *et al.* (1998) suggested that four different groups may be classified by separating the previous third group into two subgroups. Therefore, a group of IS elements including IS*4*, IS*5* and IS*10* selects genomic targets of short and well-conserved sequences, while the other group of IS elements including IS*30* prefers longer and less conserved genomic target sequences.

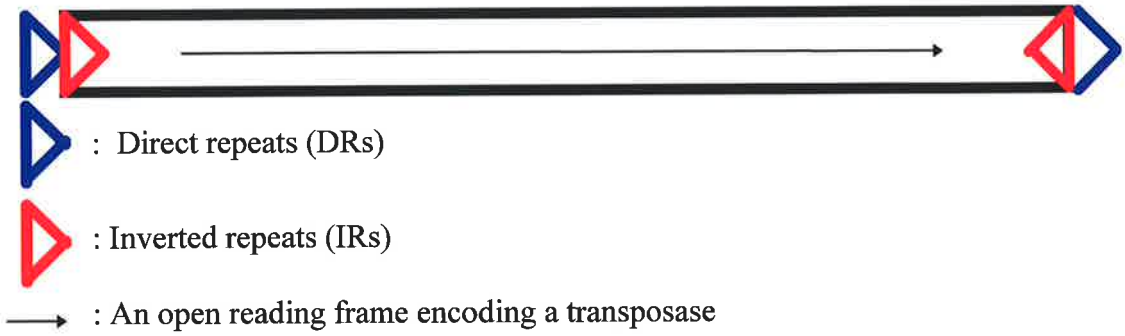


Figure 1.2 Simplified common structural features of IS elements

DR sequence and length variations are not unusual and insertion point preferences may be influenced by genomic locations with different nucleotide sequence background, helix conformation and helix flexibility (Galas and Chandler, 1989). Gerischer *et al.* (1996) observed different DR sequences of copies of IS*1236* inserted into different genes in *Acinebacter calcoaceticus* and concluded that DR sequences and lengths can be affected by different genomic target sites.

1.4.1c Transpositional process

IS elements may use one of two different mechanisms of transposition: non-replicative transposition (cut and paste mode), without increasing the copy number of IS elements in the bacterial genome (Hallet and Sherratt, 1997), and replicative transposition with a consequent increase in the copy number (Sherratt, 1989). The non-replicative mechanism involves excision of both DNA strands at the two ends of the IS element and transmission of the element into the target location (Hallet and Sherratt, 1997). Replicative transposition involves a more complicated process including the fusion of donor and target sequences, duplication of the element and separation of the cointegrate formation by recombination between the two copies of the element (Sherratt, 1989). The replicative process, therefore, requires two different enzymes, a transposase and a resolvase, to complete the transposition (Sherratt, 1989). Both mechanisms, nevertheless, require the common structural features of IS elements:

- a gene encoding a functionally active transposase and additional enzymes.
- specific nucleotide sequences in IRs at the termini to be recognised by active transposase.
- selection of specific target DNA sequences at the insertion point upon transposition.

Excision of the IS element is crucial, especially in the transpositional process, in restoration of the function of a gene that has been inactivated by IS insertion and in relief of polar mutations in genes located downstream from the insertion (Galas and Chandler, 1989). Excision of the IS element may be completed by following intra-strand pairing between IRs, leading to the expansion of the intervening DNA segment and its deletion by replication "slippage" across DR sequences (Galas and Chandler, 1989). The frequency of IS element excision may depend on the length of IR sequences, the genomic location and sequence context and the stability of the expanded structure (Galas and Chandler, 1989). Gerischer *et al.* (1996) suggested that precise target site duplications at the time of insertion are crucial for IS element excision. This suggestion was based on the observation that in two out of nine spontaneous mutants caused by

IS*1236* that failed to give revertants, no obvious target site duplications could be identified by sequencing (Gerischer *et al.*, 1996).

1.4.1d Transposition frequency

It is difficult to estimate the rate of transposition of IS elements though many IS elements are known to transpose actively and/or to cause mutations at high frequency (Galas and Chandler, 1989). Early studies of the transposition rate of IS*1* by monitoring mutations in *lacI* gene in *E. coli* suggested that the element caused approximately 4 % of spontaneous *lacI* mutations (Galas and Chandler, 1989; Calos *et al.*, 1978; Farabaugh *et al.*, 1978).

Transposition frequency varies substantially from element to element and also depends on various factors such as genomic locations of IS elements and their target sequence specificity (Galas and Chandler, 1989; Wall *et al.*, 1999; Farkas *et al.*, 1996). Other factors such as copy numbers of IS elements in genome, different strains of bacterial species and various growth conditions of bacterial culture may also influence transposition rate (Galas and Chandler, 1989; Wall *et al.*, 1999). Little is known of the constrains on accumulation of multiple copies of IS elements within a particular genome.

1.4.1e Host range

IS elements are found in eubacteria and archaea (Galas and Chandler, 1989). In eubacteria, they are found in the chromosomes and in plasmids (Galas and Chandler, 1989). In some cases, nearly identical IS elements are discovered among taxonomically distantly related bacterial species (Iida *et al.*, 1983). Some IS-like elements, especially the IS*630*/Tc*1* family, have also been found in a broad range of eukaryotic host organisms, such as fungi, plants, invertebrates, vertebrate metazoa and ciliated protozoa (Ohtsubo and Sekine, 1996).

1.4.1f Families

More than 500 IS elements have been identified in eubacteria and archaea. Many have been grouped into certain families based on similarities between amino acid sequence motifs of the gene products and also on nucleotide sequences in the IRs (Rezsöhazy *et al.*, 1993).

IS*1*, isolated from *E. coli*, is 768 bp in length and has imperfect IRs of 20 bp at its termini (Lawrence *et al.*, 1992). It generates a target site DNA duplication of 9 bp and length variations of the duplication are occasionally observed (Galas and Chandler, 1989). It carries two ORFs, named insA and insB, and the element utilises -1 translational frameshifting to produce a functionally active transposase (Sekine and Ohtsubo, 1989). Interestingly, InsA is a transposition inhibitor or transcriptional repressor when -1 frameshifting does not occur. IS*1* elements have been found in wild strains of *E. coli*, *Shigella* spp., *Salmonella typhimurium* and other related enteric bacteria (Lawrence *et al.*, 1992). The nucleotide sequence of different IS*1* family members is very similar and almost all contain two ORFs corresponding to insA and insB, suggesting that IS*1* elements might have been laterally transferred among the enteric bacteria (Lawrence *et al.*, 1992; Krishnapillai, 1996; Ohtsubo and Sekine, 1996).

The IS*5*/IS*4* and IS*3* families appear to be the largest families (Mahillon and Chandler, 1998) and these will be discussed in more detail in 1.4.1f (i) IS*5*/IS*4* family and 1.4.1f (ii) IS*3* family.

The IS*630*/Tc*1* family of elements can be found in a wide range of organisms, such as *Shigella sonnei* (IS*630*), *Pseudomonas* spp. (IS*1066*), *Caenorhabditis elegans* (Tc*1*), hypotrichous ciliated protozoa (TBE*1*), *Drosophila* (*pogo*) and rice (Tnr*1*) (Ohtsubo and Sekine, 1996). Most family members appear to insert into the dinucleotide target, 5'NTAN3', (eg. CTAG, TTAA and ATAT) (Tenzen *et al.*, 1990). Interestingly, the family members contain the well-conserved dipeptide, D (aspartic acid)+35+E (glutamic acid), which is homologous to the same motif of retroviral-retrotransposon integrase and IS*3*-like transposase (Ohtsubo and Sekine, 1996).

1.4.1f (i) IS*5*/IS*4* family

IS*4* (Fiandt *et al.*, 1972) is 1426 bp with 18 bp IRs and generates DRs of 11-13 bp (Klaer *et al.*, 1981). In contrast to IS*1* and IS*3* elements, IS*4* has a long ORF encoding a transposase. IS*5* (Engler and van Bree, 1981; Schoner and Kahn, 1981; Kröger and Hobom, 1982) is 1195 bp in length with 16 bp IRs. IS*5* has a long ORF and generates 4 bp DRs. The consensus target site sequence of the IS*5* element is C-T/A-A-G/A (Rezsöhazy *et al.*, 1993). More than 40 IS elements from diverse bacterial species belong to the IS*4* family (Rezsöhazy *et al.*, 1993). Members of the IS*4* family contain highly conserved regions within the transposase, called C1 region, corresponding to the *C*-terminal with R (arginine)+3+E (glutamic acid) motif, and N3 region at the *N*-

terminal with D (aspartic acid)+1+G (glycine)/A (alanine)+Y (tyrosine)/F (phenylalanine) (Rezsōhazy *et al.*, 1993).

IS*4* family elements have been divided into two groups: IS*4* group and IS*5* group, on the basis of sequence similarities and relative locations of the two conserved regions (N3 and C1 regions), and characteristic IR sequences (Rezsöhazy *et al.*, 1993). Within the IS*4* group, the N3 and C1 domains are separated by nearly 110 amino acids, while IS*5* group elements have more closely located N3 and C1 domains (Rezsöhazy *et al.*, 1993). Members of the IS*5*/IS*4* family contain similar configurations within their IR sequences, especially the AT-rich motif close to the inner limit of the IRs (Rezsöhazy *et al.*, 1993). However, IS*4* group elements contain CAT as a conserved external trinucleotide, while IS*5* group elements contain GGC or GAG (Rezsöhazy *et al.*, 1993; Ohtsubo and Sekine, 1996). IS*5* group elements, therefore, are significantly different from IS*4* group elements. Interestingly, the N3 region containing D (aspartic acid)+1+G (glycine)/A (alanine)+Y (tyrosine)/F (phenylalanine) of the IS*5*/IS*4* family elements displays strong similarities to the integrase domain, D (aspartic acid)+35+E (glutamic acid), shared by the IS*3* family elements and retroelements (Rezsöhazy *et al.*, 1993). Biochemical and structural studies suggested that the conserved residues, such as aspartic acid (D), may play a crucial role in catalysis of integrases (Grindley and Leschziner, 1995; Rice *et al.*, 1996).

IS*903* (Oka *et al.*, 1981), which is a member of the IS*4* family was utilised to define functional regions of its transposase by deletional and point mutational analysis (Tavakoli *et al.*, 1997). Deletion of 46 amino acid residues at the *N*-terminal region of IS*903* transposase resulted in reducing the affinity of DNA binding of the protein to the IR sequences. More precisely, IS*903* transposase 22-139 was the minimal region of the protein required for DNA binding, inducing a DNA bend and forming complexes to complete protein-protein interactions (Tavakoli *et al.*, 1997). In addition, Tavakoli *et al.* (1997) utilised site-directed mutagenesis to determine the roles of the conserved N3 and C1 domains (Rezsöhazy *et al.*, 1993). Substitutions of six highly conserved [D (aspartic acid)121, D (aspartic acid)193, Y (tyrosine)196, Y (tyrosine)252, R (arginine)255 and E (glutamic acid)259] amino acid residues with alanine (A) residue resulted in severely defective transposition, but mutant proteins were still able to bind to IRs, confirming the significance of the *N*-terminal region for DNA-binding activity (Tavakoli *et al.*, 1997). In summary, phenotypes of mutations in the *C*-terminal region and the high sequence

homology of residues related to the catalytic activity of other transposase and integrase enzymes suggested that conserved residues are responsible for catalytic steps, while the *N*-terminal region is primarily more responsible for DNA-binding ability (Tavakoli *et al.*, 1997).

1.4.1f (ii) IS*3* family

IS*3* (Fiandt *et al.*, 1972) was isolated from *E. coli* and is present in diverse species belonging to enterobacteriaceae (Lawrence *et al.*, 1992). It is 1258 bp with imperfect 39 bp IRs and generates target duplications of 3 bp upon insertion (Lawrence *et al.*, 1992). It codes for two ORFs, orfA and orfB, and the IS*3* transposase is produced by -1 translational frameshifting at a specific sequence in the overlapping region between the two ORFs. For active IS*3* transposase, OrfA-OrfB transframe protein must be produced and Sekine *et al.* (1994) found that AAAAG is the site promoting the frameshifting. A large number of IS elements found in gram-negative and gram-positive bacteria belongs to the IS*3* family. Nearly all members of this group contain two ORFs. The amino acid sequences of orfA show relatively little similarity, while those of orfB are significantly similar (Ohtsubo and Sekine, 1996).

The most conserved domain of IS*3* family elements is D (aspartic acid)+35+E (glutamic acid) at the *C*-terminus of the transposase (Ohtsubo and Sekine, 1996). The two amino acid residues have been shown to be important in site-specific cleavage and integration, and are involved in binding the metal cofactors required for the catalytic activities (Ohtsubo and Sekine, 1996). These two conserved residues in IS*3* element transposases are apparently shared by retrovirus and retrotransposon integrase (Ohtsubo and Sekine, 1996). In addition, the sequence 5'-TG……CA-3' in the IRs, which is also found at the proviral DNA ends of retroviruses and is a part of the *cis*-acting region required for integration (Khan *et al.*, 1991), is conserved in many IS*3* family elements (Ohtsubo and Sekine, 1996).

Translational frameshifting has been demonstrated in two members of the IS*3* family, IS*911* and IS*150*, (Polard *et al.*, 1991; Vögele *et al.*, 1991) and a high frequency of frameshifting was observed in both cases. Although the IS*3* family displays similar configurations to IS*1* family elements, several significant differences are present (Polard *et al.*, 1992). For example, in contrast to InsA in IS*1* element OrfA of IS*911* does not serve as an inhibitor of transposition and does not affect the overall activity of transposase (Chandler and Fayet, 1993). Rather, IS*911* transposase activity appears to

be dependent upon the level of a fusion protein OrfAB. On the other hand, OrfB of IS*150* transposase seems to affect the element-mediated cointegrate formation (Vögele *et al.*, 1991) and OrfB in IS*911* transposase may play a similar role. Therefore, the transposition process of IS*911* in general requires transcription of the two ORFs, translation initiation of OrfA, rephasing of ribosomes at the frameshifting site and completion of OrfAB to yield an active, mature transposase (Chandler and Fayet, 1993).

1.4.2 Genetic effects of IS elements

IS elements are natural mutagens, causing insertional mutations at high frequency in bacterial genomes (Galas and Chandler, 1989). They are capable of translocating by excision and insertion of themselves sometimes together with various lengths of duplicated target genomic DNA. In addition, IS elements mediate deletion and inversion of flanking genomic DNA, and homologous recombinations (Ohtsubo and Sekine, 1996; Galas and Chandler, 1989), resulting in host genomic DNA reorganisation and change.

IS elements have been shown to be transfer genes vital for pathogenicity and virulence in pathogenic bacteria horizontally among strains causing increased adaptability and compatibility (Bukhalid *et al.*, 1998; Healy *et al.*, 1999; Tyler *et al.*, 1996). In addition, IS elements form composite transposons, which are often present in plasmids, and mobilise non self-transposable genes, such as antibiotic resistant genes (Galas and Chandler, 1989; Byrne, 1989; Quintiliani and Courvalin, 1996; Tauch *et al.*, 1998).

1.4.2a IS elements affecting genetic heterogeneity

Dissemination of bacterial genes for virulence determinants followed by intragenic recombination has been shown to increase genetic plasticity by creating mosaic structures of the genes (Boyd and Hartl, 1999). Acquisition of exogenous virulence genes and non-random intragenic recombination are particularly important for bacterial antigenic molecule variation, as the variation assists pathogenic bacteria to avoid defence-immune systems of host organisms. Additionally, genetic variation is a significant factor for pathogenic bacteria to invade more diverse host organisms, resulting in expanded host ranges and ecological niches.

Pseudomonas aeruginosa expresses A- and B-band lipopolysaccharide together and the LPS complex is a main virulence factor for the species (Burrows *et al.*, 1996). Based on DNA sequencing and molecular analysis of a gene cluster for B-band O-antigen

biosynthesis, a group of genes was identified specifically in a few serotypes, while other genes were commonly present in several different serotypes tested. IS*1209* was identified at the linkage between the serotype-specific and non-specific genes (Burrows *et al.*, 1996), suggesting its involvement in transfer of the serotype-specific genes.

Salmonella enterica is a pathogenic bacterium causing gastroenteritis and typhoid fever. In this species a 17 kb DNA region of the large pathogenicity island SPI-3 contains a *mgtCB* operon encoding the macrophage survival protein, MgtC, and the Mg^{2+} transporter protein, MgtB (Blanc-Potard *et al.*, 1999). A group of genes of the SPI-3 island was present in all *Salmonella* species tested, but a central part of the island containing four genes was present only in a few subgroups. The four genes were flanked by incomplete IS elements flanked a genomic region where very low G+C content (37 %) and atypical codon usage were detected, implying the possibility of the region having been transferred from a very different genome. Further multiple lateral gene transfers might have occurred at the central region of the island consisting of a gene (*misL*) with high G+C content (53 %) and the three genes (*rmbA*, *fidL* and *marT*) with low G+C content (38 %, 45 % and 47 %, respectively). The results, therefore, suggested that the SPI-3 island has a compound structure and the four genes in the middle of the island might have originated elsewhere. Sequence analysis of encoded proteins of two of the genes, *misL* and *marT*, revealed that they are similar to AIDA-I adhesin (a virulence factor) in *E. coli* and ToxR regulatory protein in *V. cholerae* (Blanc-Potard *et al.*, 1999).

Comparative nucleotide sequence analysis of a *fim* gene cluster encoding type 1 pilin in both *E. coli* and *Salmonella* spp. indicated that it may contain another example of lateral acquisition of genes for bacterial antigens (Boyd and Hartl, 1999). Antigenic pilin adhesins are virulence determinants that enable some *Salmonella* isolates to colonise eukaryotic cells. However, the *fim* gene clusters in *E. coli* and *Salmonella* spp. were located at different genomic regions and had distinctly different organisation of the component genes. Phylogenetic and molecular genetic analyses on a few genes of the *fim* gene cluster in *Salmonella* confirmed the divergence of nucleotide sequences, G+C content and codon usages from those of *E. coli* (Boyd and Hartl, 1999). The analyses also suggested that lateral gene transfers and intragenic recombination might have occurred more than once in some genes of the cluster (Boyd and Hartl, 1999).

In summary, IS elements in several bacterial species are actively involved in transferring between strains and species genes that encode virulence and pathogenicity determinants. Intragenic and other homologous recombination followed by the lateral gene transfer, greatly increase genomic variation among bacterial populations, which has been capitalised by natural selection.

1.4.2b IS elements affecting gene expression

The insertion of IS elements may cause flanking gene activation by providing a promoter region upstream of a cryptic gene. Gene inactivation may also occur by insertional mutation (Ohtsubo and Sekine, 1996; Galas and Chandler, 1989). IS*406* and IS*407* elements were identified in *Burkholderia* (formally, *Pseudomonas*) *cepacia* strain 249 when inserted into a promoter/operator region of Tn*951* to activate the expression of the *lac* genes of Tn*951* (Wood *et al.*, 1991). Nucleotide sequence analysis of the IS*407* element determined an outwardly directed σ^{70}-like promoter in IR sequences of the element (Wood *et al.*, 1991), such that the promoter probably triggered the flanking gene expression. Similarly, insertional activation of a promoterless *pheBA* gene for phenol degradation was mediated by an outward-directed promoter from IS*1411* element in *Pseudomonas putida* (Kallastu *et al.*, 1998).

IS element insertions often cause inactivation of gene(s) (Ohtsubo and Sekine, 1996; Galas and Chandler, 1989). In *Porphyromonas gingivalis*, an agent in adult and early-onset periodontal disease, insertions of IS*195* was found to inactivate *prtP*, a gene encoding a protease that is a major virulence factor of the species, resulting in a reduction of virulence (Lewis and Macrina, 1998). IS*51* and IS*52* elements in *Pseudomonas syringae* subsp. *savastanoi* were identified as they were inserted into and inactivated an *iaaM* gene that is required for indolacetic acid production and virulence of the species (Yamada *et al.*, 1986). Gerischer *et al.* (1996) showed that IS*1236* in *Acinebacter calcoaceticus* caused spontaneous mutants which lost the ability to dissimilate *p*-hydroxybenzoate by preferential insertion into a *pobR* gene, a transcriptional activator of a structural gene for *p*-hydroxybenzoate hydroxylase.

IS elements are actively involved in bacterial phase variations and are known to cause reversible mutations by insertion and precise excision in diverse bacterial species. A spontaneous mutation in *N. meningitidis* involved an encapsulated wild-type bacterium becoming a capsule-negative mutant (Hammerschmidt *et al.*, 1996). Genetic analysis of the capsule-negative mutants revealed a unique mechanism for regulation of the capsule

expression which was based on reversible inactivation of an essential sialic acid biosynthesis gene, *siaA*, by insertion and excision of IS*1301* (Hammerschmidt *et al.*, 1996). The frequency of the reversion event was estimated to be approximately 1 in 40,000 and 13 such revertants had precise excisions of IS*1301* together with the DRs, leaving exactly the same sequences in the revertants as in wild-type.

Ziebuhr *et al.* (1999) reported an example of a phase variation in *Staphylococcus epidermidis*. Biofilm formation in *S. epidermidis* is mediated by the *ica* operon and on activation of the operon, a polysaccharide intercellular adhesin (PIA) is synthesised to initiate the production of biofilms (Ziebuhr *et al.*, 1999). PIA-negative phase variants were analysed to determine molecular mechanisms and approximately 30 % of the phase variants was found to be due to the inactivation of either *icaA* or *icaC* genes by IS*256* insertion (Ziebuhr *et al.*, 1999). Ziebuhr *et al.* (1999) also characterised a single revertant that restored the biofilm-forming ability, and found a complete excision of IS*256* including the duplicated 8 bp DRs. However, the reversion rate in *S. epidermidis* was low, less than 1 in 10^8 per cell per generation.

In summary, IS elements are involved in gene inactivation by insertional mutations and in gene activation by providing a promoter or creating a hybrid promoter. IS elements also play crucial roles in mediating molecular mechanisms of bacterial phase alternation.

1.5 Summary and aims of the project

R. solanacearum is a plant pathogenic bacterium causing a commercially devastating bacterial wilt disease and is a genetically heterogeneous species. Its genetic heterogeneity is reflected in extremely wide host range, various biochemical properties and genomic variations of isolates. However, the genetic mechanisms responsible for the genetic heterogeneity in *R. solanacearum* are poorly understood.

A significant proportion of genomic variation of many bacterial species may be caused by insertion sequence (IS) elements, which are present in the majority of bacterial genomes. IS elements cause various genomic DNA rearrangements and mediate gene regulation by insertion and excision, resulting in genetic heterogeneity of bacterial genome. They appear to have had a significant involvement in modifications of bacterial virulence and pathogenicity, because of the strength of the selective forces operating in this arena.

The specific aims of this project were to isolate and characterise IS elements from *R. solanacearum*. Characterisation of IS elements was mainly based on nucleotide and deduced amino acid sequence analyses. Genomic distribution and organisation within the species of the IS elements isolated in this project were determined by Southern hybridisations of a significant number of isolates that represented each different group of the species. In addition, the IS elements isolated in this study were investigated to clarify their possible involvement in the well-known phase variation of *R. solanacearum*, a phenotype conversion. Genomic regions flanking IS elements were also isolated and analysed by DNA sequencing and Southern hybridisation to reveal any possible association of IS elements with virulence and pathogenicity genes in *R. solanacearum*.

A quantitative-PCR method was developed with the purpose of establishing a potentially useful method of detecting and quantifying *R. solanacearum* in latently infected potato, based on species- and strain-specificity of two clones previously isolated.

CHAPTER 2

Materials and Methods

2.1 Materials

2.1.1 Bacterial strains

Bacterial strains of *Ralstonia solanacearum* (*R. solanacearum*) and *Escherichia coli* (*E. coli*) used in this study are summarised in Table 2.1.

Table 2.1 Bacterial strains

R. solanacearum	biovar	host	date	country	source[a]
AW1	1	tomato	1988	USA	Denny
CIP0007	1	banana	1964	Honduras	CIP[a]
CIP0009	1	plantain	1967	Costa Rica	CIP
CIP0066	1	potato	1972	Costa Rica	CIP
CIP0070	1	banana	1965	Colombia	CIP
CIP0120	1	potato	1979	Peru	CIP
CIP0218	1	potato	1978	Brazil	CIP
CIP0301	1	potato	1988	Peru	CIP
CIP0418	1	peanut	1991	Indonesia	CIP
CIP0419	1	peanut	1991	Indonesia	CIP
ACH0127R	2	potato	1966	Australia	ACH
ACH0158	2	potato	1966	Australia	ACH
ACH0732	atypical 2[b]	tomato	1979	Australia	ACH
ACH1061	2	potato	1990	Australia	ACH
ACH1068S	2	potato	1990	Australia	ACH
CIP0117	2	potato	1968	Nigeria	CIP

R. solanacearum

	biovar	host	date	country	source[a]
CIP0310	2	potato	1905	Columbia	CIP
CIP0358	N2[c]	potato	NA[d]	Cameroon	CIP
CIP0359	2	potato	1989	Cameroon	CIP
CIP0402	2	potato	1991	Indonesia	CIP
CIP0403	2	potato	1991	Indonesia	CIP
CIP0405	2	potato	1991	Indonesia	CIP
PD1427	2	tomato	NA	Australia	Janse
PD1435	2	potato	NA	Australia	Janse
South Africa 28	2	potato	1985	South Africa	NA
South Africa 45	2	potato	1986	South Africa	NA
South Africa 46	2	potato	1986	South Africa	NA
South Africa 49	2	potato	1988	South Africa	NA
South Africa 65	2	potato	1989	South Africa	NA
ACH0171	3	eggplant	1967	Australia	ACH
ACH0190	3	*Xanthium pungens*	1967	Australia	ACH
ACH0369	3	tomato	1969	Australia	ACH
ACH0574	3	tomato	1975	Australia	ACH
ACH0671	3	pepper	1977	Australia	ACH
ACH1024	3	*Strelitzia reginae*	1987	Australia	ACH
ACH1064	3	heliconia	1989	Australia	ACH
ACH1070	3	ornamental ginger	1990	Australia	ACH
CIP0065	3	chili	1972	Costa Rica	CIP
CIP0284	3	common olive	1988	China	CIP
PD1437	3	tomato	NA	Australia	Janse
South Africa 8	3	potato	1984	South Africa	NA
South Africa 12	3	tobacco	1985	South Africa	NA
South Africa 53	3	potato	1988	South Africa	NA
South Africa 93	3	potato	1991	South Africa	NA
South Africa 112	3	potato	1994	South Africa	NA
ACH0262	4	ginger	1967	Australia	ACH
ACH0319	4	*Solanum nigrum*	1968	Australia	ACH

R. solanacearum

	biovar	host	date	country	source[a]
ACH1066	4	heliconia	1990	Australia	ACH
CIP0121	4	potato	NA	Sri Lanka	CIP
CIP0264	4	potato	1985	Indonesia	CIP
CIP0283	4	common olive	1988	China	CIP
PD1431	4	potato	NA	The Philippine	Janse
CIP0277	5	mulberry	1988	China	CIP

E. coli

DH-5αF'	Laboratory stock	Host for cloning into pGEM®-T and general cloning	*recA*1, *endA*1, *gyrA*96, *thi*-1, *hsdR*17 (r_K^-, m_K^+), *supE*44, *relA*1, *deoR*, Δ(*lacZYA -argF*)U169 [F', F80d*lac* ZΔM15]; Hanahan, 1983
SOLR™	Stratagene, Heidelburg, Germany	Host for excision of phagemid from λZAP®II vector	e14$^-$(*mcr*A), Δ(*mcrCB -hsdSMR-mrr*)171, *sbcC*, *recB*, *recJ*, *muC*::Tn*5*(kan^r), *uvrC*, *lac*, *gyrA96*, *relA1*, *thi-1*, *endA1*, l^R, [F' *proAB*, *lacI*^{q}ZΔ*M15*]Su$^-$; Hay and Short, 1992
XL1-Blue	Stratagene, Heidelburg, Germany	Host for propagation of λZAP®II	*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac*, [F' *proAB*, *lacI*^{q}Z Δ*M15*, Tn*10*(tet^r)]; Bullock *et al.*, 1987

[a] CIP; International Potato Centre, Lima, Peru
ACH; Chris Hayward

[b] reference; Taghavi *et al.*, 1996

[c] N2; biovar N2 strain

[d] NA; data not available

2.1.2 Chemical reagents

General laboratory chemicals were of analytical research grade and were purchased from a variety of manufacturers including Aldrich Chemical Co, BDH Chemicals Australia Pty Ltd, Oxoid Pty Ltd and Sigma Chemical Company. Specialist reagents used in this study are listed below.

- Bacteriological agar No. 1 — Oxoid, Unipath, Hampsire, UK
- Bacteriological peptone — Oxoid, Unipath, Hampsire, UK
- Casamino acids — Difco Laboratories, Detroit, Michigan
- Yeast extract — Oxoid, Unipath, Hampsire, UK
- Tryptone — Oxoid, Unipath, Hampsire, UK
- Caesium chloride — BDH Chemicals, Australia
- Dextran sulphate — Progen Industries Ltd, Australia
- Polyethylene glycol 8000 — Sigma Chemical Co, St Louis, MO
- Urea — BDH Chemicals, Australia
- Agarose; low gelling temperature-type VII — Sigma Chemical Co, St Louis, MO
- Agarose; molecular biology certified agarose-ultra pure DNA grade agarose — Bio-Rad Laboratories, Richmond, CA

2.1.3 Stains

- Bromophenol blue — Sigma Chemical Co, St Louis, MO
- Ethidium bromide — Sigma Chemical Co, St Louis, MO

2.1.4 Enzymes

Restriction endonucleases (and 10 x reaction buffers) were obtained from Promega, New England Biolabs (NEB) and Boehringer Mannheim. Other enzymes used in this study were obtained from the following manufacturers.

- Calf intestinal phosphatase (CIP) — Boehringer Mannheim, Germany
- Lysozyme — Sigma Chemical Co, St Louis, MO
- RNaseA — Sigma Chemical Co, St Louis, MO
- T_4 DNA ligase — New England Biolabs, Beverley, MA
- Klenow fragment — Boehringer Mannheim, Germany
- *Taq* DNA polymerase — Bresatec, Adelaide, Australia

2.1.5 Radioactive isotopes

- [α-^{32}P]dATP (3000 Ci/mmole; ethanolic) — Bresatec, Adelaide, Australia

2.1.6 Nucleic acids

A number of plasmid clones and oligonucleotide primers to plasmid vector sequences were used in this study and are presented in Table 2.2. Oligonucleotide primers designed from insert sequence are summarised where each insert sequence is discussed. Molecular size markers and other nucleic acids were purchased from the following suppliers.

- DNA ladder, 100 bp — Bresatec, Adelaide, Australia
- Lambda DNA restricted with *Hin*dIII — Progen Industries Ltd, Australia
- Lambda ladder PFG marker — New England Biolabs, Beverley, MA
- Low range PFG marker — New England Biolabs, Beverley, MA
- Mixed random decamers — Bresatec, Adelaide, Australia
- Salmon sperm DNA — Sigma Chemical Co, St Louis, MO

Table 2.2 Plasmid clones and oligonucleotides

plasmid	relevant characteristics	location	reference
pSC15	321 bp *Sau*3AI fragment from ACH0158 cloned into pBluescript® SK±, Ap^r	Chapter 3	This study
pSV102	2.7 kb *Eco*RI fragment containing IS*Rso3* and IS*Rso4* from ACH0158 cloned into pBluescript® SK±, Ap^r	Chapter 3	This study
pSV1022	588 bp *Sal*I-*Eco*RI fragment from pSV102 cloned into pBluescript® SK±, Ap^r	Chapter 3	This study
pSV102-260	258 bp *Sal*I-*Pst*I fragment from pSV102 cloned into pBluescript® SK±, Ap^r	Chapter 3	This study
pSV102-825	825 bp *Sal*I-*Pst*I fragment from pSV102 cloned into pBluescript® SK±, Ap^r	Chapter 3	This study
pSV102-1040	1046 bp *Sal*I fragment from pSV102 cloned into pBluescript® SK±, Ap^r	Chapter 3	This study
pISBE	5.2 kb *Eco*RI fragment containing IS*Rso2* from ACH0158 cloned into pBluescript® SK-, Ap^r	Chapter 3	This study
pISBE.1v	3.9 kb *Eco*RI-*Pst*I fragment from pISBE cloned into pBluescript® SK±, Ap^r	Chapter 3	This study

plasmid	relevant characteristics	location	reference
pIS61	1.02 kb *Sau*3AI fragment containing IS*61* from ACH0171 cloned into pBluescript® SK±, Ap^r	Chapter 3	This study
pWT	1.3 kb *Eco*RI fragment containing *phcA* from ACH0158 cloned into pGEM®-T easy	Chapter 5	This study
pM81C	1.64 kb *Eco*RI fragment containing *phcA81C* from ACH0158-M81C cloned into pGEM®-T easy	Chapter 5	This study
pM3	1.15 kb *Eco*RI fragment containing *phcA3* from ACH0158-M3 cloned into pGEM®-T easy	Chapter 5	This study
pM8	1.3 kb *Eco*RI fragment containing *phcA8* from ACH0158-M8 cloned into pGEM®-T easy	Chapter 5	This study
pIS102JI	6.6 kb *Eco*RI fragment containing IS*Rso4* from ACH0158 cloned into pBluescript® SK-, Ap^r	Chapter 6	This study
pIS102JI.2v	2.2 kb *Eco*RI-*Sac*I fragment from pIS102JI cloned into pBluescript® SK±, Ap^r	Chapter 6	This study
pIS102JI.4v	4.4 kb *Eco*RI-*Sac*I fragment from pIS102JI cloned into pBluescript® SK±, Ap^r	Chapter 6	This study
pSC15K	4.0 kb *Eco*RI fragment containing IS*Rso3* from ACH0158 cloned into pBluescript® SK-, Ap^r	Chapter 6	This study
pSC15K.K2	2.0 kb *Sac*II fragment from pSC15K cloned into pBluescript® SK±, Ap^r	Chapter 6	This study
p092-03B5	282 bp fragment from ACH092 cloned into pCR™II, Ap^r, Km^r	Chapter 7	This study
p092-03B5EL	282 bp fragment from p092-03B5 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study
p092-03B5EL+170	455 bp fragment (282 bp + 173 bp) from p092-03B5 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study
pBs2.10	570 bp fragment from ACH0158 cloned into pUC19, Ap^r	Chapter 7	This study
pBs2.10+170	743 bp fragment (570 bp + 173 bp) from pBs2.10 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study

plasmid	relevant characteristics	location	reference
pIIIF2	9.3 kb *Sac*I fragment containing *lpxC* from ACH0158 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study
pIIIF2.1v	4.3 kb *Sac*I-*Eco*RI fragment from pIIIF2 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study
pIIIF2.2v	2.74 kb *Eco*RI fragment from pIIIF2 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study
pIIIF2.3v	1.4 kb *Sac*I-*Eco*RI fragment from pIIIF2 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study
pIIIF2.4v	0.88 kb *Eco*RI fragment from pIIIF2 cloned into pBluescript® SK±, Ap^r	Chapter 7	This study

oligonucleotides	sequence (5'→3')
T3	ATT AAC CCT CAC TAA AG
T7	AAT ACG ACT CAC TAT AG
SP6	TAT TTA GGT GAC ACT ATA G
RSP	AAC AGC TAT GAC CAT G
-20	GTA AAA CGA CGG CCA GT

2.1.7 Antibiotics and indicators

- Ampicillin — Boehringer Mannheim, Germany
- IPTG (isopropyl β-D-thiogalactopyranoside) — Boehringer Mannheim, Germany
- Kanamycin — Sigma Chemical Co, St Louis, MO
- Tetracycline — Sigma Chemical Co, St Louis, MO
- 2,3,5 Triphenyl-tetrazolium — BDH Chemicals, Australia
- X-gal (5-bromo-4-chloro-3-indolyl β-D-galactopyranoside) — Progen Industries Ltd, Australia

2.1.8 Cloning vectors

Plasmid and bacteriophage vectors used in this study are presented in Table 2.3.

Table 2.3 Plasmid and bacteriophage vectors

vector	source	use	reference
pBluescript® SK±	Stratagene, Heidelburg, Germany	general cloning and subcloning	Short *et al.*, 1988
pBluescript® II SK±	Stratagene, Heidelburg, Germany	general cloning and subcloning	Short *et al.*, 1988
pBluescript® SK-	Stratagene, Heidelburg, Germany (excision from λZAP®II)	general cloning and subcloning	Short *et al.*, 1988
pGEM®-T Easy	Promega, Madison, WI	cloning of PCR products	Robles and Doers, 1994
λZAP®II	λZAP®II digested with *Eco*RI and; dephosphorylated Stratagene, Heidelburg Germany	construction of genomic DNA library of strain ACH0158	Short *et al.*, 1988
ExAssist™helper filamentous phage	Stratagene, Heidelburg Germany	rescue of phagemids from λZAP®II vector	Hay and Short, 1992

2.1.9 Kits and miscellaneous materials

Kits and other materials used in this study were obtained from the following manufacturers.

- Wizard genomic DNA purification kit — Promega, Madison, WI
- Predigested λZAP®II/*Eco*RI/CIAP kit — Stratagene, La Jolla, CA
- QIAquick PCR purification kit — QIAGEN GmbH, Gemany
- QIAquick gel extraction kit — QIAGEN GmbH, Gemany
- Bio-Gel P60 (50-100 mesh) — Bio-Rad Laboratories, Richmond, CA
- Bio-Gel P60 (100-200 mesh) — Bio-Rad Laboratories, Richmond, CA
- Exassist™/SOLR™ System — Stratagene, La Jolla, CA
- Glass beads, 150-212 mm — Sigma Chemical Co, St Louis, MO

- Hybond™-N+ nylon membrane — Amersham, UK
- Packagene® lambda DNA packaging system — Promega, Madison, WI
- pGEM®-T Easy vector system — Promega, Madison, WI
- Sepharose CL-6B — Pharmacia, Uppsala, Sweden
- ABI Prism™ Big-Dye-terminator cycle sequencing reactions — Applied Biosystems-Perkin Elmer
- X-ray film — Fuji RX Medical X-ray Film, Fuji Photo Co, Ltd

2.2 Solutions, buffers and media

2.2.1 Solutions and buffers

All solutions and buffers were prepared using millipore water and autoclaved. All restriction endonucleases and most other enzymes were provided with reaction buffers. Solutions and all other buffers routinely used in this study were as follows.

Southern transfer buffers

- Denaturing solution — 1.5 M NaCl, 0.5 M NaOH
- Neutralising solution — 1.5 M NaCl, 1 M Tris-HCl (pH 7.2), 1 mM EDTA
- Alkaline blot solution — 0.4 N NaOH, 0.6 N NaCl
- 1 x SSC — 0.15 M NaCl, 0.015 M Na_3-citrate, adjusted to pH 7.2 with NaOH

Southern hybridisation buffers

- 1 x Denhardt's — 0.02 % (w/v) Ficoll, 0.02 % (w/v) PVP, 0.02 % (w/v) gelatin
- Hybridisation buffer — 5 x SSPE, 5 x Denhardt's, 0.5 % (w/v) SDS, 100 mg/ml denatured, sonicated salmon sperm DNA
- 1 x SSPE — 0.18 M NaCl, 10 mM NaH_2PO_4, 1 mM EDTA (pH 7.2)
- Stop buffer — 10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), 0.2 % (w/v) SDS

Minipreparations of plasmid DNA by alkaline lysis method

- Solution I — 50 mM Glucose, 25 mM Tris.Cl (pH 8.0), 10 mM EDTA (pH 8.0)

- Solution II — 0.2 N NaOH, 1 % (w/v) SDS
- Solution III — 3 M potassium, 5 M acetate

Electrophoresis Buffers

- 1 x TAE (Tris-acetate) — 0.04 M Tris-acetate, 0.001 M EDTA, pH 8.0
- 1 x TBE (Tris-borate) — 0.09 M Tris-borate, 0.002 M EDTA, pH 8.0
- Gel-loading buffer type IV — 0.25 % (w/v) bromophenol blue, 40 % (w/v) sucrose

Agarose plugs for use in pulsed-field gel electrophoresis (PFGE)

- 1 x SE — 75 mM NaCl, 25 mM EDTA (pH 8.0)
- 1 x ES + Proteinase K — 0.44 M EDTA (pH 9.5), 1 % (w/v) Sarcosyl, 0.5 % (w/v) Proteinase K

Commonly used buffers

- 1 x TE — 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)
- 1 x STE — 10 mM Tris-HCl (pH 8.0), 0.1 M NaCl, 1 mM EDTA
- 1 x STET — 0.1 M NaCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), 5 % (w/v) Triton X-100
- Lysing solution — 10 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA, 4 mg/ml lysozyme
- Phenol/Chloroform — 50 % (v/v) phenol, 48 % (v/v) chloroform, 2 % (v/v) isoamylalcohol, buffered with an equal volume of Tris-HCl (pH 8.0), 0.2 % (v/v) β-mercaptoethanol
- SM buffer — 100 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris-HCl (pH 7.5), 2 % (w/v) gelatin

2.2.2 Media

All liquid and solid media were prepared using millipore filtered RO water and sterilised by autoclaving. The compositions of various media were as follows.

- LB broth (Luria-Bertani broth) — 1 % (w/v) tryptone, 1 % (w/v) NaCl, 0.5 % (w/v) yeast extract, adjusted to pH 7.2 with 1 M NaOH
- LB agar — LB broth with the addition of 1 % (w/v) bacteriological agar No. 1
- Modified Kelman's TZC agar — 0.1 % (w/v) casamino acids, 0.5 % (w/v) glucose, 1 % (w/v) bacteriological peptone, 1.5 % (w/v) bacteriological agar No. 1, 0.1 % (w/v) yeast extract. After autoclaving the basal medium,

	0.005 % (w/v) TZC (2,3,5 triphenyl-tetrazolium)
• BG broth (Boucher *et al.*, 1985)	0.1 % (w/v) yeast extract, 1 % (w/v) bacteriological peptone, 0.1 % (w/v) casamino acids, 0.5 % (w/v) glucose
• SOC medium	2 % (w/v) bacto-tryptone, 0.5 % (w/v) bacto-yeast extract, 0.05 % (w/v) NaCl, 2.5 mM KCl, adjust pH 7.0 with 5 N NaOH. After autoclaving the basal medium, 10 mM $MgCl_2$, 20 mM glucose
• NZY agar	1 % (w/v) NZ amine (casein hydrolysate), 0.5 % (w/v) yeast extract, 0.5 % (w/v) NaCl, 1 % (w/v) bacteriological agar No. 1
• NZY top agar/agarose	1 % (w/v) NZ amine (casein hydrolysate), 0.5 % (w/v) yeast extract, 0.5 % (w/v) NaCl, 10mM $MgSO_4$, 0.7 % (w/v) bacteriological agar No. 1 or agarose
• TB agar	1 % (w/v) tryptone, 0.5 % (w/v) NaCl, 1 % (w/v) bacteriological agar No. 1
• TB top agar/agarose	1 % (w/v) tryptone, 0.5 % (w/v) NaCl, 10 mM $MgSO_4$, 0.7 % (w/v) bacteriological agar No. 1 or agarose

2.3. Methods

2.3.1 General techniques

2.3.1a Restriction enzyme digests

All reactions involving restriction endonucleases were performed under conditions as recommended by manufacturers. In general, 5 units of enzyme were added per 1 μg of DNA to a maximum enzyme volume of 10 % of the total reaction volume and the reaction was allowed to proceed for at least 2 hours. Samples of bacterial genomic DNA were restricted with 10 units of enzyme per 1 μg of DNA overnight. If required, RNaseA was added to restrictions to a final concentration of 0.5 mg/ml.

2.3.1b Electrophoretic separation of DNA on agarose gels

Restriction enzyme digested bacterial genomic DNA was size fractionated by electrophoresis on 0.6 % (w/v) agarose gels in submarine gel tanks with 1 x TAE as the running buffer. *Spe*I-digested bacterial genomic DNA was resolved using pulsed-field

gel electrophoresis as described in section 2.3.11. PCR-amplified DNA samples were analysed by electrophoresis on 1.2-1.5 % (w/v) agarose gels. Samples were loaded with 1 x gel-loading buffer type IV and electrophoresed until the bromophenol blue marker dye had migrated a sufficient distance to allow adequate separation of the DNA fragments. Following electrophoresis, the DNA was stained in 5 μg/ml of ethidium bromide and viewed under UV light. Sizes of DNA fragments were estimated according to Duggleby *et al.* (1981).

2.3.1c Recovery of DNA from agarose gels

DNA was eluted from agarose gels using either the freeze-squeeze method adapted from Thuring *et al.* (1975) or the QIAquick gel extraction kit. For the freeze-squeeze method, the frozen gel slice was centrifuged at 12,000 x g at room temperature for 10 mins. The supernatant was collected and DNA recovered by precipitation with ethanol. Recovery of DNA using the QIAquick gel extraction kit was performed under conditions recommended by the manufacturer.

2.3.1d Concentrating nucleic acids with ethanol

DNA was precipitated by the addition of one tenth of the volume of 3 M Na-acetate (pH 5.2) followed by 2-2.5 volumes of cold redistilled ethanol. For concentrating short pieces (< 100 nucleotides) and small amounts of nucleic acid, $MgCl_2$ was added to a final concentration of 0.01 M. The DNA was left to precipitate at -20 °C for a period of 15 mins-overnight depending on its molecular weight. The DNA was recovered by centrifugation at 12,000 x g for at least 15 mins at 4 °C. The pellet was washed with cold 70 % (v/v) ethanol, dried under vacuum and resuspended in a minimal volume of 1 x TE or water.

2.3.1e Purification of nucleic acids

In general, DNA was purified from proteins by the addition of an equal volume of phenol/chloroform and vigorous mixing in a microfuge tube. The emulsion was centrifuged at 12,000 x g for 2 mins and the aqueous layer transferred to a new tube before the extraction was repeated. Contaminating phenol was removed by the addition of an equal volume of chloroform, mixing and centrifugation as above, and the aqueous phase was transferred to a new tube and DNA precipitation.

Plasmid DNA utilised in the experiments described in chapter 7 was purified by equilibrium centrifugation in CsCl-ethidium bromide gradients followed by ethanol

precipitation after removal of the ethidium bromide with water-saturated butanol (Sambrook *et al.*, 1989).

Plasmid DNA for automated sequencing was purified by precipitation of DNA with 13 % (w/v) polyethylene glycol (PEG 8000) containing 1.5 M NaCl on ice for 30 mins following RNaseA treatment. DNA was then pelleted by centrifugation at 12,000 x g at 4 °C for 30 mins and washed with 95 % (v/v) ethanol before being dried and resuspended in 1 x TE. The DNA solution was extracted sequentially with phenol, phenol/chloroform and chloroform and precipitated with one half the volumes of 7.5 M NaAc and 2 volumes of absolute ethanol. The pellet was washed with 70 % (v/v) ethanol, dried and resuspended in 1 x TE.

2.3.1f Quantitation of DNA

The concentration of purified DNA in solution was determined by spectrophotometric measurement of the amount of ultraviolet irradiation absorbed 260 nm. For low amounts of DNA or DNA samples contaminated with other substances, samples were coelectrophoresed with λDNA restricted with *Hin*dIII and comparative estimates made after staining with ethidium bromide. The quantity of DNA in the sample was estimated by comparing the fluorescent amount of the sample with that of a series of standard markers.

2.3.2 Preparation of DNA Samples

2.3.2a Isolation of genomic DNA from *R. solanacearum*

R. solanacearum was grown on modified Kelman's TZC agar at 30 °C for 1-several days and well-separated colonies showing typical characteristics of wild-type or phenotype conversion (PC)-type for the experiments described in chapter 4 were selected. *R. solanacearum* was cultured in 10 ml of BG broth (Boucher *et al.*, 1985) at 30 °C overnight with gentle shaking. 1.5 ml aliquots were then used to isolate and purify genomic DNA samples using a wizard genomic DNA purification kit (Promega).

2.3.2b Preparations of plasmid DNA; lysis of *E. coli* by alkali

Mini- and large-scale preparation of plasmid DNA was carried out by alkaline lysis (Sambrook *et al.*, 1989). Purifications were completed by either phenol/chloroform extraction, equilibrium centrifugation in CsCl-ethidium bromide gradients or precipitation with polyethylene glycol (see section 2.3.1e).

2.3.3 Cloning strategies

2.3.3a Preparation of insert DNA and plasmid vector

DNA fragments for ligation were generated in two ways. DNA was restricted to completion with the appropriate restriction enzyme and then recovered from agarose gels by the freeze-squeeze method or by utilising the QIAquick gel extraction kit (see section 2.3.1c). Alternatively, insert DNA was prepared by direct purification of the DNA sample that was PCR-amplified by *Taq* DNA polymerase using the QIAquick PCR purification kit.

In order to prevent vector DNA self-ligation vector DNA was digested with the appropriate restriction enzyme. Linearised plasmid DNA was treated with calf intestinal phosphatase (CIP) to remove 5'-phosphate groups from vector DNAs, essentially as described by Sambrook *et al.* (1989),. For dephosphorylation of linearized plasmid DNA, a 50 µl reaction containing the digested plasmid DNA, the appropriate amount of CIP, and 1 x CIP buffer was incubated at 37 °C for 30 mins, after which 10 µl 10 x STE, 5 µl 10 % (w/v) SDS and 40 µl dH_2O were added and the mixture was incubated at 68 °C for 15 mins to inactivate the CIP. The TA cloning vector pGEM®-T Easy (Table 2.3) was utilised in ligation of insert DNA after PCR-amplification using *Taq* DNA polymerase.

2.3.3b Ligations

Ligations of 50-100 ng of vector DNA with a 3-5 molar excess of insert DNA were catalysed by 1 Weiss unit of T_4 DNA ligase in 10 µl reactions. Ligation reactions were placed in a 7 L insulated waterbath at 16 °C and left overnight in a 4 °C coldroom, except for ligation reactions involving pGEM®-T Easy vectors which were carried out according to manufacturer's instructions.

2.3.3c Preparation and transformation of competent *E. coli*

Preparation of *E. coli* competent cells for transformation was by one of two methods. Calcium chloride method was used to prepare the cells for heat shock transformation at 42 °C and the simpler method was used to prepare the cells for electrotransformation as per Sambrook *et al.* (1989). The transformed competent cells were allowed to recover and express the antibiotic resistance marker in either LB or SOC medium (see section 2.2.2) at 37 °C for 30 mins. Colonies containing recombinant plasmids were selected by

plating the transformation mixture onto LB agar medium containing appropriate antibiotic with an overlayer of LB agar plus appropriate antibiotic, 200 μg/ml X-gal (5-bromo-4-chloro-3-indolyl β-D-galactopyranoside) and 160 μg/ml IPTG (isopropyl β-D-thiogalactopyranoside) to test bacteria for α-complementation.

2.3.4 Phage manipulation

Bacterial cells for plating of phage were prepared as per Stratagene instructions for the predigested λZAP®II/*Eco*RI/CIAP cloning kit and diluted to OD_{600} = 0.5 with 10 mM $MgSO_4$ before use.

Phage and host cells were mixed and incubated at 37 °C for 15-20 mins prior to plating on the appropriate medium. Phage for the screening of libraries was plated on 150 mm plates using 600 μl of plating cells and 8 ml top agarose, melted and cooled to 48 °C. Phage for secondary and tertiary screening was plated on 85 mm plates using 200 μl of plating cells and 3 ml of molten top agar or agarose at 48 °C. Plates were incubated at 37 °C for between 8 and 16 hours, not to allow the plaques to get larger than 1-2 mm.

Positive plaques identified by autoradiography were identified on the agar plates. Plaques were plugged into 0.5 ml SM buffer and 20 μl chloroform using the large end of a sterile pasteur pipette and the phage were allowed to elute at room temperature for 2 hours or at 4 °C for overnight. Phage were plated at low density on 85 mm plates for secondary screening and the procedure repeated for a tertiary round of purification. Following tertiary screening, recombinant pBluescript®SK(-) was rescued from λZAP®II using the ExAssist™/SOLR™ system, following manufacturer's instructions.

2.3.5 Transfer of DNA to membranes

2.3.5a Southern blotting

Following electrophoretic separation of DNA on agarose gels, DNA was transferred and alkali-fixed onto a Hybond™-N+ nylon membrane, following the instructions of the manufacturer.

2.3.5b Colony hybridisation

Colony hybridisation was achieved using the method described by Grunstein and Hogness (1975). Recombinant bacterial colonies were patch plated onto an LB agar plate supplemented with appropriate antibiotic and replica plated onto a Hybond™-N+

membrane which overlayed another such plate. Following growth of bacterial colonies at 37 °C overnight, the nylon filter was placed colony side up on a pad of three pieces of Whatman® 3MM paper soaked in lysing solution (see section 2.2.1) for 20 mins, with blotting onto paper towel every 5 mins. The filter was transferred to a fresh pad of 3MM paper soaked in denaturing solution plus 1 % (w/v) SDS for a total of 20 mins with blotting every 7 mins. This procedure was repeated with 3MM paper soaked in neutralising solution before the filter was partially air-dried. Bacterial debris was removed from the filter using absorbent cotton wool and washing in 2 x SSC. DNA was alkali-fixed onto the filters.

2.3.6 Radiolabelling of DNA probes

DNA fragments requiring radioactive labelling were purified by size fractionation on an agarose gel and freeze-squeezing as previously described (see section 2.3.1c). Labelling of 50-200 ng of double-stranded DNA was performed by primer extension of random oligonucleotides (Feinberg and Vogelstein, 1983) using the klenow fragment and [α-^{32}P]dATP.

Unincorporated nucleotides were removed from oligolabeled probes by spun column chromatography through Bio-Gel P60 resin columns made from 100 µl G50-100 beads and 400 µl G100-200 beads. Columns were packed and equilibrated with 100 µl stop buffer (see section 2.2.1) by centrifugation at 1,000 x g for 2 mins. Labelled DNA fragments were eluted by adding the labelling reaction plus the stop buffer (to a total volume of 100 µl) to the top of the column and recentrifuging at 1,000 x g for 2 mins.

2.3.7 Prehybridisation, hybridisation and washing

Filters were prehybridised in hybridisation buffer (see section 2.2.1) at 65 °C for a minimum of 2 hours. The volume of prehybridisation solution was approximately 0.1 ml/cm^2 of filter.

Following the addition of 300 µg sonicated salmon sperm DNA to a total volume of 100 µl of the radiolabelled probe, the probe was denatured by the addition of 35 µl 1 M NaOH and incubated at room temperature for 10 mins. The solution was neutralised by adding 35 µl 1 M HCl. Filters were hybridised with 1 x 10^6 cpm of radiolabelled probe per ml of hybridisation solution for a minimum of 16 hours at 65 °C.

Filters were washed sequentially in 2 x SSC, 0.1 % (w/v) SDS at 65 °C for 30 mins, 1 x SSC, 0.1 % (w/v) SDS at 65 °C for 30 mins and 0.5 x SSC, 0.1 % (w/v) SDS at 65 °C for a further 30 mins. Southern blots subject to a high stringency wash were incubated in 0.1 x SSC, 0.1 % (w/v) SDS at 65 °C for 2 or more hours.

Filters were covered in plastic wrap to prevent drying and exposed to X-ray film backed by an intensifying screen (DuPont Hi-Plus) at -70 °C for an appropriate period of time. Southern blots and plaque lifts were stripped of probe by incubating for 10 mins in 0.1 x SSC, 1 % (w/v) SDS at 90 °C. Stripped filters were rinsed in 2 x SSC and stored in plastic wrap at 4 °C for rehybridisation.

2.3.8 DNA oligonucleotides

Oligonucleotide primers were synthesised commercially (Bresatec, Adelaide). In general, all oligonucleotides were adjusted to a concentration of 10 mM for PCR and automated sequencing. Annealing temperatures were calculated according to the formula $T_A = [2(A+T)+4(G+C)]-5$ °C.

Oligonucleotide primers were designed to previously sequenced regions. They were generally synthesised as 20-mers and contained a GC composition of at least 50 %. Efficient priming at high temperatures was maximised by the inclusion of two G or C residues at the 3' end of the primers. In addition, primer pairs were checked using computer programs for the possibility of primer-dimer formation and intramolecular bonding.

2.3.9 Sequencing of double-stranded DNA templates

Plasmid DNA to be used for automated sequencing reactions was prepared using the alkaline lysis method. Purification of the DNA was performed by precipitation of DNA with 13 % (w/v) polyethylene glycol (PEG 8000) containing 1.5 M NaCl, followed by sequential extraction with phenol, phenol/chloroform and chloroform and precipitation with 2 volumes of 100 % ethanol.

Automated sequencing reactions were performed using ABI Prism™ Big-Dye-terminator cycle sequencing reactions (Perkin Elmer), according to the manufacturer's instructions. All automated sequencing (ABI Prism Model 377) was done at the Institute of Medical and Veterinary Science (IMVS), Adelaide, South Australia. Compilation and comparison of multiple nucleotide sequences was performed using the

facilities provided by the Australian National Genomic Information Service (ANGIS). Multiple sequence alignments were done using the Eclustalw program (Thompson *et al.*, 1994). Blast was used to search the nucleotide databases (Altschul *et al.*, 1990; Altschul, 1991). For nucleotide and amino acid sequence alignment, the program Gap (Needleman and Wunsch, 1970) and Bestfit (Smith and Waterman, 1980) in the GCG package (Genetics Computer Group, Madison, Wis.) were used.

2.3.10 Polymerase Chain Reaction (PCR)

Reaction mixtures (25 µl) for PCR contained 0.4 mM each of dATP, dCTP, dGTP and dTTP, 1 x PCR reaction buffer (Bresatec, Adelaide), 3 mM $MgCl_2$, 100 ng of each primer, 1.0 U of *Taq* DNA polymerase (Bresatec, Adelaide) and 100 ng of template DNA. DNA was amplified in a PTC-200 DNA engine (MJ Research, Inc.) generally programmed as follows: initial template denaturation was at 94 °C for 2 mins followed by 25 PCR cycles of denaturation at 94 °C for 1 min, primer annealing at 55 °C for 1 min and extension at 72 °C for 1 min. If the product was to be cloned, an additional cycle was carried out with an extension time of 7 mins. Annealing temperature varied, depending on the primer pairs, between 45 and 55 °C.

2.3.11 Pulsed-field gel electrophoresis (PFGE)

Pulsed-field gel electrophoresis was performed as described by Ratnaningsih *et al.* (1990).

2.3.11a Manipulations of bacterial DNA embedded in agarose plugs

Bacterial cells were collected by centrifugation at 5,000 rpm at room temperature for 15 minutes, then washed twice with SE buffer (see section 2.2.1) and diluted to OD_{570} = 0.2 (approximately equal to 3 x 10^9 cells/ml). Agarose plugs were prepared by molding a mixture of 1 volume of bacterial cell solution maintained at 55 °C and 1 volume of 2 % (w/v) low gelling agarose maintained at 65 °C. Agarose plugs containing bacterial cells were incubated on ice for 30 minutes, transferred into freshly prepared ES + Proteinase K buffer (see section 2.2.1) and incubated at 56 °C overnight.

After the overnight incubation, agarose plugs were washed with 1 x TE buffer at least three times at room temperature prior to *Spe*I restriction digestion of the embedded bacterial genomic DNA.

The agarose plug was pre-treated with 1 x reaction buffer of *Spe*I enzyme at room temperature for 15 minutes and the bacterial DNA sample was restriction-digested with 9 U *Spe*I with 0.1 mg/ml acetylated BSA at 4 °C overnight followed by an additional incubation at 37 °C overnight. Upon completion of the restriction digest the agarose plug was washed with 1 x TE and left at 4 °C for 1 hour for hardening.

2.3.11b Pulsed-field gel electrophoresis

PFGE apparatus utilised in this study was the contour clamped homogeneous electric field electrophoresis (CHEF) system (Pharmacia). The CHEF system generates uniform electric fields across the gel by a hexagonal array of electrodes of predetermined ("clamped") electric potential. Approximately 1 hour prior to the gel electrophoresis, 2.5 l of 0.2 (or 0.1) x TBE buffer was circulated in the electrophoresis tank to equilibrate the refrigeration. Agarose gels in appropriate TBE buffer were prepared using ultra pure DNA grade agarose (Bio-Rad Laboratories).

Agarose plugs containing *Spe*I-digested bacterial genomic DNA were inserted into the wells which were then filled and sealed with molten 1.2 % (w/v) agarose at 65 °C. Two different PFG DNA size markers were used, depending on the size range of *Spe*I-digested bacterial DNA fragments to be resolved. Lambda ladder PFG marker was used for the size range 50-1000 kb fragments and low range PFG markers (New England Biolabs) were used for the size range 0.1-200 kb fragments.

Four different gel electrophoresis conditions were specifically designed, depending on the size range of *Spe*I fragments of *R. solanacearum* genomic DNA to be separated. To resolve large *Spe*I fragments (> 97.0 kb), 1.2 % (w/v) agarose gel, 0.2 x TBE buffer, 40 seconds-pulse time and 44 hours gel running were applied. To resolve medium sized *Spe*I fragments (> 23. 1 kb), 1.2 % (w/v) agarose gel, 0.1 x TBE buffer, 15 seconds-pulse time and 30 hours gel running were applied. To resolve small *Spe*I fragments (between 145.5 kb and 4.26 kb), 1.0 % (w/v) agarose gel, 0.1 x TBE buffer, 6 seconds-pulse time and 20 hours gel running were applied. To resolve *Spe*I fragments smaller than 48.5 kb, 1.3 % (w/v) agarose gel, 0.1 x TBE buffer, 2 seconds-pulse time and 22 hours gel running were applied. Following electrophoresis, DNA was stained in 5 mg/ml of ethidium bromide and viewed under UV light.

CHAPTER 3

Identifications and characterisations of insertion sequence (IS) elements from *Ralstonia solanacearum*

3.1 Introduction

In contrast to many other bacterial species, the presence of insertion sequence (IS) elements in *Ralstonia solanacearum* had not been reported prior to the isolation of IS elements described in this chapter. Previously a 321 bp genomic DNA clone pSC15 from the *R. solanacearum* (~~Knipe, 1994~~) was shown to contain sequences that were highly repetitive within the strain ACH0158. pSC15 was a subclone of an initial clone p20 which was present in the genome of ACH0158 but absent in strain ACH0171 (Fig. 3.1B). These experiments were designed to isolate repetitious DNA fragments that showed various degrees of differentiation between the genomes of different strains of the bacterium. Sequencing of pSC15 revealed high homology to IS elements from many bacterial species. This chapter describes the full characterisation of pSC15 as part of an IS element and the subsequent molecular cloning of three additional IS elements from strains ACH0158 and ACH0171 that appear to be associated with genomic diversity of strains within the species. The sequences of the identified IS elements were analysed and characterised to classify them and determine their unique structural features.

3.2 Cloning of IS elements from strains ACH0158 and ACH0171

Cloning of IS elements from ACH0158 was initiated after a comparison of the Southern hybridisation patterns of *Eco*RI-digested genomic DNA of two biovar 2 strains, ACH0158 and ACH1068S probed with the SC15. Both ACH0158 and ACH1068S are biovar 2 strains which were isolated from potato plants in 1966 and 1990, respectively.

In general, *Eco*RI is a useful restriction endonuclease to digest the genomic DNA of *R. solanacearum* [approximately 67 % GC content (Palleroni, 1975)] as it produces relatively few fragments. In addition the IS elements isolated contained no internal *Eco*RI recognition sequences, such that digestion with this enzyme was useful to estimate the minimum copy number of each IS element in a particular bacterial genome. Numerous (more than 40) *Eco*RI fragments of ACH0158 and ACH1068S were hybridised with the SC15 probe and a very similar spectrum of hybridising fragments was present in both strains (Fig. 3.2B). However, a 2.7 kb *Eco*RI fragment was present in ACH0158, but absent in ACH1068S. This fragment was cloned for further characterisation because it was likely to contain a full length IS element and because it differentiated two closely related bacterial isolates (Fig. 3.2B).

To clone the 2.7 kb *Eco*RI fragment, *Eco*RI-digested genomic DNA of ACH0158 was size fractionated by electrophoresis on 0.6 % (w/v) agarose gels and a gel block containing 2 to 4 kb fragments was excised from a gel. The DNA was eluted and ligated into *Eco*RI-digested pBluescript II®SK$^{(-)}$. The clone pSV102 was selected after colony hybridisation with the SC15 probe and an *Eco*RI digest confirmed the existence of the expected 2.7 kb insert (result not shown).

The simple method of isolating IS elements by isolating repetitious genomic clones was extended. Genomic DNA of ACH0158 strain was restricted with *Sau*3AI and ligated into *Bam*HI-digested pBluescript II®SK$^{(-)}$. DNA from recombinant colonies was duplicated on nylon membranes and the replicas hybridised with [α-^{32}P]dATP labelled total genomic DNA from ACH0158 and ACH0171. Colonies showing strong hybridisation to the ACH0158 probe and little or no hybridisation to the ACH0171 probe, were selected for further characterisation because they were likely to carry plasmids containing repetitious DNA that was specific to ACH0158. Several such clones were sequenced and pISB was identified as a clone carrying a partial IS element. As predicted, the insert DNA of pISB hybridised to several *Eco*RI fragments from ACH0158 but none from ACH0171 (Fig. 3.1C). The insert DNA of pISB was utilised to screen a genomic library of ACH0158 to yield a larger clone (pISBE), containing a 5.2 kb *Eco*RI fragment, carrying a full length IS element, designated as IS*Rso2*.

The reciprocal of the previously described differential hybridisation strategy was also applied to clone repetitious DNA sequences that were unique to ACH0171. In this case,

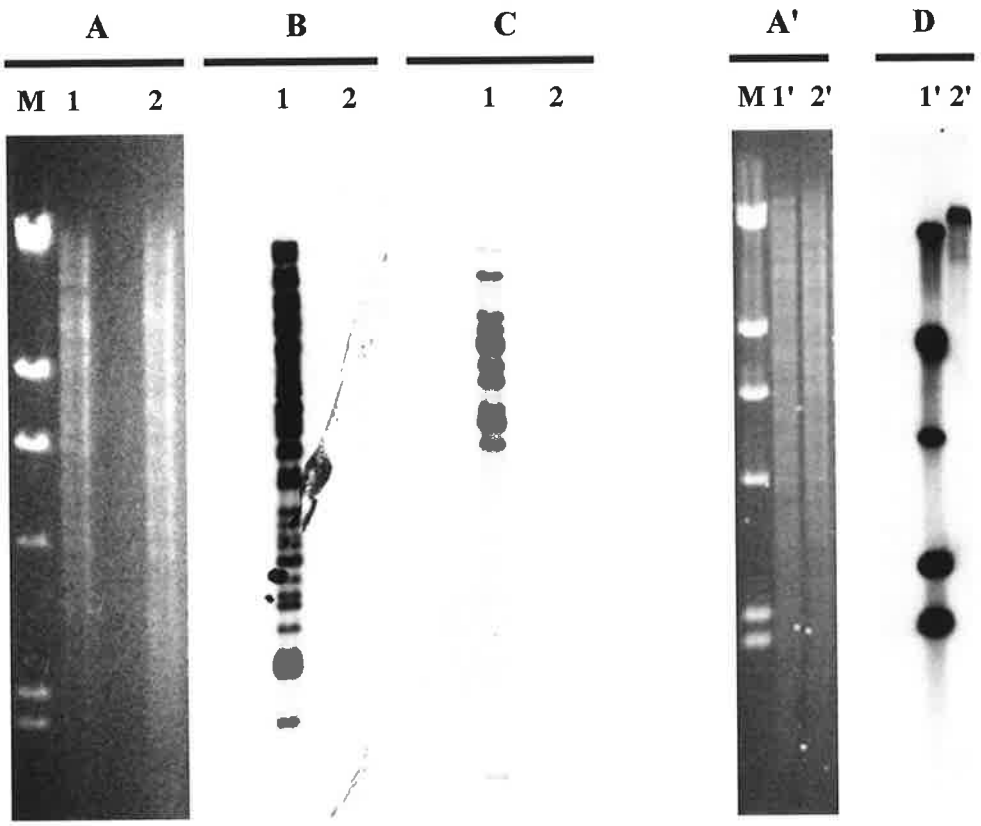


Figure 3.1

A. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of *R. solanacearum*. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.6, 4.4, 2.3, 2.0 and 0.56 kb was included as a size marker (M).

1. ACH0158 strain.
2. ACH0171 strain.

B. Southern hybridisation of the gel in A with the SC15 insert of pSC15.

C. Southern hybridisation of the gel in A with the ISB (a partial IS*Rso2*) insert of pISB.

A'. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of *R. solanacearum*. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.6, 4.4, 2.3, 2.0 and 0.56 kb was included as a size marker (M).

1'. ACH0171 strain.
2'. ACH0158 strain.

D. Southern hybridisation of the gel in A' with the IS61 insert of pIS61.

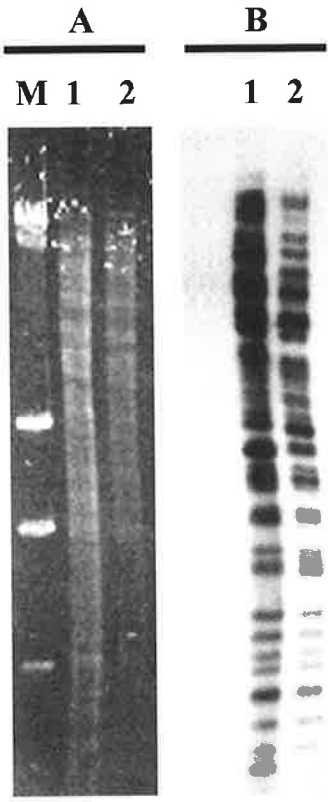


Figure 3.2

A. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of *R. solanacearum*. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.6, 4.4, 2.3, 2.0 and 0.56 kb was included as a size marker (M).
    1. ACH0158 strain.
    2. ACH1068S strain.

B. Southern hybridisation of the gel in A with the SC15 insert of pSC15.

duplicate DNA samples from colonies containing *Sau*3AI fragments of ACH0171 DNA were hybridised with the same genomic probes and colonies showing strong preferential hybridisation to the ACH0171 probe were selected for further characterisation. Sequencing of two such clones identified pIS61 that contained a full length IS element within a 1.02 kb *Sau*3AI insert, designated IS*61*. The labelled insert of pIS61 revealed multiple *Eco*RI fragments in Southern analysis of ACH0171 genomic DNA, while a single high molecular weight band was hybridised in ACH0158 (Fig. 3.1D).

3.3 Analysis of IS elements from ACH0158

3.3.1 Analysis of the SV102 insert

Sequencing of the clone pSV102 was carried out initially with the M13 forward (-20) and reverse (RSP) primers and the approximately 1 kb of DNA sequences obtained did not contain the SC15 sequence used as a probe. Nevertheless, almost 500 bp of the sequence obtained by the RSP showed extensive similarity to several known IS elements including IS*1031* and IS*12528* of *Acetobacter xylinum* and *Gluconobacter suboxydans*, respectively (Coucheron, 1993; Kondo and Horinouchi, 1997) while the remaining 500 bp sequence showed no significant similarity to any sequences in the database. Convenient *Sal*I sites were utilised to generate two subclones, pSV1022 and pSV102-1040 and a single *Pst*I site was used to generate two additional subclones, pSV102-825 and pSV102-260 (Fig. 3.3A). Sequencing of the four subclones (Fig. 3.4A, 3.4B, 3.5A, 3.5B and 3.7) revealed that the SV102 insert DNA was 2698 bp in length and contained two different putative IS elements, designated as IS*Rso3* and IS*Rso4* (Fig. 3.3A and 3.7). The 520 bp upstream of IS*Rso3* in pSV102 showed no significant homology with any known sequences from the databases. The 1209 bp long IS*Rso3* and 855 bp long IS*Rso4* were contiguous and the two IS elements were separated by a CTAG sequence which appeared to be a direct repeat (DR) shown by both IS elements (Fig. 3.7). 112 bp sequences located downstream of IS*Rso4* showed no apparent sequence similarity to any known sequences. The SC15 sequences, utilised as a probe to obtain pSV102, were located at the 3' end of IS*Rso3* (Fig. 3.4A).

3.3.2 Sequence analysis of IS*Rso3*

IS*Rso3* was 1209 bp in length and its 3' region contained the SC15 sequences (Fig. 3.4A). IS*Rso3* carried imperfect 18 bp IRs with 4 nucleotide mismatches (Fig. 3.4A). Sequence comparisons of the 5' end, between IRs of IS*Rso3* and homologous IS elements such as IS*1384* (GenPep accession no. AAC98743),, ISPSMC (GenPep accession no. BAA75460) and IS*1646* (GenPep accession no. AAC27326) of

Pseudomonas species and *Xanthomonas campestris* pv. *vesicatoria*, identified a consensus sequence within the IR (Fig. 3.9A). However, the consensus sequence contained neither GGC nor GAG which are the conserved outer triplets of the IRs for the IS*5* family (Rezsöhazy *et al.*, 1993; Fig. 3.9A). IS*Rso3* contained the A/T rich internal region and the relatively G/C rich external region, which are typical of the IS*5* family (Rezsöhazy *et al.*, 1993).

In addition to the IR features, IS*Rso3* contained characteristic 4 bp CTAG DRs immediately flanked by the IRs of IS*Rso3* (Fig. 3.7). DRs are direct duplications of genomic DNA targeted specifically by the IS element at the point of insertion. The 3' DR for IS*Rso3* was also the 5' DR of IS*Rso4* (Fig. 3.7). In pSC15K (discussed in chapter 6), containing another genomic copy of IS*Rso3*, the corresponding DR sequences were CTAG and TTAG. Subsequent DR analysis confirmed frequent mismatches in DR sequences between different copies of IS*Rso3* (Table 3.1). Variations in the DR sequence may influence the excision rate of IS*Rso3* as proven for IS*1236* (Gerischer *et al.*, 1996).

IS*Rso3* contained a long ORF encoding a putative transposase of 372 aa (Fig. 3.4A). A search in the protein database revealed several homologous transposases including those of IS*1384* (GenPep accession no. AAC98743), ISPSMC (GenPep accession no. BAA75460) and IS*1646* (GenPep accession no. AAC27326), from distantly related *Pseudomonas* species and *Xanthomonas campestris* pv. *vesicatoria.* Amino acid sequence comparisons clearly identified two highly conserved regions and overall sequence identity and similarity were greater than 50 % and 70 % respectively in all cases (Fig. 3.4D). IS*Rso3* and the homologous IS elements shared two domains characteristic and diagnostic of the IS*5* family: N3 with [D (aspartic acid)+1+G (glycine)/A (alanine)+Y (tyrosine)/F (phenylalanine)] as highly conserved sequences and C1 with [Y (tyrosine)+2R (arginine)+3E (glutamic acid)+6K (lysine)] (with only [R (arginine)+3E (glutamic acid)]) as invariant sequences (Fig. 3.4C). Therefore, it is evident that IS*Rso3* is a member of the IS*5* family with [R (arginine)+3E (glutamic acid)+6K (lysine)] as C1 core sequences.

3.3.3 Sequence analysis of IS*Rso4*

The clone pSV102 also contained IS*Rso4* which was immediately adjacent to IS*Rso3* (Fig. 3.7). IS*Rso4* was readily identified as an IS element because of its extensive nucleotide and derived amino acid sequence similarities to several known IS elements

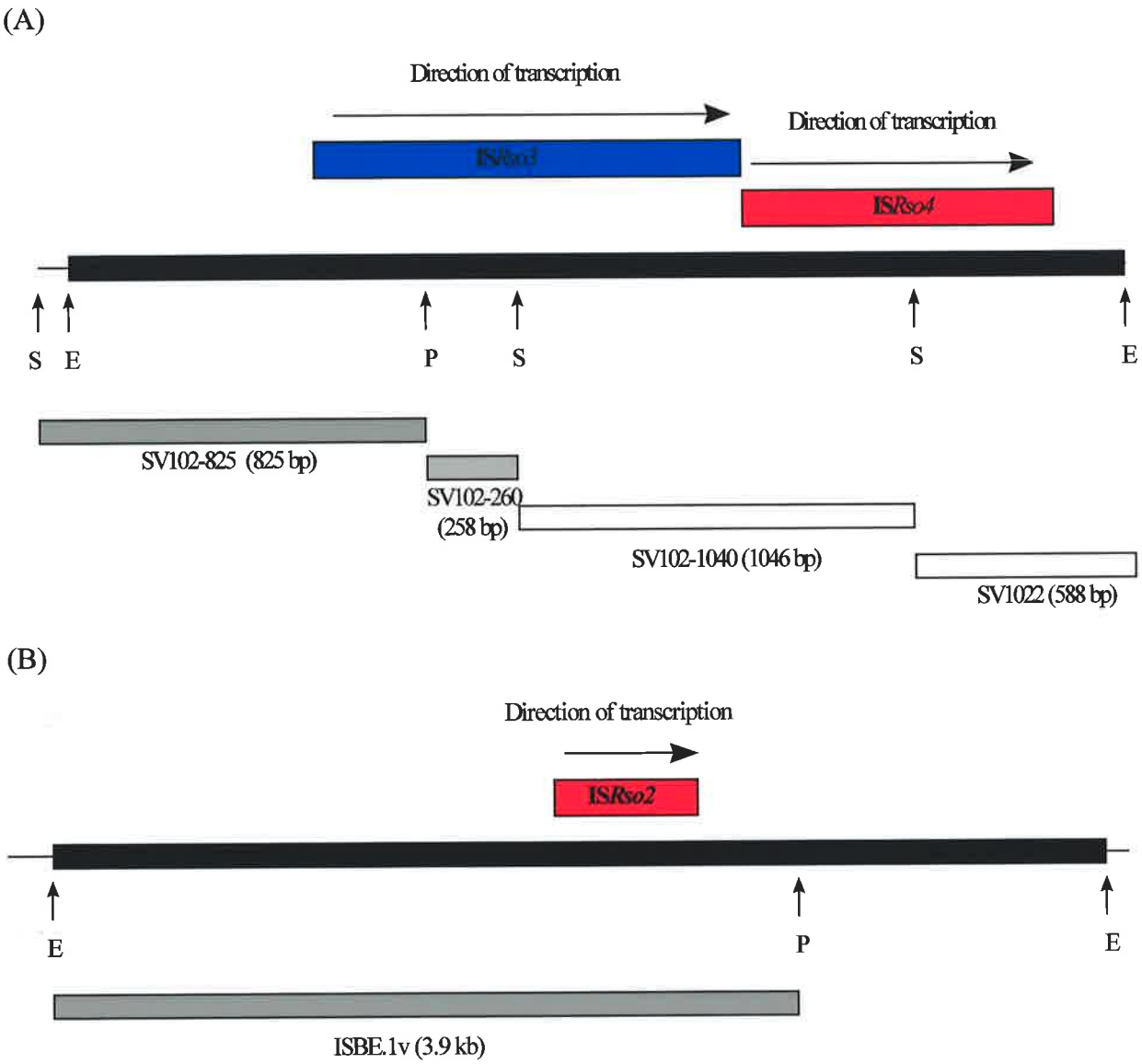


Figure 3.3

A. A physical map of SV102 with restriction sites for the endonucleases *Eco*RI (E), *Sal*I (S) and *Pst*I (P) marked. The fragment sizes for the four subclones are given in bp. The locations and the transcriptional directions of IS*Rso3* (blue box) and IS*Rso4* (red box) are indicated.

B. A physical map of ISBE with restriction sites for the endonucleases *Eco*RI (E) and *Pst*I (P) marked. The fragment size for the subclone of ISBE.1v is given in kb. The location and the transcriptional direction of IS*Rso2* (red box) are indicated.

of *Acetobacter xylinum* and *Gluconobacter suboxydans*, such as IS*1031* and IS*12528* (Coucheron, 1993; Kondo and Horinouchi, 1997). IS*Rso4* is 855 bp long with perfectly matched 17 bp terminal IRs (Fig. 3.5A). The sequence comparison of the left-hand ends with homologous IS elements, showed the characteristic external triplet GAG although IS*Rso4* contained only AG (Fig. 3.9B) which was directly flanked by the CTAG DR sequence. The GAG sequence, or GGC in some cases are recognised as IS*5* family-specific trinucleotides (Rezsöhazy *et al.*, 1993; Ohtsubo and Sekine, 1996; Fig. 3.9B). In addition to the presence of the AG sequences, the internal part of the IRs of IS*Rso4* and its homologues is relatively A/T rich while the external part is G/C rich. Therefore it appears that IS*Rso4* and its the homologous IS elements show higher conservation and more typical IR features than the IS*Rso3*-like elements.

IS*Rso4* contained characteristic 4 bp CTAG DR sequences flanked by IRs and, as previously described, the 5' DR of IS*Rso4* was also the 3' DR of IS*Rso3* (Fig. 3.7). Clear DR duplications were observed in two additional clones (pM81C and pIS102JI, discussed in chapter 4 and 6 respectively) carrying different copies of IS*Rso4* (Table 3.1). However, the insertion of IS*Rso4* found in pM81C had created a CTGAG motif (Table 3.1), instead of the CTAG found in pSV102 and pIS102JI. This would suggest that IS*Rso4* has DR sequence and length variations, although the insertion event seemed to create perfectly paired DRs. In fact, the variations of the target sequence and its length are influenced by the genomic background of the target locations as well as the consensus target sequences of each IS element (Galas and Chandler, 1989). Nevertheless, the target sequence and length variation, as well as the precise duplications upon the insertion, would influence the ability of IS*Rso4* to transpose. IS*Rso4* contained a long ORF encoding a putative transposase of 281 aa. A protein database search with this amino acid sequence identified several homologous sequences of elements belonging to the IS*5* family from various bacterial species. They included the conceptual translation products of IS*1031* in *Acetobacter xylinum* (35 % identical, 57 % similar), IS*12528* from *Gluconobacter suboxydans* (35 % identical, 57 % similar), and ISpRm*220-12-1* from *Sinorhizobium meliloti* (35 % identical, 58 % similar) (Fig. 3.5D). Multiple sequence alignment of the amino acid sequences of IS*Rso4* and those available from the IS*5* family revealed the well conserved core sequences of two highly characteristic domains of the IS*5* family (Rezsöhazy *et al.*, 1993; Ohtsubo and Sekine, 1996) which are the N3 domain with [D (aspartic acid)+1+G (glycine)/A (alanine)+Y (tyrosine)/F (phenylalanine)] sequences and the C1 domain with [Y (tyrosine)+2R (arginine)+3E (glutamic acid)+6K (lysine)] motifs (Fig. 3.5C). However, IS*Rso4* and its

homologues appeared to contain only the invariant sequence of [R (arginine)+3E (glutamic acid)] for the C1 region (Fig. 3.5C). The distance between the two main domains of the IS*Rso4*-like elements was shorter than that of the IS*Rso3*-like sequences.

3.3.4 Sequence analysis of IS*Rso2*

The insert of the clone pISBE was an *Eco*RI fragment of approximately 5.2 kb derived from ACH0158 genomic DNA containing a full length copy of IS*Rso2* (Fig. 3.3B). IS*Rso2* was 864 bp in length with 15 bp IRs that differ from each other by only a single base pair (Fig. 3.6A). A sequence comparison of IRs of IS*Rso2* and homologous IS elements, such as IS*427*, IS*711* and IS*1301*, identified particular consensus sequences (Fig. 3.9C). IS*Rso2* in pISBE contained a mismatched DR sequence pair of CTAA and TTAA (Table 3.1). Unfortunately, sequencing of IS*Rso2* in the pISB clone showed that the clone did not include full-length element and it contained only a single copy of the putative DR (CTAA) directly flanked by an IR similar to those in pISBE. Therefore only 3 examples of the DR in IS*Rso2* are available. Nevertheless, DR sequence pairs of CTAA and TTAA in pISBE indicated that insertion of IS*Rso2* does not always create precise DR duplications.

IS*Rso2* revealed a relatively short ORF for a putative transposase of 134 aa (Fig. 3.6A) which identified significant similarities to amino acid sequences of transposases of several IS elements in a protein database search. These IS elements included IS*427* (36 % identity and 56 % similarity), IS*711* (33 % identity and 53 % similarity) and IS*1301* (28 % identity and 49 % similarity) from *Agrobacterium tumefaciens*, *Brucella melitensis* and *Neisseria meningitidis*, respectively (Fig. 3.6D). Multiple amino acid sequence alignment revealed that this group of IS elements contained specific conserved sequences (Fig. 3.6C), but the group has not as yet been classified into a particular family of IS elements. Some of the bacterial species in which elements similar to IS*Rso2* were found are taxonomically distant from *R. solanacearum*, with *Neisseria meningitidis* and *Brucella melitensis* being human pathogens causing meningococcal meningitis and brucellosis, respectively.

3.4 Sequence analysis of IS*61* from strain ACH0171

The insert of the clone pIS61 derived from ACH0171 genomic DNA was a 1.02 kb *Sau*3AI fragment. Sequencing of pIS61 revealed a novel IS element, named IS*61* that was 781 bp in length with 28 bp IRs containing 12 nucleotide mismatches (Fig. 3.8A). A protein database search with the major derived amino acid sequence identified

proteins from several elements belonging to the IS*3* family. There was little similarity between IR sequences of IS*61* and these homologous elements (result not shown), consistent with the observation that IR sequences of IS*3* family are typically heterogeneous. Sequencing also revealed putative DRs, being 3 bp TGG flanked immediately by the IR sequences. Target sequence specificity, as well as DR sequence analysis, were not possible because only a single copy of IS*61* was cloned and characterised.

IS*61* contained a short ORF encoding 122 aa which showed no significant similarity to any known sequences in a protein database search. However, a -1 translational frame-shift could link part of the protein encoded by this ORF to a protein encoded by a second ORF (Fig. 3.8A). The conceptual translation product showed extensive similarity to amino acid sequences of several IS elements including IS*1236* from *Acinetobacter calcoaceticus* (Gerischer *et al.*, 1996) and IS*904* from *Lactococcus lactis* (Dodd *et al.*, 1990) of the IS*3* family. The second ORF did not have a start codon but showed the invariant amino acid sequences of D (aspartic acid)+35+E (glutamic acid) characteristic of the IS*3* family close to the *C*-terminus (Fig. 3.8A). The -1 translational frame-shift is typical of members of the IS*3* family and allows the synthesis of a functional transposase. The first ORF contained the start codon and the nucleotide sequence AGAAAG that probably causes the -1 translational frameshift (Ohtsubo and Sekine, 1996) (Fig. 3.8A). However, the AGAAAG motif found in IS*61* is slightly different from the sequence (AAAAG) which is known to induce translational frame-shift in IS*3* in *E. coli* (Sekine *et al.*, 1994) and the putative frame-shifting signal (AAAAAAA) found in IS*1236* from *A. calcoaceticus* (Gerischer *et al.*, 1996). A helix-turn-helix (HTH) motif was not found in the first ORF of IS*61* although it is present in several members of the IS*3* family. The amino acid sequence of the putative fusion transposase is 199 aa in length and it shows high sequence similarity to ORF sequences for transposase of several IS elements from diverse bacterial species (Fig. 3.8B). These include the IS*3*-like element from *Agrobacterium tumefaciens* (60 % similar, 44 % identical), IS*1236* from *Acinetobacter calcoaceticus* (49 % similar, 30 % identical), and IS*904* from *Lactococcus lactis* (45 % similar, 28 % identical) (Fig. 3.8C). It was concluded that IS*61* is a member of the IS*3* family but that contains some unique characteristics.

3.5 Summary and discussion

In summary, sequencing of the three genomic clones, pSV102, pISBE from ACH0158 and pIS61 from ACH0171 revealed four novel IS elements: IS*Rso3*, IS*Rso4*, IS*Rso2* and IS*61*. The nomenclatures of IS*Rso4*, IS*Rso3* and IS*Rso2* were registered through http://www-is.biotoul.fr/is.html while IS*61* is not yet registered. Sequence analyses of IS*Rso3* and IS*Rso4* indicated that they are classified as IS*5* family members. IS*61* belongs to the IS*3* family on the basis of both its nucleotide and derived amino acid sequence characteristics. As yet IS*Rso2* has not been classified into a particular IS family although it readily groups with a cluster of known IS elements.

The IS elements characterised in this chapter share significant nucleotide and derived amino acid sequence characteristics with IS elements from a diverse range of bacterial species. Most of the bacterial species for which sequence similarity was demonstrated are grouped within the alpha- and gamma-subdivisions of the proteobacteria, whereas *R. solanacearum* is placed in the beta-subdivision, suggesting that horizontal transfer of the IS elements may have occurred. The most extreme example IS*904* from *Lactococcus lactis*, which is similar to IS*61* because the species belongs to the Bacillus/Clostridium group of the Streptococcaceae family. The ability of *R. solanacearum* to be transformation competent (Bertolla *et al.*, 1999) during infection of a host plant would enhance the chance for the pathogen to take up foreign DNAs. The consequences of horizontal gene transmissions to obtain exotic genetic material could well be that the pathogen is very adaptable to new environments and more capable of infecting new host plants. Both these characteristics are pronounced features of *Ralstonia solanacearum*. In the future, by testing an extensive range of diverse bacterial species with the four IS element probes, it may be possible to draw conclusions about the frequency and mode of horizontal gene transfer of the elements among different bacterial species. A rapidly growing number of complete bacterial genome sequences will potentially clarify the possibility of horizontal gene transfer of IS elements.

Directly repeated (DR) nucleotide sequences are created upon insertions of IS elements. However, IS*Rso3* and IS*Rso2* elements in different locations in the ACH0158 genome were frequently flanked by unmatched DR sequences. It is not clear whether unmatched nucleotide sequences were actually generated upon insertions of the two IS elements or that the DR sequences were modified during DNA replications after the insertion events. Nevertheless, nucleotide sequence variations in the DR may influence the transposition frequency of IS*Rso3* and IS*Rso2*.

Figure 3.4

A. Complete nucleotide and deduced amino acid sequences of IS*Rso3*. Inverted repeat sequences are shaded. The nucleotide sequences of IS*Rso3*-derived primers for sequencing and PCR are underlined. The identities and directions of the primer are also indicated. The invariant amino acid sequences of N3 region with [D+1(G/A)(Y/F)] and C1 with [R+3E+6K] are in bold and underlined. Nucleotide sequences of the SC15 are in bold and two *Sau*3AI sites utilised to clone pSC15 are indicated.

B. The identities and nucleotide sequences of IS*Rso3*-derived primers for sequencing and PCR.

C. Alignment of amino acid sequences of putative transposases encoded by IS*Rso3*, IS*1384* (GenPep accession no. AAC98743), ISPSMC (GenPep accession no. BAA75460) and IS*1646* (GenPep accession no. AAC27326). The asterisk (*) indicates the amino acid sequences which are identical. Amino acid sequences of D+1GY and R+3E for N3 and C1 regions respectively are in bold. Numbering refers to the amino acid residues of the putative transposase of IS*Rso3*.

D. The percentages of amino acid sequence identity and similarity of IS*Rso3* to IS*1384*, ISPSMC and IS*1646*.

(A)

1 GGAAACGCTGATTTATCAGTCTCAGCCGTCATCGCTGACGCAGCTGAGGCCGTTGTAGCA

61 GGTAGCTCACGGAACGGAGCCATAGTGATGAAGCGGCAGATGAGCTTTGCAGAAGCGGAA

← BD1

M K R Q M S F A E A E

121 AGCGCAGGTAAGAAGCGCGTGACCAAACGCCAGCGCTTCCTGGCAGAGATGGAGAAGGTC

S A G K K R V T K R Q R F L A E M E K V

181 GTGCCGTGGCAGCGCTTGCTGTCGGCGATTGGGCCCCACTATCCGAGAGGCGAGCGAGGT

V P W Q R L L S A I G P H Y P R G E R G

241 CGCCCGCCGATTGGTCTGGAGAGGATGCTGCGGATCTACTTCCTGCAGCAGTGGTACGGG

R P P I G L E R M L R I Y F L Q Q W Y G

301 CTGTCGGACGAAGGTCTGGAAGACGCGCTGCACGACAGTATGGCGATGCGAGCCTTCGCC

L S D E G L E D A L H D S M A M R A F A

361 GGCATCGATCTGGCGGTCGAGGACGTGCCGGATGCGACCACGCTGTTGAAGTTCCGACGC

G I D L A V E D V P D A T T L L K F R R

421 CTGCTCAACGAACACGACTTGACGCGAAAGCTGTTCGACGAGATCGGCATCATGCTGTGC

L L N E H D L T R K L F D E I G I M L C

481 CGAGCGGGGGCTGATGATGAAGGAAGGCACGATCGTGGATGCCACCATCATTGAAGCGCC

E R G L M M K E G T I V D A T I I E A P

541 CCGTCGACGAAGAACGCCGACAAGAGCCGCGACCCGGAGATGCATCAGACCAAGAAAGGC

P S T K N A D K S R D P E M H Q T K K G

601 AACGAGTGGCACTTCGGAATGAAGGCGCATATCGGAGTCGATGCGTCGTCGGGCTTGGTG

N E W H F G M K A H I G V D A S S G L V

661 CACAGCGTGGTTGGCACGGCAGCCAATGCGTCCGATGTGTCGCAGGCCCACGCGCTGCTG

H S V V G T A A N A S D V S Q A H A L L

721 CATGGCCACGAGACGGATGCGTTTGGCGATGCGGGCTACACCGGCGTTGAGAAGCGCGAC

H G H E T D A F G D A G Y T G V E K R D

781 GCGATGCAAGGCAAGCGAGCGACGTGGCATGTCGCGATCAAGCGCGGCAAGATCAAAGCG

A M Q G K R A T W H V A I K R G K I K A

841 ATGCGCGAAGGCCCAATCAAGGACTTGCTGATCGCGGTGGAGCGAACCAAGGCACAGATT

*Sau*3AI

M R E G P I K D L L I A V E R T K A Q I

901 CGGGCTCGGGTCGAACATCCGTTCCATGTCATCAAGAATCTGTTTGGTCATCGCAAGGTT

← PSC3

R A R V E H P F H V I K N L F G H R K V

961 CGTTACAAGGGCTTGGCCAAGAACACGGCGCAACTGTTCAGCCTGTTTGGTTTGGCCAAT

BD2 →

R Y K G L A K N T A Q L F S L F G L A N

1021 CTGGTGCTCGCCAGGAGGCAGTTGCTGGCCAGTCCTGGGAGCATTGCGTCCTGAGCGCGC

L V L A R R Q L L A S P G S I A S *

1081 AAAGCGCGCCAAGATAGGTGCGAAAGCAGCGAAAAACCGTGCTGAATCGAAGCATCCCTT

1141 CCTCAATCCGAAAAACCAAGATGGCATGACTCGCGATTCCATTCGCCGACTTCATTGATC

*Sau*3AI

1201 AGCGGTTCC

(B)

oligo names	sequences (5′ → 3′)
BD1	CTG CCG CTT CAT CAC TAT GG
BD2	GCA AGG TTC GTT ACA AGG GC
PSC3	ATG ACA TGG AAC GGA TGT TCG

(C)

```
                1                   21                  41
ISRso3          MKRQMSFAEAESAGKKRVTKRQRFLAEMEKVVPWQRLLSAIGPHYPR-GERGRPPIGLER
IS1384          -MKQMTFADAEYAGKRKQTRKELFLIEMDRVVPWKGLIALIEPHYPK-GEGGRPAYPLMA
ISPSMC          -MKQMTFADAEYAGKRKHTRRERFLIEMDQVVPWKGLIALIEPHYPK-GEGGRPAYPLMA
IS1646          --MQLTFGDAEGLGKRKQTRREIFLAEMEQVVPWQQLLGLVAPHYPVSGRPGRQPYALAT
                   *  *  **  **   *    ** **  ****  *     ****  *  **    *

                61                  81                  101
ISRso3          MLRIYFLQQWYGLSDEGLEDALHDSMAMRAFAGIDLAVEDVPDATTLLKFRRLLNEHDLT
IS1384          MLRVHLLQNWFGYSDPAMEEALYETTILRQFAGLN--LERIPDETTIFNFRRLLEKHELA
ISPSMC          MLRIHLMQNWFGYSDPAMEEALYETTILRQFAGLS--LDRIADETTILNFRRLLEKHELA
IS1646          MLRIHLLQQWYALSDPAMEEALHEIPTLRRFAQLGG-LDNVPDETTILNFRRLLETHGLA
                ***    * *   **   * **      * **          * **   *****  * *

                121                 141                 161
ISRso3          RKLFDEIGIMLCERGLMMKEGTIVDATIIEAPPSTKNADKSRDPEMHQTKKGNEWHFGMK
IS1384          AGILAVINGYLGDRGLSLRQGTIVDATLINAPSSTKNKDGKRDPEMHQTKKGNQYYFGMK
ISPSMC          GGILQVINGYLGDRGLIRRQGTVVDATIIHAPSSTKNKDGKRDPEMHQTKKGNQYFFGMK
IS1646          ARMLEAVNAHLARKGQSLRSGTIVDATLIAAPSSTKNPDHARDPEMHQTKKGNQWYFGMK
                         *   *     ** *****  ** **** *  ************   ****

                181                 201                 221
ISRso3          AHIGVDASSGLVHSVVGTAANASDVSQAHALLHGHETDAFGDAGYTGVEKRDAMQGKRAT
IS1384          AHIGADDESGLVHSVVGTAANVADVTQVDKLLHGDENVVCADAGYTGVEKRPEHEGREVI
ISPSMC          AHIGVDAESGLVHSLVGTAANVADVTQVDQLLHGEETYVSGDAGYTGVDKRAELQDRQMI
IS1646          AHIGVDEFSGLVHHVHCTAANVADVTVTHALLHGKEDSVFGDSGYTGADNREELQTCKAA
                **** *  *****    ****  **     **** *     * ****   *

                241                 261                 281
ISRso3          WHVAIKRGKIKAMRE-GPIKDLLIAVERTKAQIRARVEHPFHVIKNLFGHRKVRYKGLAK
IS1384          WQVAARRSTYKKLDKRSVLYKAKRKIEKAKAQVRAKVEHPFRVIKRQFGYTKVRFRGLAK
ISPSMC          WSIAARPSRYKKHGEKSLIARVYRKIEFTKAQLRAKVEHPFRVIKRQFGYTKVRFRGLAK
IS1646          FFIAARRSVLQAIGNKRERAREQR-WEHFKASVRAKVEHPFRVIKRQFGYTKVRYRGLAK
                   *                       *  **  ** ***** ***  **  ***  ****

                301                 321                 341
ISRso3          NTAQLFSLFGLANLVLARRQLLASPGSIAS-ARKARQDRCESSEKPC-IEASLPQSEKPR
IS1384          NTAQLVTLFALSNLWMARRHLLTNAGEVRL-------------------------------
ISPSMC          NTAQQATLFALSNLWMVRKRLLA-MGEVRL-------------------------------
IS1646          NTAQVLTLFALSNLWMKRKQLLPAMGSVRL-------------------------------
                ****   ** * **   *  **   *

                   361
ISRso3          WHDSRFHSPTSLISGS
IS1384          ----------------
ISPSMC          ----------------
IS1646          ----------------
```

(D)

		aa identity (%)	aa similarity (%)
ISRso3	IS1384 (AAC98743)	54	72
	ISPSMC (BAA75460)	54	72
	IS1646 (AAC27326)	52	72

Figure 3.5

A. Complete nucleotide and deduced amino acid sequences of IS*Rso4*. Inverted repeat sequences are shaded. The nucleotide sequences of IS*Rso4*-derived primers for sequencing and inverse-PCR (see chapter 5) are underlined. The identities and directions of each primer are also indicated. The invariant amino acid sequences of N3 region with [D+1(G/A)(Y/F)] and C1 with [R+3E] are in bold and underlined.

B. The identities and nucleotide sequences of IS*Rso4*-derived primers for sequencing and inverse-PCR (see chapter 5).

C. Alignment of amino acid sequences of putative transposases encoded by IS*Rso4*, IS*1031* (GenPep accession no. AAA25029), IS*12528* (GenPep accession no. BAA19758) and ISpRm*220-12-1* (GenPep accession no. AAD28750). The asterisk (*) indicates the amino acid sequences which are identical. Amino acid sequences of D+1GY and R+3E for N3 and C1 regions respectively are in bold. Numbering refers to the amino acid residues of the putative transposase of IS*Rso4*.

D. The percentages of amino acid sequence identity and similarity of IS*Rso4* to IS*1031*, IS*12528* and ISpRm*220-12-1*.

(A)

```
1    AGCGTGTTAGGCACTTTCGGAAAGGCTCGTTATGCTGTGTGCATGAGCCCCCGCAAGCCT
                                                  M  S  P  R  K  P

61   TACCCGACAGACGTATCCGATGAAGAATGGAGCTTCGCAGCGCCATACCTGACGCTGATG
     Y  P  T  D  V  S  D  E  E  W  S  F  A  A  P  Y  L  T  L  M

121  CGCGAGGACGCGCCGCAACGCACACATGACCTACGCGAGATGTTCAACGCACTGCGCTGG
     R  E  D  A  P  Q  R  T  H  D  L  R  E  M  F  N  A  L  R  W

181  ATGGCGCGTGCCGGAGCTGCGTGGCGCATGCTGCCGACGAACTTTCCTCCGTGGGAGCTG
                     ←  IS102RR
     M  A  R  A  G  A  A  W  R  M  L  P  T  N  F  P  P  W  E  L

241  GTGTACCAGCAAACACAACGCTGGCTCAATGCGGGTTGCTTTGAGGCCATGGTCAATGAT
     V  Y  Q  Q  T  Q  R  W  L  N  A  G  C  F  E  A  M  V  N  D

301  CTGCGTTCGGTGATACGGGTCGCCCAAGAGAGGCAAGGTCAACCCAGTGCCGTCATTCTG
     L  R  S  V  I  R  V  A  Q  E  R  Q  G  Q  P  S  A  V  I  L

361  GATGGTCGAACCTTGCAGTCGACCTGCGAGAGTGGCCCGCGCGCGGGCTACGACGGCTAC
     D  G  R  T  L  Q  S  T  C  E  S  G  P  R  A  G  Y  D  G  Y

421  AAACGCAAACGGGGCAGCAAGGTTCACATGGCAGTGGATACGCTGGGCCATCTGCTGGCG
     K  R  K  R  G  S  K  V  H  M  A  V  D  T  L  G  H  L  L  A

481  GTGCATGTCACGCCGGCCAATGAACAAGAGCGTGCGCAGGTAGCGGAATTGGCTCGTCAG
     V  H  V  T  P  A  N  E  Q  E  R  A  Q  V  A  E  L  A  R  Q

541  GTTCAGCAGGCCACAGGCCAAACCGTGAAGGTAGCGTTTGCCGACCAGGGGTATACCGGT
     V  Q  Q  A  T  G  Q  T  V  K  V  A  F  A  D  Q  G  Y  T  G

601  GAAGCGCCTGCACAAGCCGCGCTGGATGAAGGAATCGACCTTCAAGTCATCAAGCTATCA
                                                  PSC1  →
                                  .....
                                  ←  PSC2
     E  A  P  A  Q  A  A  L  D  E  G  I  D  L  Q  V  I  K  L  S

661  GAGGCCAAGAAAGGCTTTGTGCTGCTGCCCAGGCGCTGGGTCGTTGAGCGTAGCTTTGGC

     E  A  K  K  G  F  V  L  L  P  R  R  W  V  V  E  R  S  F  G

721  TGGCTCAACCGCTTCAGGCGACTCGCCCGTGACTATGAACGCCTTCCAGAAACCCTTGCG
     W  L  N  R  F  R  R  L  A  R  D  Y  E  R  L  P  E  T  L  A

781  GGTGTGCACTTTGTTGTCTTCGCCATGATCATGTTGGTTCATGCCGTGCCGATCATGCAA
     G  V  H  F  V  V  F  A  M  I  M  L  V  H  A  V  P  I  M  Q

841  AGTGCCTAACACGCT
        Ja  →
     S  A  *
```

(B)

oligo names	sequences (5′ → 3′)
IS102RR	CAG CAT GCG CCA CGC AGC
PSC1	AGT CAT CAA GCT ATC AGA GG
PSC2	ATG ACT TGA AGG TCG ATT CC
Ja	GCA AAF TGC CTA ACA CG

(C)

```
                              1                         21                        41
ISRso4              -----------MS----PRKPYPTDVSDEEWSFAAPYLTLMREDAPQRTHDLREMFNALR
IS1031              MVTWTGIARREHS---REGLRYPSDMMDGEWALIMPFVPPAKRGGRPRTTDMREVVNAML
IS12528             --MWTPAQRGRMAGITRKTKRYPSDLTDEEWERIAPLMSPANRRGRKRTTDFREIINALR
ISpRm220-12-1       -MGWTDFTRRQYG---RRTGRYASDLTDREWSLIAPFMPMPPQLGRPRKTELREVLNALL
                                     *  *  * **    *                  **  **

                              61                        81                        101
ISRso4              WMARAGAAWRMLPTNFPPWELVYQQTQRWLNAGCFEAMVNDLRSVIRVAQERQGQPSAVI
IS1031              YIASAGCAWRLLPKCFPPVSTVRRYFYAWRDTGLFEVMNTVLVMSLREIEGREASPSAGV
IS12528             YLVRSGCGWEMLPVHFGPWQTVYWWFRRLMRRFLFQTIHDVCLMLDREAAGRETSPSGGV
ISpRm220-12-1       YIASTGCQWRMLPKDFPPYSTVQGYFYEWRSTGLWLRINYHLVMETRELEGKEASPTAGV
                         *  *  **  * *   *                        *        *

                              121                       141                       161
ISRso4              LDGRTLQSTCESGPRAGYDGYKRKRGSKVHMAVDTLGHLLAVHVTPANEQERAQVAELAR
IS1031              IDSQSVKTTE-SGGLSGYDAGKKVKGRKRHIVTDTCGFLIFLLVHAADIQDRDGAVDVLA
IS12528             IDSQSIKAP--HAKTRGYDAGKKIVGRKRHIAVDTDGRLLLVQLTTADISDSAGGQMILD
ISpRm220-12-1       IDSQSVKTTE-SGGIAGYDAGKKIKGRKRHIVVDTLGLMVGLVVHSADIQDRDGAVAALK
                     *              ***  *   * * *   ** *         *

                              181                       201                       221
ISRso4              QVQQATGQTVKVAFADQGYTGEAPAQAAL-DEGIDLQVIKLSEAKKGFVLLPRRWVVERS
IS1031              AIRRR-FPWLRHIFADGGYAGEKLRSALASMGKWTVEIIRRSDTVKGFQILPRRWVVERT
IS12528             AIRKR-WPWMKHLFADGAYDRLQLMDKAT-FLDFTVEIIRRSETAKGFEILPRRWVVERT
ISpRm220-12-1       TVLRR-WPWLRNIFADGGYAGPKLRASLRNVAKFTLQIVKRTDKAKGFEVLPRRWVVERT
                             ***  *                              ***  *********

                              241                       261
ISRso4              FGWLNRFRRLARDYERLPETLAGVHFVVFAMIMLVHAVPIMQSA-HA
IS1031              FAWLGRCRRLAKDWEQSIASSTAWTLIASIRMLTRRTARHCQA----
IS12528             FGWMIRWRRLVKDYEQRIDVAEAMIHIAMGSLMLRRNAHP-------
ISpRm220-12-1       FAWLGRCRRLAKDWEKSIASAEAWMLIAHIRILTRRLARYHVY----
                    * *  * ***  * *
```

(D)

		aa identity (%)	aa similarity (%)
ISRso4	IS1031 (AAA25029)	35	57
	IS12528 (BAA19758)	35	57
	ISpRm220-12-1 (AAD28750)	34	58

Figure 3.6

A. Complete nucleotide and deduced amino acid sequences of IS*Rso2*. Inverted repeat sequences are shaded. The nucleotide sequences of IS*Rso2*-derived primers for sequencing are underlined. The identities and directions of each primer are also indicated.

B. The identities and nucleotide sequences of IS*Rso2*-derived primers for sequencing.

C. Alignment of amino acid sequences of putative transposases encoded by IS*Rso2*, IS*427* (GenPep accession no. AAA22086), IS*711* (GenPep accession no. AAC98620) and IS*1301* (GenPep accession no. CAA88914). The asterisk (*) indicates the amino acid sequences which are identical. Numbering refers to the amino acid residues of the putative transposase of IS*Rso2*.

D. The percentages of amino acid sequence identity and similarity of IS*Rso2* to IS*427*, IS*711* and IS*1301*.

(A)

```
1    GAGCCGCTAACACAACTGGCCCTACGGCAAGAAGCGGGTTATCGTACGAGGCATGTCGAG
                                        ISB1  →
                                                    M  S  R

61   AAGAAAAGTCAGCAAAGAGTTGTGGTTGGCGCTGGAGCCGCTGATCCCGGAGTTTGTCGC
     R  K  V  S  K  E  L  W  L  A  L  E  P  L  I  P  E  F  V  A

121  CTCGCCCAAAGGTGGCCGGCGGCGCTCGGTCAATGATCGTGCAGCGTTGAGCGGCATCCT
      S  P  K  G  G  R  R  R  S  V  N  D  R  A  A  L  S  G  I  L

181  GTATGTGTTGCATAAGGGTATTCCGTGGGAAGACCTTCCCCAGGAACTTGGCTTTGGCAG

      Y  V  L  H  K  G  I  P  W  E  D  L  P  Q  E  L  G  F  G  S

241  CGGCATGACCTGCTGGCGCCGCCTGCGAGACTGGCAGGCAGGCGGCGTGTGGGACAAGCT
       B4  →
      G  M  T  C  W  R  R  L  R  D  W  Q  A  G  G  V  W  D  K  L

301  GCATCGGGCCATGCTGGTTCGACTGCGCGAGCACGACCAAATTGATTGGAGTCGGGCCAG
      H  R  A  M  L  V  R  L  R  E  H  D  Q  I  D  W  S  R  A  S

361  TATCGACGGGGCAAGCGTGCCCAGCCCCCGGGGGGGCGAACAAACCGGGCCGAGCCCGACG
      I  D  G  A  S  V  P  S  P  R  G  A  N  K  P  G  R  A  R  R

421  GATCGCGGCAAACTCGGGAGCAAGCGGCACCTCGTAGTCGATGCCCGAGGCGTTCCCCTG
      I  A  A  N  S  G  A  S  G  T  S  *

481  GCCATCACCGTAACAGGGGCGAATCGACATGACTCGATCGCATTCGAATCCACCTTGGAT
                                       Sau3AI              ←  B3

541  GCGATTCCCGCGATACGTGGCCTGGATGGCCGGCCGCGTAAGCGCCCCGACAAGCTGCAT

601  GCCGACAAGGCTTACGACTGCCGCCGATGTCGGCAGTATCTTAAGCGGCACGGCATAAGG

661  GCCCGGATCGCCCGCAAGGGCATCGAGAGCAGGGAGCGCCTGGGCCGTTACCGGTGGGTG

721  GTTGAGCGCACGCACGCTTGGTTTGCCGGCTTTGGAAAGATACGCGTTCGTTTCGAACGA

781  CGGCTCGATATCCATTGCGCGCTCCTCTCGCTTGCTGCCGCCATCATCTGCGCTCGATTC
                                                      ←  ISB2

841  GTGGATGACTTGTGTTAGCGACTC
```

(B)

oligo names	sequences (5′ → 3′)
ISB1	GCG GGT TAT CGT ACG
ISB2	TCC ACT AAT CGA GCG
B3	AAT CGC ATC CAA GGT GG
B4	TGG CAG CGG CAT GAC C

(C)

```
                  1                    21                   41
ISRso2          --MSRRKVSKELWLALEPLIPEFVASPKGGRRRSVNDRAALSGILYVLHKGIPWEDLPQE
IS427           --MSRYDLTDFEWRVIEPLLPN---KPRGVPRVD--DRRVLNGIFWVLRSGAPWRDLPER
IS711           MTRRRYELTDHEWSIISPLLPN---KPRGVARVD--DRRVLNGILWRFRTGSPWAEVPER
IS1301          --MARTAITDNIWEQLQTTM-----KAHGCHQWKN-DRTVMEAILWKLRTGAPWRDIPIE
                    *       *               *       **     *      * **   *

                  61                   81                          101
ISRso2          LGFGSGMTCWRRLRDWQAGGVWDKLHRAMLVRLREHD------QIDWSRASIDG----AS
IS427           Y--GPRTTCYNRFIRWRKAGVWDRMMDAITAAYDGDIQMIDSTSVRAHQQAATAKRGIEI
IS711           Y--GPAAACYNRFVRWRKAGVWDRLFETVPKAYDGDIVMIDSTCVRVHQHAATGKKGMET
IS1301          L--GSWKTAYNRFNRWSKKGLWQNFFLIYEKKLTKNG--YSSTEVMYGVINMQVELGVVS
                  *       *   *   * *

                             121
ISRso2          VPS---PRGANKPGRARRIAANSGASGTS--------------
IS427           IVS-AVPEAGSRPKSTRSSMGRVSRSG----------------
IS711           MVAWDVPAAGLRAKSTRSLMPKVARSIYV--------------
IS1301          IEQLDKAVAETRQKYTYVWTRMEIRSILKSLGVTCTTVKLQTT
                                       *
```

(D)

		aa identity (%)	aa similarity (%)
ISRso2	IS427 (AAA22086)	36	56
	IS711 (AAC98620)	33	52
	IS1301 (CAA88914)	27	49

Figure 3.7

Complete nucleotide sequences of SV102 insert of pSV102. IS*Rso3* sequences are shaded and IS*Rso4* sequences are shaded darker. The CTAG sequences which are the putative direct repeat duplications of IS*Rso3* and IS*Rso4* are in bold. The restriction sites of the endonucleases *Eco*RI, *Sal*I and *Pst*I used for the cloning and subcloning are marked. Nucleotide sequences and the direction of PSC4 primer for sequencing are indicated.

```
   1  GAATTCCTTCGATTCAGACGGGATTCCGTTCTATATTGATAGCCTGCCAGAGGATGATCT
      EcoRI
  61  CTATAACAATTGCGGCTAATTCTTTCTACGTCGCTCCAATCTTCGCACGCGAGCAGCGTA

 121  CCGTAATAGGTTACAACGTTCACAAAGCCTCCTTTGGGAGAACCGCTCTTCAGCTCGGTG

 181  ATGAGGGGCAGATGGTTGAGAAGTCGCAGCGAGCGCAATGGTGACCGACGATTGGACTGT

 241  TGTCCCCAGCGAGCTGGCCGTATCGCGGCGCCGTACGACTTCGTTGTCTCGATCCTCTTA

 301  GCGCCGGGGGCGTCGTCTACCTTGTCGGAAAGAATGTAGATCCGAGTCAGCTTCTTTGTG

 361  CCGGAGACGCTGAGTCCAAGAATCAGGGGCTTGAGCACCGTGGCAAGGTCCGACACACGT

 421  CGGCGTGCGAACCGTACCATCGCATACTCTCGCACATACCTGTTTTCCTGGATCAGGAAA
                                                    PSC4   →
 481  CTATCGACCATGCCGGCAAGTTTCTGAAGGCAGTGCTTTGCTAGGGAAACGCTGATTTAT
 541  CAGTCTCAGCCGTCATCGCTGACGCAGCTGAGGCCGTTGTAGCAGGTAGCTCACGGAACG
 601  GAGCCATAGTGATGAAGCGGCAGATGAGCTTTGCAAAAAGCGGAAAGCGCANGTAAGAAG
 661  CGCGTGACCAAACGCCAGCGCTTCCTGGCATAGATGGAGAAGTCGTGGCAGCGCTTGCTG
 721  TCGGCGATTGGGCCCCACTATCCGAGAGGCGAGCGAAGTCGCCCGCCGATTGGTCTGGAG
 781  AGGATGCTGCGGATCACTTCCTGCAGCAGTGGTACGGGCTGTCGGACGAAGTTCTGGAAT
                          PstI
 841  ACGCGCTGCACGACAGTATGGCGATGCGAGCCTTCGCCGGCATCGATCTGGCGGTCGAGA
 901  CGTGCCGGATGCGACCACGCTGTTGAAGTCCGACGCCTGCTCAACGAACACGACTTGACG
 961  CGAAAGCTGTTCGACGAGATCGGCATCATGCTGTGCGAGCGGGGGCTGATGATGAAGGAA
1021  GGCACAATCTTGGATGCCACCATCATTGAAGCGCCGCCGTCGACAATAACGCCGACAATA
                                              SalI
1081  ACCTCGACCCGGAGATCCTCTTACCAATAAAGGCAACGAGTGGCACTTCGGAATGAAGGC
1141  GCATATCGGAGTCGATGCGTCGTCGGGCTTGGTGCACAGCGTGGTTGGCACGGCAGCCAA
1201  TGCGTCCGATGTGTCGCAGGCCCACGCGCTGCTGCATGGCCACGAGACGGATGCCTTTGG
1261  CGATGCGGGCTACACCGGCGTTGAGAAGCGCGACTCGATGCAAGGCAAGCGAGCGACGTG
1321  GCATGTCCCGATCAAGCGCGGCAAGATCAAAGCGATGCCCGAAGGCCCAATCAAGGACTT
1381  GCTGATCGCGGTGGAGCGAACCAAGGCACAGATTCGGGCTCGGGTCGAACATCCGTTCCA
1441  TGTCATCAAGAATCTGTTTGGTCATCGCAAGGTTCGTTACAAGGGCTTGGCCAAGAACAC
1501  GGCGCAACTGTTCAGCCTGTTTGGTTTGGCCAATCTGGTGCTCGCCAGGAGGCAGTTGCT
1561  GGCCAGTCCTGGGAGCATTGCGTCCTGAGCGCGCAAAGCGCGCCAAGATAGGTGCGAAAG
1621  CAGCGAAAAACCGTGCTGAATCGAAGCATCCCTTCCTCAATCCGAAAAACCAAGATGGCA
1681  TGACTCGCGATTCCATTCGCCGACTTCATTGATCAGCGGTTCCCTAGAGCGTGTTAGGCA
1741  CTTTCGGAAAGGCTCGTTATGCTGTGTGCATGAGCCCCCGCAAGCCTTACCCGACAGACG
1801  TATCCGATGAAGAATGGAGCTTCGCAGCGCCATACCTGACGCTGATGCGCGAGGACGCGC
1861  CGCAACGCACACATGACCTACGCGAGATGTTCAACGCACTGCGCTGGATGGCGCGTGCCG
1921  GAGCTGCGTGGCGCATGCTGCCGACGAACTTTCCTCCGTGGGAGCTGGTGTACCAGCAAA
1981  CACAACGCTGGCTCAATGCGGGTTGCTTTGAGGCCATGGTCAATGATCTGCGTTCGGTGA
2041  TACGGGTCGCCCAAGAGAGGCAAGGTCAACCCAGTGCCGTCATTCTGGATGGTCGAACCT
2101  TGCAGTCGACCTGCGAGAGTGGCCCGCGCGCGGGCTACGACGGCTACAAACGCAAACGGG
        SalI
2161  GCAGCAAGGTTCACATGGCAGTGGATACGCTGGGCCATCTGCTGGCGGTGCATGTCACGC
2221  CGGCCAATGAACAAGAGCGTGCGCAGGTAGCGGAATTGGCTCGTCAGGTTCAGCAGGCCA
2281  CAGGCCAAACCGTGAAGGTAGCGTTTGCCGACCAGGGGTATACCGGTGAAGCGCCTGCAC
2341  AAGCCGCGCTGGATGAAGGAATCGACCTTCAAGTCATCAAGCTATCAGAGGCCAAGAAAG
2401  GCTTTGTGCTGCTGCCCAGGCGCTGGGTCGTTGAGCGTAGCTTTGGCTGGCTCAACCGCT
2461  TCAGGCGACTCGCCCGTGACTATGAACGCCTTCCAGAAACCCTTGCGGGTGTGCACTTTG
2521  TTGTCTTCGCCATGATCATGTTGGTTCATGCCGTGCCGATCATGCAAAGTGCCTAACACG
2581  CTCTAGAATGCATTGTGNTGGCCTATGCTTGGTCAGCACTNGCAAGAATGCAGAATCAGT

2641  AAAATCAGAATTATTTTGGGGCTGACGATGGCTGTGTATACCGACAACGCCCGAATTC
                                                        EcoRI
```

Figure 3.8

A. Complete nucleotide and deduced amino acid sequences of IS*61*. Inverted repeat sequences are shaded and the putative direct repeat sequences of TGG are in bold. AGAAAG sequence at the overlapping region of the first and second ORFs is underlined. The filled circle (•) at the beginning of the second ORF indicates the possible position of -1 translational frame-shift. The characteristic amino acid sequences of D+35E of IS*3* family members are in bold.

B. Alignment of amino acid sequences of putative transposases encoded by IS*61*, IS*3*-like (GenPep accession no. AAC25918), IS*1236* (GenPep accession no. AAC97541) and IS*904* (GenPep accession no. AAA25194). The asterisk (*) indicates the amino acid sequences which are identical. Numbering refers to the amino acid residues of the putative transposase of *IS61*.

C. The percentages of amino acid sequence identity and similarity of IS*61* to IS*3*-like, IS*1236* and IS*904*.

(A)

```
  1   TCTAGAACTAGTGGATCGGGGACACCTCGAAGATGCAGTGATCGGCCGCGACCACGCGCA

 61   GCAGGAAATCACAGTCGCCCGCCATCAGGTGGCACTCCACCACCTCCGGCAACTGCTGGA

121   TGGCCGCCGTGAAGGCATCCACCGTGGCCGCGTCCTGTCCCTTGAGCCAGATGCGCGCGA

181   ACACCGATAGCCCCTTGCCGACCTTGGCCGGATTGAGCACGGCGACATAGCGCTCGATCG

241   ACATCGGCATTGTTGCGACCACGGCCATGCACGGTGCGAACGCGCCGGTCGAGAAGACCG
                                  M  H  G  A  N  A  P  V  E  K  T  D

301   ACTGGTCGCCGCTGGCCGGCAAGGCCGTGCTGATCTCGGTGCCGCGTGCCGGCGTCCGTC
        W  S  P  L  A  G  K  A  V  L  I  S  V  P  R  A  G  V  R  P

361   CGCCTGAAACCGAACGTCACCCATGTCCGCCAATCCTTCCTACGTCGACGAGCAGACGCT
        P  E  T  E  R  H  P  C  P  P  I  L  P  T  S  T  S  R  R  S

421   CATGCGCCGGCTGCTGTCCGGCAAGACCATCGCCGTGGTCGGGCTGTCGCCGCGGCCGAC
        C  A  G  C  C  P  A  R  P  S  P  W  S  G  C  R  R  G  R  R

481   GCGGCCCAGCTACGACGTGGCGCGCTATCTGCAGCAGGCCGGCTACCGGATCGTTGAGAA
        G  P  A  T  T  W  R  A  I  C  S  R  P  A  T  G  S  L  R  K
                                                              *  E  •

541   AGCGCTGGCCCAGTTCGGTCGGCCGGATCAGGGTGCCAGTTCACCGGCCACACTTGGCAG
        R  W  P  S  S  V  G  R  I  R  V  P  V  H  R  P  H  L  A  E
      S  A  G  P  V  R  S  A  G  S  G  C  Q  F  T  G  H  T  W  Q

601   AGCTTCCTGCGCGAGCACAACCTGCTGTGTAGCATGAGCCGGCGCGGCAACTGCCACGAC
        L  P  A  R  A  Q  P  A  V  *
      S  F  L  R  E  H  N  L  L  C  S  M  S  R  R  G  N  C  H  D

661   AATGCCGTGGCTGAGAGCTTCTTCCAGTTGCTCAAGCGCGAGCGAGTGCGGCGGCAAATC
      N  A  V  A  E  S  F  F  Q  L  L  K  R  E  R  V  R  R  Q  I

721   TATGTCACCCGGCAGCAGGCCAAGTCCGATGTCTTCAACTACATCGAGATGTTCTACAAC
      Y  V  T  R  Q  Q  A  K  S  D  V  F  N  Y  I  E  M  F  Y  N

781   CCAACACGGCGACATTCGAGCGCCAACGGACTATCGCCGGTAGAGTTCGAACAACGCCAT
      P  T  R  R  H  S  S  A  N  G  L  S  P  V  E  F  E  Q  R  H

841   TCCCAACGGCTCGCGGGTGTCTAGAAAAACCGGGGCGATTCAACGAGGGCGGCGACGGCG
                          XbaI
      S  Q  R  L  A  G  V  *

901   TCGGTGGTGACGGCGTCGGTGGTGACGGCGACGGTAGTGGCGCTGCGGGTGGCACCGGTG

961   AAGGTGGCGCCAATGCAGGTGCAGCCGGCACAACGCCACCGCGACGAGATCCCCCGGGCT
```

(B)

```
IS61            ------------------------------------------------------------
IS3-like        ------------------------------------------------------------
IS1236          KSACVLSDRSCQEKYTVIQDLDVNEVTVSSACKCLGVSTSGYYAWRKRQTNLAQKYNDLK
IS904           ---------------------------------------MHRRPSKQQVEREILSEKIK

                                                      1                  20
IS61            ----------------------------------MHGANAPVEKTDWSPLAGKAVLIS
IS3-like        ----------------------MCCPNRVARLTRLAGIKAQIG-YKRRPGIYGGRPSV
IS1236          AVYWQHHARLGAPSLVHDMHDLGYSMSERTVGRMLKKLGLRSKIA-RKYKHTTDSNHRLP
IS904           AVFHEHKGRYGAVRITKVLHNTGIMTNTKRVGKLMHLMGLYAKGSRYKYKHYNRKG-ASL
                                                     *

                               40                  60
80
IS61            VPRAGVRPPETERHPCPPILPTSTSRRSCAGCCPARPSPWSGCRRGRRGPATTWR----A
IS3-like        VVDNKLDRQFDVAAPDKAWVTDITYIRTYEGFAYLAVVIDLYSRR-VIGWAVQSRQTTDV
IS1236          TAPNLLDRQFTVNEPNKIWTTDITYIRTKQGWLYLCVMLDLFSRR-IVGWQTSHRIDRQL
IS904           SRPNLINQIFKATAPNKVWLGDMTYIPTKEGTLYLAVNIDVFSRK-IVGWSMSSRMQDKL
                              *       *      *           *    *     *

                           •          100                 120
IS61            ICSRPATGSLRS---AGPVRSAGSGCQFTGHTWQSFLREHNLLCSMSRRGNCHDNAVAES
IS3-like        VLQALLMAVWRRKPKDKALIHSDQGSQFTSMDWASFLRHHNLVHSMSRRGNCHDNAVAES
IS1236          VCDAFHYAMARQGYPMGVMVHSDQGSQYCSRDFRALLLTNNCVQSMSRRGNCWDNAVTES
IS904           VRDCFLQACGKEHPQPGLIVHTDQGSQYTSSRYQSTLRQVGAQSSMSRKGNPYDNAMMES
                                       * *         *       **** **  ***  **

                 140                 160                 180
IS61            FFQLLKRERVRRQIYVTRQQAKSDVFNYIEMFYNPTRRHSSANGLSPVEFEQRHSQRLAG
IS3-like        FFNLLKRERIRRRVYRSRDEARQDVFDYIEMFYNPKRKHVRNGMLSPVEFEKQQKI----
IS1236          FFHTLKGHMVHGSVFATRKEANAVLFDYIEIYYNRIRRHSTNGWLSPEAFEQKYFKNLEG
IS904           FYKTLKRELINDAHFETRAEATQEIFKYIETYYNTKRMHSGLDYKSPKDFEKYNS-----
                *   **           *  *    * ***  **  * *      **  **

                198
IS61            V------
IS3-like        -------
IS1236          FVVHDTV
IS904           -------
```

(C)

		aa identity (%)	aa similarity (%)
IS*61*	IS*3*-like (AAC25918)	44	59
	IS*1236* (AAC97541)	30	49
	IS*904* (AAA25194)	27	45

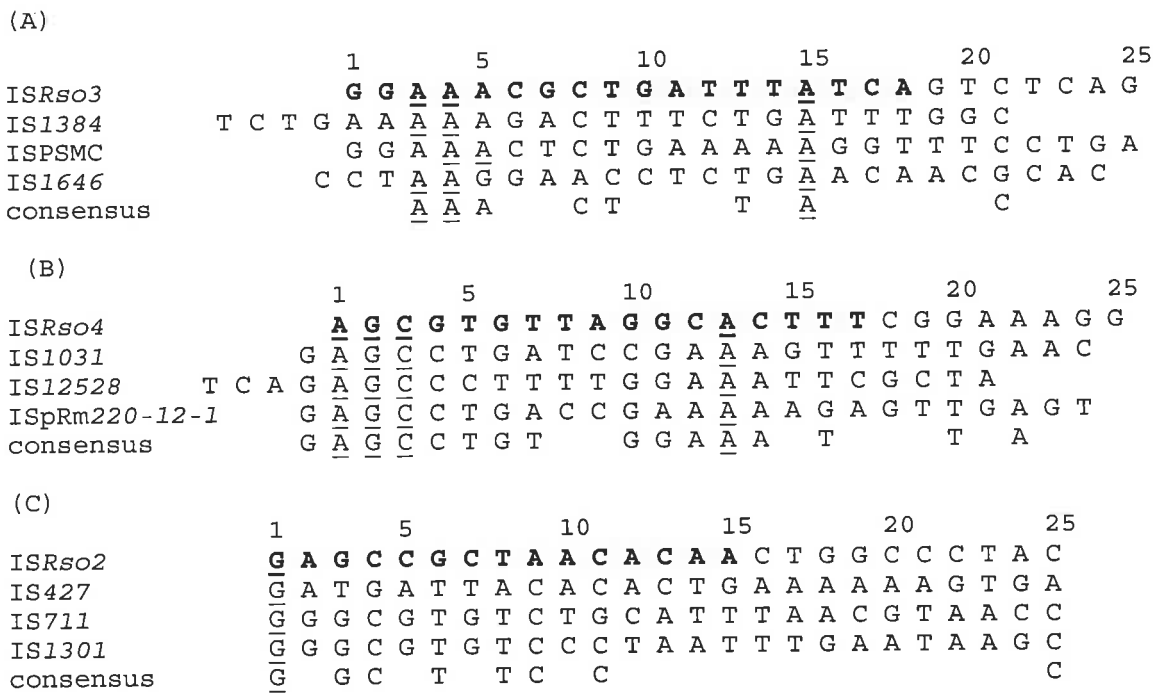


Figure 3.9

Nucleotide sequence comparisons of the 5' end regions containing inverted repeat sequences of IS*Rso3*, IS*Rso4* and IS*Rso2* with their homologs. Inverted repeat sequences of IS*Rso3*, IS*Rso4* and IS*Rso2* are in shade.

A. IS*Rso3* (GenBank accession no. AF183890) with IS*1384* (GenBank accession no. AF052751), ISPSMC (GenBank accession no. AB023075) and IS*1646* (GenBank accession no. AF077016).

B. IS*Rso4* (GenBank accession no. AF079849) with IS*1031* (GenBank accession no. INS1031A), IS*12528* (GenBank accession no. D86631) and ISpRm*220-12-1* (GenBank accession no. AF126536).

C. IS*Rso2* (GenBank accession no. AF186082) with IS*427* (GenBank accession no. ATUIS427), IS*711* (GenBank accession no. AF043474) and IS*1301* (GenBank accession no. NMISORFS).

IS elements	clones	DR sequences (left to right)
IS*Rso3*	pSV102	CTAG + CTAG
IS*Rso3*	pSC15K	CTAG + TTAA
IS*Rso3*	pSC15J	CTAG + CTAG
IS*Rso3*	pSC15S	CTAG + TTAG
IS*Rso3*	pSC15H	CTAG + TTAG
IS*Rso3*	pSC15R	CTTG + CTAG
IS*Rso3*	pSC15E	TTAG + ATAG
IS*Rso4*	pSV102	CTAG + CTAG
IS*Rso4*	pIS102JI	CTAG + CTAG
IS*Rso4*	pM81C	CTGAG + CTGAG
IS*Rso2*	pISB	CTAA + n.a.
IS*Rso2*	pISBE	CTAA + TTAA

Table 3.1

Nucleotide sequence comparisons of direct repeats of IS*Rso3*, IS*Rso4* and IS*Rso2* of several clones.

CHAPTER 4

IS*Rso4* and phenotype conversion in *Ralstonia solanacearum*

4.1 Introduction

Insertion sequence elements mediate gene activation and inactivation (Galas and Chandler, 1989) and promote genomic DNA rearrangements by deletions and insertions. Several IS elements have been isolated because of their involvement in virulence and pathogenicity-related gene regulation. Ziebuhr *et al.* (1999) described a molecular mechanism that accounted for a polysaccharide intercellular adhesin (PIA)-negative phase variation in *Staphylococcus epidermidis*. Synthesis of the PIA in *S. epidermidis* is initiated by *ica* operon activation and the PIA synthesised assists biofilm formation that contributes to the virulence of the species. IS*256* was isolated in PIA-negative phase variants, as IS*256* insertions inactivated either *icaA* or *icaC* genes (Ziebuhr *et al.*, 1999). More importantly, IS*256* insertion into the *ica* operon was a reversible event and restored the biofilm-forming phenotype of PIA-negative mutant. Similarly, IS*1301* in *Neisseria meningitidis* was isolated as the element was involved in a spontaneous mutation in which an encapsulated wild-type bacterium became capsule-negative (Hammerschmidt *et al.*, 1996). Genetic analysis of a capsule-negative mutant confirmed that regulation of *siaA* (a gene involved in sialic acid biosynthesis) by insertion/excision of IS*1301* was responsible for capsule expression which was in turn related to virulence in the species (Hammerschmidt *et al.*, 1996). Phenotype conversion (PC) is a well-known phase variation related to the virulence of *R. solanacearum* (Brumley and Denny, 1990). It was therefore considered necessary to investigate the possible involvement of IS elements isolated from strain ACH0158 in PC.

Phenotype conversion (PC) (Brumley and Denny, 1990) is a spontaneous phenomenon in which a wild-type *R. solanacearum* concomitantly loses the ability to produce extracellular polysaccharide (EPS) and a variety of secreted proteins and extracellular enzymes, resulting in a much less virulent mutant. Mutations in *phcA* were implicated

in PC (Brumbley and Denny, 1990). This gene encodes a transcriptional regulator that controls several genes involved in synthesis and regulation of virulence factors of *R. solanacearum* (Schell, 1996). Reversion of PC events in *R. solanacearum* has yet to be reliably documented in contrast to several reversible phase variations in diverse bacterial species (Brumbley and Denny, 1990). This chapter describes the isolation and characterisation of a PC-type ACH0158-M81C caused by IS*Rso4* insertion into *phcA* and it also discusses reversion of PC phenomena.

4.2 Isolation of a PC-type ACH0158-M81C involving transposition of IS*Rso4*

ACH0158 undergoes phenotype conversion at a very high rate. Nine spontaneous PC-types of strain ACH0158 were isolated by plating a small amount of a liquid stock of bacterial culture and by selecting colonies showing typical PC-type characteristics. Genomic DNA of the nine PC-types was prepared and *Eco*RI-digested. Southern blots of these samples were probed with the IS*Rso4*, IS*Rso3* and IS*Rso2*-specific probes prepared as described in the chapter 5. No obvious difference was observed in Southern hybridisations using the IS*Rso3* and IS*Rso2* probes (Fig. 4.1B and C). However, the Southern blot probed with the IS*Rso4* revealed a PC-type (ACH0158-M8) showing an additional band of 2.1 kb (Fig. 4.2A-M8) compared with the six fragments in wild-type (Fig. 4.2A-WT) and the other PC-types (Fig. 4.2A). The amount of IS*Rso4* hybridisation to the 2.1 kb fragment however was lower than that to the other six fragments in the PC-type ACH0158-M8. Two explanations for the observation were possible. One was based on the possibility of late transposition of IS*Rso4* during short time liquid culture of the PC-type. In this case, IS*Rso4* would not be a cause of the PC-type. However event would be important to investigate because it reflects a known transposition of IS*Rso4*. Alternatively, two PC-types caused by two different mutations were accidentally inoculated into the liquid culture for DNA preparation. An involvement of IS*Rso4* in a PC event could not be discounted if the latter was the reason for the unexpected observation. Therefore further characterisation of the PC-type ACH0158-M8 was clearly necessary.

A small amount of the liquid culture of the PC-type ACH0158-M8 was re-plated and 11 well-isolated colonies showing PC-type characteristics were selected. Purified genomic DNA of the 11 ACH0158-M8 derivatives was *Eco*RI-digested, Southern blotted and probed with the IS*Rso4*. The Southern analysis with the IS*Rso4* probe clearly showed presence of three different hybridisation patterns (Fig. 4.2B). Five PC-types including a PC-type ACH0158-M82 showed the identical hybridisation pattern to

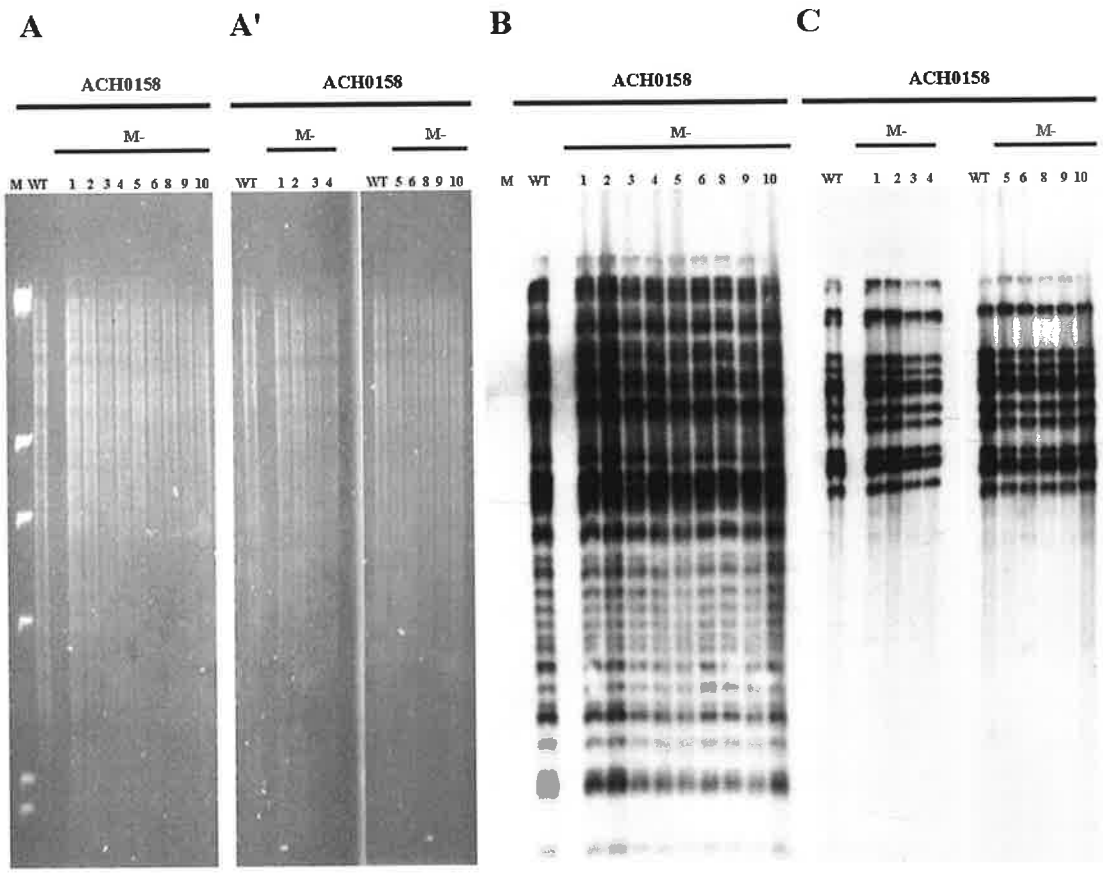


Figure 4.1

A and A'. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of wild-type and PC-types of ACH0158. λ bacteriophage DNA digested with *Hin*dIII with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was included as a size marker (M).
WT; wild-type ACH0158, M1; PC-type ACH0158-M1, M2; PC-type ACH0158-M2, M3; PC-type ACH0158-M3, M4; PC-type ACH0158-M4, M5; PC-type ACH0158-M5, M6; PC-type ACH0158-M6, M8; PC-type ACH0158-M8, M9; PC-type ACH0158-M9 and M10; PC-type ACH0158-M10.

B. Southern hybridisation of the gel in A with IS*Rso3* probe

C. Southern hybridisation of the gel in A' with IS*Rso2* probe

Figure 4.2

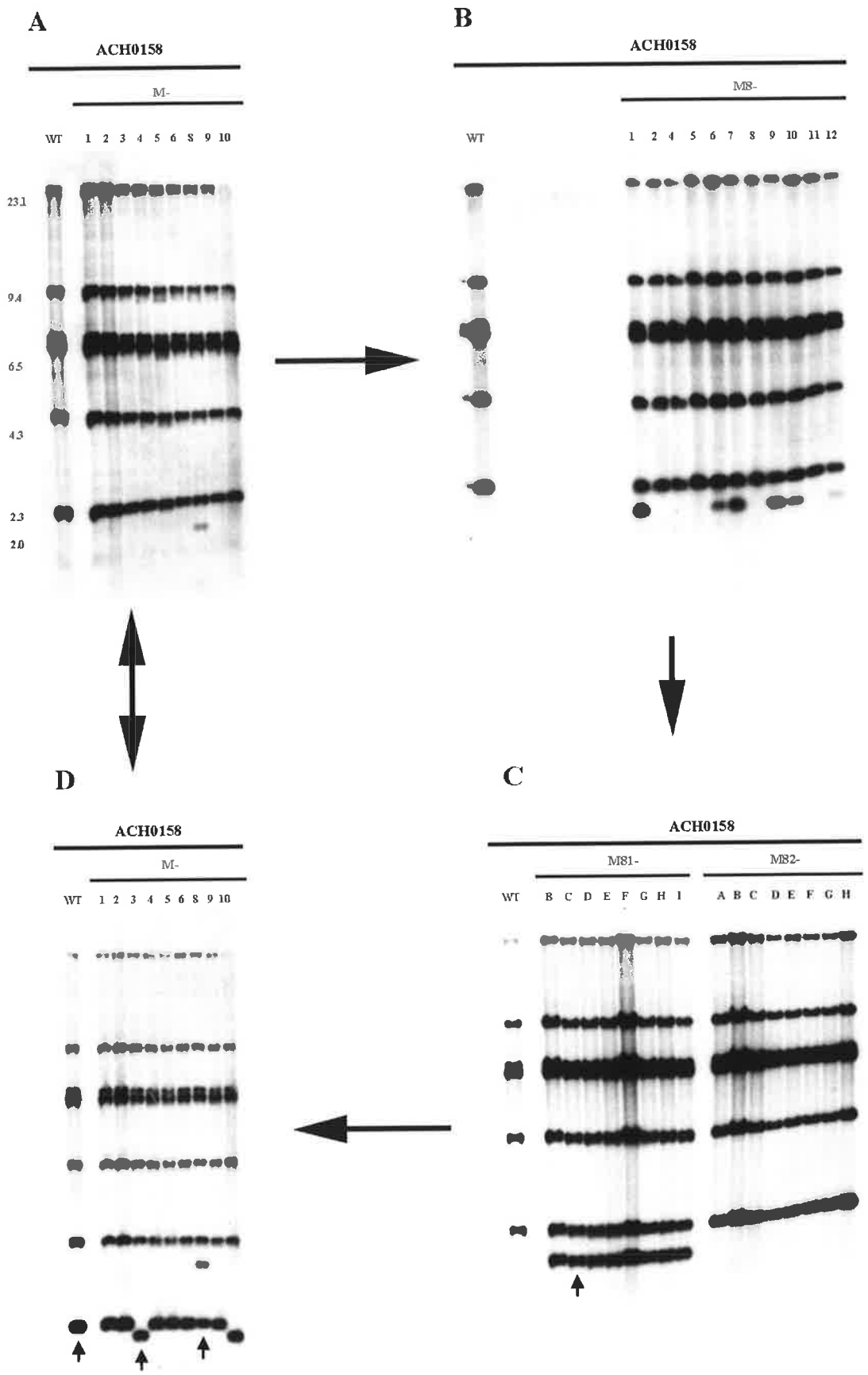
A. Southern hybridisations of *Eco*RI-digested genomic DNA of wild-type and PC-types of ACH0158 with IS*Rso4* probe. WT; wild-type ACH0158, M1; PC-type ACH0158-M1, M2; PC-type ACH0158-M2, M3; PC-type ACH0158-M3, M4; PC-type ACH0158-M4, M5; PC-type ACH0158-M5, M6; PC-type ACH0158-M6, M8; PC-type ACH0158-M8, M9; PC-type ACH0158-M9 and M10; PC-type ACH0158-M10.

B. Southern hybridisations of *Eco*RI-digested genomic DNA of wild-type and 11 ACH0158-M8 derivatives with IS*Rso4* probe. WT; wild-type ACH0158, M81; PC-type ACH0158-M81, M82; PC-type ACH0158-M82, M84; PC-type ACH0158-M84, M85; PC-type ACH0158-M85, M86; PC-type ACH0158-M86, M87; PC-type ACH0158-M87, M88; PC-type ACH0158-M88, M89; PC-type ACH0158-M89, M810; PC-type ACH0158-M810, M811; PC-type ACH0158-M811 and M812; PC-type ACH0158-M812.

C. Southern hybridisations of *Eco*RI-digested genomic DNA of wild-type and 8 ACH0158-M81 derivatives and 8 ACH0158-M82 derivatives with IS*Rso4* probe. WT; wild-type ACH0158, M81B; PC-type ACH0158-M81B, M81C; PC-type ACH0158-M81C, M81D; PC-type ACH0158-M81D, M81E; PC-type ACH0158-M81E, M81F; PC-type ACH0158-M81F, M81G; PC-type ACH0158-M81G, M81H; PC-type ACH0158-M81H, M81I; PC-type ACH0158-M81I, M82A; PC-type ACH0158-M82A, M82B; PC-type ACH0158-M82B, M82C; PC-type ACH0158-M82C, M82D; PC-type ACH0158-M82D, M82E; PC-type ACH0158-M82E, M82F; PC-type ACH0158-M82F, M82G; PC-type ACH0158-M82G and M82H; PC-type ACH0158-M82H.

D. Southern hybridisation of *Eco*RI-digested genomic DNA of wild-type and PC-types of ACH0158 with IS*Rso4* and *phcA* probes. WT; wild-type ACH0158, M1; PC-type ACH0158-M1, M2; PC-type ACH0158-M2, M3; PC-type ACH0158-M3, M4; PC-type ACH0158-M4, M5; PC-type ACH0158-M5, M6; PC-type ACH0158-M6, M8; PC-type ACH0158-M8, M9; PC-type ACH0158-M9 and M10; PC-type ACH0158-M10.

Small arrows indicate *Eco*RI fragments cloned and characterised.


A
ACH0158
M-
WT 1 2 3 4 5 6 8 9 10
23.1
9.4
6.5
4.3
2.3
2.0
B
ACH0158
M8-
WT
1 2 4 5 6 7 8 9 10 11 12
C
ACH0158
M81-
M82-
WT B C D E F G H I
A B C D E F G H
D
ACH0158
M-
WT 1 2 3 4 5 6 8 9 10

the wild-type strain ACH0158 that has six hybridising fragments (Fig. 4.2B). A second group of three PC-types including a PC-type ACH0158-M81 showed seven hybridised fragments. These all included the extra 2.1 kb fragment (Fig. 4.2B). The remaining PC-types of ACH0158-M86, ACH0158-M810 and ACH0158-M812 showed a faintly hybridising 2.1 kb fragment similar to the pattern originally observed in ACH0158-M8. In a third experiment, eight derivatives of the PC-type ACH0158-M81 and the other eight of the ACH0158-M82 were isolated and a Southern blot containing their *Eco*RI-digested genomic DNA was probed with the IS*Rso4* (Fig. 4.2C). As expected, DNA from the eight PC-types derived from the PC-type ACH0158-M81 showed identical hybridisation pattern of 7 fragments including a prominent 2.1 kb fragment when hybridised with the probe. The result therefore confirmed presence of a specific PC-type in the liquid culture of the PC-type ACH0158-M81 (Fig. 4.2B). Southern hybridisation of the other eight ACH0158-M82 derivatives demonstrated identical hybridisation pattern of 6 *Eco*RI fragments. However, it was still unknown what caused the faintly hybridised 2.1 kb fragment in ACH0158-M8 (Fig. 4.2A).

The seven hybridising fragments in the genome of the PC-type ACH0158-M81 clearly showed that the copy number of IS*Rso4* was increased and a possibly new copy of IS*Rso4* inserted into an *Eco*RI fragment of approximately 1.25 kb, resulting in the 2.1 kb *Eco*RI fragment hybridised with the IS*Rso4* probe. Therefore, cloning of the 2.1 kb *Eco*RI fragment from the genomic DNA of the PC-type ACH0158-M81C was initiated to define genetic content of the IS element insertion.

4.3 Cloning of the genomic region flanking IS*Rso4* in ACH0158-M81C

Cloning of the genomic region in the PC-type ACH0158-M81C into which IS*Rso4* was inserted utilised an inverse-PCR approach (Triglia *et al.*, 1988). The inverse-PCR method required nested primers within IS*Rso4* to amplify flanking but unknown sequence of the element. Details of the cloning strategy are illustrated in Figure 4.3. Genomic DNA of ACH0158-M81C was *Eco*RI-digested and a self-ligation reaction of the digested genomic DNA was set up after heat-inactivation of *Eco*RI at 65 °C for 15 minutes to prevent any re-digestion after ligation. Purification of the ligated molecules prior to inverse PCR, was performed either by phenol/chloroform and chloroform extractions or by Sepharose CL-6B column followed by ethanol precipitation (Fig. 4.4A).

Initial PCR reactions were performed with PSC1 and PSC2 primers derived from IS*Rso4* sequence (Fig. 3.5A). Purified template DNA of self-ligated *Eco*RI genomic fragments of the ACH0158-M81C was denatured prior to the PCR reactions. Thermal conditions for the PCR reactions were a single stage of 95 °C for 1 minute for initial denaturation of the template DNA, followed by 94 °C for 30 seconds, 45 °C for 60 seconds and 72 °C for 90 seconds repeated 30 times for amplification of target fragment. A final step at 72 °C for 7 minutes was included for final polymerase extension. Reaction mixtures in total 25 µl contained 0.4 mM each of dATP, dCTP, dGTP and dTTP, 1 x PCR reaction buffer (Bresatec, Adelaide), 1.5 mM $MgCl_2$, 0.8 mM of each primer (PSC1 and PSC2), and 1 U of *Taq* DNA polymerase (Bresatec, Adelaide). The PCR-amplified products were loaded onto an 1.2 % (w/v) agarose gel and the 2.1 kb expected fragment was clearly visible (Fig. 4.4A). A small amount of the PCR-amplified products was utilised as a source to provide template DNA for subsequent PCR reactions with primers IS102RR and Ja nested within IS*Rso4* (Fig. 3.5A). The expected 1.6 kb fragment was PCR-amplified with the nested primers (Fig. 4.4B), demonstrating that the initial 2.1 kb fragment contained IS*Rso4* and its flanking regions. Consequently, the 1.6 kb fragment was cloned into pGEM®-T Easy vector and the clone was designated pM81C. The 1.6 kb insert in clone pM81C was sequenced on both DNA strands and an open reading frame (ORF) of 347 aa was detected. A standard protein database search with the amino acid sequence derived from the flanking sequence of IS*Rso4* closely matched (99 % identical) to PhcA sequences (Brumbley *et al.*, 1993). PhcA is the product of *phcA* originally isolated from strain AW1 of *R. solanacearum* (Brumbley and Denny, 1990). Therefore, the gene containing the IS*Rso4* insert in the clone pM81C was designated as *phcA81C*.

The *phcA* gene from *R. solanacearum* strain AW1 was primarily isolated as a locus responsible for phenotype conversion (Brumbley and Denny, 1990) and the protein product was subsequently determined to be a global regulator of several genes involved in virulence of *R. solanacearum* (Schell, 1996). PhcA protein is a member of the LysR family of transcriptional regulators containing a characteristic helix-turn-helix (HTH) DNA binding motif toward the *N*-terminal of the protein (Brumbley *et al.*, 1993). Functionally active PhcA protein controls, either directly or indirectly, expression of several genes encoding and regulating virulence factors (Schell, 1996). Expression of *phcA* is regulated by volatile extracellular signalling molecules (Clough *et al.*, 1994, 1995), the levels of which reflect bacterial numbers in an infected plant. Brumbley *et al.*

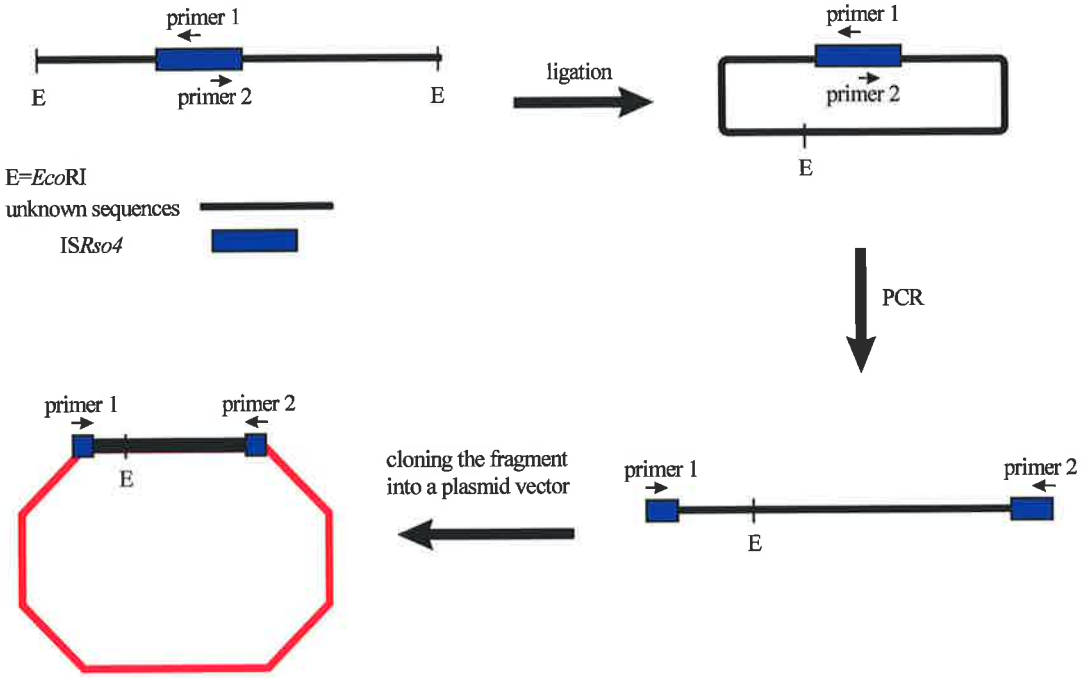


Figure 4.3

A schematic diagram illustrating the inverse-PCR strategies.

Figure 4.4

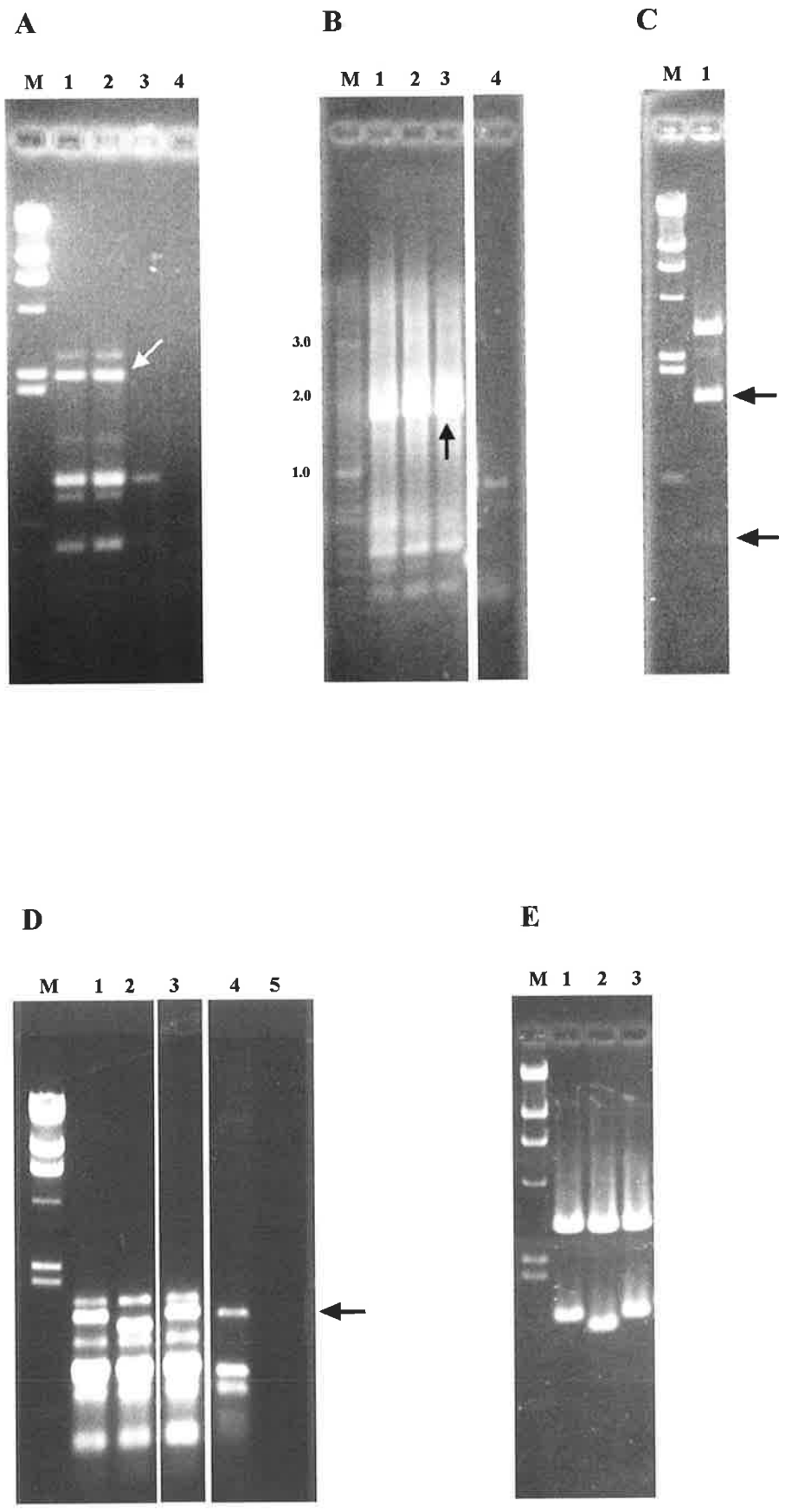
A. Agarose gel electrophoresis of initially amplified inverse-PCR products with PSC1 and PSC2 primers. λ bacteriophage DNA digested with *Hin*dIII with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was included as a size marker (M). 1. PCR products of ligates purified by phenol/chloroform and chloroform extractions followed by ethanol precipitation. 2. PCR products of ligates purified by direct precipitation with 2 volumes of ethanol. 3. PCR products of ligates purified by Sepharose CL-6B column followed by the ethanol. 4. Negative control with no DNA added to the PCR reaction. The arrow indicates the expected fragment.

B. Agarose gel electrophoresis of amplified PCR products with IS102RR and Ja primers using initial PCR-products of Fig. 4.4A-lane 1 as template DNA. M. Molecular size marker of 100 bp ladder with 1 kb, 2 kb amd 3 kb fragments indicated. 1. PCR products using 1μl of Fig. 4.4A-lane 1 reaction. 2. PCR products using 1μl of 1/10 dilution of Fig. 4.4A-lane 1 reaction. 3. PCR products using 1μl of 1/100 dilution of Fig 4.4A-lane 1 reaction. 4. Negative control with no DNA added to the PCR reaction. The arrow indicates the expected fragment.

C. Agarose gel electrophoresis of *Eco*RI-digested pM81C clone. M. Molecular size marker of λDNA restricted with *Hind*III. 1. *Eco*RI-digested pM81C. The arrows indicate two *Eco*RI fragments.

D. Agarose gel electrophoresis of PCR products with PPHC3 and PHC4 primers using 100 ng of genomic DNAs from wild-type and PC-types of ACH0158. M. Molecular size marker of λDNA restricted with *Hind*III. 1. PCR products using 100 ng of wild-type ACH0158 genomic DNA. 2. PCR products using 100 ng of a PC-type ACH0158-M3 genomic DNA. 3. PCR products using 100 ng of a PC-type ACH0158-M8 genomic DNA. 4. PCR products using 100 ng of the AW1 genomic DNA. 5. Negative control with no DNA added to the PCR reaction. The arrow indicates the expected fragment.

E. Agarose gel electrophoresis of *Eco*RI-digested pWT, pM3 and pM8 clones. M. Molecular size marker of λDNA restricted with *Hind*III. 1. *Eco*RI-digested pWT. 2. *Eco*RI-digested pM3. 3. *Eco*RI-digested pM8.


A
M 1 2 3 4
B
M 1 2 3 4
3.0
2.0
1.0
C
M 1
D
M 1 2 3 4 5
E
M 1 2 3

(1993) isolated a spontaneous PC-type AW1-PC containing a 2 bp insertion in the ORF of *phcA1* and two other PC-types appeared to contain larger insertions in the *phcA* gene (Brumbley *et al.*, 1993).

Cloning of *phcA* from two other PC-types that showed weak hybridisation of IS*Rso4* to a 2.1 kb fragment was undertaken to determine whether the element was inserted into the same gene. The two independent PC-types ACH0158-M810 and ACH0158-M812G were selected and *phcA810* and *phcA812G* were cloned by the inverse-PCR method described previously. Nucleotide sequence comparisons between *phcA81C*, *phcA810* and *phcA812G* verified that IS*Rso4* was inserted into the same genomic location within the *phcA* gene and flanked by CTGAG DRs. Only *phcA81C* was extensively characterised to determine the consequence of IS*Rso4* insertion in ACH0158-M81C. The integration of a new copy of IS*Rso4* into its new genomic position within *phcA81C* created 5 bp CTGAG DR (Fig. 4.6A). The DR sequence was different from the 4 bp CTAG DR found in pSV102 and in pIS102JI (discussed in chapter 3 and chapter 6, respectively) (Table 3.1). IS*Rso4* was identified previously as a member of the IS*5* family whose members have target sequence specificities with conserved sequences (Iida *et al*, 1983). Transpositional frequency and efficiency of IS*Rso4* may depend on its tolerance of target sequence and length variations. Target duplication sequence and length variation at the point of IS element insertion are not unusual (Galas and Chandler, 1989). The variation created by IS*Rso4* in the *phcA81C* may be influenced by different helix conformation of nucleotide sequences of the *phcA* gene during the initial cleavage. Other unknown factors which are not related to nucleotide sequence confirmation may also affect the DR variations of IS*Rso4*, like those shown in IS*30* (Olasz *et al.*, 1998). However, additional length and sequence variations of the DR motif between different genomic copies of IS*Rso4* cannot be ruled out, as we have isolated and characterised only three out of seven insertion sites of the element. Nevertheless, three copies of IS*Rso4* in three different genomic clones seemed to create the obvious target sequence duplications that may be important for the excision of the element. Gerischer *et al.* (1996) suggested that creation of obvious target site duplications might be crucial for the excision of IS elements, as they observed that two spontaneous mutants caused by IS*1236* insertion with the absence of apparent DR motif in *Acinetobacter calcoaceticus* failed to become revertants.

In *phcA81C*, IS*Rso4* was inserted 2 bp upstream of the ribosome binding site (RBS) motif and 82 bp downstream of the -10 site (Fig. 4.5 and 4.6B). The direction of

transcription for the IS*Rso4* transposase gene was in the opposite direction to that of the PhcA gene. This result suggests that the IS*Rso4* insertion was the likely cause of the *phcA* gene inactivation. A small part of the new copy of IS*Rso4* was cloned in *phcA81C* and sequenced. Nucleotide sequence identity between the two different copies of IS*Rso4* in the clones pSV102 and pM81C is not known, although approximately 200 nucleotides of IS*Rso4* obtained from *phcA81C* were 100 % identical to the corresponding region of IS*Rso4* found in pSV102.

The appearance of the new copy of IS*Rso4* in the *phcA81C* suggested that the element was duplicated from one of the six genomic locations resulting in a consequent increase in IS*Rso4* number in ACH0158-M81C. Therefore IS*Rso4* transposition in PC-types might utilise a replicative transposition mechanism (Sherratt, 1989). Replicative transposition of IS elements involves complex processes including cointegrate formation, IS element duplication and resolution of the cointegrate formation between the two copies of the element (Sherratt, 1989). The sequences of the IS*Rso4* analysed in chapter 3 determined that IS*Rso4* encodes a transposase. However, it is not certain whether IS*Rso4* encodes a putative resolvase (Sherratt, 1989). Both transposase and resolvase are required for the replicative transposition mechanism of IS elements.

4.4 Cloning of *phcA* from wild-type ACH0158 and from several PC-types

The reason that the 2.1 kb fragment hybridised faintly with the IS*Rso4* probe in the PC-type ACH0158-M8 was still to be determined. Cloning of *phcA* from wild-type ACH0158 was initiated to compare nucleotide sequences of the *phcA* and *phcA81C*. In addition *phcA8* from the culture of the ACH0158-M8 was cloned to differentiate between the two possible origins proposed previously. Genomic DNA of a PC-type ACH0158-M3 was also used to clone *phcA3* to clarify the nature of a putative deletion identified by the Southern analysis (Fig. 4.2D).

Cloning of *phcA* variants was initiated by designing a primer pair PPHC3 and PHC4 (Fig. 4.5A). PPHC3 was designed on the basis of the *phcA* sequence of the strain AW1 (Brumbley *et al.*, 1993). PHC4 primer was a degenerate primer based on *phcA* sequences (Brumbley *et al.*, 1993) and *phcA81C* sequences (Fig. 4.5A). The nucleotide sequence for PHC4 primer is TCGACGGCGARATCGAGC, where R refers to alternative A or G nucleotides which were polymorphic between *phcA* from the AW1 and *phcA81C* from ACH0158-M8. PCR reactions were performed using 100 ng of genomic DNA of wild-type ACH0158 and the PC-types of ACH0158-M8 and

ACH0158-M3. In addition, *phcA* was PCR-amplified from the AW1 as a positive control for the PCR reactions. PCR products were resolved on 1 % (w/v) agarose gels and both 1.5 kb and 1.3 kb fragments were detected (Fig. 4.4D). The 1.5 kb fragment was amplified from wild-type ACH0158 (Fig. 4.4D-lane 1), the PC-type ACH0158-M8 (Fig. 4.4D-lane 3) and wild-type AW1 (Fig. 4.4D-lane 4) while the 1.3 kb fragment was amplified only from the PC-type ACH0158-M3 (Fig. 4.4D-lane 2). These DNA fragments were recovered from the gels and cloned into the pGEM®-T Easy vector. A plasmid clone containing *phcA* from wild-type ACH0158 was designated as pWT. Likewise pM8 and pM3 plasmid clones accommodated the *phcA8* and *phcA3*, respectively. The DNA inserts were sequenced in one direction.

Derived amino acid sequence alignments between the two *phcA* genes from the strains AW1 and ACH0158 revealed 99 % identity (Fig. 4.5C). The sequence of the *phcA* isolated from the ACH0158 contained the helix-turn-helix (HTH) motif characteristic of this class of transcriptional regulators (Brumbley *et al.,* 1993). A polymorphic nucleotide sequence between two *phcA* genes was related to an *Eco*RI site (GAATTC) adjacent to the -35 site of the *phcA* from the ACH0158 (Fig. 4.5B). This GAATTC sequence was replaced by GAATTT in the *phcA* gene of the AW1 (Fig. 4.5B). The *Eco*RI site was present in *phcA3* and *phcA8* variants isolated from the PC-types ACH0158-M3 and ACH0158-M8, respectively. This *Eco*RI site was utilised to generate the 1.3 kb DNA fragment containing the full-length *phcA* from ACH0158 while the *phcA* from the strain AW1 was present in a 4.0 kb *Eco*RI fragment in the clone pGA91 (Brumbley and Denny, 1990).

Nucleotide sequence alignment between *phcA* from wild-type ACH0158 and *phcA81C* revealed no difference with an exception of the IS*Rso4* insertion and resulting CTGAG duplications (Fig. 4.6A). This result strongly suggested that the PC-type ACH0158-M81C was caused by IS*Rso4* insertion. Additionally, the *phcA8* and *phcA* sequence alignment identified that a 2 bp (TG) insertion had occurred in *phcA8* (Fig. 4.7A). The TG insertion was in a short run of TG doublets in the coding region of PhcA (Fig. 4.7A) and the insertion was not present in the *phcA81C*. Therefore, it was concluded that *phcA8* and *phcA81C* were isolated from two independent PC events that were present in the mixed culture used to prepare DNA for ACH0158-M8. In addition to *phcA8*, *phcA82* from ACH0158-M82 was cloned and an identical 2 bp insertion was confirmed. The 2 bp insertion in the ORF of *phcA8* would cause a frame shift, resulting in an early stop codon and a truncated protein (Fig. 4.7B). A similar mutation has previously been

described in *phcA1* from a spontaneous PC mutant of AW1-PC (Brumbley *et al.*, 1993). This *phcA1* gene contained a 2 bp GC insertion occurred as part of a GCGC doublet adjacent to the *C*-terminal of *phcA1* (Brumbley *et al.*, 1993). This kind of insertional mutation may be a result of slipped-strand mispairing, which typically occurs in sequences with repeated bases or base doublets (Roth, 1974; Streisinger and Owen, 1985).

phcA3 gene isolated from the PC-type ACH0158-M3 had a 132 bp deletion identified by nucleotide sequence alignment with the *phcA* from the ACH0158 (Fig. 4.8). This region contains the -10 site, the RBS and the start codon (Fig. 4.8).Therefore, a range of different mutations including a small insertion, IS element insertion and a large deletion caused phenotype conversion in the mutants of *R. solanacearum* analysed in these experiments.

4.5 Summary and discussion

In summary, the results obtained here indicated that in one out of three cases fully analysed IS*Rso4* insertion into *phcA81C* caused PC, a characteristic phase variation in *R. solanacearum*. In addition to the IS element insertion, two other mutations at the *phcA* locus were detected. A reversible mechanism of active phase variation related to virulence of *R. solanacearum* has not yet been reported. However, the insertional mutations in the *phcA* gene described here may be reversible. Revertants may be caused either by precise excision of the insertion sequences or removal of the 2 bp doublets in the *phcA8* allowing the resumption of the ability to produce a large amount of extracellular polysaccharide (EPS) and virulence. However, extensive plating of the ACH0158-M81C could not identify any revertants that produced large amounts of EPS. The culture of the ACH0158-M81C was also inoculated onto tomato seedlings and no revertant was recovered. Therefore, *in vitro* and *in vivo* experiments to obtain revertants of the IS insertional mutation in this work were unsuccessful.

There are many examples of active phase variations characterised in several pathogenic bacterial species. More importantly, typical examples of phase variations have been shown underlying genomic rearrangements, transpositions or insertions of IS element at detectable frequencies. Capsule (a virulence factor) expression regulation in *N. meningitidis* was based on activation/inactivation of the *siaA* gene by IS*1301* transposition (Hammerschmidt *et al.*, 1996). The *siaA* gene is an essential gene

involved in sialic acid biosynthesis and its inactivation by the IS element was reversible at a frequency of approximately 0.0025 % (Hammerschmidt *et al.*, 1996). Ziebuhr *et al.* (1999) reported a phase variation in *Staphylococcus epidermidis*. In the species, polysaccharide intercellular adhesin (PIA)-negative phase variants do not induce a biofilm production and approximately 30 % of the PIA-negative phase variants analysed were found to be due to the inactivation of either *icaA* or *icaC* genes by insertion of IS*256* (Ziebuhr *et al.*, 1999).

Figure 4.5

A. Complete nucleotide and deduced amino acid sequence of *phcA* isolated from wild-type ACH0158. The -35, -10 and RBS sites are shaded. Nucleotide sequences of *phcA*-derived primers for sequencing and PCR experiments are underlined. The identities and the directions of the primers are also indicated.

B. Nucleotide sequence comparison of the promoter region of *phcA* gene from strains ACH0158 (upper line) and AW1 (lower line). The asterisk (*) indicates polymorphic nucleotide sequences. *Eco*RI site (GAATTC) from ACH0158 and polymorphic GAATTT from AW1 are shaded. The initial part of the coding region for PhcA protein is underlined.

C. Alignment of deduced amino acid sequence of *phcA* from ACH0158 (upper line) and AW1 (lower line) strains. The mark (|) indicates the identical sequence. The asterisk (*) indicates polymorphic amino acid sequences. Amino acid sequences for the helix-turn-helix (HTH) motif are in bold.

(A)

```
1     CAATTCAACACGCCGCCTTTTGTTATGCACTAAAACGAAAACGTTTGCGCATAAACAAAT
        PPHC3  →

61    TTCGCAGCGCTGGCGCTGGAATTCGCTTTGAAAATCGCCTGGCAATCCGTACACTGAACA
                              -35                    -10
                              PPHC  →

                        ......
                        EcoRI
121   ATTACTACATTTGTGACGCAGCCGCCGTTTCGTGCTGCGTCGAGGCCGTTGCCTCGGCAA

181   CGTCTGCCTTTTCACTTTCTGTCTGAGCCAAGGCGTCCCATGGTCAACGTCGATACCAAG
                              RBS                PHC3  →
                                                  M  V  N  V  D  T  K

241   CTGTTGGTGATTTTTGTGGAGTTGCTCAGTAAGCGGAACGCTACCTATGTGGCGGAGAAG
      L  L  V  I  F  V  E  L  L  S  K  R  N  A  T  Y  V  A  E  K

301   ATGCACATGACAGCGCCTGCGGTATCGCATTCGCTGGGTCGCCTGCGCGAAATCTTCGAC
      M  H  M  T  A  P  A  V  S  H  S  L  G  R  L  R  E  I  F  D

361   GATCCCTTGTTCATCCGTGTGCCGCACGGTTTGACGCCAACGCCCAAGGCGCTCGAACTC
      D  P  L  F  I  R  V  P  H  G  L  T  P  T  P  K  A  L  E  L

421   GGTCCCAAAGTGCGCGAGATGCTCGACCTGTGGGCGGCGATCAACGAGGGCGATATCGCC
      G  P  K  V  R  E  M  L  D  L  W  A  A  I  N  E  G  D  I  A

481   ACGTTCGATCCCGCGGAAGCCGCCGGCACCTTCAACGTCAGCTTCGCCGGTACGTTGGGC
                                    ←  PHC5
      T  F  D  P  A  E  A  A  G  T  F  N  V  S  F  A  G  T  L  G

541   GATGCCCTGTTCGACCGTTTCCTGCTGCGCGTCAAGCGCCTGGCGCCGGGCCTGCAGGTG
                  PHC1  →
      D  A  L  F  D  R  F  L  L  R  V  K  R  L  A  P  G  L  Q  V

601   CGCCTGACCGAGTCCTCCTCGTGGGAAGCCGATGTGGCGGCGATGCGTTCCAACGAGTTG
      R  L  T  E  S  S  S  W  E  A  D  V  A  A  M  R  S  N  E  L

661   GACCTGGCGTTCTCGCCGTTCCCGACGCGGCATCCGGAAATCGTGGAGGAGGTCGTCACC
      D  L  A  F  S  P  F  P  T  R  H  P  E  I  V  E  E  V  V  T

721   TCCTTCAACATGTGGGTCTGCGCGCGCAAGGACCATCCGGTTCTGAAGGACGGTTGCTCG
      S  F  N  M  W  V  C  A  R  K  D  H  P  V  L  K  D  G  C  S

781   CTCGATCAGTACCTTGAGTGCGAGCATATCTTCATCGCGCAGGGCAATCCCGGCAGCCGT
      L  D  Q  Y  L  E  C  E  H  I  F  I  A  Q  G  N  P  G  S  R

841   GCGGCACCGTCGCTGATTCCGCTCGATTACGCGCTGCAGCAGCGCGGCCTGAAACGCCAC
      A  A  P  S  L  I  P  L  D  Y  A  L  Q  Q  R  G  L  K  R  H

901   TCCACCATGACCGTGCACGCGTGGCGCACCCAGGCCGAAGTGGCGGCGCAGACGGATTTG
      S  T  M  T  V  H  A  W  R  T  Q  A  E  V  A  A  Q  T  D  L

961   ATCTTCACGGTCAATTCGCTGATGAAGGATCTGGTGTGCGAGGCCTACAGCCTCAACGCC
                                    PHC6  →
      I  F  T  V  N  S  L  M  K  D  L  V  C  E  A  Y  S  L  N  A

1021  TTCCCGCTGCCGTCGGAACTGGAAACCGTGCTCGGCCTGAACATGCTGTGGCACCGCAGT
      F  P  L  P  S  E  L  E  T  V  L  G  L  N  M  L  W  H  R  S

1081  CGCAACACGCATCCGATGCTGGTGTGGGCGCGCAACCTGTTCAAGCAGGTGGTGGCCGAA
                                                ←  PHC2
      R  N  T  H  P  M  L  V  W  A  R  N  L  F  K  Q  V  V  A  E

1141  TACACCGGCAAGGCTTCCAACGCGCCGATGCACCCGCCGATGCTGACCGACGATTCCGGC

      Y  T  G  K  A  S  N  A  P  M  H  P  P  M  L  T  D  D  S  G
```

```
1201    AAGGCGGGCAAGACCGGCAAGGGCGATGCTGAAAAGGAGGATGAGTCGCGGCTGTCCGCC
        K  A  G  K  T  G  K  G  D  A  E  K  E  D  E  S  R  L  S  A

1261    TGATGTCCGCAGCGTCGCGCTCCAACAAAAACGGCCGCAGAGGATGCGGCCGTTTTCATT
        *

1321    GATGAGCCTGTCGCGTCAGCGCTTGGGGCGGTGCGGGCAAGGGTCCTTGGTGCAATTGCC

1381    GTACAGCGACAGGGCATGCTCTTGCAGCGCGAATCCGCGGGCGCGGGCGATGCTTTGTTG

1441    GCGCTGCTCGATCTCGCCGTCGA
                ←   PHC4
```

(B)

```
ACH0158 CAATTCAACACGCCGCCTTTTGTTATGCACTAAAACGAAAACGTTTGCGCATAAACAAAT
AW1     CAATTCAACACGCCGCCTTTTGTTATGCACTAAAACGAAAACGTTTGCGCATAAACAAAT

ACH0158 TTCGCAGCGCTGGCGCTGGAATTCGCTTTGAAAATCGCCTGGCAATCCGTACACTGAACA
AW1     TTCGCAAGACTGTCGTCGGAATTTGCTTTGAAAATCGCCCGGCAATCCGTACACTGAACA
              ***   *  **      *                *

ACH0158 ATTACTACATTTGTGACGCAGCCGCCGTTTCGTGCTGCGTCGAGGCCGTTGCCTCGGCAA
AW1     ATTACTACATTTGTGACGCAGTCGCCGTTTCTTGCTGCGTCGAGGCCGTTGCCTCGGCAA
                               *        *

ACH0158 CGTCTGCCTTTTCACTTTCTGTCTGAGCCAAGGCGTCCCATGGTCAACGTCGATACCAAG
AW1     CGTCTGCCTTTTCACTTTCTGTCTGAGCCAAGGCGTCCCATGGTCAACGTCGATACCAAG
```

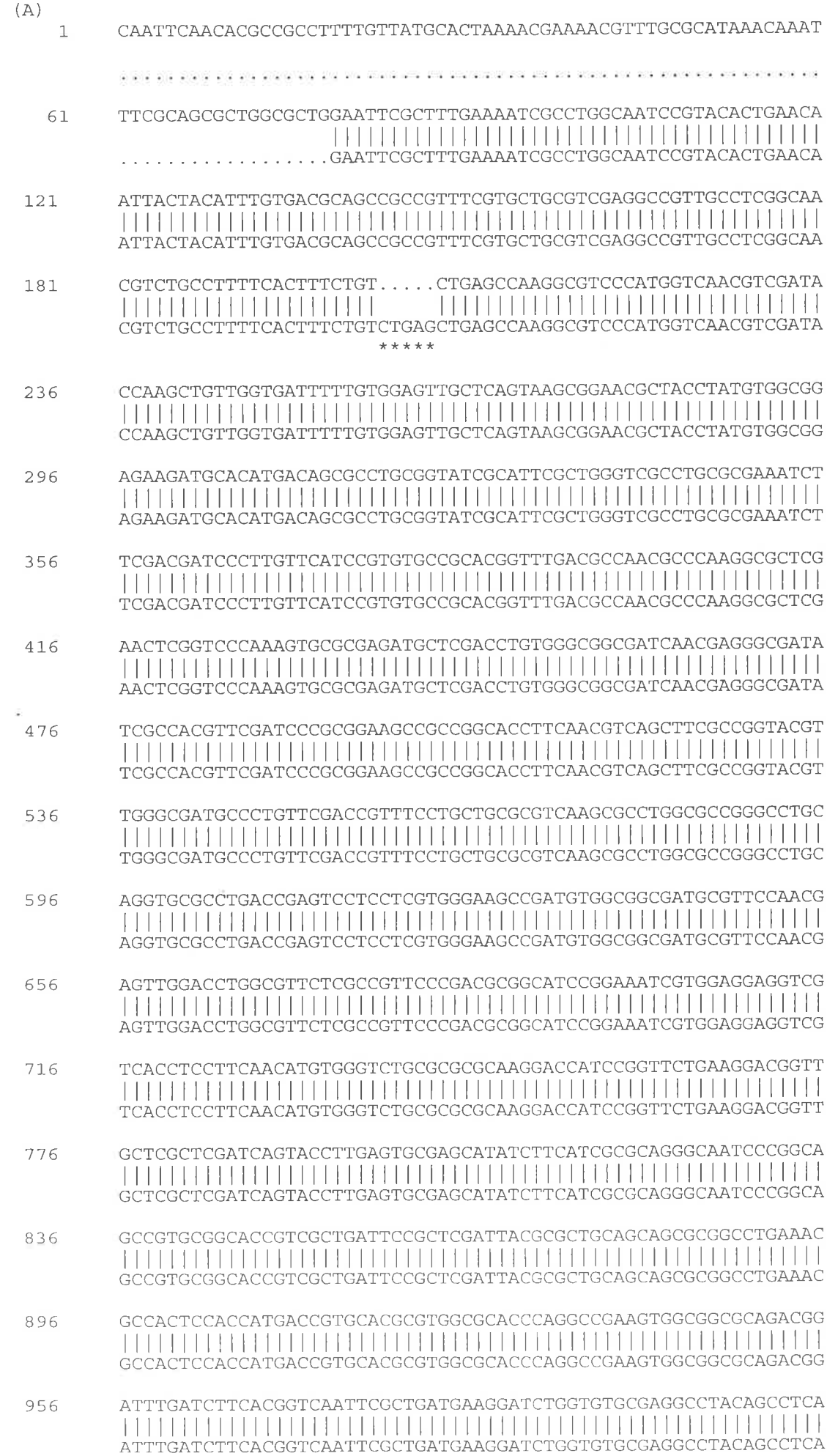
(C)

```
ACH0158 MVNVDTKLLVIFVE**LLSKRNATYVAEKMHMTAPAVSHSLGRLREIFDDPLFIRVP**HGLTP
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AW1     MVNVDTKLLVIFVE**LLSKRNATYVAEKMHMTAPAVSHSLGRLREIFDDPLFIRVP**HGLTP

ACH0158 TPKALELGPKVREMLDLWAAINEGDIATFDPAEAAGTFNVSFAGTLGDALFDRFLLRVKR
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AW1     TPKALELGPKVREMLDLWAAINEGDIATFDPAEAAGTFNVSFAGTLGDALFDRFLLRVKR

ACH0158 LAPGLQVRLTESSSWEADVAAMRSNELDLAFSPFPTRHPEIVEEVVTSFNMWVCARKDHP
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AW1     LAPGLQVRLTESSSWEADVAAMRSNELDLAFSPFPTRHPEIVEEVVTSFNMWVCARKDHP

ACH0158 VLKDGCSLDQYLECEHIFIAQGNPGSRAAPSLIPLDYALQQRGLKRHSTMTVHAWRTQAE
        |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AW1     VLKDRCSLDQYLECEHIFIAQGNPGSRAAPSLIPLDYALQQRGLKRHSTMTVHAWRTQAE
            *

ACH0158 VAAQTDLIFTVNSLMKDLVCEAYSLNAFPLPSELETVLGLNMLWHRSRNTHPMLVWARNL
        |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
AW1     VAAQTDLIFTVNSLMKDLVCEAYNLNAFPLPSELETVLGLNMLWHRSRNTHPMLVWARNL
                                    *

ACH0158 FKQVVAEYTGKASNAPMHPPMLTDDSGKAGKTGKGDAEKEDESRLSA
        |||||||||||||||||||||||||||||||||||||||||||||||
AW1     FKQVVAEYTGKASNAPMHPPMLTDDSGKAGKTGKGDAEKEDESRLSA
```

Figure 4.6

A. Nucleotide sequence comparison of *phcA* from wild-type ACH0158 (upper line) and *phcA81C* from the PC-type ACH0158-M81C (lower line). The mark (|) indicates the identical sequences. The asterisk (*) indicates CTGAG sequences duplicated upon the IS*Rso4* insertion.

B. A schematic diagram of the *phcA* (black box) with IS*Rso4* (blue box) insertion. *Eco*RI (E) sites are indicated. Directions of transcriptions for IS*Rso4* transposase and PhcA protein are indicated. Black arrows indicate the positions of IS*Rso4* insertion, -35, -10, RBS and ATG start codon. Yellow arrows indicate the locations of primers utilised for the inverse-PCR experiment (see details in Figure 3.5A for identities and sequences of the primers). Red box indicates the location of the HTH motif.


(A)
1 CAATTCAACACGCCGCCTTTTGTTATGCACTAAAACGAAAACGTTTGCGCATAAACAAAT
61 TTCGCAGCGCTGGCGCTGGAATTCGCTTTGAAAATCGCCTGGCAATCCGTACACTGAACA
GAATTCGCTTTGAAAATCGCCTGGCAATCCGTACACTGAACA
121 ATTACTACATTTGTGACGCAGCCGCCGTTTCGTGCTGCGTCGAGGCCGTTGCCTCGGCAA
ATTACTACATTTGTGACGCAGCCGCCGTTTCGTGCTGCGTCGAGGCCGTTGCCTCGGCAA
181 CGTCTGCCTTTTCACTTTCTGT.....CTGAGCCAAGGCGTCCCATGGTCAACGTCGATA
CGTCTGCCTTTTCACTTTCTGTCTGAGCTGAGCCAAGGCGTCCCATGGTCAACGTCGATA
*****
236 CCAAGCTGTTGGTGATTTTTGTGGAGTTGCTCAGTAAGCGGAACGCTACCTATGTGGCGG
CCAAGCTGTTGGTGATTTTTGTGGAGTTGCTCAGTAAGCGGAACGCTACCTATGTGGCGG
296 AGAAGATGCACATGACAGCGCCTGCGGTATCGCATTCGCTGGGTCGCCTGCGCGAAATCT
AGAAGATGCACATGACAGCGCCTGCGGTATCGCATTCGCTGGGTCGCCTGCGCGAAATCT
356 TCGACGATCCCTTGTTCATCCGTGTGCCGCACGGTTTGACGCCAACGCCCAAGGCGCTCG
TCGACGATCCCTTGTTCATCCGTGTGCCGCACGGTTTGACGCCAACGCCCAAGGCGCTCG
416 AACTCGGTCCCAAAGTGCGCGAGATGCTCGACCTGTGGGCGGCGATCAACGAGGGCGATA
AACTCGGTCCCAAAGTGCGCGAGATGCTCGACCTGTGGGCGGCGATCAACGAGGGCGATA
476 TCGCCACGTTCGATCCCGCGGAAGCCGCCGGCACCTTCAACGTCAGCTTCGCCGGTACGT
TCGCCACGTTCGATCCCGCGGAAGCCGCCGGCACCTTCAACGTCAGCTTCGCCGGTACGT
536 TGGGCGATGCCCTGTTCGACCGTTTCCTGCTGCGCGTCAAGCGCCTGGCGCCGGGCCTGC
TGGGCGATGCCCTGTTCGACCGTTTCCTGCTGCGCGTCAAGCGCCTGGCGCCGGGCCTGC
596 AGGTGCGCCTGACCGAGTCCTCCTCGTGGGAAGCCGATGTGGCGGCGATGCGTTCCAACG
AGGTGCGCCTGACCGAGTCCTCCTCGTGGGAAGCCGATGTGGCGGCGATGCGTTCCAACG
656 AGTTGGACCTGGCGTTCTCGCCGTTCCCGACGCGGCATCCGGAAATCGTGGAGGAGGTCG
AGTTGGACCTGGCGTTCTCGCCGTTCCCGACGCGGCATCCGGAAATCGTGGAGGAGGTCG
716 TCACCTCCTTCAACATGTGGGTCTGCGCGCGCAAGGACCATCCGGTTCTGAAGGACGGTT
TCACCTCCTTCAACATGTGGGTCTGCGCGCGCAAGGACCATCCGGTTCTGAAGGACGGTT
776 GCTCGCTCGATCAGTACCTTGAGTGCGAGCATATCTTCATCGCGCAGGGCAATCCCGGCA
GCTCGCTCGATCAGTACCTTGAGTGCGAGCATATCTTCATCGCGCAGGGCAATCCCGGCA
836 GCCGTGCGGCACCGTCGCTGATTCCGCTCGATTACGCGCTGCAGCAGCGCGGCCTGAAAC
GCCGTGCGGCACCGTCGCTGATTCCGCTCGATTACGCGCTGCAGCAGCGCGGCCTGAAAC
896 GCCACTCCACCATGACCGTGCACGCGTGGCGCACCCAGGCCGAAGTGGCGGCGCAGACGG
GCCACTCCACCATGACCGTGCACGCGTGGCGCACCCAGGCCGAAGTGGCGGCGCAGACGG
956 ATTTGATCTTCACGGTCAATTCGCTGATGAAGGATCTGGTGTGCGAGGCCTACAGCCTCA
ATTTGATCTTCACGGTCAATTCGCTGATGAAGGATCTGGTGTGCGAGGCCTACAGCCTCA

```
1016      ACGCCTTCCCGCTGCCGTCGGAACTGGAAACCGTGCTCGGCCTGAACATGCTGTGGCACC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ACGCCTTCCCGCTGCCGTCGGAACTGGAAACCGTGCTCGGCCTGAACATGCTGTGGCACC

1076      GCAGTCGCAACACGCATCCGATGCTGGTGTGGGCGCGCAACCTGTTCAAGCAGGTGGTGG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GCAGTCGCAACACGCATCCGATGCTGGTGTGGGCGCGCAACCTGTTCAAGCAGGTGGTGG

1136      CCGAATACACCGGCAAGGCTTCCAACGCGCCGATGCACCCGCCGATGCTGACCGACGATT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CCGAATACACCGGCAAGGCTTCCAACGCGCCGATGCACCCGCCGATGCTGACCGACGATT

1196      CCGGCAAGGCGGGCAAGACCGGCAAGGGCGATGCTGAAAAGGAGGATGAGTCGCGGCTGT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CCGGCAAGGCGGGCAAGACCGGCAAGGGCGATGCTGAAAAGGAGGATGAGTCGCGGCTGT

1256      CCGCCTGATGTCCGCAGCGTCGCGCTCCAACAAAAACGGCCGCAGAGGATGCGGCCGTTT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CCGCCTGATGTCCGCAGCGTCGCGCTCCAACAAAAACGGCCGCAGAGGATGCGGCCGTTT

1316      TCATTGATGAGCCTGTCGCGTCAGCGCTTGGGGCGGTGCGGGCAAGGGTCCTTGGTGCAA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TCATTGATGAGCCTGTCGCGTCAGCGCTTGGGGCGGTGCGGGCAAGGGTCCTTGGTGCAA

1376      TTGCCGTACAGCGACAGGGCATGCTCTTGCAGCGCGAATCCGCGGGCGCGGGCGATGCTT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTGCCGTACAGCGACAGGGCATGCTCTTGCAGCGCGAATCCGCGGGCGCGGGCGATGCTT

1436      TGTTGGCGCTGCTCGATCTCGCCGTCGA
          ||||||||||||||||||||||||||||
          TGTTGGCGCTGCTCGATCTCGCCGTCGAAGAATTC
```

(B)

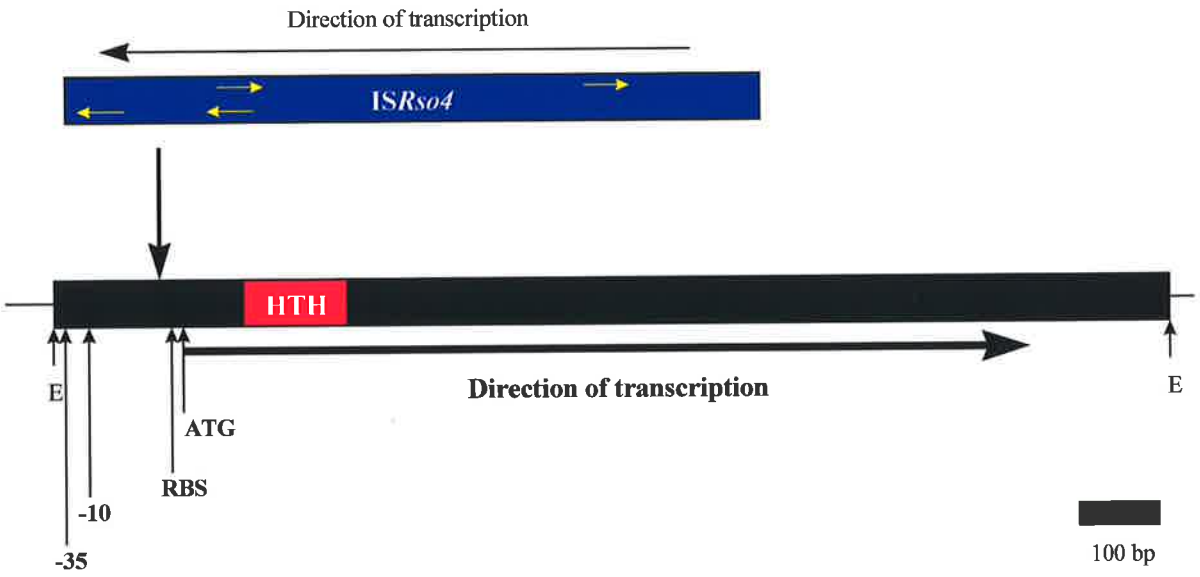
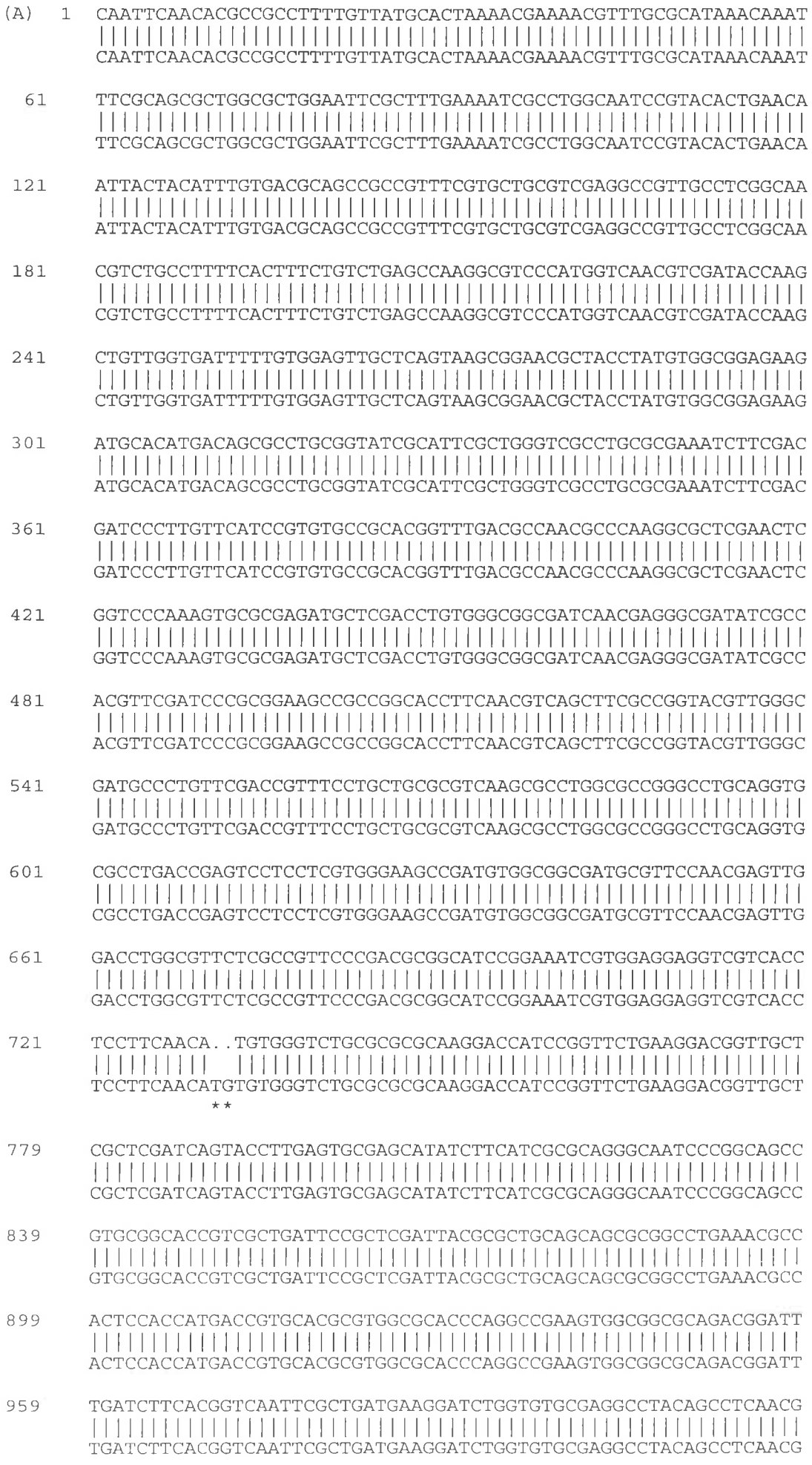


Figure 4.7

A. Nucleotide sequence comparison of the *phcA* from wild-type ACH0158 (upper line) and *phcA8* from the PC-type ACH0158-M8 (lower line). The mark (|) indicates the identical sequence. The asterisk (*) indicates the TG insertion.

B. Alignment of deduced amino acid sequence of *phcA* from wild-type ACH0158 (upper line) and PC-type ACH0158-M8 (lower line). The mark (|) indicates the identical sequence.



```
1019    CCTTCCCGCTGCCGTCGGAACTGGAAACCGTGCTCGGCCTGAACATGCTGTGGCACCGCA
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        CCTTCCCGCTGCCGTCGGAACTGGAAACCGTGCTCGGCCTGAACATGCTGTGGCACCGCA

1079    GTCGCAACACGCATCCGATGCTGGTGTGGGCGCGCAACCTGTTCAAGCAGGTGGTGGCCG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        GTCGCAACACGCATCCGATGCTGGTGTGGGCGCGCAACCTGTTCAAGCAGGTGGTGGCCG

1139    AATACACCGGCAAGGCTTCCAACGCGCCGATGCACCCGCCGATGCTGACCGACGATTCCG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        AATACACCGGCAAGGCTTCCAACGCGCCGATGCACCCGCCGATGCTGACCGACGATTCCG

1199    GCAAGGCGGGCAAGACCGGCAAGGGCGATGCTGAAAAGGAGGATGAGTCGCGGCTGTCCG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        GCAAGGCGGGCAAGACCGGCAAGGGCGATGCTGAAAAGGAGGATGAGTCGCGGCTGTCCG

1259    CCTGATGTCCGCAGCGTCGCGCTCCAACAAAAACGGCCGCAGAGGATGCGGCCGTTTTCA
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        CCTGATGTCCGCAGCGTCGCGCTCCAACAAAAACGGCCGCAGAGGATGCGGCCGTTTTCA

1319    TTGATGAGCCTGTCGCGTCAGCGCTTGGGGCGGTGCGGGCAAGGGTCCTTGGTGCAATTG
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
        TTGATGAGCCTGTCGCGTCAGCGCTTGGGGCGGTGCGGGCAAGGGTCCTTGGTGCAGTTG

1379    CCGTACAGCGACAGGGCATGCTCTTGCAGCGCGAATCCGCGGGCGCGGGCGATGCTTTGT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        CCGTACAGCGACAGGGCATGCTCTTGCAGCGCGAATCCGCGGGCGCGGGCGATGCTTTGT


1439    TGGCGCTGCTCGATCTCGCCGTCGA
        |||||||||||||||||||||||||
        TGGCGCTGCTCGATCTCGCCGTCGA
```

(B)

```
WT  MVNVDTKLLVIFVELLSKRNATYVAEKMHMTAPAVSHSLGRLREIFDDPLFIRVPHGLTP
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
M8  MVNVDTKLLVIFVELLSKRNATYVAEKMHMTAPAVSHSLGRLREIFDDPLFIRVPHGLTP

WT  TPKALELGPKVREMLDLWAAINEGDIATFDPAEAAGTFNVSFAGTLGDALFDRFLLRVKR
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
M8  TPKALELGPKVREMLDLWAAINEGDIATFDPAEAAGTFNVSFAGTLGDALFDRFLLRVKR

WT  LAPGLQVRLTESSSWEADVAAMRSNELDLAFSPFPTRHPEIVEEVVTSFNMWVCARKDHP
    ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||
M8  LAPGLQVRLTESSSWEADVAAMRSNELDLAFSPFPTRHPEIVEEVVTSFNMCGSARARTI

WT  VLKDGCSLDQYLECEHIFIAQGNPGSRAAPSLIPLDYALQQRGLKRHSTMTVHAWRTQAE

M8  RF*

WT  VAAQTDLIFTVNSLMKDLVCEAYSLNAFPLPSELETVLGLNMLWHRSRNTHPMLVWARNL


WT  FKQVVAEYTGKASNAPMHPPMLTDDSGKAGKTGKGDAEKEDESRLSA*
```

Figure 4.8

Nucleotide sequence comparison of the *phcA* from wild-type ACH0158 (upper line) and *phcA3* from the PC-type ACH0158-M3 (lower line). The mark (|) indicates the identical sequence. The asterisk (*) indicates the 132 bp region deleted.

```
1     CAATTCAACACGCCGCCTTTTGTTATGCACTAAAACGAAAACGTTTGCGCATAAACAAAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CAATTCAACACGCCGCCTTTTGTTATGCACTAAAACGAAAACGTTTGCGCATAAACAAAT

61    TTCGCAGCGCTGGCGCTGGAATTCGCTTTGAAAATCGCCTGGCAATCCGTACACTGAACA
      |||||||||||||||||||||||||||||||||||||||||||||||||||
      TTCGCAGCGCTGGCGCTGGAATTCGCTTTGAAAATCGCCTGGCAATCCG..........
                                                     **********

121   ATTACTACATTTGTGACGCAGCCGCCGTTTCGTGCTGCGTCGAGGCCGTTGCCTCGGCAA

      ..............................................................
      **************************************************************

181   CGTCTGCCTTTTCACTTTCTGTCTGAGCCAAGGCGTCCCATGGTCAACGTCGATACCAAG

      ..............................................................
      **************************************************************

241   CTGTTGGTGATTTTTGTGGAGTTGCTCAGTAAGCGGAACGCTACCTATGTGGCGGAGAAG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      ..GTTGGTGATTTTTGTGGAGTTGCTCAGTAAGCGGAACGCTACCTATGTGGCGGAGAAG
      **

301   ATGCACATGACAGCGCCTGCGGTATCGCATTCGCTGGGTCGCCTGCGCGAAATCTTCGAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      ATGCACATGACAGCGCCTGCGGTATCGCATTCGCTGGGTCGCCTGCGCGAAATCTTCGAC

361   GATCCCTTGTTCATCCGTGTGCCGCACGGTTTGACGCCAACGCCCAAGGCGCTCGAACTC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GATCCCTTGTTCATCCGTGTGCCGCACGGTTTGACGCCAACGCCCAAGGCGCTCGAACTC

421   GGTCCCAAAGTGCGCGAGATGCTCGACCTGTGGGCGGCGATCAACGAGGGCGATATCGCC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GGTCCCAAAGTGCGCGAGATGCTCGACCTGTGGGCGGCGATCAACGAGGGCGATATCGCC

481   ACGTTCGATCCCGCGGAAGCCGCCGGCACCTTCAACGTCAGCTTCGCCGGTACGTTGGGC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      ACGTTCGATCCCGCGGAAGCCGCCGGCACCTTCAACGTCAGCTTCGCCGGTACGTTGGGC

541   GATGCCCTGTTCGACCGTTTCCTGCTGCGCGTCAAGCGCCTGGCGCCGGGCCTGCAGGTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GATGCCCTGTTCGACCGTTTCCTGCTGCGCGTCAAGCGCCTGGCGCCGGGCCTGCAGGTG

601   CGCCTGACCGAGTCCTCCTCGTGGGAAGCCGATGTGGCGGCGATGCGTTCCAACGAGTTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CGCCTGACCGAGTCCTCCTCGTGGGAAGCCGATGTGGCGGCGATGCGTTCCAACGAGTTG

661   GACCTGGCGTTCTCGCCGTTCCCGACGCGGCATCCGGAAATCGTGGAGGAGGTCGTCACC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GACCTGGCGTTCTCGCCGTTCCCGACGCGGCATCCGGAAATCGTGGAGGAGGTCGTCACC

721   TCCTTCAACATGTGGGTCTGCGCGCGCAAGGACCATCCGGTTCTGAAGGACGGTTGCTCG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      TCCTTCAACATGTGGGTCTGCGCGCGCAAGGACCATCCGGTTCTGAAGGACGGTTGCTCG

781   CTCGATCAGTACCTTGAGTGCGAGCATATCTTCATCGCGCAGGGCAATCCCGGCAGCCGT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CTCGATCAGTACCTTGAGTGCGAGCATATCTTCATCGCGCAGGGCAATCCCGGCAGCCGT

841   GCGGCACCGTCGCTGATTCCGCTCGATTACGCGCTGCAGCAGCGCGGCCTGAAACGCCAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GCGGCACCGTCGCTGATTCCGCTCGATTACGCGCTGCAGCAGCGCGGCCTGAAACGCCAC

901   TCCACCATGACCGTGCACGCGTGGCGCACCCAGGCCGAAGTGGCGGCGCAGACGGATTTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      TCCACCATGACCGTGCACGCGTGGCGCACCCAGGCCGAAGTGGCGGCGCAGACGGATTTG
```

```
961     ATCTTCACGGTCAATTCGCTGATGAAGGATCTGGTGTGCGAGGCCTACAGCCTCAACGCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        ATCTTCACGGTCAATTCGCTGATGAAGGATCTGGTGTGCGAGGCCTACAGCCTCAACGCC

1021    TTCCCGCTGCCGTCGGAACTGGAAACCGTGCTCGGCCTGAACATGCTGTGGCACCGCAGT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        TTCCCGCTGCCGTCGGAACTGGAAACCGTGCTCGGCCTGAACATGCTGTGGCACCGCAGT

1081    CGCAACACGCATCCGATGCTGGTGTGGGCGCGCAACCTGTTCAAGCAGGTGGTGGCCGAA
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        CGCAACACGCATCCGATGCTGGTGTGGGCGCGCAACCTGTTCAAGCAGGTGGTGGCCGAA

1141    TACACCGGCAAGGCTTCCAACGCGCCGATGCACCCGCCGATGCTGACCGACGATTCCGGC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        TACACCGGCAAGGCTTCCAACGCGCCGATGCACCCGCCGATGCTGACCGACGATTCCGGC

1201    AAGGCGGGCAAGACCGGCAAGGGCGATGCTGAAAAGGAGGATGAGTCGCGGCTGTCCGCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        AAGGCGGGCAAGACCGGCAAGGGCGATGCTGAAAAGGAGGATGAGTCGCGGCTGTCCGCC

1261    TGATGTCCGCAGCGTCGCGCTCCAACAAAAACGGCCGCAGAGGATGCGGCCGTTTTCATT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        TGATGTCCGCAGCGTCGCGCTCCAACAAAAACGGCCGCAGAGGATGCGGCCGTTTTCATT

1321    GATGAGCCTGTCGCGTCAGCGCTTGGGGCGGTGCGGGCAAGGGTCCTTGGTGCAATTGCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        GATGAGCCTGTCGCGTCAGCGCTTGGGGCGGTGCGGGCAAGGGTCCTTGGTGCAATTGCC

1381    GTACAGCGACAGGGCATGCTCTTGCAGCGCGAATCCGCGGGCGCGGGCGATGCTTTGTTG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        GTACAGCGACAGGGCATGCTCTTGCAGCGCGAATCCGCGGGCGCGGGCGATGCTTTGTTG

1441    GCGCTGCTCGATCTCGCCGTCGA
        |||||||||||||||||||||||
        GCGCTGCTCGATCTCGCCGTCGA
```

CHAPTER 5

Analysis of genomic diversity of *Ralstonia solanacearum* isolates as determined by the insertion sequence elements

5.1 Introduction

Variation in many bacterial species is initiated by mobile DNA which is common amongst bacterial genomes (Healy *et al.*, 1999) and the variation generated has been shown in several instances to have been influenced in the evolution of the host/pathogen relationship. Healy *et al.* (1999) identified IS*1629* as a causal element to prompt genomic variation of gram-positive filamentous *Streptomyces* spp.. IS*1629* was flanked to the *nec1* gene (Bukhalid *et al.*, 1998), which is a putative pathogenicity or virulence gene in *Streptomyces* spp.. Southern hybridisation results of three *Streptomyces* species with the IS*1629-nec1* region as a probe suggested that the region was recently transferred from *S. scabies* to *S. acidiscabies* and *S. turgidiscabies*, both of which have been lately reported as pathogenic and are only found in a limited geographical region. These results prompted the following examination of genomic variation in *R. solanacearum* using IS elements as probes.

The four IS elements described in chapter 3 were isolated from a biovar 2 strain ACH0158 and a biovar 3 strain ACH0171 of *R. solanacearum*. Three of these, IS*Rso3*, IS*Rso4* and IS*Rso2*, derived from ACH0158 appeared to be absent from ACH0171 (Fig. 3.1B and 3.1C). Likewise, an ACH0171-derived IS*61* was not present in multiple copies in ACH0158 (Fig. 3.1D). This chapter describes experiments that include Southern hybridisations of genomic DNA samples from several isolates with the four IS

element probes. This chapter also discusses the possibility of lateral transfer of the IS elements among less closely related groups of strains.

5.2 Southern analysis of the four IS elements in 52 *R. solanacearum* isolates

Southern blots containing *Eco*RI-digested genomic DNA samples of *R. solanacearum* isolates were hybridised with the four IS element probes to determine copy numbers and identify genomic polymorphisms in the distribution of the IS elements. DNA samples of 52 isolates were tested that were isolated from various host plants and which represented the five different biovar groups of the pathogen (Table 2.1).

The IS*Rso3* probe was prepared by using *Spe*I and *Sma*I enzymes to remove the 307 bp long insert from the clone pSC15. Preparation of the IS*Rso4* probe utilised the two restriction enzymes *Sal*I and *Eco*RI to recover a 588 bp fragment from the clone pSV102 (Fig. 3.7) and the IS*Rso2* probe was prepared by recovering a 554 bp insert from the clone pISB using *Eco*RI and *Xba*I. Similarly, a 1020 bp insert was obtained from the clone pIS61 using *Eco*RI and *Xba*I. The IS61 insert also contained a 781 bp long IS*61* element and a 233 bp flanking sequence, which searches of the sequence database revealed as having no similarities to any previously characterised sequences.

5.2.1 Southern analysis in nine biovar 1 strains

A Southern blot containing *Eco*RI-digested genomic DNA of nine biovar 1 strains was hybridised with IS*Rso3* probe (Fig. 5.1B). Multiple hybridising fragments existed in the CIP0066 strain (Fig. 5.1B-lane 4) and a prominent single band showed in the CIP0218 strain (Fig. 5.1B-lane 7) with little or no hybridisation to DNA of other strains. The same blot was hybridised with IS*Rso4* probe after it was stripped to remove the previous hybrids. CIP0066, CIP0120 and CIP0218 strains each of which was isolated from the potato plant contained multiple *Eco*RI fragments hybridising the probe (Fig. 5.1C-lane 4, 6 and 7) (Table 5.1A). When the Southern blot was hybridised with the IS*Rso2* probe, more than three *Eco*RI fragments in all nine biovar 1 strains strongly hybridised with the probe (Fig. 5.1D). CIP0066 and CIP0070 strains contained a minimum 11 and 12 copies of IS*Rso2*, respectively (Fig. 5.1D-lane 4 and 5). Hybridisation patterns of CIP0418 and CIP0419 strains of Indonesian origin with IS*Rso2* probe were very similar to those of CIP0007 and CIP0009 isolated from Central America. The IS61 sequence that was initially though to be specific to biovar 3 genomes, hybridised to a high molecular weight band in most biovar 1 strains and to at least two *Eco*RI fragments in CIP0066 and CIP0218 (Fig. 5.1E-lane 4 and 7). It is

possible that the hybridisation result with the IS61 sequence was influenced by a 233 bp flanking sequences included to prepare the IS61 probe.

In general, biovar 1 strains are though to be most similar to biovar 2 strains. Extensive RFLP analysis has placed the two biovars in division II (Cook *et al.*, 1989; Cook and Sequeira, 1994). It was therefore anticipated that biovar 2 strain-derived IS elements may be present in some biovar 1 strains. All the biovar 1 strains tested showed multiple hybridisation signals with at least one of the three IS element probes that originated from ACH0158. In addition, most biovar 1 strains showed low copy numbers of sequences hybridising the IS61 probe. However, there was no clear indication with regard to the presence of particular IS elements in association with the host and geographical origins of the biovar 1 strains examined. Based on PCR-RFLP analysis of *hrp* region of several biovar 1 strains Poussier *et al.* (1999) recently suggested the possibility that there could be two different biovar 1 populations in Africa, both endemic and recently imported from the Americas. This possibility appears to be applicable only to African biovar 1 isolates at present. Yet, it is clearly necessary to test more biovar 1 strains to deduce any conclusive evidence in addition to the results described here.

5.2.2 Southern analysis in seventeen biovar 2 strains

*Eco*RI-digested genomic DNA samples of seventeen biovar 2 strains were resolved on an agarose gel (Fig. 5.2A and A'-lane 1-17) and Southern blotted. The genomic DNA blot was initially hybridised with IS*Rso3* probe and subsequently with IS*Rso4* and IS*Rso2* probes. A large number (> 40) of *Eco*RI fragments was hybridised with IS*Rso3* probe from all biovar 2 strains tested (Fig. 5.2B-lane 1-17). In contrast, a relatively small number of 5 or 6 and 12 genomic fragments respectively were hybridised with IS*Rso4* and IS*Rso2* probes (Fig. 5.2C-lane 1-17 and 5.2D-lane 1-17). Despite the possibility of minor in copy numbers of three IS elements, the overall similarity of the within probe-Southern hybridisation patterns in biovar 2 strains tested is remarkably high and consistent with a high level of genetic uniformity in biovar 2 strains. Based on RFLP analysis of several isolates of *R. solanacearum*, Cook and Sequeira (1991 and 1994) suggested that biovar 2 (race 3) strains have almost identical genomes, commensurate with the similarity of their metabolic and other biological characteristics, regardless of the geographical region of origin. Taken together, the results obtained here support the suggestion that biovar 2 strains have a clonal origin in South America (Cook and Sequeira, 1994). However, there is a clear and potentially useful possibility to use

the low level of variation in restriction pattern revealed here to subdivide this group of strains in a way that was formerly impossible.

For IS*Rso3* and IS*Rso2*, the copy numbers may be underestimated because of the possibility of multiple copies of the elements in the more intensely hybridised some or all of these multiple fragments that are clearly visible (Fig. 5.2B- and 5.2D-lane 1-17). Of course fragments may reflect the presence of truncated elements or relics that have diverged in sequence after old insertion events. Another possible explanation for the observed differences in hybridisation intensity is the presence of other functional IS elements that are closely related to the probe but sufficiently different in sequence to reduce the level of hybridisation at high stringency. As only a single example of IS*Rso3* and IS*Rso2* was sequenced in this study, distinguishing between these possibilities is not possible.

Comparisons of the Southern hybridisation patterns of the seventeen biovar 2 strains with the three IS elements revealed significant genomic variations between the strains, suggesting that the IS elements may have been involved in generating genetic diversity among strains. The most divergent biovar 2 strain was CIP0310 (Fig. 5.2-lane 3) which originated uniquely from South America. Cook and Sequeira (1994) previously showed that the CIP0310 strain had different RFLP patterns for the probes in their experiments. The strain was also proven to have a slightly different metabolic phenotype (Hayward, 1994), compared with typical biovar 2 strains. Genomic variations in CIP0117, CIP0403, PD1427, ACH1061 and ACH1068 strains are revealed by the occurrence of *Eco*RI polymorphisms. However, deletions and additions of hybridising fragments were detected in CIP0359, South Africa 28, South Africa 45 and South Africa 65 strains (Fig. 5.2-lane 4, 8, 9 and 12, respectively) that reflect the transpositional activity of the IS elements.

A second Southern blot containing sixteen biovar 2 strains, not including strain ACH1061 was hybridised with the IS61 probe (Fig. 5.2E). The probe hybridised to high molecular weight (> 23.1 kb) fragments in several biovar 2 strains tested. Yet, most of the strains did not show multiple hybridising fragments and, in fact, there was none containing more than three *Eco*RI fragments hybridised with the probe (Fig. 5.2E-lane 1-17). The analysis therefore suggested that the biovar 2 strains tested possessed a low level of the IS*61* element as a repetitive DNA sequence. The identity of the high molecular weight fragments detected is still unknown and it is a possibility that the

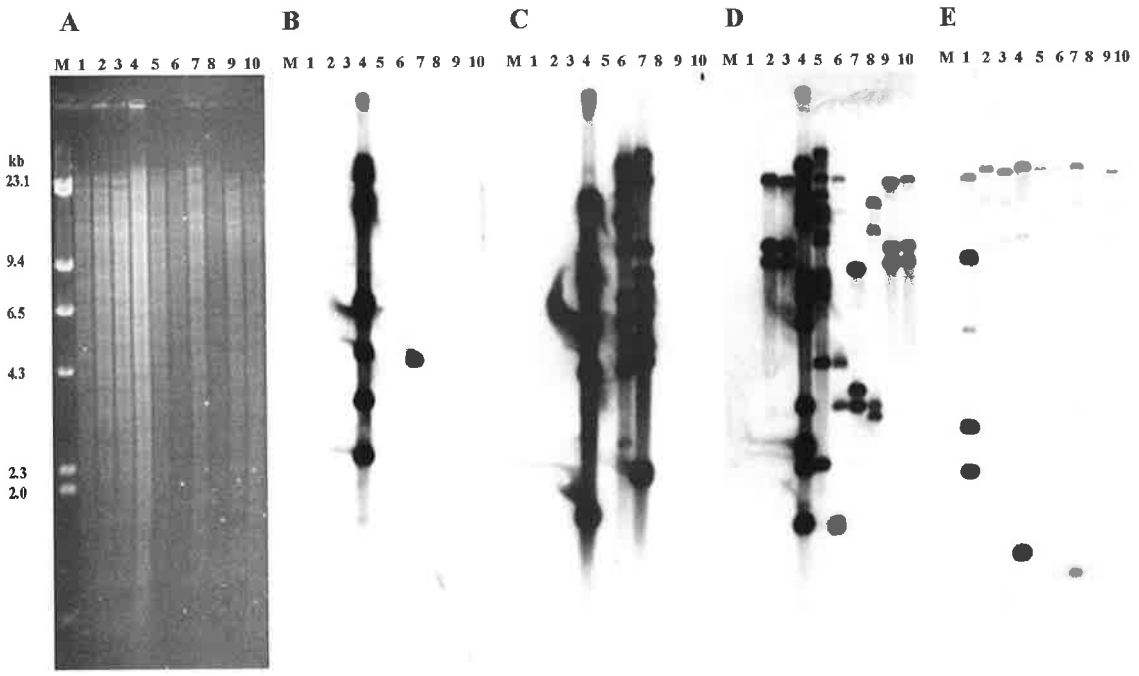


Figure 5.1

A. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of nine biovar 1 strains and ACH0171 of *R. solanacearum.*

M. Molecular size marker of λ DNA restricted with *Hind*III

1. ACH0171 strain (biovar 3)
2. CIP0007 strain (biovar 1)
3. CIP0009 strain (biovar 1)
4. CIP0066 strain (biovar 1)
5. CIP0070 strain (biovar 1)
6. CIP0120 strain (biovar 1)
7. CIP0218 strain (biovar 1)
8. CIP0301 strain (biovar 1)
9. CIP0418 strain (biovar 1)
10. CIP0419 strain (biovar 1)

B. Southern hybridisation of the gel in A with IS*Rso3* probe.

C. Southern hybridisation of the gel in A with IS*Rso4* probe.

D. Southern hybridisation of the gel in A with IS*Rso2* probe.

E. Southern hybridisation of the gel in A with IS61 probe.

Figure 5.2

A and A'. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of seventeen biovar 2 strains, a biovar N2 strain and an atypical biovar 2 strain of *R. solanacearum.* The asterisk (*) in A' indicates *Eco*RI-digested genomic DNA of ACH0171.

M. Molecular size marker of λ DNA restricted with *Hind*III

1. ACH0158 strain (biovar 2)
2. CIP0117 strain (biovar 2)
3. CIP0310 strain (biovar 2)
4. CIP00359 strain (biovar 2)
5. CIP0402 strain (biovar 2)
6. CIP0403 strain (biovar 2)
7. CIP0405 strain (biovar 2)
8. South Africa 28 strain (biovar 2)
9. South Africa 45 strain (biovar 2)
10. South Africa 46 strain (biovar 2)
11. South Africa 49 strain (biovar 2)
12. South Africa 65 strain (biovar 2)
13. PD1427 strain (biovar 2)
14. PD1435 strain (biovar 2)
15. ACH0127R strain (biovar 2)
16. ACH1061 strain (biovar 2)
17. ACH1068 strain (biovar 2)
18. CIP0358 strain (biovar N2)
19. ACH0732 strain (atypical biovar 2)

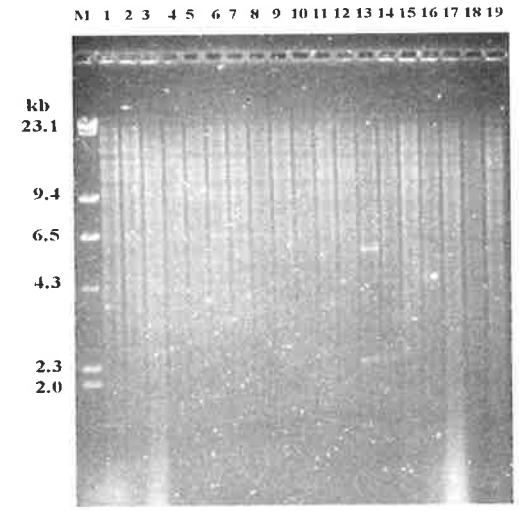
B. Southern hybridisation of the gel in A with IS*Rso3* probe.

C. Southern hybridisation of the gel in A with IS*Rso4* probe.

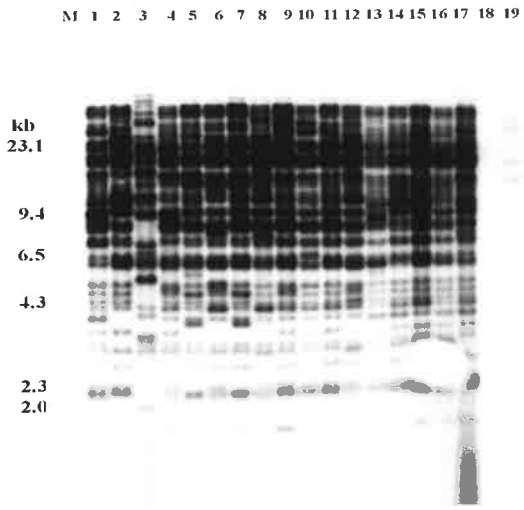
D. Southern hybridisation of the gel in A with IS*Rso2* probe.

E. Southern hybridisation of the gel in A' with IS61 probe. The asterisk (*) indicates ACH0171 DNA hybridisation.

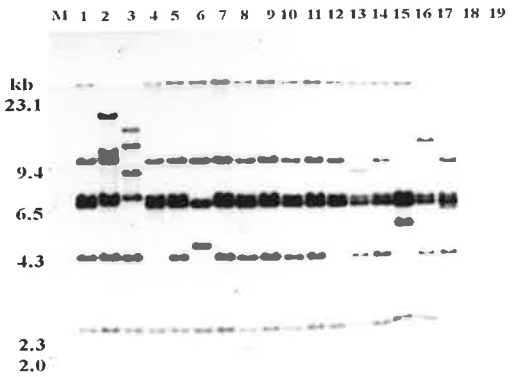
A


M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
kb
23.1
9.4
6.5
4.3
2.3
2.0

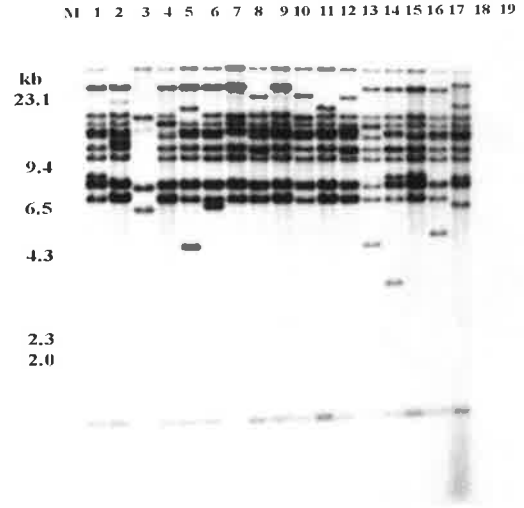
B


M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
kb
23.1
9.4
6.5
4.3
2.3
2.0

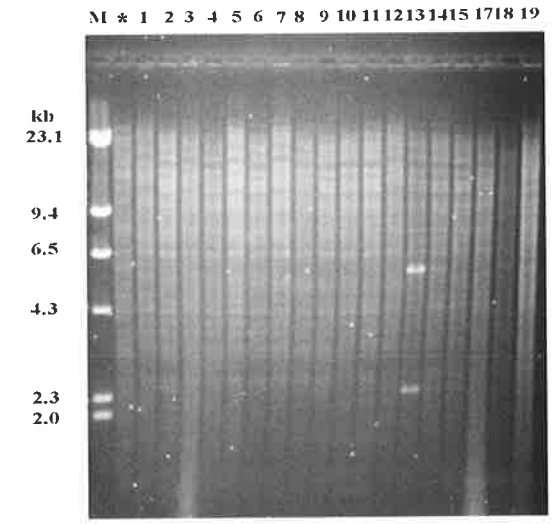
C


M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
kb
23.1
9.4
6.5
4.3
2.3
2.0

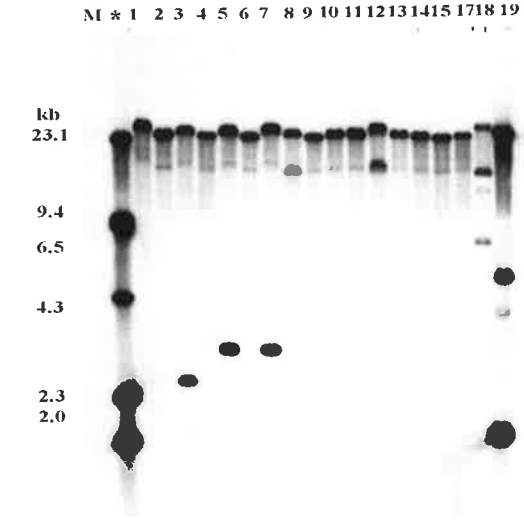
D


M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
kb
23.1
9.4
6.5
4.3
2.3
2.0

A'


M * 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 17 18 19
kb
23.1
9.4
6.5
4.3
2.3
2.0

E


M * 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 17 18 19
kb
23.1
9.4
6.5
4.3
2.3
2.0

hybridisation result with the IS61 sequence was caused by the additional 233 bp sequence that was present in the IS61 probe. Therefore, it is clearly necessary to prepare the IS*61*-element specific probe and to test it against biovar 1 and biovar 2 strains.

5.2.3 Southern analysis in CIP0358 and ACH0732 strains

CIP0358 is an African biovar N2 strain (Hayward, 1994) originating from Cameroon (Table 2.1). In general, biovar N2 strains are isolated from the tropical lowland of Peru and Brazil in South America (French, *et al.,* 1995). Therefore, CIP0358 is a biovar N2 strain of unusual origin. ACH0732 is an atypical biovar 2 strain isolated from tomato in the Northern Territory in Australia (Taghavi *et al.*, 1996; Li and Hayward, 1994). Southern hybridisations of these strains with the three ACH0158-derived IS elements showed little or no hybridisation (Fig. 5.2B, 5.2C and 5.2D-lane 18 and 19). Previously, Cook and Sequeira (1991) suggested that biovar N2 strains originating from the lowland South America are distinct from typical biovar 2-race 3 strains based on RFLP analysis. The result obtained here supports that suggestion. However, CIP0358 and ACH0732 DNA contained three *Eco*RI fragments hybridising the IS61 probe. Stronger hybridisations in ACH0732 than those in CIP0358 DNA were due to genomic DNA loading variations (Fig. 5.2A'-lane 19). Obviously it is not possible to draw any conclusions from this result because only two strains were tested and it is necessary to test more of typical and atypical biovar N2 and atypical biovar 2 strains. Nevertheless, the biovar N2 and atypical biovar 2 strains tested here did not appear to be similar to the typical biovar 2 strains tested.

5.2.4 Southern analysis in sixteen biovar 3 strains

Fragments of *Eco*RI-digested genomic DNA samples of sixteen biovar 3 strains were resolved on an agarose gel (Fig. 5.3A) and blotted onto a nylon membrane. In Southern blot hybridisation with IS*Rso3* probe, seven biovar 3 strains showed positive hybridisations and CIP0284 genomic DNA strain appeared to contain at least 9 copies of IS*Rso3* (Fig. 5.3B-lane 15 and Table 5.1). ACH1064 DNA also showed multiple less strongly hybridising fragments, whose reduced intensity was clearly not due to DNA loading differences (Fig. 5.3A-lane 7). The same genomic DNA blot was hybridised with IS*Rso4* probe after the previous probe was stripped and no hybridisation was observed (Fig. 5.3C). In contrast, Southern hybridisation with IS*Rso2* probe disclosed multiple, weak hybridising bands, which were particularly in CIP0284 DNA (Fig. 5.3D-lane 15), suggesting the existence of divergent IS*Rso2*-like sequences in these biovar 3 strains.

As expected, most of the biovar 3 strains tested, with the exception of the ACH0190 showed strong hybridisation to multiple *Eco*RI fragments with the IS61 probe in Southern analysis (Fig. 5.3E). The other 15 of the 16 biovar 3 strains, including ACH0171, from which IS*61* was derived, showed quite similar hybridisation profiles, although some genomic rearrangements probably caused by deletions and insertions were clearly identified. Highly similar Southern hybridisation patterns were particularly characteristic in several strains isolated in Australia (Fig. 5.3E-lane 1-8 and -lane 16). South Africa 12 and South Africa 53 strains appeared to contain the lowest copy number of IS*61*, whilst CIP0065 originated from Costa Rica contained at least 9 copies of IS*61* in the genome. CIP0065 and CIP0284 strains from China showed quite different hybridisation patterns (Fig. 5.3E-lane 14 and 15). However, no hybridisation on ACH0190 DNA with the IS61 probe was not clearly understood. In summary, almost all the biovar 3 strains tested contained IS*61* in their genomes, and in some strains the hybridisation patterns were very different from others. However, only few strains appeared to contain IS*Rso3* and IS*Rso2*-like sequences.

5.2.5 Southern analysis in seven biovar 4 strains and one biovar 5 strain

Genomic DNA samples of seven biovar 4 strains and a single biovar 5 strain, CIP0277 were *Eco*RI-digested, fragments were separated on an agarose gel (Fig. 5.4A) and Southern blotted. Southern hybridisation showed that the genomes of six biovar 4 strains and CIP0277 contained multiple copies of IS*Rso3* (Fig. 5.4B). The same genomic DNA blot was hybridised with IS*Rso4* probe and no hybridisation was observed (Fig. 5.4C). In a third hybridisation of the same filter, DNA from the Indonesian strain CIP0264 (Fig. 5.4D-lane 2) appeared to contain at least 16 copies of IS*Rso2* element with a lower level of hybridisation in the other biovar 4 strains tested. Most biovar 4 strains, with the exception of ACH0262 and CIP0277, contained highly polymorphic bands consistent with the presence of multiple copies of IS*61* (Fig. 5.4E). Reason for no hybridisation on ACH0262 DNA with the IS61 probe was not fully understood. The hybridisation pattern of CIP0277 (Fig. 5.4E-lane 8) was similar to those of some biovar 4 strains.

Taken together, the presence of multiple copies of IS*61* was observed in the biovar 4 and biovar 5 strains tested, while the appearance of highly repetitive IS*Rso3* in several strains and IS*Rso2* in some strains was also detected. The results described here contrast with previous observations that ACH0158-derived IS elements were rarely present as repetitious DNA in biovar 3 strains tested (see chapter 5.2.4). This study supports the

Figure 5.3

A. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of sixteen biovar 3 strains of *R. solanacearum.*

M. Molecular size marker of λ DNA restricted with *Hind*III

1. ACH0171 strain (biovar 3)
2. ACH0190 strain (biovar 3)
3. ACH0369 strain (biovar 3)
4. ACH0574 strain (biovar 3)
5. ACH0671 strain (biovar 3)
6. ACH1024 strain (biovar 3)
7. ACH1064 strain (biovar 3)
8. ACH1070 strain (biovar 3)
9. South Africa 8 strain (biovar 3)
10. South Africa 12 strain (biovar 3)
11. South Africa 53 strain (biovar 3)
12. South Africa 93 strain (biovar 3)
13. South Africa 112 strain (biovar 3)
14. CIP0065 strain (biovar 3)
15. CIP0284 strain (biovar 3)
16. PD1437 strain (biovar 3)

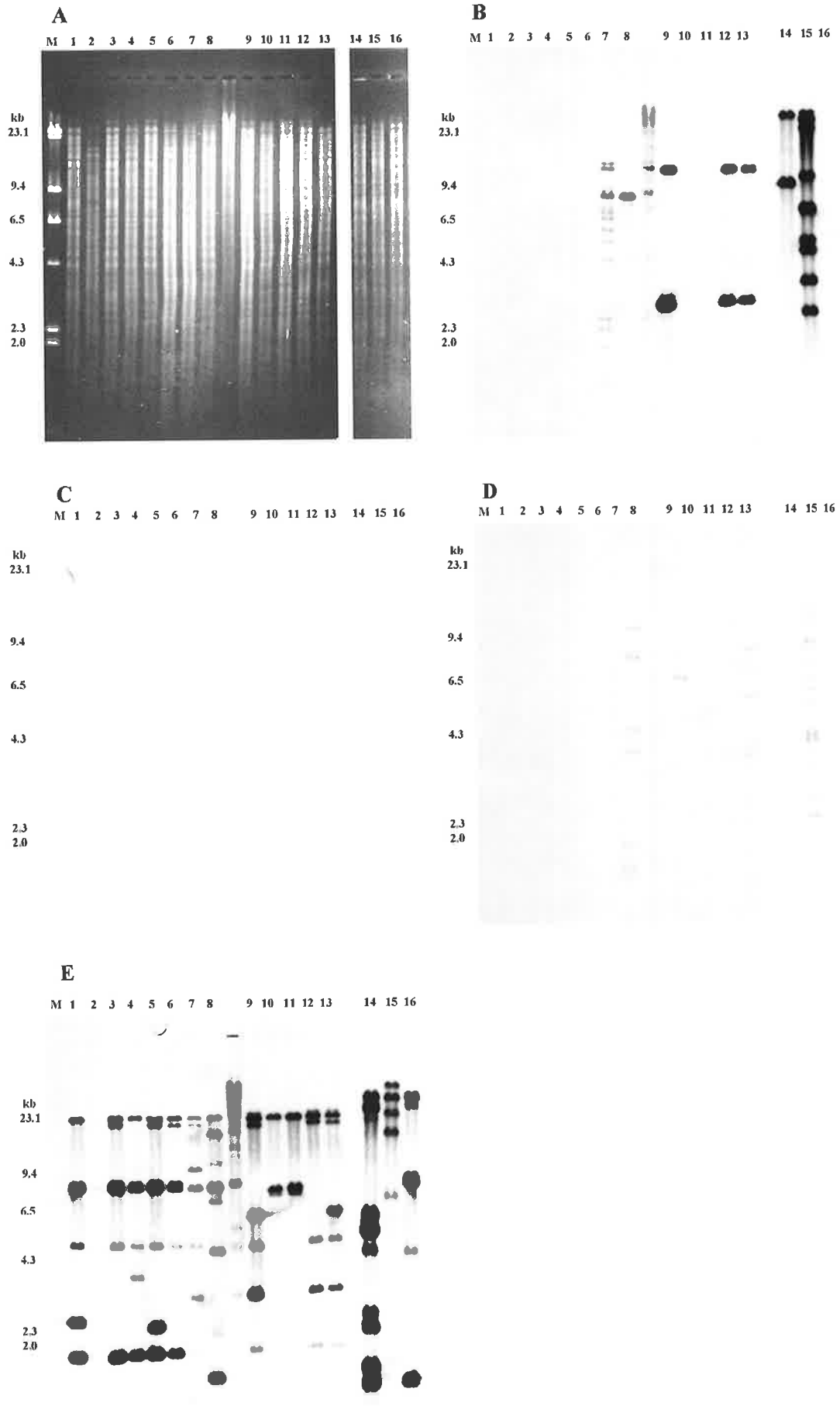
B. Southern hybridisation of the gel in A with IS*Rso3* probe.

C. Southern hybridisation of the gel in A with IS*Rso4* probe.

D. Southern hybridisation of the gel in A with IS*Rso2* probe.

E. Southern hybridisation of the gel in A with IS61 probe.

Not numbered lane between lane 8 and 9 is to be disregarded because DNA sample is partially digested.


A
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
kb
23.1
9.4
6.5
4.3
2.3
2.0
B
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
kb
23.1
9.4
6.5
4.3
2.3
2.0
C
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
kb
23.1
9.4
6.5
4.3
2.3
2.0
D
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
kb
23.1
9.4
6.5
4.3
2.3
2.0
E
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
kb
23.1
9.4
6.5
4.3
2.3
2.0

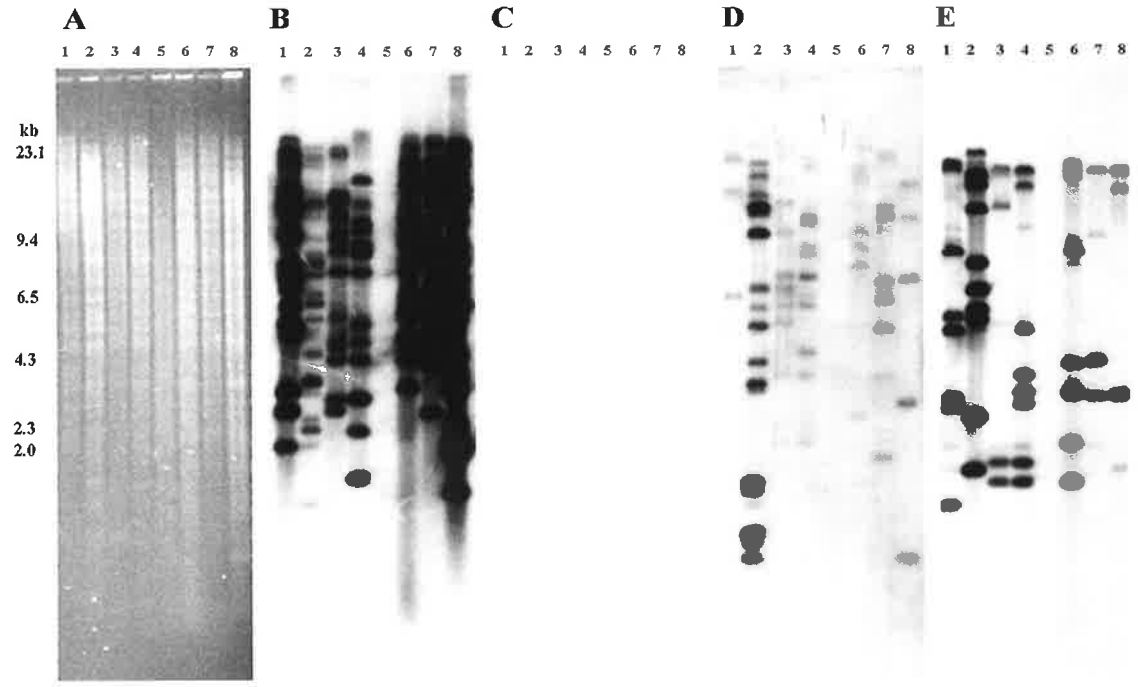


Figure 5.4

A. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of seven biovar 4 strains and a biovar 5 strain of *R. solanacearum*. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.6, 4.4, 2.3, 2.0 and 0.56 kb was included as a size marker (M).

1. CIP0121 strain (biovar 4)
2. CIP0264 strain (biovar 4)
3. CIP0283 strain (biovar 4)
4. PD1431 strain (biovar 4)
5. ACH0262 strain (biovar 4)
6. ACH0319 strain (biovar 4)
7. ACH1066 strain (biovar 4)
8. CIP0277 strain (biovar 5)

B. Southern hybridisation of the gel in A with IS*Rso3* probe.

C. Southern hybridisation of the gel in A with IS*Rso4* probe.

D. Southern hybridisation of the gel in A with IS*Rso2* probe.

E. Southern hybridisation of the gel in A with IS61 probe.

notion (Cook and Sequeira, 1994) that biovar 3, 4 and 5 strains are quite closely related to each other, and separated taxonomically from biovar 1 and 2 strains. However, it appears that IS element transmission has occurred between the more distantly related groups of biovar 2 and biovar 4 strains based on the hybridisation results with IS*Rso3* and IS*Rso2* probes.

ACH0190 (Fig. 5.3B, C, D and E-lane 2) and ACH0262 (Fig. 5.4B, C, D and E-lane 5) were thought to be biovar 3 and biovar 4 strains, respectively but they contained no sequences that hybridised with any of the four different IS element probes including the biovar 3 strain ACH0171-derived IS61 probe. It is possible, though unlikely, that DNA samples of the two strains were degraded during DNA preparation and manipulation or that the two strains are not *R. solanacearum*. This second possibility could be tested by amplifying the species-specific 282 bp fragment using primers AU759f and AU760r (Opina *et al.*, 1997).

In summary, extensive Southern hybridisation results of *Eco*RI-digested genomic DNA of 52 *R. solanacearum* isolates suggested that all the biovar 1 and 2 strains tested contained multiple copies of one of the three IS elements isolated from ACH0158. In addition, more than 90 % of the biovar 3, biovar 4 and biovar 5 strains tested contained repetitive IS*61*-like sequences in their genomes. The results, therefore, ~~strongly~~ supported the notion that biovar 1 and 2 strains are closely related to each other and more distantly related to biovar 3, biovar 4 and biovar 5 which also group together. A similar suggestion that two divisions (division I and division II) of *R. solanacearum* strains are clearly separated was ~~already~~ made on the basis of restriction fragment length polymorphism (Cook *et al.*, 1989; Cook and Sequeira, 1994) (Fig. 5.5A) and 16S rRNA gene sequence comparisons (Li *et al.*, 1993; Taghavi *et al.*, 1996) (Fig. 5.5B). However, in the results presented here, there are several exceptions to this general rule and some anomalies that are difficult to explain without further knowledge and/or experimentation.

5.3 Transmission of IS elements in *R. solanacearum*

IS*Rso4* was present predominantly in biovar 2 strains tested but was also found in three out of nine biovar 1 strains, such that the frequency of occurrence of IS*Rso4* in non-biovar 2 strains was only 9 % (Table 5.1B). It is possible therefore that IS*Rso4* was laterally transferred from biovar 2 strains to biovar 1 strains probably by transformation

rather than conjugation (Bertolla *et al.*, 1999) although the direction of movement is difficult to establish and the mode of transfer is unknown.

In contrast to IS*Rso4*, IS*Rso2* existed not only in all the biovar 2 strains, but also in all the biovar 1 strains tested, although the biovar 1 strains appeared to have fewer copy numbers of the element. Therefore, with the exception of a high level of hybridisation to CIP0264, IS*Rso2* was confined to biovar 1 and biovar 2 (Table 5.1B). The observation of predominant occurrence IS*Rso2* element in biovar 1 and 2 strains suggests that the element has been transmitted within a common lineage of division II strains of *R. solanacearum* (Fig. 5.5C).

IS*Rso3* revealed the greatest copy numbers (> 40 copies) in the genomes of the biovar 2 strains tested, and IS*Rso3*-like sequences were also present in all non biovar 2 strains tested, represented in the biovar 4 strains, as highly repetitive sequences. The presence of IS*Rso3* in every major biovar group of *R. solanacearum* suggests that the element was present in a common ancestral strain before separation of the two divisions (Fig. 5.5C).

The IS*61* element that was present in all biovar 3 strains with one exception (ACH0190), was also detected in all biovar 4 strains (except ACH0262) and in CIP0277 (a biovar 5 strain). The frequency of occurrence of the IS*61* element in the division I strains tested would be 100 % if ACH0190 and ACH0262 strains are proved to be other bacterial species and consequently discarded in the analysis. In addition, IS*61* existed as repetitive sequences in the biovar N2 strain, CIP0358, and the atypical biovar 2 strain, ACH0732, suggesting that these two strains are related to the division I strains. Overall, IS*61* was present in the majority of the biovar 3 strains tested and at a lower frequency (25.7 %) in non-biovar 3 strains (Table 5.1B).

Cook and Sequeira (1994) determined that a race 3 (biovar 2)-specific 2 kb DNA fragment derived from a race 3 strain was present in a minority of non-race 3 strains (5.5 %) and suggested the occurrence of lateral gene transfer of genomic DNA between distantly related *R. solanacearum* strains. The analysis of hybridisation results using the four different IS elements as probes discussed in this chapter may suggest the possibility of horizontal transmission of the IS elements between the distantly related strains as well as vertical transmission the IS elements from an ancestral strain. Misclassification is possible by any taxonomic method and in the interpretation of these

experiments depends on totally reliable labelling of strains as well as the methodology itself.

R. solanacearum has been proven to develop transformation competent growth stage *in planta* for horizontal exchanges of genetic material (Bertolla *et al.*, 1999). This particular ability of the species may enable it to accelerate dynamic horizontal transmission of IS elements, possibly accompanied by flanking genomic regions containing genes involved in virulence and pathogenicity of *R. solanacearum*. If this has occurred, it would provide a powerful mechanism for the generation of strain variation within the species. In other studies, several IS elements carrying virulence and pathogenicity genes between different bacterial strains and species have already been identified (Bukhalid *et al.*, 1998; Byrne, 1989; Healy *et al.*, 1999; Quintiliani and Courvalin, 1996; Tyler *et al.*, 1996).

In summary, the nearly ubiquitous presence of IS*Rso2* in biovar 1 and biovar 2 strains suggested that the IS element originated in a common ancestor of the division II strains. Similarly, IS*61* was probably vertically transferred from an ancestor of the division I strains. In contrast, IS*Rso4* may have been horizontally transmitted between biovar 1 and biovar 2 strains because only small number of biovar 1 strains tested possessed the particular IS element that probably originated in biovar 2. IS*Rso3* was present in all the major biovar groups although transmission patterns and genetic mechanism to generate the wide distribution of this IS element are not yet clearly understood.

5.4 Genomic organisation of the four IS elements in four different biovar strains

Pulsed-field gel electrophoresis (PFGE) of *Spe*I-digested genomic DNA samples of five *R. solanacearum* strains representing four different biovars was performed to determine the genomic distribution and organisation of the four IS elements. *Spe*I was selected particularly because it cuts the high G+C *R. solanacearum* genome relatively infrequently to give large fragments (Römling *et al.*, 1994).

In this study, *Spe*I-digested genomic DNA fragments were resolved using four different electrophoresis conditions (see details in chapter 2.3.11), Southern blotted and hybridised with the four different IS element probes. The five strains tested were AW1 (biovar 1), ACH0158 (biovar 2), ACH1068 (biovar 2), ACH1064A (biovar 3) and ACH092 (biovar 4). Major restriction fragment length polymorphisms were observed between the strains (Fig. 5.6A, F, K and P). The two biovar 2 strains each contained

nineteen *Spe*I fragments, their size ranging from larger than 630.0 kb to approximately 5 kb and the restriction patterns were very similar to each other (Fig. 5.6A, F, K and P-lane 2 and 3).

Hybridisation with IS*Rso3* indicated that 15 out of 19 of these contained the particular IS element (Fig. 5.6B, G, L and Q-lane 2 and 3). Several *Spe*I fragments hybridising IS*Rso3* in the biovar 2 strains were anticipated because of the large number of *Eco*RI fragments hybridising the probe observed in previous Southern analyses (Fig. 5.2B). The results also suggested that significant *Spe*I polymorphisms were present in the two closely related biovar 2 strains. The result suggested that IS*Rso3* was dispersed throughout the genome of the strains tested because most of the large *Spe*I fragments were hybridised with the IS*Rso3* probe. Some fragments showed increased levels of hybridisation with the probe, implying that, as expected, these genomic regions contained multiple copies of IS*Rso3* element.

Hybridisation of 10 out of 25 *Spe*I fragments in ACH092 strain with IS*Rso3* probe was consistent with the previous Southern hybridisation result of *Eco*RI-digested genomic DNA of biovar 4 strains showed the prominent hybridisation with IS*Rso3* probe.

The four Southern blots containing *Spe*I-digested bacterial genomic DNA were stripped and rehybridised with IS*Rso4* (Fig. 5.6C, H, M and R-lane 2 and 3). Three large *Spe*I fragments (> 630.5 kb, > 582.0 kb and > 339.5 kb) (Fig. 5.6C, H, M and R) were hybridised. The hybridisation result showed the slightly different hybridisation patterns between ACH0158 and ACH1068 (Fig. 5.6C, H, M and R), confirming genomic variation between the strains that was used to isolate this element (Fig. 3.2).

The same genomic DNA membranes were hybridised with IS*Rso2* (Fig. 5.6D, I, N and S-lane 2 and 3). IS*Rso2* appeared to be present in *Spe*I fragments ranged in size from > 339.5 kb to > 97.0 kb. Previous Southern hybridisation results of *Eco*RI-digested genomic DNA of nine biovar 1 strains showed that many biovar 1 strains contained IS*Rso2*-like sequences but AW1 showed no sign of any hybridisation.

The IS61 insert in pIS61 was also used as a probe to the four PFGE Southern blots. The probe hybridised strongly to several *Spe*I fragments from ACH1064A (biovar 3) and ACH092 (biovar 4), but not to any fragments from biovar 1 and biovar 2 strains (Fig. 5.6E, J, O and T-lane 5 and 6). Hybridisation results were consistent with the previous

Southern hybridisation analysis of *Eco*RI-digested genomic DNA (Fig. 5.3.E and 5.4.E).

An approximately 339.5 kb *Spe*I fragment of ACH0158 appeared to contain all three IS*Rso4*, IS*Rso3* and IS*Rso2* elements (Table 5.2). In addition, the other three *Spe*I fragments appeared to contain IS*Rso4* and IS*Rso3*. Similarly, two *Spe*I fragments showed the co-presence of IS*Rso3* and IS*Rso2* (Table 5.2). It is not known whether those *Spe*I fragments containing several IS elements may consist of particular genomic regions carrying genes related to bacterial pathogenicity (pathogenicity island) (Hacker *et al.*, 1997) and virulence (virulence island).

5.5 Summary and discussion

In summary, the results obtained from the extensive Southern hybridisation analyses of the four IS elements suggested a discontinuous distribution with each element type having a particular distribution within particular groups of strains in a pattern that is generally consistent with previously existing taxonomic groups. In addition, the results implied that vertical transmission of the IS elements from a common ancestor and lateral transfer of the IS elements between less closely related strains.

In general, the distribution of different types of IS elements within the collection of strains used in this thesis, parallels previous systems that have been used to classify *R. solanacearum*. The major grouping into two divisions (Cook and Sequeira, 1994) is fully supported by IS element Southern analysis and previously recognised subdivisions within these major groups are also generally compatible with the current results. However, as in the RFLP scheme of Cook and Sequeira (1994), there are anomalies that require further investigation because they appear to reflect rare recombination between taxonomically different strains or to result from horizontal transfer. Only IS*Rso3* appears in most of the strains in this collection suggesting that the sequence is as old as the species itself. In contrast IS*Rso4*, IS*Rso2* and IS*61* show various degrees of specificity to particular groups of strains suggesting that they originated by horizontal transfer from other species. Whilst there is general compatibility of the Southern analyses using IS elements as probes with previous classification systems, there is clear evidence that the method reveals varying degree of genomic heterogeneity within what were considered homogeneous groups. For example, biovar 2 was previously thought to be essentially clonal, but analyses of IS element organisation reveal a significant level of genomic diversity. Such heterogeneity is much more pronounced in other groups of

the pathogen. Novel diagnostic methods may be developed on the basis of the Southern hybridisation results described in this chapter. In particular, IS*Rso4* and IS*Rso2* elements are appropriate to determine division II (biovar 1 and biovar 2 groups) strains while the IS*Rso3* element is very specific to group division I (biovar 3, 4 and 5 groups) strains. This result suggests that there has been little genetic exchange between these two groups despite the observation that there has been a large admixture of strains on a world scale and there are several common hosts and soil environments. Therefore, sporadic occurrences of some IS elements across the two groups are most likely to be explained by horizontal transfer of the IS elements. The only alternative to this explanation is that there is some strains or species misidentification within the stocks utilised in the experiments.

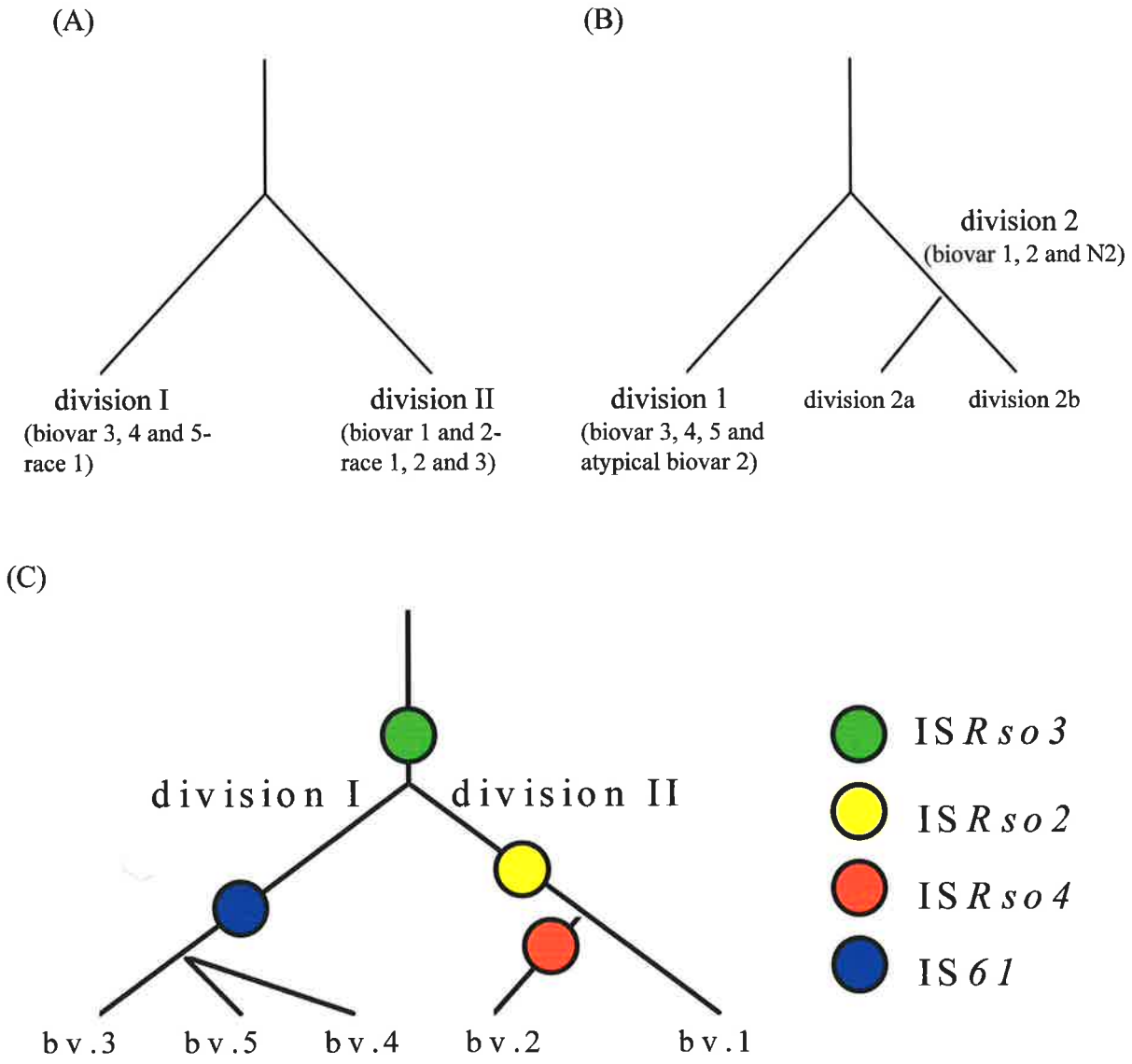


Figure 5.5

Simplified diagrams of phylogenetic relationships of *R. solanacearum* groups.

A. The diagram was drawn based on the restriction fragment length polymorphism analysis. Adapted from Cook and Sequeira (1994).

B. The diagram was drawn based on the 16S rRNA gene sequence comparison. Adapted from Taghavi *et al.* (1996).

C The diagram is drawn based on the Southern hybridisation results of 52 *R. solanacearum* strains described in this study with the coloured circles indicating the possible evolutionary origins of the IS elements.

Figure 5.6

A, F, K and P. Pulsed-field gel electrophoresis of *Spe*I-digested genomic DNA of *R. solanacearum*. Lambda ladder PFG marker (NEB) and Low range PFG marker (NEB) were used as size markers (M). The numbers on the left of the gels represent the molecular sizes (kb) of each fragment.

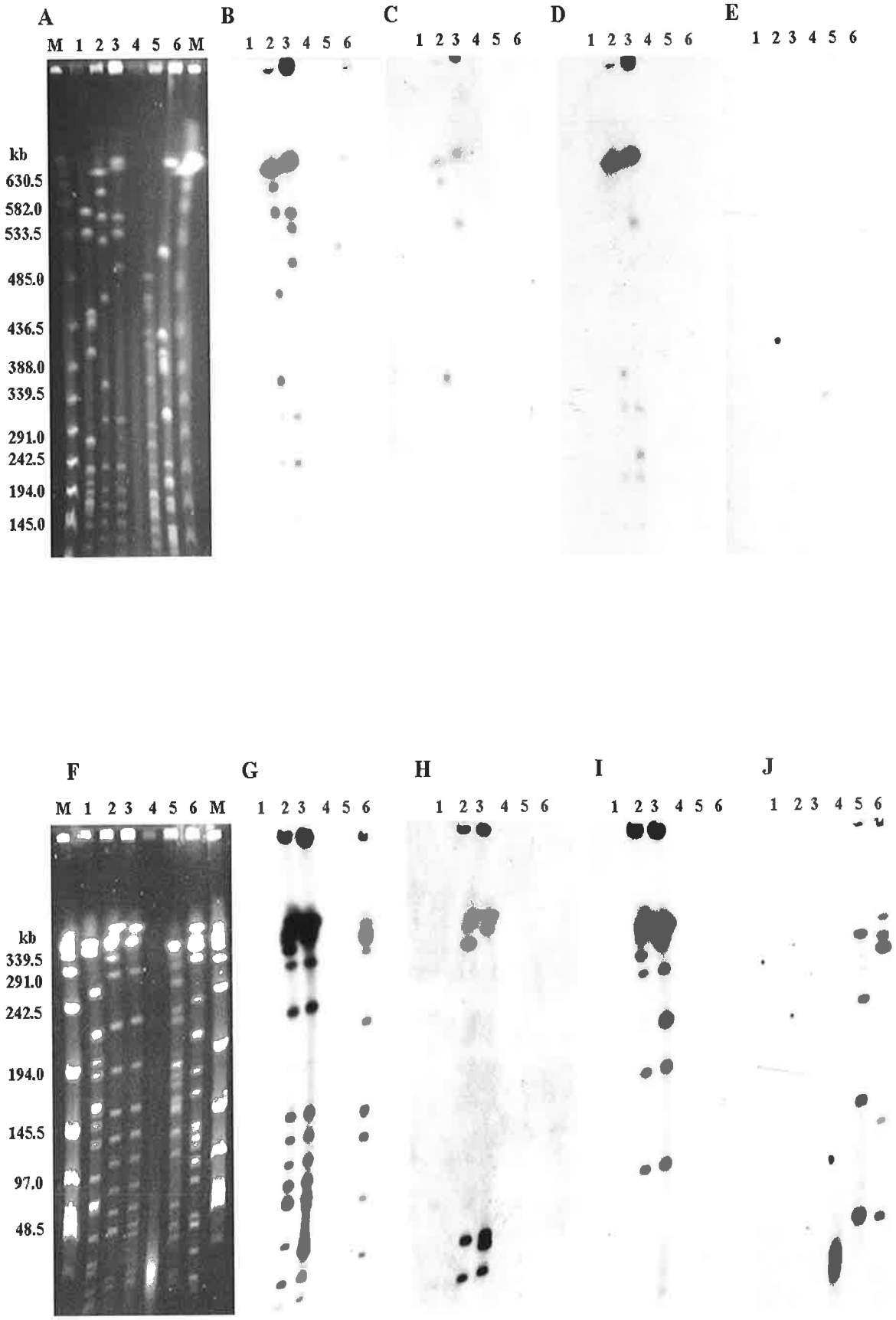
1. AW1 strain (biovar 1)
2. ACH0158 strain (biovar 2)
3. ACH1068 strain (biovar 2)
4. ACH0171 strain (biovar 3)
5. ACH1064A strain (biovar 3)
6. ACH092 strain (biovar 4)

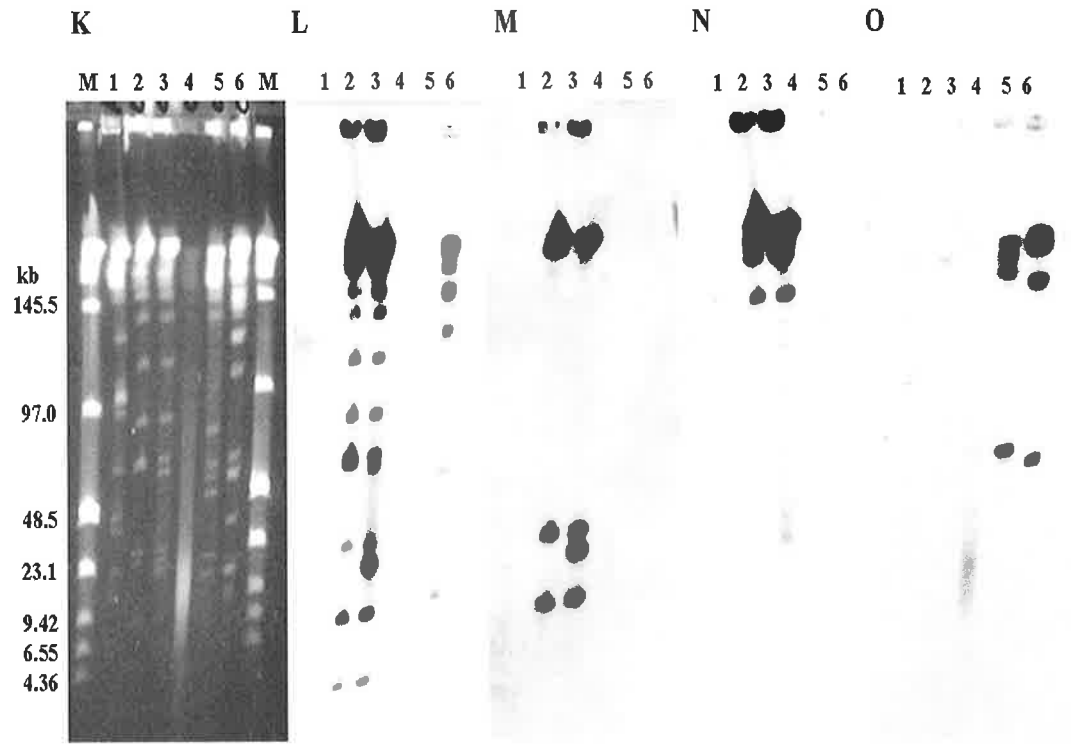
B, G, L and Q. Southern hybridisations of the gels in A, F, K and P with IS*Rso3* probe.

C, H, M and R. Southern hybridisations of the gel in A, F, K and P with IS*Rso4* probe.

D, I, N and S. Southern hybridisations of the gel in A, F, K and P with IS*Rso2* probe.

E, J, O and T. Southern hybridisations of the gel in A, F, K and P with IS61 probe.


A
B
C
D
E
M 1 2 3 4 5 6 M
1 2 3 4 5 6
kb
630.5
582.0
533.5
485.0
436.5
388.0
339.5
291.0
242.5
194.0
145.0
F
G
H
I
J
M 1 2 3 4 5 6 M
1 2 3 4 5 6
kb
339.5
291.0
242.5
194.0
145.5
97.0
48.5


K
L
M
N
O
M 1 2 3 4 5 6 M
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
kb
145.5
97.0
48.5
23.1
9.42
6.55
4.36

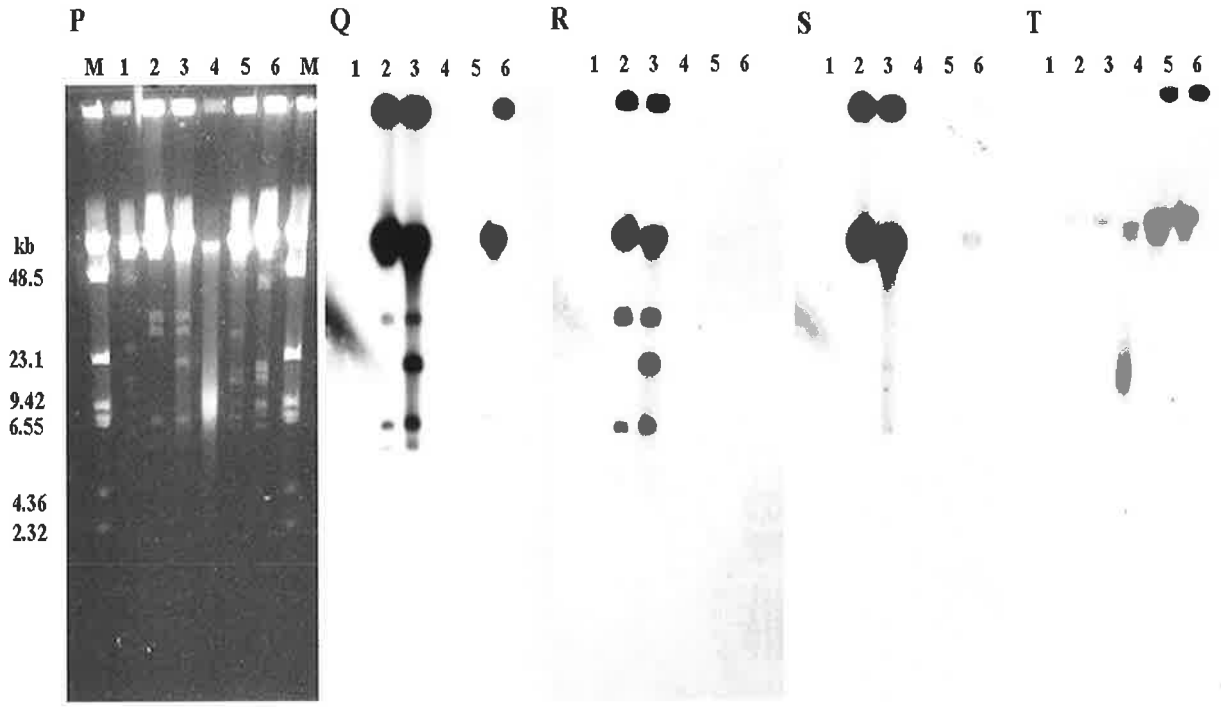

P
Q
R
S
T
M 1 2 3 4 5 6 M
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
kb
48.5
23.1
9.42
6.55
4.36
2.32

Table 5.1

Summary of the Southern hybridisation results of *Eco*RI-digested genomic DNA of 52 strains of *R. solanacearum* with IS*Rso4*, IS*Rso3*, IS*Rso2* and IS61 probes.

A. A list of strains tested. In addition, host plants and biovars of the strains are listed. Numbers of *Eco*RI fragments hybridised are indicated.

B. Frequency of the occurrence of each IS element in five different biovars, a biovar N2 strain and an atypical biovar 2 strain.

C. A graph of the data that were summarised in Table 4.1B. Data, which indicate the percentages of number of isolates with more than 3 fragments clearly hybridised, were only utilised to draw the graph.

(A)

Ralstonia solanacearum			probes			
strains	**biovar**	**host**	**IS*Rso3***	**IS*Rso4***	**IS*Rso2***	**IS*61***
CIP0007	1	banana	-	-	+ 3	+ 1
CIP0009	1	plantain	-	-	+ 3	+ 1
CIP0066	1	potato	+ 6	+ 7	+ 11	+ 2
CIP0070	1	banana	-	-	+ 12	+/- 1
CIP0120	1	potato	-	+ 8	+ 4	+/- 1
CIP0218	1	potato	+ 1	+ 9	+ 3	+ 2
CIP0301	1	potato	-	-	+ 4	-
CIP0418	1	peanut	-	-	+ 3	+/- 1
CIP0419	1	peanut	-	-	+ 3	-
ACH0127R	2	potato	+ >30	+ 5	+ 12	+ 1
ACH0158	2	potato	+ >30	+ 6	+ 12	+ 1
ACH0732	atypical 2	tomato	+/-[a] 5	-	-	+ 3
ACH1061	2	potato	+ >30[b]	+ 6	+ 11	NA[c]
ACH1068	2	potato	+ >30	+ 6	+ 12	+ 1
CIP0117	2	potato	+ >30	+ 7	+ 12	+ 1
CIP0310	2	potato	+ >25	+ 6	+ 4	+ 2
CIP0358	N2[d]	potato	-	+/- 1	-	+ 3
CIP0359	2	potato	+ >30	+ 5	+ 10	+ 1
CIP0402	2	potato	+ >30	+ 6	+ 12	+ 2
CIP0403	2	potato	+ >30	+ 5	+ 12	+ 1
CIP0405	2	potato	+ >30	+ 6	+ 9	+ 2
PD1427	2	tomato	+ >30	+ 6	+ 11	+ 1
PD1435	2	potato	+ >30	+ 5	+ 12	+ 1
South Africa 28	2	potato	+ >30	+ 7	+ 11	+ 2
South Africa 45	2	potato	+ >30	+ 7	+ 11	+ 1
South Africa 46	2	potato	+ >30	+ 6	+ 11	+ 1
South Africa 49	2	potato	+ >30	+ 6	+ 11	+ 1
South Africa 65	2	potato	+ >30	+ 5	+ 11	+ 2
ACH0171	3	eggplant	-	-	-	+ 6
ACH0190	3	*Xanthium pungens*	-	-	-	-
ACH0369	3	tomato	-	-	-	+ 6
ACH0574	3	tomato	-	-	-	+ 6
ACH0671	3	pepper	-	-	-	+ 7
ACH1024	3	*Strelitzia reginae*	-	-	-	+ 6
ACH1064	3	heliconia	+ 4	-	-	+ 6
ACH1070	3	ornamental ginger	+ 1	-	-	+ 6
CIP0065	3	chilli	+ 2	-	-	+ 9
CIP0284	3	common olive	+ 9	-	+/- 14	+ 5
PD1437	3	tomato	-	-	-	+ 6
South Africa 8	3	potato	+ 2	-	-	+ 6
South Africa 12	3	tobacco	-	-	-	+ 3
South Africa 53	3	potato	-	-	-	+ 3
South Africa 93	3	potato	+ 2	-	-	+ 4
South Africa 112	3	potato	+ 2	-	-	+ 5

(A) continued

strains	biovar	host	IS*Rso3*	IS*Rso4*	IS*Rso2*	IS*61*
ACH0262	4	ginger	-	-	-	-
ACH0319	4	*Solanum nigrum*	**+ 13**	-	+/- 6	**+ 8**
ACH1066	4	heliconia	**+ 11**	-	+/- 7	**+ 4**
CIP0121	4	potato	**+ 14**	-	+/- 4	**+ 7**
CIP0264	4	potato	**+ 11**	-	**+ 16**	**+ 10**
CIP0283	4	common olive	**+ 11**	-	+/- 7	**+ 4**
PD1431	4	potato	**+ 13**	-	+/- 10	**+ 8**
CIP0277	5	mulberry	**+ 19**	-	+/- 4	**+ 3**

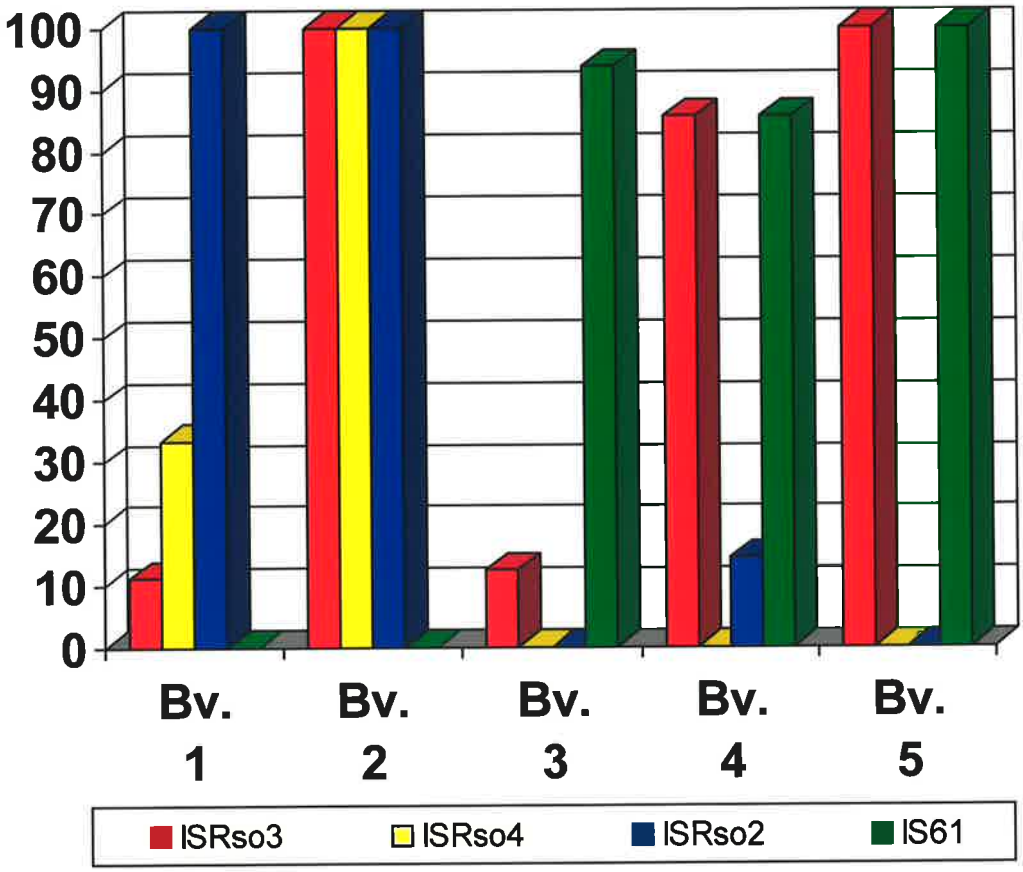
+, positive hybridisation signal; -, no hybridisation signal
[a] very weak hybridisations
[b] more than 30 *Eco*RI fragments hybridised
[c] data not available
[d] biovar N2 from lowland

(B)

	IS*Rso3*	IS*Rso4*	IS*Rso2*	IS*61*
Biovar 1	**11.1 (22.2)**	**33.3 (33.3)**	**100 (100)**	0[a] (44.4)[b]
Biovar 2	**100 (100)**	**100 (100)**	**100 (100)**	0 (100)
Biovar N2	0	0	0	**100 (100)**
Biovar 2 (atypical)	0	0	0	**100 (100)**
Biovar 3	12.5 (43.8)	0	0	**93.8 (93.8)**
Biovar 4	**85.7 (85.7)**	0	14.3 (14.3)	**85.7 (85.7)**
Biovar 5	**100**	0	0	**100**
Other than Biovar 2	28.6 (45.7)	8.6 (8.6)	28.6 (28.6)	
Other than Biovar 3				25.7 (82.9)

[a] % of number of isolates with *more than 3 fragments* clearly hybridised per number of isolates tested.
[b] % of number of isolates with *more than 1 fragment* clearly hybridised per number of isolates tested.

(C)


100
90
80
70
60
50
40
30
20
10
0
Bv. 1
Bv. 2
Bv. 3
Bv. 4
Bv. 5
ISRso3
ISRso4
ISRso2
IS61

*Spe*I fragments	IS*Rso3*	IS*Rso4*	IS*Rso2*
A	+	+	-
B	+	-	-
C	-	-	-
D	+	-	-
E	+	+	+
F	+	-	+
G	+	-	-
H	-	-	+
I	-	-	-
J	+	-	-
K	+	-	+
L	+	-	-
M	+	-	-
N	+	-	-
O	+	+	-
P	-	-	-
Q	+	+	-
R	+	-	-

Table 5.2

Summary of the Southern hybridisation results of *Spe*I-digested genomic DNA of ACH0158 with the IS*Rso3*, IS*Rso4* and IS*Rso2* probes. Each *Spe*I fragment was designated by letters of the English alphabet based on the proposal nomenclature by Tümmler *et al.* (1992).

CHAPTER 6

Isolation of flanking genes of IS*Rso3* and IS*Rso4* in pSC15K and pIS102JI

6.1 Introduction

Several IS elements have been identified as associated with antibiotic resistance and virulence genes in diverse bacterial species (Hacker *et al.*, 1997) and it is known that they are involved in transmission of, and the generation of genetic variation in, bacterial virulence and pathogenicity genes. Transmission and non-random intragenic recombination within virulence genes are particularly important in antigenic molecule variations of many bacterial species because the variations can assist pathogens to avoid the defence systems of host organisms. Additionally, the diversity of bacterial antigens is a crucial factor in enabling pathogens to invade various host organisms, expanding host ranges and colonising new ecological habitats. *R. solanacearum* has been strikingly successful in all these aspects.

Vibrio cholerae serotype O139 was identified as a causal agent for a recent epidemic of cholera, while previously serotype O1 was the only serotype capable of causing epidemic cholera (Stroeher and Manning, 1997). Silver staining of the lipopolysaccharide (LPS) complex of *V. cholerae* serotype O139 showed a distinct form of LPS, compared to the LPS complex of serotype O1. Genetic studies of a genomic region containing genes involved in O-antigen/capsular biosynthesis suggested that *V. cholerae* serotype O139 might have acquired DNA regions for alternative O-antigen biosynthesis by IS*1358* (Stroeher and Manning, 1997).

The four IS elements isolated from *R. solanacearum* were involved in genetic heterogeneity of isolates tested (discussed in chapter 5) and virulence gene regulation (discussed in chapter 4). This chapter discusses the possible associations of IS*Rso4* and

IS*Rso3* with genes responsible for virulence and pathogenicity. Experimental procedures included sequencing of genomic regions adjacent to IS*Rso4* and IS*Rso3* and Southern blot hybridisations.

6.2 Isolation of pIS102JI and pSC15K containing IS*Rso4* and IS*Rso3*, respectively

To obtain genomic clones containing IS*Rso4* and IS*Rso3*, a genomic DNA library of *Eco*RI-digested genomic DNA of wild-type ACH0158 was prepared in λZAP®II (Table 2.2). Screening of plaques used IS*Rso4* and IS*Rso3* probes yielded a clone designated pIS102JI that contains IS*Rso4* and six other clones carrying IS*Rso3* for further characterisations. The six genomic clones containing IS*Rso3* were pSC15E, pSC15H, pSC15J, pSC15K, pSC15R and pSC15S.

A primer designed from sequence of the IS*Rso4* and two primers from IS*Rso3* were utilised to obtain sequences from adjacent regions in the seven genomic clones. The primers utilised were Ja from IS*Rso4* and BD1 and BD2 derived from IS*Rso3* (Fig. 3.4 and 3.5). DNA sequence obtained from initial sequencing reactions was used in a standard database search. Sequences adjacent to IS*Rso4* in pIS102JI were similar to genes encoding eubacterial porin proteins. Sequences near IS*Rso3* in the clone pSC15K showed strong similarities to genes encoding UDP-*N*-acetylenolpyruvoylglucosamine reductase, an enzyme involved in synthesis of a constituent unit of the bacterial peptidoglycan layer. However, nucleotide sequences obtained from all other genomic clones carrying IS*Rso3* showed no significant similarity to any known sequences in either nucleotide or protein databases. Nevertheless, the nucleotide sequences obtained from the five clones were utilised to investigate target site specificity of IS*Rso3* (Table 3.1).

6.3 IS*Rso4* was adjacent to a gene encoding porin protein in pIS102JI

The clone pIS102JI contained an approximately 6.6 kb *Eco*RI genomic fragment from ACH0158 and two subclones, pIS102JI.2v and pIS102JI.4v, were prepared using a unique *Sac*I site (Fig. 6.1A and B and Fig. 6.2A). The nucleotide sequence of both strands of a gene encoding porin protein in pIS102JI was obtained. However, IS*Rso4* was not a subject of sequencing reactions and therefore it is unknown whether the nucleotide sequences of the two copies of IS*Rso4* in clones pSV102 and pIS102JI are identical. In subclone pIS102JI.2v, IS*Rso4* was located at the 3' end of a gene encoding eubacterial surface porin protein (Fig. 6.1A). DNA sequences at the junction between IS*Rso4* and the gene for porin protein contained the CTAG direct repeat (DR) of IS*Rso4*

(Fig. 6.2A). The other copy of the CTAG DR in pIS102JI.2v was located between IS*Rso4* and an unknown gene sequence. Perfectly matched CTAG DRs were previously identified in the clone pSV102 (Fig. 3.7 and Table 3.1), suggesting that the CTAG sequence may be a specific target sequence for IS*Rso4* transposition.

A gene encoding porin protein in ACH0158 adjacent to IS*Rso4* in pIS102JI might have been a highly variable locus, as low G+C content (62 %) on the coding region and the atypical codon usage of the gene were determined (Table 6.1). Generally, G+C composition of genes in *R. solanacearum* is approximately 67 % (Palleroni, 1975), thus 62 % in the gene encoding porin protein is significantly low. In addition, atypical codon usages for the porin protein were clearly identified with an example of much less frequent CGC codon usage to encode arginine (A) residue in comparisons with other genes (Table 6.1).

Approximately 2.0 kb of DNA sequence was obtained for pIS102JI.2v and it contained an ORF potentially encoding 381 amino acid residues (Fig. 6.2A) which showed strong similarity to several known bacterial porin protein sequences including OpcP1 isolated from *Burkholderia cepacia* (47 % identical and 61 % similar, GenPep BAA09892) (Tsujimoto *et al.*, 1997) in a standard protein database search. However, the amino acid sequence of the ORF for porin protein in pIS102JI or of OpcP1 contained none of the characteristic motifs of porins known to be active in membrane transport (Jeanteur *et al.*, 1991). Therefore the *R. solanacearum* porin and OpcP1 may belong to a very unique porin protein family and may also play different roles.

Porin proteins are bacterial cell surface-exposed, antigenic molecules that are known to serve as phage and bacteriocin receptors (Hammond *et al.*, 1984). As a result of rapid evolutionary changes, large variations in surface-exposed molecules of many pathogenic bacteria are common, as heterogeneous antigenic molecules assist in pathogens evading host defence systems, and avoiding infection and damage by bacteriophages and bacteriocins (Meyer, 1991; Nikaido, 1992).

A 950 bp *Sac*I-*Pst*I fragment prepared from pIS102JI.2v was used as a porin gene probe (Fig. 6.2A) in Southern analyses of a wide range of *R. solanacearum* strains. Southern blots containing *Eco*RI-digested genomic DNA from 51 isolates were hybridised with the probe (Fig. 6.5B, E and H). Partially digested DNA sample of a biovar 3 strain (Fig. 6.5G) was not counted for the analysis. 51 strains contained representatives of the five

different biovars, a biovar N2 strain and an atypical biovar 2 strain (Table 4.1). Three equally loaded Southern membranes were hybridised, washed together and exposed to X-ray films for an equal amount of time.

Nine biovar 1 strains (Fig. 6.5B-lane 1-lane 9), six biovar 4 strains (Fig. 6.5B-lane 10-lane 16) and a biovar 5 strain CIP0277 (Fig. 6.5B-lane 17) were hybridised strongly with the probe, but significant *Eco*RI fragment length polymorphisms are observed. However, a ACH0262 strain (Fig. 6.5B-lane 14) showed no hybridisation with the probe. The sixteen biovar 2 strains tested showed highly similar hybridisation patterns (Fig. 6.5E-lane 18-lane 33). A biovar N2 strain CIP0358 (Fig. 6.5E-lane 34) showed weak hybridisation and the atypical biovar 2 strain ACH0732 (Fig. 6.5E-lane 35) showed a relatively small *Eco*RI fragment strongly hybridised with the probe, suggesting that those strains contain divergent porin genes. Most of the biovar 3 strains tested, with the exception of a ACH0190 strain, showed very similar hybridisation patterns of a single *Eco*RI fragments with the porin gene probe (Fig. 6.5H-lane 36-lane 51).

6.4 IS*Rso3* was adjacent to a gene encoding UDP-*N*-acetylenolpyruvoylglucosamine reductase in pSC15K

The clone pSC15K contained an approximately 4.0 kb *Eco*RI fragment and a subclone, pSC15K.K2, was prepared using two convenient *Sac*II sites (Fig. 6.3A and B and Fig. 6.4A). The nucleotide sequence of a gene flanking IS*Rso3* in pSC15K.K2 was determined but IS*Rso3* itself in the clone was not sequenced. Therefore, it is not known whether the clones pSV102 and pSC15K contained identical copies of IS*Rso3*. An approximately 1.9 kb of DNA sequence was obtained. It contained an ORF of 342 amino acid residues and IS*Rso3* was located downstream of the ORF. At the junction between IS*Rso3* and the ORF, a CTAG motif was found (Fig. 6.4A). However, the anticipated CTAG DR was replaced by TTAA (Table 3.1) at the other junction between IS*Rso3* and an unknown gene. Additionally mismatched DR sequences of IS*Rso3* were common in other genomic clones (Table 3.1), and perfectly matched CTAG DRs were detected only in pSV102 and pSC15J (Table 3.1). ~~However, as summarised in Table 3.1, at least one copy of the CTAG motif was found in most genomic clones carrying IS*Rso3* suggesting that CTAG may be a specific target sequence for IS*Rso3* insertions~~.

The ORF in pSC15K encoded a 342 aa peptide with strong similarities to several known UDP-*N*-acetylenolpyruvoylglucosamine reductase enzyme from various bacterial

species, including *Bordetella pertussis* (GenPep CAB41011) to which it showed 59 % sequence identity and 76 % similarity in a sequence alignment (Fig. 6.4C). A putative promoter region was identified and a stop codon for the ORF was found only 29 bp distant from the CTAG DR of IS*Rso3* (Fig. 6.4A).

A gene encoding UDP-*N*-acetylenolpyruvoylglucosamine reductase adjacent to IS*Rso3* in pSC15K contained significantly higher G+C content (70 %), and showed similar codon usages to an ORF encoding UDP-3-*O*-acyl-GlcNAc deacetylase identified in pIIIF2 (discussed in chapter 7) (Table 6.1).

A 310 bp *Sal*I-*Ava*I fragment from pSC15K was used as the gene-specific probe (Fig. 6.4A). Southern blots, that were utilised for porin gene probe, were hybridised with the probe after the blots were stripped to remove the previous hybridisations. All the biovar 1 strains (Fig. 6.5C-lane 1-lane 9), six biovar 4 strains (Fig. 6.5C-lane 10-lane 16) with an exception of a ACH0262 strain (Fig. 6.5C-lane 14) and a biovar 5 strain CIP0277 (Fig. 6.5C-lane 17) were hybridised strongly with the probe derived from pSC15K. Those strains also showed slight *Eco*RI restriction fragment length polymorphisms (Fig. 6.5C). The fifteen biovar 2 strains showed very similar hybridisation patterns (Fig. 6.5F-lane 18-lane 33) but a CIP0310 strain (Fig. 6.5F-lane 20) contained a larger *Eco*RI fragment hybridised. A biovar N2 strain CIP0358 (Fig. 6.5F-lane 34) showed weak hybridisation and the atypical biovar 2 strain ACH0732 (Fig. 6.5F-lane 35) contained a large *Eco*RI fragment strongly hybridised. Most of the biovar 3 strains, with exceptions of three strains (Fig. 6.5I-lane 43, 49 and 50) and a ACH0190 strain (Fig. 6.5I-lane 37), contained approximately 2 kb *Eco*RI fragments hybridised with the probe (Fig. 6.5I-lane 36-lane 51). As expected, the hybridisation results suggested that most of the strains tested contained a gene that was highly homologous to the probe because the signals survived stringent washes.

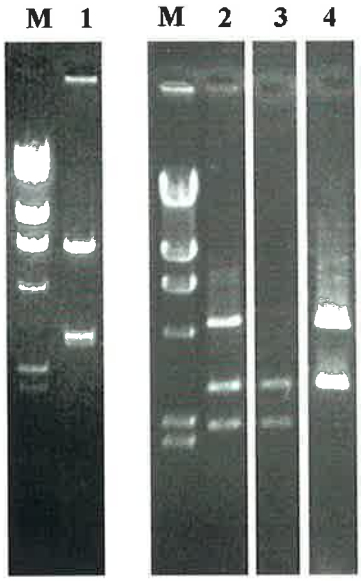
UDP-*N*-acetylenolpyruvoylglucosamine reductase is involved in reduction of the phosphoenolpyruvate moiety in order to yield UDP-*N*-acetylmuramic acid. *N*-acetylmuramic acid is a consisting unit of linear polysaccharide chains (glycan strands) to construct peptidoglycan layer, a crucial cell wall component in both gram-positive and gram-negative bacteria (Hammond *et al.*, 1984).

6.5 Summary and discussion

Sequencing of genomic regions adjacent to IS*Rso4* and IS*Rso3* in pIS102JI and pSC15K identified two genes encoding a porin protein and an enzyme involved in peptidoglycan layer construction. Genomic distributions of the genes were determined using Southern hybridisations of 51 diverse strains. Comparative analyses of G+C contents and codon usages of four genes isolated in ACH0158 suggested that genes encoding the porin protein and PhcA (see chapter 4) contained relatively low G+C contents (62 and 63 %, respectively) and atypical codon usages.

It is not very clear whether IS*Rso4* and IS*Rso3* flanked by genes encoding the porin protein and the enzyme required for the peptidoglycan layer formation play any roles and have associations, influencing to virulence and pathogenicity of *R. solanacearum*. Therefore, additional investigations are required to determine potential interaction and association between IS elements and the genes encoding putative virulence and pathogenicity factors of the bacterial species.

(A)



(B)

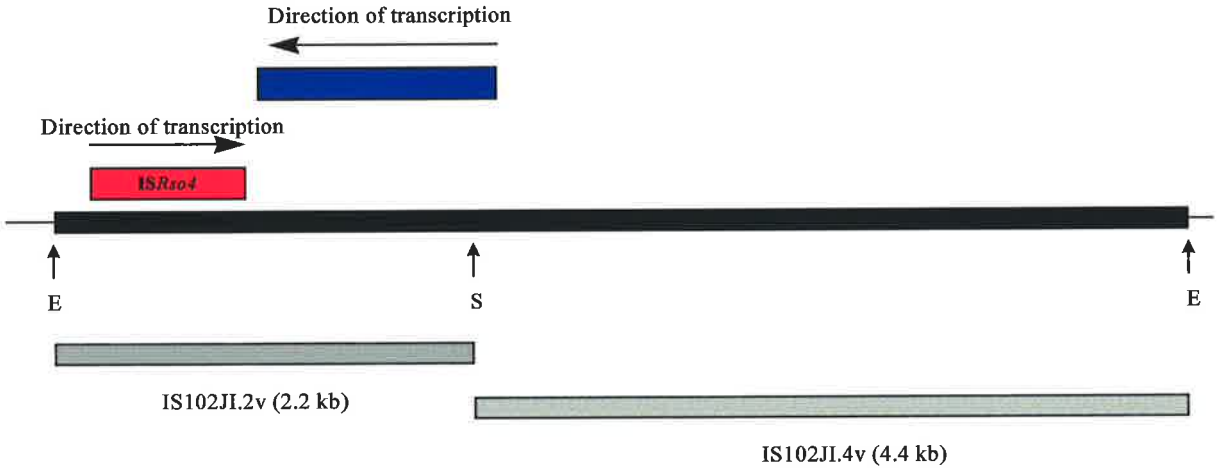


Figure 6.1

A. Agarose gel electrophoresis of a variety of endonucleases-digested pIS102JI and its subclones. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was included as a size marker (M).
   1. *Eco*RI-digested pIS102JI
   2. *Eco*RI and *Sac*I double-digested pIS102JI
   3. *Eco*RI and *Sac*I double-digested pIS102JI.2v
   4. *Eco*RI and *Sac*I double-digested pIS102JI.4v

B. A physical map of IS102JI with restriction sites for the endonuclease *Eco*RI (E) and *Sac*I (S) marked. The fragment sizes for the two subclones are given in kb. The locations and transcriptional directions of IS*Rso4* (red box) and the gene for encoding a porin protein (blue box) are indicated.

Figure 6.2

A. Complete nucleotide and deduced amino acid sequences of a gene encoding porin protein. The nucleotide sequences of the gene-derived primers for sequencing are underlined. The identity and direction of each primer are indicated. The putative -35, -10, RBS and terminator are underlined. A *Sac*I site utilised for the subcloning and the gene-specific probe preparation is marked. A *Pst*I site utilisd for the gene-specific probe preparation is marked. The CTAG DR of IS*Rso4* is in bold.

B. The identities and nucleotide sequences of the gene-derived primers for sequencing.

C. Alignment of amino acid sequences. Rs refers to amino acid sequences of the ORF encoding porin protein in pIS102JI. Bc refers to amino acid sequences of OpcP1 (GenPep accession no. BAA09892).

(A)

```
   1  GTATTGCACCGTCACCAACCGGATCGCGGAAGGCAAGAACGGCATCACGCTGTTCTTCAT

  61  GCTGACGGCGCTGGCGCTGTGGATCAAGTACTTCGCCACGGGCGGCGGCGCGATGCATTG

 121  GAGTTACGAGGCATCATCTATCTCGCCCTGCGGCTTCGGCGATGGGGTTCGGCTATGCGG

 181  CCTGGAATGTGGGCATCCTGCATGGCAACGTGACTGTCCTTGCCGGCGCTTCGTACTTCA

 241  TCCCAGTGTTTTCCGCTGCCCTGGCGGCCACGCTGTTGCACACGCCGTTGTCGTTCGCCT

 301  TTTGGCGGGGCGCATCCATGGTCTGTGCGGGGTCGATCCTGTGTTGGCTCGCAACACGCG
                                                      Porin2  →

 361  GGCAGCGTTCACGGGGGCCTTCCGCGCCGAGCCGTTCCCAACCGCGGGAGCGAGCCCATG

 421  GACAGTGAGTTCTGAGCTGGCGTCCGTGTGCCAGGGGCACCTGCCCCCAATCGATTGACT
                                                            -35

 481  TCACGCAAGCGGCCGGCTGTCCGCTGAGATGGCCCGCTTGCAACCATCACGTTGACCTCG
                                  -10

 541  CCGAGCGATCAGCGCGGGCGACGTCGTACCCAAGCATCATCAAACCATCAGGACAAGCCG

 601  ATGTCCGAGGAGACCGATATGAAGACGCGGTTGTTTGTGGCATTCGCTTCGTGCGCCATC
           RBS      M  K  T  R  L  F  V  A  F  A  S  C  A  I

 661  GCTGCGCCGACATTCGCGCAGAGCTCCGTGACCTTGTATGGCGTGCTTGACGAAGGGCTC
                          SacI
      A  A  P  T  F  A  Q  S  S  V  T  L  Y  G  V  L  D  E  G  L

 721  AACTACACGACCAACGTCGGGGGGCACAGTCAGGTCGCGATGGCCAGCGGGTTTCCGCAC
      N  Y  T  T  N  V  G  G  H  S  Q  V  A  M  A  S  G  F  P  H

 781  GGCAGTCGATGGGGCGTGAAGGGGGCCGAAGACCTTGGCGGCGGTGCCAAGACCGTCTTC
      G  S  R  W  G  V  K  G  A  E  D  L  G  G  G  A  K  T  V  F

 841  CAGCTTGAGAACGGGTTCGATGTGGATACCGGCCGGGCCTTTCAAGGCGGACTGCTGTTC
      Q  L  E  N  G  F  D  V  D  T  G  R  A  F  Q  G  G  L  L  F

 901  GGGCGTCAGGCGTACATGGGCTTGTCCAGCAACACGGTGGGCACGCTCACCGTCGGGCGT
      G  R  Q  A  Y  M  G  L  S  S  N  T  V  G  T  L  T  V  G  R

 961  CAGTACGATTCCGTGGTGGACTACCTTGCACAGACCTCCGCCGGCGGTAGTTGGGGCGGC
      Q  Y  D  S  V  V  D  Y  L  A  Q  T  S  A  G  G  S  W  G  G

1021  TACATGTTTGCCCACCCGTACGACAACGACAACCTGATCAACACGTTCCGGGCCAACAAC
      Y  M  F  A  H  P  Y  D  N  D  N  L  I  N  T  F  R  A  N  N

1081  ACGGTCAAGTACACCAGCCCGACACTGGGCGGCCTGAAGTTCGGGGCCACGTACAGCTTC
      T  V  K  Y  T  S  P  T  L  G  G  L  K  F  G  A  T  Y  S  F

1141  AGCAACGACGCGGGCTTTGCGAACAATCGGCTGGTGAGCGTGGGAAGCCAGTACACGGTT
      S  N  D  A  G  F  A  N  N  R  L  V  S  V  G  S  Q  Y  T  V

1201  GGCGGCTTGCTGCTGGCTGCCGCCTACCTTCAGGCTGACCATCCTTCCGCCACCGCATAC
      G  G  L  L  L  A  A  A  Y  L  Q  A  D  H  P  S  A  T  A  Y

1261  GGCGCGATCAACAACAGCGGCGACCAGAACCTGCTTGGCAGCCGGCTGAGGATCTTTGGC
                        Porin1  →
      G  A  I  N  N  S  G  D  Q  N  L  L  G  S  R  L  R  I  F  G

1321  GCCGGCGCTACGTACACCTTCGGTCAGGCAATCGTGGGCCTCACCTATGCCGATACGAAT
      A  G  A  T  Y  T  F  G  Q  A  I  V  G  L  T  Y  A  D  T  N

1381  GTCGCCGACCCGCAAAGCTCCGGCTACGTCGGGCCGATCACTGCGCCTGCGGGACGGCTC
      V  A  D  P  Q  S  S  G  Y  V  G  P  I  T  A  P  A  G  R  L
```

```
1441    TCTTCGCTGCGCTTCCGGAACATCGAGGTCAATGCCAAATATCAGGTGGGCGCATCGTAT
        S  S  L  R  F  R  N  I  E  V  N  A  K  Y  Q  V  G  A  S  Y

1501    TGGCTGGGCGCGATGTACACCTATACGCGTGCCAACTTCAGCGCAACGCCGGGAGCCAGG
        W  L  G  A  M  Y  T  Y  T  R  A  N  F  S  A  T  P  G  A  R

1561    CATCCCACCTACCATTCGGTCGGGCTGATGGCGGATTACCTGTTGTCGAAGCGGACCGAT
        H  P  T  Y  H  S  V  G  L  M  A  D  Y  L  L  S  K  R  T  D

1621    GTTTACCTGCAGGGCATGTATCAGCATGTCGGCGGTGATGCCACCGGCTCCATTCTGGAT
              PstI
        V  Y  L  Q  G  M  Y  Q  H  V  G  G  D  A  T  G  S  I  L  D

1681    GCCGCCTACGTAGCCGGCGCGGAGAATGTCTCATCCAATCGCAATCAGCTATTGCTCCGC
        A  A  Y  Y  V  A  G  A  E  N  V  S  S  N  R  N  Q  L  L  L  R

1741    GCAGGCGTTCGCCATTTCTTTTAAGGTTTTCGGCGAATGCCTTTTCTGGGCCGCACTCAC
        A  G  V  R  H  F  F  *

1801    CTCCTTGAGGGAGAACGGCTTGACCAGGCCATCATCCGCGCCGTGCGTAAAGCTGTTGAG

1861    GCACGGTGGACACAGTGAAATTACGGATCAAGGCAAGGAAAAAAAGCGCATGGCCGGAAC
                              terminator

1921    CGAGGAATCTGGTCCGTTCTGAGCGTTGAACCGGTTGGAGCGCCTGAAAAATAGGCCTCG

1981    TAAGCTAG
```

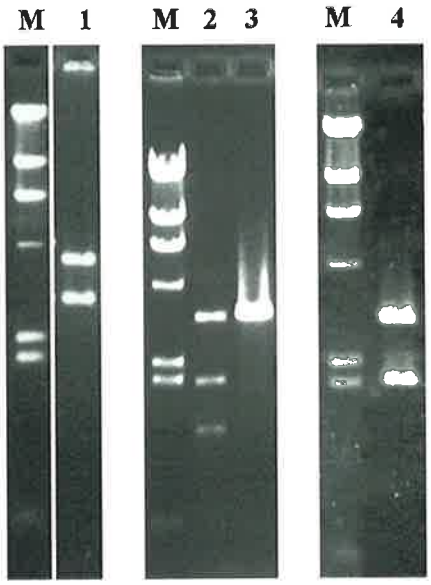
(B)

oligo names	sequences (5′ → 3′)
Porin1	CAACAGCGGCGACCAGAACC
Porin2	GGCTCGCAACACGCGGG

(C)

```
Rs      MKTRLFVAFASCAIAAPTFAQSSVTLYGVLDEGLNYTTNVGGHSQVAMASGFPHGSRWGV
        |..  |:||  |...:|...  |||||||||||.|:.|  .||||.|   .|:||.  ..||:|:
Bc      MNKTLIVAAAAASFATVAHAQSSVTLYGVLDAGITYQSNVGGKSLWSMGSGV.YQSRFGL

Rs      KGAEDLGGGAKTVFQLENGFDVDTGRAFQGGLLFGRQAYMGLSSNTVGTLTVGRQYDSVV
        :|.||||||  |.:|  ||.||:::.|:   ..|  :|.|||::||||.   ||:|:|:|||..
Bc      RGSEDLGGGLKAIFTLESGFNIGNGKLGNNGGMFNRQAFVGLSSQ.YGTVTLGKQYDATQ

Rs      DYLAQTSAGGSWGGYMFAHPYDNDNLI..NTFRANNTVKYTSPTLGGLKFGATYSFSNDA
        ||||.  .|.|||||   ||||.:||.|    ...   ||.:||||:..:||.||:||||||:.
Bc      DYLAPLTATGSWGGTYFAHPINNDRLSTNGDVALNNSIKYTSANYAGLQFGGTYSFSNNT

Rs      GFANNRLVSVGSQYTVGGLLLAAAYLQADHPSATAYGAINNSGDQNLLGSRLRIFGAGAT
        .|:|||   |.|..| ..||  |||||  ||: ..:|.  ::  ..  ..|     :|:|.:||:|.
Bc      NFGNNRRYSAGAAYQFQGLKLAAAYSQANLGDGTNTSGASSINAQ....GRVRTYGAAAG

Rs      YTFGQAIVGLTYADTNVADPQSSGYVGPITAPAGRLSSLRFRNIEVNAKYQVGASYWLGA
        |.||.|  ||  .:.:              :.|...|.  ::.||   |.||||||.:.:..  |||
Bc      YAFGPAQVGAAWTQ.............SRIDNQANGVGTLRADNYEVNAKYNLTPALGLGA

Rs      MYTYTRANFSATPGARHPTYHSVGLMADYLLSKRTDVYLQGMYQHVGGDATGSILDAAYV
         ||||.|......: ::.     .|: ||| ||||||||  |::||:   |. ...|::.:.
Bc      AYTYTNAKVNNGSSHWNQ....FGVQADYALSKRTDVYAQAVYQR..GAKGNNIVGTGIY

Rs      AGAENV..SSNRNQLLLRAGVRHFF
        .|.:..  ||. ||    .|:|| |
Bc      NGDNTTASSSSINQTAATVGLRHRF
```

(A)



(B)

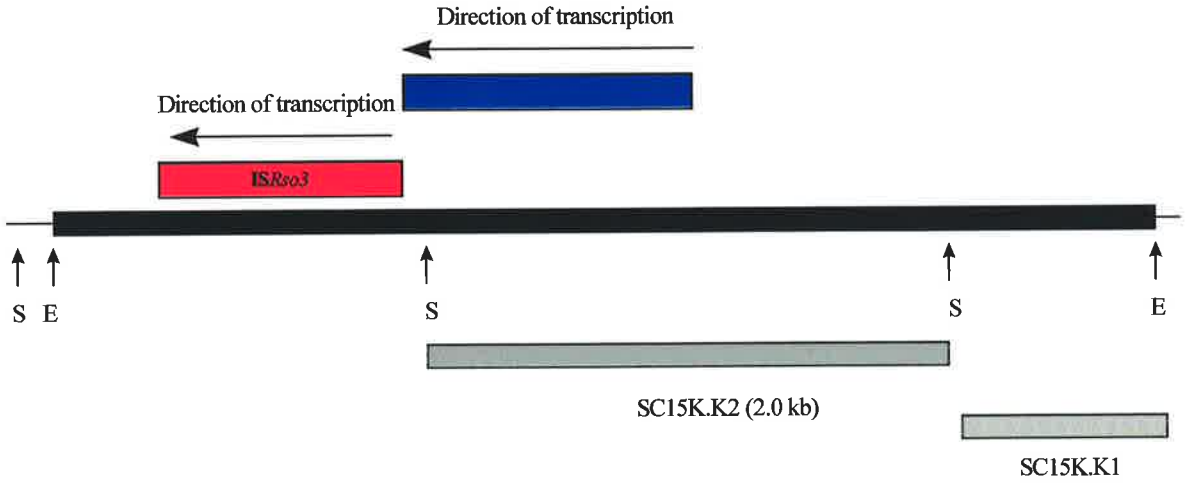


Figure 6.3

A. Agarose gel electrophoresis of a variety of restriction endonucleases-digested pSC15K and its subclones. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was included as a size marker (M).
    1. *Eco*RI-digested pSC15K
    2. *Sac*II-digested pSC15K
    3. *Sac*II-digested pSC15K.K1
    4. *Sac*II-digested pSC15K.K2

B. A physical map of SC15K with restriction sites for the endonuclease *Eco*RI (E) and *Sac*II (S) marked. The fragment size for the subclone is given in kb. The locations and transcriptional directions of IS*Rso3* (red box) and the gene for encoding the UDP-*N*-acetylenolpyruvoylglucosamine reductase (blue box) are indicated.

Figure 6.4

A. Complete nucleotide and deduced amino acid sequence of a gene encoding UDP-*N*-acetylenolpyruvoylglucosamine reductase. The nucleotide sequences of the gene-derived primers used for sequencing are underlined. The identity and direction of each primer are indicated. The putative -35, -10, RBS and terminator are underlined. *Sac*II sites utilised for the subcloning are marked. *Sal*I and *Ava*I sites utilisd for the gene-specific probe preparation are marked. The CTAG DR of IS*Rso3* is in bold.

B. The identities and nucleotide sequences of the gene-derived primers for sequencing.

C. Alignment of amino acid sequences. Rs refers to amino acid sequences of the ORF encoding the UDP-*N*-acetylenolpyruvoylglucosamine reductase in pSC15K. Bp refers to amino acid sequences of an ORF encoding the homologous protein in *Bordetella pertussis* (GenPep accession no. CAB41011).

(A)

```
   1  CCGCGGCAATATTGCCAAGCTGCTGGCGGGCAAGGAAAGCCGCATCGGCGAGAAGAAAAA
      SacII

  61  GCCCACCTGACCCGCGCAGTTCCGCGCGCACCGGATGCGGCCGTCCGCGCCGCCGTGCTC

 121  AAAGCGCGCAAGGACCGGATGCGCGCCGGCTCCGCCAAGCGCCGGTTCCGCCTGCCGTTC

 181  GCGGAGGCCACCATCGCCGCCGATCCAGCGCAGAAGGTGAAGAGGCTGTCGCAGTGAGCG

 241  CAGCCCGCGCGCCAAAAGCAAACGCCCCGCGCGTGCGGGGCGTTCTTGCCTGAACCCGGG

 301  CGTGGCCGGCGACTCAATCGCGGAAGTTGTTGAAATCCAGCGGCGTGTCGGTCACGTCCT

 361  TGCGCAGCATGGCGATCACGCTTTGCAGGTCGTCGCGCTTGGTGCCGGACACACGCACCG

 421  CGTCGCCCTGGATGCTGCCCTGGACCTTGATCTTGCTGTCCTTGATCAGCTTGACGATCT

 481  TCTTGGCCAGATCGCCCGAGACGCCCTTCTTGATCTTGACGACCTGCTTCATCTTGTCGC
                                                    END2  →

 541  CGCCGATCTTCTGCTTGTCCTGGTAGTCCAGGAAGCGCACGTCGACGTTGCGCTTGGCGA

 601  GCTTGGCCAGCAACACTTCCTTGACCTGGCCGAGCTTGAAATCGTCGTCGGCGAACAGCG

 661  TCAGCTCCTGCTCCTTGTGCTCGACACGGGCGTCCGAGCCCTTGAAGTCGAAACGCGTCG

 721  AGATTTCCTTGTTGGCCTGCTCCACGGCGTTCTTCAGTTCCACCATGTTGGCCTCGCATA
                                      -35                      -10

 781  CCACGTCGAACGACGGCATCGTTTTCTCCTTGCAGACTTGAATAGGGGTTCCTGCGGCGG

 841  CCCGCGCTCGGGCCACCTGACCGCCATGGCTCTTATAATTCCAGCTTTTCCAGCCCACCC
                                                    RBS

 901  TGCGCCATGGCGTTGCTCGACCCGCATTATCCCCTAGGCCGGCACAATACGTTCCGTTTC
            M  A  L  L  D  P  H  Y  P  L  G  R  H  N  T  F  R  F

 961  GAGGCCGCCGCCCGCTATGCCGCGCACGTGCGCGCACCGCAGGACATTGCCGAGGCACTG
      E  A  A  A  R  Y  A  A  H  V  R  A  P  Q  D  I  A  E  A  L

1021  GCCGACCCGCGCGTGCGGGGCCTGCCGGTGCTGGTGCTGGGCGGCGGCAGCAACATCGTG
      A  D  P  R  V  R  G  L  P  V  L  V  L  G  G  G  S  N  I  V

1081  CTCACGCGCGACTTCGACGGTCTGGTACTGCTGATGGAGATTCCCGGCGTGCAGGTCGGC
      L  T  R  D  F  D  G  L  V  L  L  M  E  I  P  G  V  Q  V  G

1141  CGGGCGACCCTGGAAGGACGCACCGTCCACACCGTCACCGCCGGCGCGGGTGAATCGTGG
      R  A  T  L  E  G  R  T  V  H  T  V  T  A  G  A  G  E  S  W

1201  CACGGCCTCGTCGCCTACACCGTGTCCAACGGGCTGCCCGGGCTGGAAAACCTGGCGCTG
      H  G  L  V  A  Y  T  V  S  N  G  L  P  G  L  E  N  L  A  L

1261  ATCCCCGGCACGGTCGGTGCCGCCCCCATCCAGAACATCGGCGCATACGGTGTCGAGATC
      I  P  G  T  V  G  A  A  P  I  Q  N  I  G  A  Y  G  V  E  I

1321  AAGGACCGCTTCCACTCCCTGCGCGCCTACGACCGCCACGCTGGTGAGTTCGTCACGCTG
      K  D  R  F  H  S  L  R  A  Y  D  R  H  A  G  E  F  V  T  L

1381  GATGCCGCCGACTGCGCCTTCGGCTACCGCGACAGCCTCTTCAAGCGCGCCGGCGCCGAC
      D  A  A  D  C  A  F  G  Y  R  D  S  L  F  K  R  A  G  A  D

1441  CGCTACGTCATCACCGAAGTCACCTTCGCGCTGCCGGTGGACTGGCAGCCCGACACGCAT
      R  Y  V  I  T  E  V  T  F  A  L  P  V  D  W  Q  P  D  T  H

1501  TACGCGGAGCTGGCCCGCGAACTGGCCGCGCGGGCCATCGCCGCGCCGACGGCCCAGGAC
      Y  A  E  L  A  R  E  L  A  A  R  A  I  A  A  P  T  A  Q  D
```

```
1561    ATCTTCGACGCGGTGGTCGCCATCCGCCGCCGCAAGCTGCCGGACCCGGCCGAGATCGGC
        I  F  D  A  V  V  A  I  R  R  R  K  L  P  D  P  A  E  I  G

1621    AACGCGGGCAGCTTCTTCAAGAACCCCATCGTCGACGCCGCCACGCGCGATGCGCTGCTG
                ← END1                    SalI
        N  A  G  S  F  F  K  N  P  I  V  D  A  A  T  R  D  A  L  L

1681    GCACGGTTCCCCGGCCTGGTCGGCTACGCGCAGCCAGACGGCGGCTACAAGCTCGCCGCC
        A  R  F  P  G  L  V  G  Y  A  Q  P  D  G  G  Y  K  L  A  A

1741    GGCTGGATGATCGACCAGTGCGGTTTCAAGGGCCGGCAGAGCGGCGCGGTGGGCGTGTAC
        G  W  M  I  D  Q  C  G  F  K  G  R  Q  S  G  A  V  G  V  Y

1801    GACAAGCAGGCGCTGGTGCTGGTGCACCGGGGCGGCGGCAGCGCCGTGCAGCTGATGACG
        D  K  Q  A  L  V  L  V  H  R  G  G  G  S  A  V  Q  L  M  T

1861    CTGGCACGCGAGATCCAGGACACGGTCCATGCACGCTTCGGCGTGCGTATCGAGCCCGAG
        L  A  R  E  I  Q  D  T  V  H  A  R  F  G  V  R  I  E  P  E

1921    CCCGTCGTCATCTGACGGAGCGGCACCGCGGGGCCCCCGGGGAGCCTAG
                                  SacII      AvaI
        P  V  V  I  *
```

(B)

oligo names	sequences (5′ → 3′)
END1	GTTCTTGAAGAAGCTGCC
END2	ACCTGCTTCATCTTGTCGCC

(c)

```
Rs      ....MALLDPHYPLGRH.......NTFRFEA.AARYAAHVRAPQDIAEALADPRVRGLPV
Bp      MSTVPARIEPVAPLAPQAQDLRCFNTLGLASHAPAFVALTEPSQLPALSALAPRFRQL..

Rs      LVLGGGSNIVLTRDFDGLVLLMEIPGVQVGRATLEGRTVHTVTAGAGESWHGLVAYTVSN
Bp      VVLGGGSNVVLPASIDGLVAQVRLPGVRLVGQCADAWVVE...AAAGENWHGFVTACVDN

Rs      GLPGLENLALIPGTVGAAPIQNIGAYGVEIKDRFHSLRAYDRHAGEFVTLDAADCAFGYR
Bp      GWDGLENLALIPGTVGAAPVQNIGAYGVELADRFHSLTAWDVKGGRWVEMGAAECRFAYR

Rs      DSLFKRAGADRYVITEVTFALPVDWQPDTHYAELARELAARAIAAPTAQDIFDAVVAIRR
Bp      DSFFKHQEPGAWVIGSVRFALPRPWQPVLDYPDLQRH.AALDGAAPTARAVYDAVCAIRR

Rs      RKLPDPAEIGNAGSFFKNPIVDAATRDALLARFPGLVGYAQPDGGYKLAAGWMIDQCGFK
Bp      AKLPDPAVVGNAGSFFKNPLVDAGTRQALLGRFPGLVSYPQPDGRYKLAAGWLIDQCGWK

Rs      GRQSGAVGVYDKQALVLVHRGGGSAVQLMTLAREIQDTVHARFGVRIEPEPVVI
Bp      GRQLGAAGVHDRQALVLVNRGGAQARDIMALAAAIQGDVERRYGVRLEPEPVVVPAR
```

Table 6.1

A summary of codon usages of four genes isolated from *R. solanacearum* strain ACH0158. LIP refers to the codon usage of an ORF encoding UDP-3-*O*-acyl-GlcNAc deacetylase found in pIIIF2 (discussed in chapter 7). PhcA refers to the codon usage of an ORF encoding PhcA found in pWT (discussed in chapter 4). PEP refers to the codon usage of an ORF encoding UDP-*N*-acetylenolpyruvoylglucosamine reductase found in pSC15K (discussed in this chapter). POR refers to the codon usage of an ORF encoding porin protein found in pIS102JI (discussed in this chapter). The overall G+C contents of the coding regions of the four genes is also shown.

		FRACTION			
AmAcid	Codon	LIP	PhcA	PEP	POR
Gly	GGG	0.21	0.00	0.06	0.22
Gly	GGA	0.00	0.00	0.03	0.08
Gly	GGT	0.08	0.28	0.17	0.08
Gly	GGC	0.71	0.72	0.75	0.62
Glu	GAG	0.55	0.57	0.67	0.71
Glu	GAA	0.45	0.43	0.33	0.29
Asp	GAT	0.24	0.65	0.09	0.50
Asp	GAC	0.76	0.35	0.91	0.50
Val	GTG	0.52	0.60	0.48	0.44
Val	GTA	0.00	0.04	0.03	0.04
Val	GTT	0.00	0.04	0.00	0.11
Val	GTC	0.48	0.32	0.48	0.41
Ala	GCG	0.41	0.58	0.30	0.26
Ala	GCA	0.08	0.03	0.12	0.15
Ala	GCT	0.03	0.08	0.02	0.11
Ala	GCC	0.49	0.31	0.56	0.48
Arg	AGG	0.00	0.00	0.00	0.11
Arg	AGA	0.00	0.00	0.00	0.00
Arg	CGG	0.10	0.15	0.26	0.44
Arg	CGA	0.00	0.00	0.00	0.06
Arg	CGT	0.05	0.20	0.07	0.17
Arg	CGC	0.85	0.65	0.67	0.22
Ser	AGT	0.00	0.09	0.00	0.10
Ser	AGC	0.31	0.14	0.62	0.39
Ser	TCG	0.23	0.41	0.12	0.16
Ser	TCA	0.00	0.00	0.00	0.03
Ser	TCT	0.00	0.00	0.00	0.03
Ser	TCC	0.46	0.36	0.25	0.29
Lys	AAG	0.77	0.88	1.00	0.86
Lys	AAA	0.23	0.12	0.00	0.14
Asn	AAT	0.00	0.15	0.14	0.27
Asn	AAC	1.00	0.85	0.86	0.73
Met	ATG	1.00	1.00	1.00	1.00
Ile	ATA	0.00	0.00	0.00	0.00
Ile	ATT	0.15	0.20	0.12	0.12
Ile	ATC	0.85	0.80	0.88	0.88
Thr	ACG	0.47	0.38	0.56	0.40
Thr	ACA	0.00	0.05	0.00	0.07
Thr	ACT	0.00	0.00	0.00	0.03
Thr	ACC	0.53	0.57	0.44	0.50
Trp	TGG	0.00	1.00	1.00	1.00
Opal	TGA	1.00	1.00	1.00	0.00
Cys	TGT	0.00	0.00	0.00	0.00
Cys	TGC	1.00	1.00	1.00	1.00
End	TAG	0.00	0.00	0.00	0.00
End	TAA	0.00	0.00	0.00	1.00
Tyr	TAT	0.38	0.20	0.18	0.25
Tyr	TAC	0.62	0.80	0.82	0.75

		FRACTION			
AmAcid	Codon	LIP	PhcA	PEP	POR
Leu	TTG	0.19	0.17	0.03	0.18
Leu	TTA	0.00	0.00	0.00	0.00
Leu	CTG	0.72	0.61	0.78	0.47
Leu	CTA	0.00	0.00	0.03	0.03
Leu	CTT	0.03	0.02	0.00	0.18
Leu	CTC	0.06	0.20	0.16	0.15
Phe	TTT	0.12	0.07	0.00	0.37
Phe	TTC	0.88	0.93	1.00	0.63
Gln	CAG	1.00	1.00	1.00	0.88
Gln	CAA	0.00	0.00	0.00	0.12
His	CAT	0.25	0.45	0.30	0.62
His	CAC	0.75	0.55	0.70	0.38
Pro	CCG	0.75	0.67	0.42	0.70
Pro	CCA	0.00	0.05	0.05	0.00
Pro	CCT	0.00	0.05	0.00	0.20
Pro	CCC	0.25	0.24	0.53	0.10
G+C contents		65 %	63 %	70 %	62 %

Figure 6.5

A, D and G. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of 51 *R. solanacearum* strains, stained with ethidium bromide.

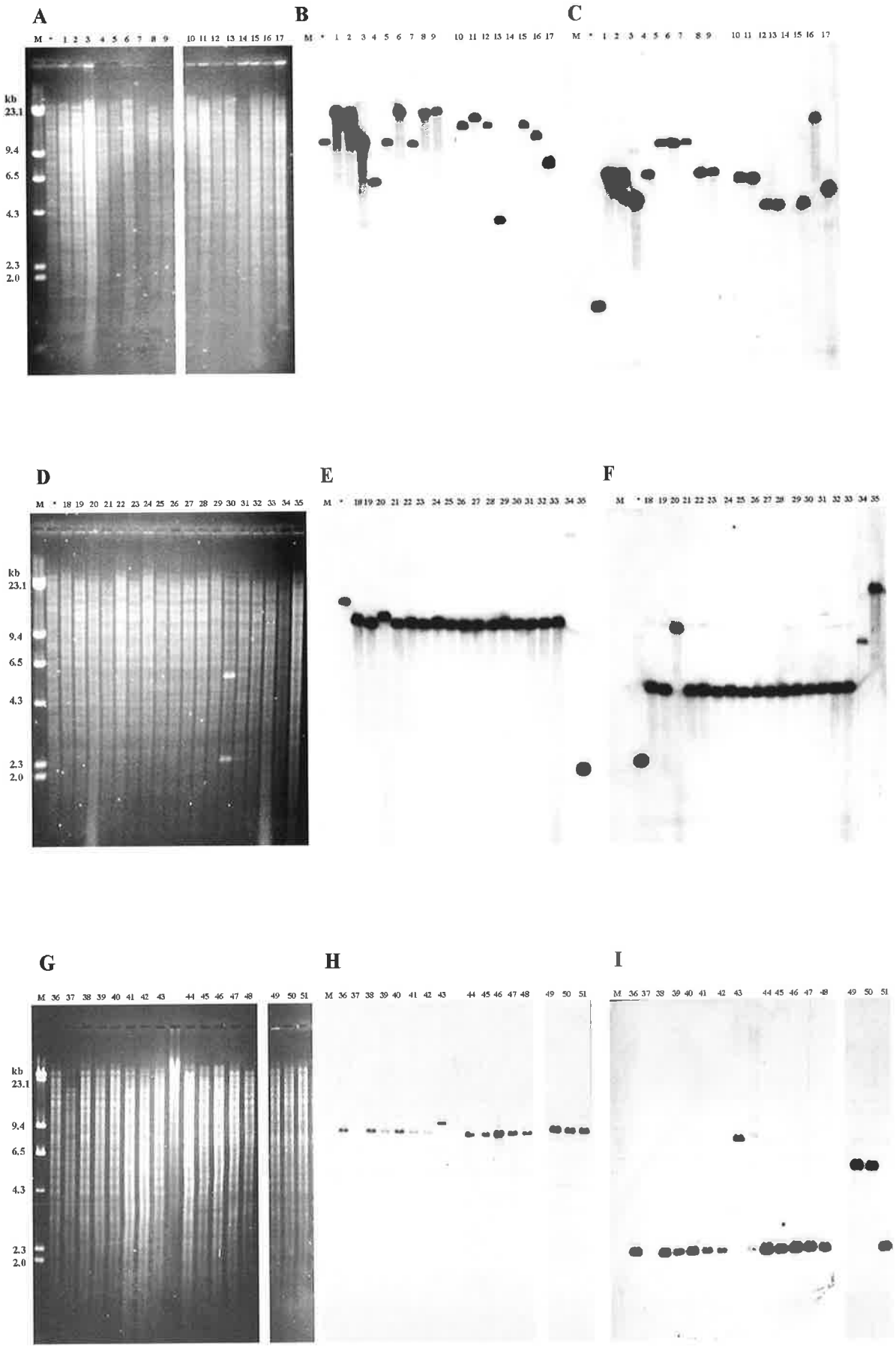
M. Molecular size marker of λDNA digested with *Hind*III

1. CIP0007 strain
2. CIP0009 strain
3. CIP0066 strain
4. CIP0070 strain
5. CIP0120 strain
6. CIP0218 strain
7. CIP0301 strain
8. CIP0418 strain
9. CIP0419 strain
10. CIP0121 strain
11. CIP0264 strain
12. CIP0283 strain
13. PD1431 strain
14. ACH0262 strain
15. ACH0319 strain
16. ACH1066 strain
17. CIP0277 strain
18. ACH0158 strain
19. CIP0117 strain
20. CIP0310 strain
21. CIP00359 strain
22. CIP0402 strain
23. CIP0403 strain
24. CIP0405 strain
25. South Africa 28 strain
26. South Africa 45 strain
27. South Africa 46 strain
28. South Africa 49 strain
29. South Africa 65 strain
30. PD1427 strain
31. PD1435 strain
32. ACH0127R strain
33. ACH1068 strain
34. CIP0358 strain
35. ACH0732 strain
36. ACH0171 strain
37. ACH0190 strain
38. ACH0369 strain
39. ACH0574 strain
40. ACH0671 strain
41. ACH1024 strain
42. ACH1064 strain
43. ACH1070 strain
44. South Africa 8 strain
45. South Africa 12 strain
46. South Africa 53 strain
47. South Africa 93 strain
48. South Africa 112 strain
49. CIP0065 strain
50. CIP0284 strain
51. PD1437 strain

B, E and H. Southern hybridisation of the gels in A, D and G with a *Sac*I-*Pst*I fragment prepared from pIS102JI.2v containing a part of the gene encoding the porin protein.

C, F and T. Southern hybridisation of the gels in A, D and G with a *Sal*I-*Ava*I fragment prepared from pSC15K containing a part of the gene encoding the UDP-*N*-acetylenolpyruvoylglucosamine reductase.

The lane between lane 43 and 44 was disregarded because the DNA sample was partially digested. The asterisk (*) indicates the lane containing ACH0171 DNA sample.


A
B
C
D
E
F
G
H
I
kb
23.1
9.4
6.5
4.3
2.3
2.0

CHAPTER 7

Detection and quantification of *Ralstonia solanacearum* DNA by quantitative-PCR method

7.1 Introduction

In potato, latently infected tubers by *R. solanacearum* have been a major cause of crop loss and means of dispersal of bacterial wilt (Ciampi and Sequeira, 1980; Ciampi *et al.*, 1980; Nyangeri *et al.*, 1984; Sunaina *et al.*, 1989; Kelman *et al.*, 1994; Hayward, 1991). Recent outbreaks occurred on potato in Europe (Caruso *et al.*, 1997) emphasised the urgent need for a sensitive and specific detection method to facilitate monitoring of overt and latent infections of *R. solanacearum* in crop planting material, soil and water samples. Latent infections are known to occur in susceptible and in some resistant cultivars of tomato, eggplant and pepper (Liao *et al.*, 1998). Therefore, selection to prevent latent infection has been included as a new parameter for breeding efforts (Liao *et al.*, 1998), but the detection of latent infection is difficult and time-consuming.

If suitable target sequences can be identified, PCR technology offers a sensitive and simple approach to the identification and quantification of a pathogenic species and in some cases, specific strains of the pathogen. In the quantitative PCR approach adopted here, a standard DNA molecule is constructed and added in known amounts to PCR reactions containing unknown amounts of target DNA. The standard DNA construct competes with the target DNA and therefore is often called competitor DNA. The standard construct contains the same primer binding sites as the target DNA molecule. However, the size of the PCR products of the target and competitor DNA molecules are different, so that they may be resolved by gel electrophoresis (Celi *et al.*, 1993). Provided the amplification efficiency of target and standard is similar, the ratio of the two will remain constant throughout the PCR process. Thus, determining the ratio of target to standard product at the end of the reaction, and knowing the starting amount of

standard DNA spiked in, allows the amount of unknown target DNA to be accurately and sensitively estimated (Zachar *et al.*, 1993).

The clone p092-03B5, derived from ACH092 of *R. solanacearum*, contained sequences that were present and highly conserved in all the strains of *R. solanacearum* examined and, with 2 exceptions, in no other species of bacteria tested (Opina *et al.*, 1997). The primer pair, AU759f and AU760r, designed from the cloned sequence amplified a 281 bp region of DNA in all the strains of *R. solanacearum* tested and in the closely related *Pseudomonas syzygii* and *P. celebensis* (Opina *et al.*, 1997). However, the 281 bp fragment was not amplified in several closely and distantly related bacterial species tested, including *P. aeruginosa*, *P. corrugata*, *P. putida*, *P. fluorescens* and *Burkholderia cepacia* (Opina *et al.*, 1997). In addition, Southern blot experiments on a diverse range of plant pathogenic bacterial species, with the 281 bp fragment as a probe, showed universal hybridisation to *R. solanacearum* but none to DNA of any other species, suggesting complete specificity of the 281 bp fragment to the *R. solanacearum* genome (Opina *et al.*, 1997). The bacterial species tested in the Southern experiments were *Xanthomonas campestris* pv. *vesicatoria*, *X. campestris* pv. *campestris*, *X. campestris* pv. *glycenia*, *Erwinia chrysanthemi*, *Burkholderia cepacia* and *B. gladioli* (Opina *et al.*, 1997). The specificity of the AU759f/AU760r primer pair was further confirmed when DNA samples of 92 isolates of soil bacteria, including 12 isolates of *R. solanacearum*, were tested (Ito *et al.*, 1998). Ito *et al.* (1998) reported that the 281 bp fragment was PCR-amplified only in the DNA samples of the *R. solanacearum* isolates. 14 isolates that were not *R. solanacearum* did amplify a variety of PCR products but these did not hybridise with the 281 bp fragment in Southern hybridisations (Ito *et al.*, 1998).

A second clone, pBs2.10, which was isolated from ACH0158, was shown to contain genomic sequences diagnostic of biovar 2 strains of *R. solanacearum*. Sequence of Bs2.10 showed no homology with any known gene within the data base. The primer pair, AU630f and AU631r, derived from the clone amplified a 307 bp fragment from ACH0158 DNA (Opina *et al.*, 1997). Fegan *et al.* (1998) tested over 200 *R. solanacearum* strains and all but one biovar 2 strain amplified the 307 bp fragment. In addition, three biovar 1 strains, that showed the same rep-PCR profiles as shown by biovar 2 strains and one of these has been previously recognised as being atypical, yielded the amplified fragment (Fegan *et al.*, 1998).

Therefore, with very few exceptions, these two sequences are species- and strain-specific respectively and they have the potential to be developed as useful probes in the detection and diagnosis of *R. solanacearum* in seeds, vegetative propagules, irrigation water and soil.

This chapter describes the construction of two standard DNA molecules from p092-03B5 and pBs2.10, and PCR experiments testing their use in quantifying *R. solanacearum* DNA. The nature of the 281 bp sequence and its genomic context is fully described.

7.2 092-03B5 is a promoter region of *lpxC* gene for encoding UDP-3-*O*-acyl-GlcNAc deacetylase in *R. solanacearum*

A *Sac*I genomic clone pIIIF2 was isolated from a genomic DNA library of ACH0158 (Fig. 7.1A'). Three internal *Eco*RI sites in the 9.3 kb insert were used to generate four subclones, pIIIF2.1v, pIIIF2.2v, pIIIF2.3v and pIIIF2.4v (Fig. 7.1 and Fig. 7.3). The 880 bp insert of pIIIF2.4v was sequenced and identified as containing a 282 bp region that was very similar to the 281 bp sequence of 092-03B5. The 092-03B5 insert was re-sequenced and found to be 282 bp long. The 282 bp regions from p092-03B5 and pIIIF2.4v were aligned and found to be 91 % similar with 24 bp nucleotide differences (Fig. 7.4). Surprisingly, given the ubiquitous amplification of *R. solanacearum* strains by AU759f/AU760r, nucleotide sequence polymorphisms were identified with the primer binding sites (Fig. 7.4), but these were confined to the 5' region of the primers. The sequence variations in the 282 bp regions from the two clones reflect the divergent origin of the two strains, ACH092 (biovar 4) and ACH0158 (biovar 2).

The AU759f and AU760r primers were used as sequencing primers to characterise the DNA flanking the 282 bp fragment in pIIIF2. Approximately 700 bp was obtained with the AU760r primer containing the 282 bp region and flanking DNA with high homology to the coding region of *lpxC*, a gene encoding UDP-3-*O*-acyl-GlcNAc deacetylase in *Pseudomonas aeruginosa* (Hyland *et al.*, 1997). Another 700 bp sequence obtained by the AU759f primer contained the 282 bp region and a DNA region showing similarity to the 3' end of a gene for a Fnr-type transcriptional regulator in *Rhizobium etli*. The results suggested that the 282 bp region was the intergenic space between the two contiguous genes and probably contained the promoter region of *lpxC*. Further sequencing reactions and analysis were carried out on the insert DNA fragments of pIIIF2.4v and pIIIF2.2v (Fig. 7.3). The approximately 2.3 kb long double stranded

sequence assembled contained the 282 bp region and an ORF of 306 amino acid residues (Fig. 7.3 and 7.5A). The 282 bp region contained the clear candidate -35, -10 and RBS motifs and a putative terminator was identified just 3' to the stop codon of the ORF (Fig. 7.5A). A standard protein database search with the amino acid sequences of the ORF confirmed the most similar protein to be UDP-3-*O*-acyl-GlcNAc deacetylase in *P. aeruginosa* (55 % identical and 75 % similar, GenPep AAC44974) (Hyland *et al.*, 1997) (Fig. 7.5B). However, the promoter region of *lpxC* in *P. aeruginosa* (GenBank PAU67855) did not show any significant nucleotide sequence similarity to the corresponding 282 bp region in *R. solanacearum* (Fig. 7.5C), consistent with the high specificity of the AU759f/AU760r primer pair for *R. solanacearum* DNA and not for closely related bacterial species (Opina *et al.*, 1997; Ito *et al.*, 1998).

UDP-3-*O*-acyl-GlcNAc deacetylase is an essential metalloamidase involved in lipid A biosynthesis (Hammond *et al.*, 1984), which is crucial for bacterial viability (Beall and Lutkenhaus, 1987). Lipid A is a major component of the lipopolysaccharide (LPS) complex, that is a unique outer membrane protein of gram-negative bacterial. To complete the LPS complex, the addition of a group of oligosaccharides called core and O-side chain polysaccharides, is absolutely necessary. The O-side chain is heterogeneous, due to its sugar composition and the configuration of the glycosidic linkage, which it may explain the diverse O-serotypes observed in several gram-negative bacteria (Hammond *et al.*, 1984). However, the lipid A protein itself and the core region of the LPS structure are relatively invariable.

A 340 bp *lpxC* gene specific probe was prepared using *Eco*RI and *Bam*HI sites in pIIIF2.4v (Fig. 7.3) and hybridised to Southern blots containing the *Eco*RI-digested genomic DNA of 51 strains representing five biovar groups of *R. solanacearum* (Fig. 7.6). As expected, the probe hybridised to the same 880 bp *Eco*RI fragment from all the strains tested (Fig. 7.6B, D and F), with the exceptions of a larger fragment hybridised in ACH0732 (an atypical biovar 2 strain) (Taghavi *et al.*, 1996) and no hybridisation in ACH0190 (biovar 3) and ACH0262 (biovar 4). The reason of no hybridisation with the probe in the two strains is not clearly understood. However, it is possible that they are not true *R. solanacearum* since these two strains had no hybridisation with the four IS element probe (Fig. 5.4B, C, D and E-lane 5; Fig. 5.3B, C, D and E-lane 2) (discussed in chapter 5), the porin gene and the UDP-*N*-acetylenolpyruvoylglucosamine reductase gene probe (Fig. 6.5B and C-lane 14; Fig. 6.5H and I-lane 37) (discussed in chapter 6) and the UDP-3-*O*-acyl-GlcNAc deacetylase gene probe (Fig. 7.6B-lane 14; Fig. 7.6F-

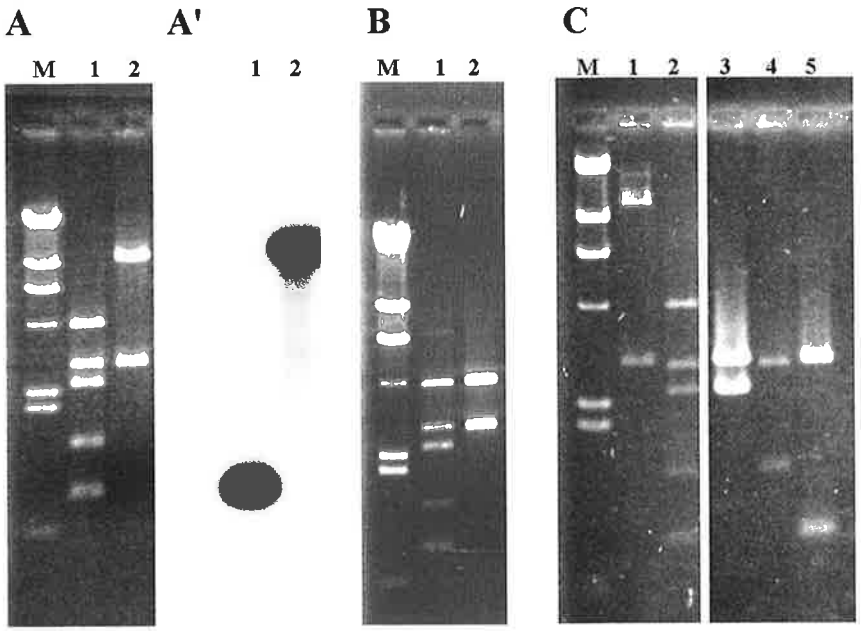


Figure 7.1

A. Agarose gel electrophoresis of *Eco*RI and *Sac*I double-digested pIIIF2 (lane 1) and *Sac*I-digested pIIIF2 (lane 2).

A'. Southern hybridisation of the gel in A with 092-03B5 insert of p092-03B5.

B. Agarose gel electrophoresis of *Eco*RI and *Sac*I double-digested pIIIF2 (lane 1) and *Eco*RI and *Sac*I double-digested pIIIF2.1v (lane 2).

C. Agarose gel electrophoresis of *Sac*I-digested pIIIF2 (lane 1), *Eco*RI and *Sac*I double-digested pIIIF2 (lane 2), *Eco*RI-digested clone pIIIF2.2v (lane 3), *Eco*RI and *Sac*I double-digested pIIIF2.3v (lane 4) and *Eco*RI-digested pIIIF2.4v (lane 5).

λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was included as a size marker (M).

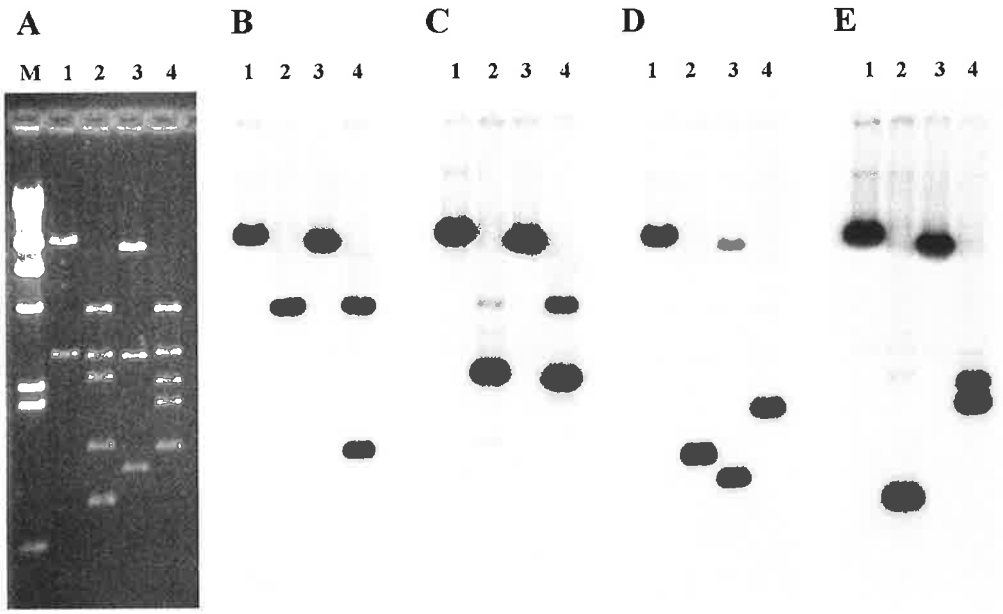


Figure 7.2

A. Agarose gel electrophoresis of pIIIF2 digested with a variety of restriction endonucleases. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was included as a size marker (M).
    1. *Sac*I-digested pIIIF2
    2. *Sac*I and *Eco*RI double-digested pIIIF2
    3. *Sac*I and *Apa*I double-digested pIIIF2
    4. *Sac*I and *Bam*HI double-digested pIIIF2

B. Southern hybridisation of the gel in A with IIIF2.1v insert of pIIIF2.1v.

C. Southern hybridisation of the gel in A with IIIF2.2v insert of pIIIF2.2v.

D. Southern hybridisation of the gel in A with IIIF2.3v insert of pIIIF2.3v.

E. Southern hybridisation of the gel in A with IIIF2.4v insert of pIIIF2.4v.

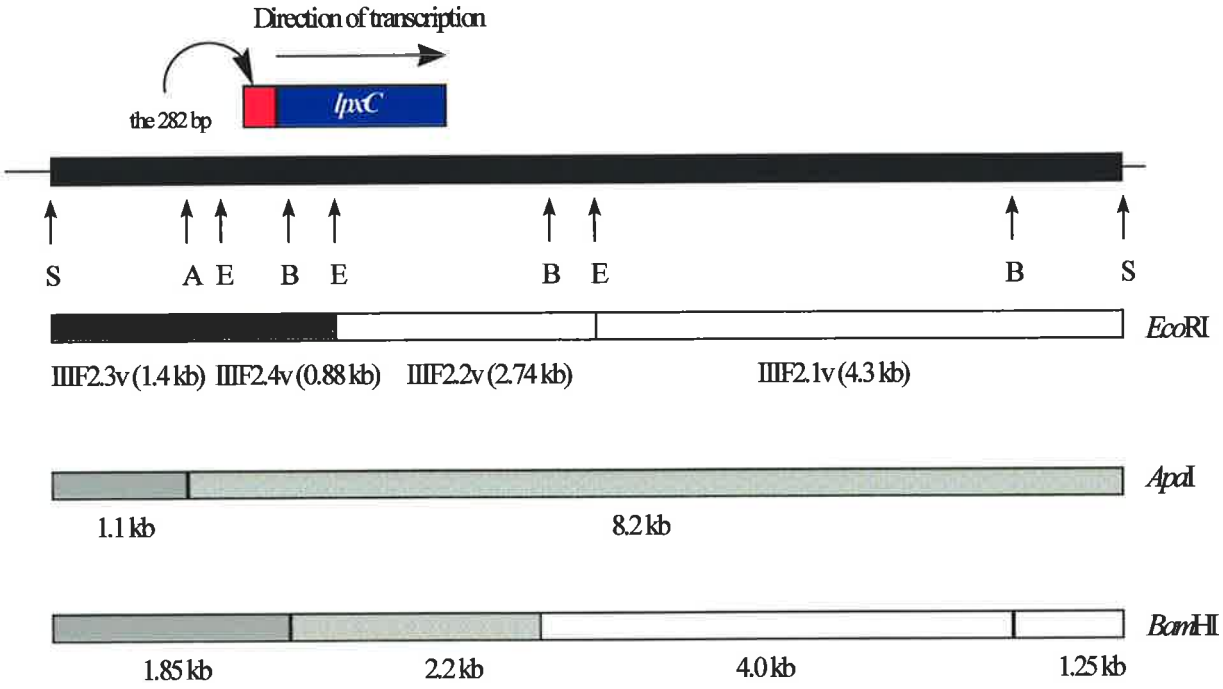


Figure 7.3

A physical map of IIIF2 insert with restriction sites for the endonucleases *Eco*RI (E), *Sac*I (S), *Apa*I (A) and *Bam*HI (B) marked. The fragment sizes for the four subclones are given in kb. The location and the transcriptional direction of the *lpxC* gene are indicated. The 282 bp region (red box) is indicated.

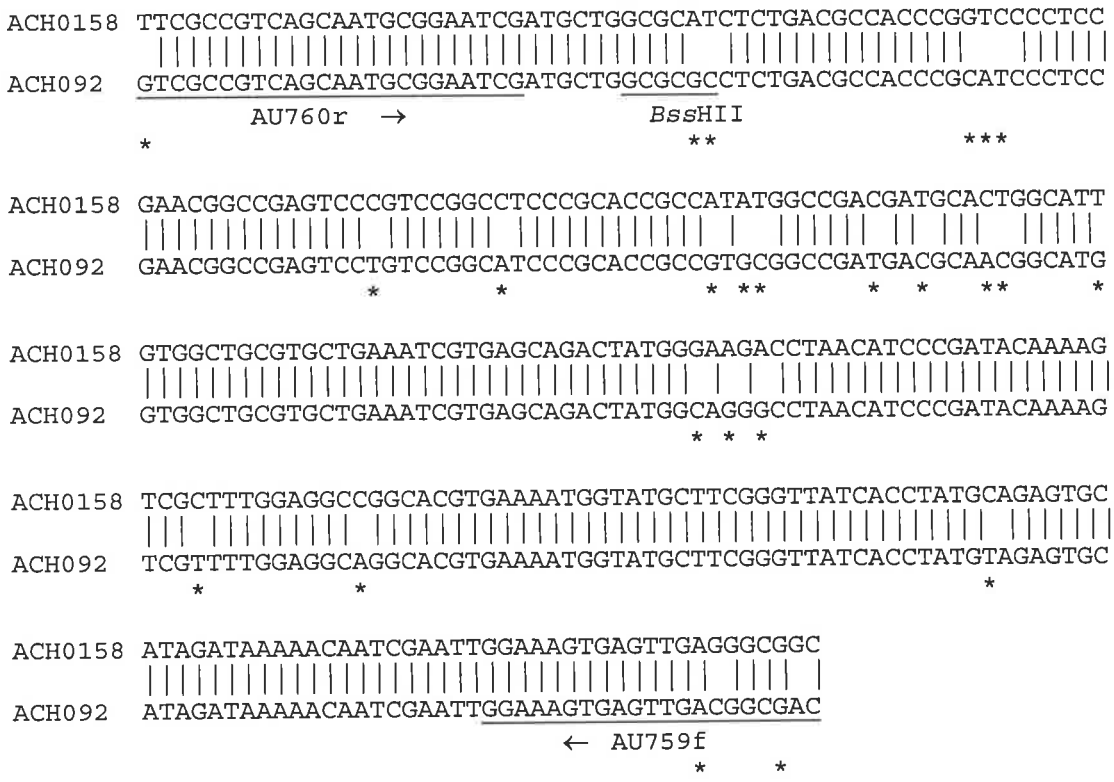


Figure 7.4

Alignment of nucleotide sequences of 092-03B5 insert from clone p092-03B5 and the corresponding region from clone pIIIF2. ACH0158 refers to the 282 bp sequence cloned in pIIIF2. ACH092 refers to the 282 bp sequence cloned in p092-03B5. Sequence and locations of AU759f and AU760r primers are indicated. The *Bss*HII site is marked. Asterisks (*) indicate nucleotide sequence differences.

Figure 7.5

A. Complete nucleotide and deduced amino acid sequences of the *lpxC* gene identified in pIIIF2. The nucleotide sequences of the 282 bp region are in bold. Two *Eco*RI sites utilised for subcloning are marked. AU759f and AU760r primers are indicated and the putative -35, -10 and RBS sites in the 282 bp region are also marked. The putative terminator and a location of a secondary structure formation are indicated. LIPO1 primer utilised for the sequencing reaction is indicated.

B. Alignment of amino acid sequences. Rs refers to the deduced amino acid sequences of the *lpxC* gene from *R. solanacearum* cloned in pIIIF2. Pa refers to the amino acid sequences of UDP-3-*O*-acyl-GlcNAc deacetylase encoded by the *lpxC* gene in *Pseudomonas aeruginosa* (Hyland *et al.*, 1997) (Genpep accession no. AAC44974).

C. Alignment of nucleotide sequences of the promoter region of the *lpxC* gene from *R. solanacearum* and the corresponding region from *P. aeruginosa*. Rs refers to the promoter region (the 282 bp region) from *R. solanaceraum* cloned in pIIIF2. Pa refers to the promoter region of the *lpxC* gene from *P. aeruginosa* (GenBank PAU67855). Nucleotide sequences of the 282 bp region cloned in pIIIF2 are in bold. Three asterisks (***) indicate the start codon of UDP-3-*O*-acyl-GlcNAc deacetylase in *R. solanacearum* and *P. aeruginosa*.

(A)

```
   1  TCACCCTGCTCGGCGGGGGTGCCCGTTCGTCGCGCCGCCTGCGGTGCGATCGCGCTGATG

  61  TGATGGTTGTCTCCCGGGACCGCCCGCGTGGATCTCCCACGCGGGCGGTCTCGTTTTATG

 121  ATGTGGGTATCCCGCGTGAATGCTTGCCACGATCATGATCCAGCCCGGCCAGCCGCTGCC

 181  TGACGCGACGCTCTACGAATACTTCGAAGTGGAAAAGGATGGCTGCGCGCTCGGGCCCAA

 241  TGCCTTCTCGGCAGCGGACCTGGCGCAGGGCAAGACGATTCTCATCTTCGGCCTGCCCGG

 301  CGCTTTCACGCCGACCTGTTCGGCACGCCACGTGCCGGGCTATCTCGCCCACTACGATGC

 361  GCTGCGCGCCAAAGGTGTCGACGAAATCTGGTGCGTGTCGGTCAACGATGCTTTCGTGAT

 421  GGGAGCATGGGCGCGTGCACAGGACACCGATGGCAGGGTTCGCATGCTGGCCGACGGCAG

 481  TGCCGAATTCACGCGCAAGCTCGGCCTTGAGCAAGACTTGTCCAAGCGCGGCATGGGCGT
           EcoRI

 541  GCGTTCGCAGCGCTACGCCATGATCGTCAGAAACGGTGTGGTAACCGCGCTACAGGTGGA

 601  GGCTCCAGGGCAGTTCGCCGTCAGCAATGCGGAATCGATGCTGGCGCATCTCTGACGCCA
                           AU760r  →

 661  CCCGGTCCCCTCCGAACGGCCGAGTCCCGTCCGGCCTCCCGCACCGCCATATGGCCGACG

 721  ATGCACTGGCATTGTGGCTGCGTGCTGAAATCGTGAGCAGACTATGGGAAGACCTAACAT
                                -35                          -10

 781  CCCGATACAAAAGTCGCTTTGGAGGCCGGCACGTGAAAATGGTATGCTTCGGGTTATCAC

 841  CTATGCAGAGTGCATAGATAAAAACAATCGAATTGGAAAGTGAGTTGAGGGCGGCCATGT
                                          ←  AU759f
                                               *****     M  L
                                              RBS

 901  TGAAACAGCGCACCATCAAGTCCGTGGTCAAAACCGTCGGCATCGGGTTGCACTCGGGCC
        K  Q  R  T  I  K  S  V  V  K  T  V  G  I  G  L  H  S  G  R

 961  GCAAGGTGACGTTGACGTTGCGTCCGGCCGCTCCCGGTACCGGCATCATTTTCACGCGGG
        K  V  T  L  T  L  R  P  A  A  P  G  T  G  I  I  F  T  R  V

1021  TGGACCTGGATCCGGCCGTCGAGATTCCCGCCACCGCCAGCGCCATCGGGGATACCCGCC
        D  L  D  P  A  V  E  I  P  A  T  A  S  A  I  G  D  T  R  L

1081  TGGCCTCCGTGCTGCAGAAGGACGGCGCCCGCGTGTCCACGGTCGAGCACCTGATGTCCG
        A  S  V  L  Q  K  D  G  A  R  V  S  T  V  E  H  L  M  S  A

1141  CCTGCGCGGGGCTTGGCATCGACAACCTGTATGTCGATGTCGACGCCGAGGAAATCCCCA
        C  A  G  L  G  I  D  N  L  Y  V  D  V  D  A  E  E  I  P  I

1201  TCATGGACGGCAGCGCGGCGTCGTTCGTGTTCCTGCTGCAGTCGGCCGGCATCGAGGAGC
        M  D  G  S  A  A  S  F  V  F  L  L  Q  S  A  G  I  E  E  Q

1261  AGGGTGCGGCCAAGCGCTTCATCCGCGTGACCAAGCCGGTGGAAGTGCTCGACGGCGACA
        G  A  A  K  R  F  I  R  V  T  K  P  V  E  V  L  D  G  D  K

1321  AGCTGGCGCGCCTGGCCCCGTACTTCGGCTTTAAGTTGGCGTTCACCATCGAATTCCGCC
                                          LIPO1  →
                                                      ......
                                                      EcoRI
        L  A  R  L  A  P  Y  F  G  F  K  L  A  F  T  I  E  F  R  H

1381  ATCCGGCCGTCGACAAAACGGGGCAGACGTTCGAGATCGACTTCGCCGACACCAGCTACA
        P  A  V  D  K  T  G  Q  T  F  E  I  D  F  A  D  T  S  Y  T
```

```
1441    CCCGCGAAATCGCCCGCGCGCGCACCTTCGGCTTTGCCCACGAGGTCGAGATGCTGCGGG
          R  E  I  A  R  A  R  T  F  G  F  A  H  E  V  E  M  L  R  E

1501    AGATCGGGCTGGCGCGCGGCGGCAGCCTGGACAACGCGATCGTGCTCGACGAGCACCGCA
          I  G  L  A  R  G  G  S  L  D  N  A  I  V  L  D  E  H  R  M

1561    TGCTGAACAACGATGAACTGCGCTATGGCGACGAGTTCGTCCGCCACAAGATTCTGGACG
          L  N  N  D  E  L  R  Y  G  D  E  F  V  R  H  K  I  L  D  A

1621    CGATCGGCGACCTGTACGTGGTCGGCTATCCGCTGATCGCAGCGTACACGGCACACAAGT
          I  G  D  L  Y  V  V  G  Y  P  L  I  A  A  Y  T  A  H  K  S

1681    CCGGCCATGGCCTGAACAACGCGCTGCTGCGCGCGCTGCTGGCCGATGAAACGGCGTACG
          G  H  G  L  N  N  A  L  L  R  A  L  L  A  D  E  T  A  Y  E

1741    AAATCGTCACGTTCGACAAGGTGGAAGAAGCGCCGCGCGCCTTCCTGCCGCAGTTGCAGC
          I  V  T  F  D  K  V  E  E  A  P  R  A  F  L  P  Q  L  Q  P

1801    CGGCATTCTCCTGATCCCGTGTTGCGGTTGCACGCATGAAAAAGCGCCCGAAAGGGCGCT
          A  F  S  *                    terminator

1861    TTTCTTATTGGGGCCGGCGGTGCCGTGCCACCATGTCCGCCACCGCCCGCCTGAGCGGCG
                        putative secondary structure formation

1921    AATCGGGCAGGCCGTCGGCCAGCGACGTCAGGCTATTGAGCCCAACTGCCGTCATCTTGG

1981    CGCGTTTGGGCGGCTTTTCGGTCTCTTGTGGCCATAAGCTGTCATTTCCGCCCTGCGGTT

2041    GTACCCGTACACGAATTGCGGTAACCTGCGATCCCCGGCGCTGCAGACGTTCCAGCAAGG

2101    TGGGTACCACCTGCCGTACGCGCGCCGCAGCTGCACTGTTTGCAGTGAGCAGCACCAGGG

2161    TCGCTTCCCGGCTGTCGCGCGCATCGTTCCGGATGCCGCCGATCGCCACACCGCCACCCA

2221    GTCCGGGTGGCAGCAGCGACAACACCTCCGCTTCCAGAGACGCCAGTTGCTTGGCCGCCT

2281    GCATCAGCGGGCCGACCGAGCCGGCACCCGTGAGCCAGTCCTGAACCGGCTTTGCCACGG

2341    GCGTTTTGAGAGCGGGATGAACGAAGATGCGCATGCCGGCATTGTAGGGCGCTTCCCCGC
```

(B)

```
Rs  MLKQRTIKSVVKTVGIGLHSGRKVTLTLRPAAPGTGIIFTRVDLDPAVEIPATASAIGDT
    |:||||:|.:::..|:|||||  ||  |||:||:.:|||:|.|.||||.|||||  |..:|:|
Pa  MIKQRTLKNIIRATGVGLHSGEKVYLTLKPAPVDTGIVFCRTDLDPVVEIPARAENVGET

Rs  RLASVLQKDGARVSTVEHLMSACAGLGIDNLYVDVDAEEIPIMDGSAASFVFLLQSAGIE
     :...|  |::.:|.|||||:||  |||||||  ||:::.|.|:|||||||:.||||:||||::
Pa  TMSTTLVKGDVKVDTVEHLLSAMAGLGIDNAYVELSASEVPIMDGSAGPFVFLIQSAGLQ

Rs  EQGAAKRFIRVTKPVEVLDGDKLARLAPYFGFKLAFTIEFRHPAVDKTGQTFEIDFADTS
    ||:|||:  ||:.:.|.|  :|||  |  :.|:  |||:.|.|:|  ||..   .|  .:||..||
Pa  EQEAAKKFIRIKREVSVEEGDKRAVFVPFDGFKVSFEIDFDHPVFRGRTQQASVDFSSTS

Rs  YTREIAPARTFGFAHEVEMLREIGLARGGSLDNAIVLDEHRMLNNDELRYGDEFVRHKIL
    :.:|:. ||||||  :::|  ||. .||  |||::||||:||:|:||:|:|||:||||:||||
Pa  FVKEVSPARTFGFMRDIEYLRSQNLALGGSVENAIVVDENRVLNEDGLRYEDEFVKHKIL

Rs  DAIGDLYVVGYPLIAAYTAHKSGHGLNNALLRALLADETAYEIVTFDKVEEAPRAFLPQL
    ||||||  ::|  .||:.: : ||||:|||.|||.|:||..|:|:|||:.. .|| .::..
Pa  DAIGDLYLLGNSLIGEFRGFKSGHALNNQLLRTLIADKDAWEVVTFEDARTAPISYMRPA

Rs  QPA
    .:.
Pa  AAV
```

(C)

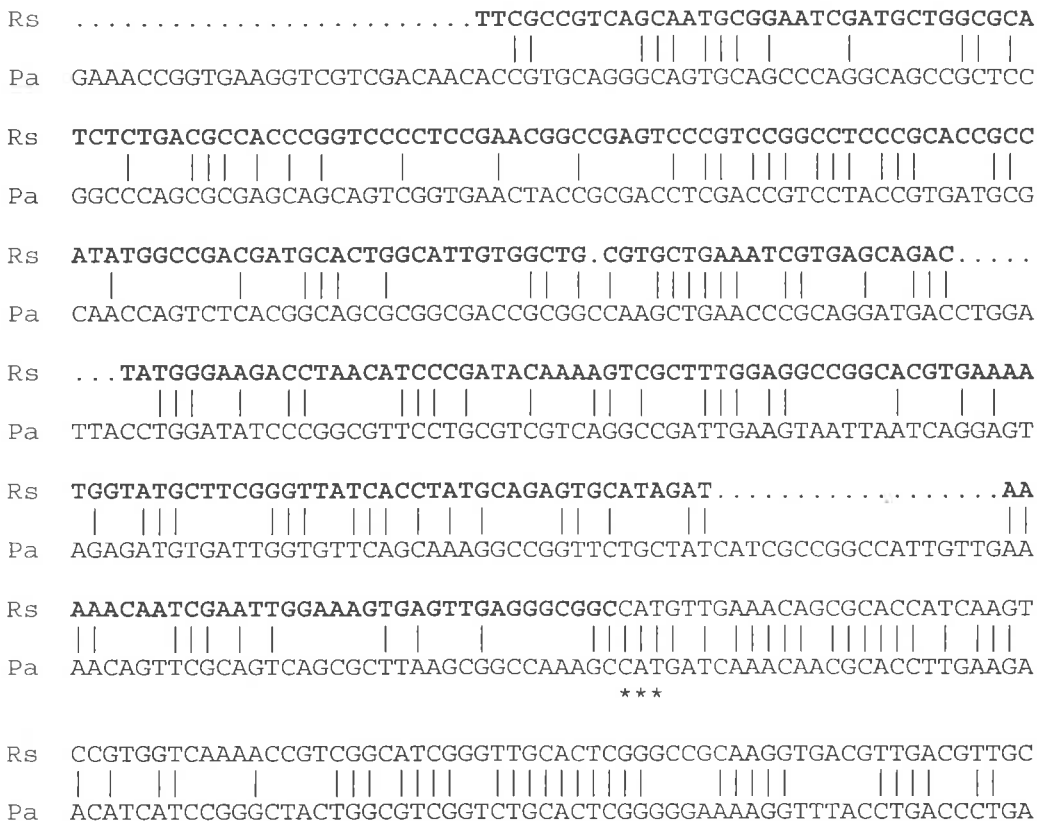

Rs .....TTCGCCGTCAGCAATGCGGAATCGATGCTGGCGCA
Pa GAAACCGGTGAAGGTCGTCGACAACACCGTGCAGGGCAGTGCAGCCCAGGCAGCCGCTCC
Rs TCTCTGACGCCACCCGGTCCCCTCCGAACGGCCGAGTCCCGTCCGGCCTCCCGCACCGCC
Pa GGCCCAGCGCGAGCAGCAGTCGGTGAACTACCGCGACCTCGACCGTCCTACCGTGATGCG
Rs ATATGGCCGACGATGCACTGGCATTGTGGCTG.CGTGCTGAAATCGTGAGCAGAC.....
Pa CAACCAGTCTCACGGCAGCGCGGCGACCGCGGCCAAGCTGAACCCGCAGGATGACCTGGA
Rs ...TATGGGAAGACCTAACATCCCGATACAAAAGTCGCTTTGGAGGCCGGCACGTGAAAA
Pa TTACCTGGATATCCCGGCGTTCCTGCGTCGTCAGGCCGATTGAAGTAATTAATCAGGAGT
Rs TGGTATGCTTCGGGTTATCACCTATGCAGAGTGCATAGAT.....AA
Pa AGAGATGTGATTGGTGTTCAGCAAAGGCCGGTTCTGCTATCATCGCCGGCCATTGTTGAA
Rs AAACAATCGAATTGGAAAGTGAGTTGAGGGCGGCCATGTTGAAACAGCGCACCATCAAGT
Pa AACAGTTCGCAGTCAGCGCTTAAGCGGCCAAAGCCATGATCAAACAACGCACCTTGAAGA
***
Rs CCGTGGTCAAAACCGTCGGCATCGGGTTGCACTCGGGCCGCAAGGTGACGTTGACGTTGC
Pa ACATCATCCGGGCTACTGGCGTCGGTCTGCACTCGGGGGAAAAGGTTTACCTGACCCTGA

Figure 7.6

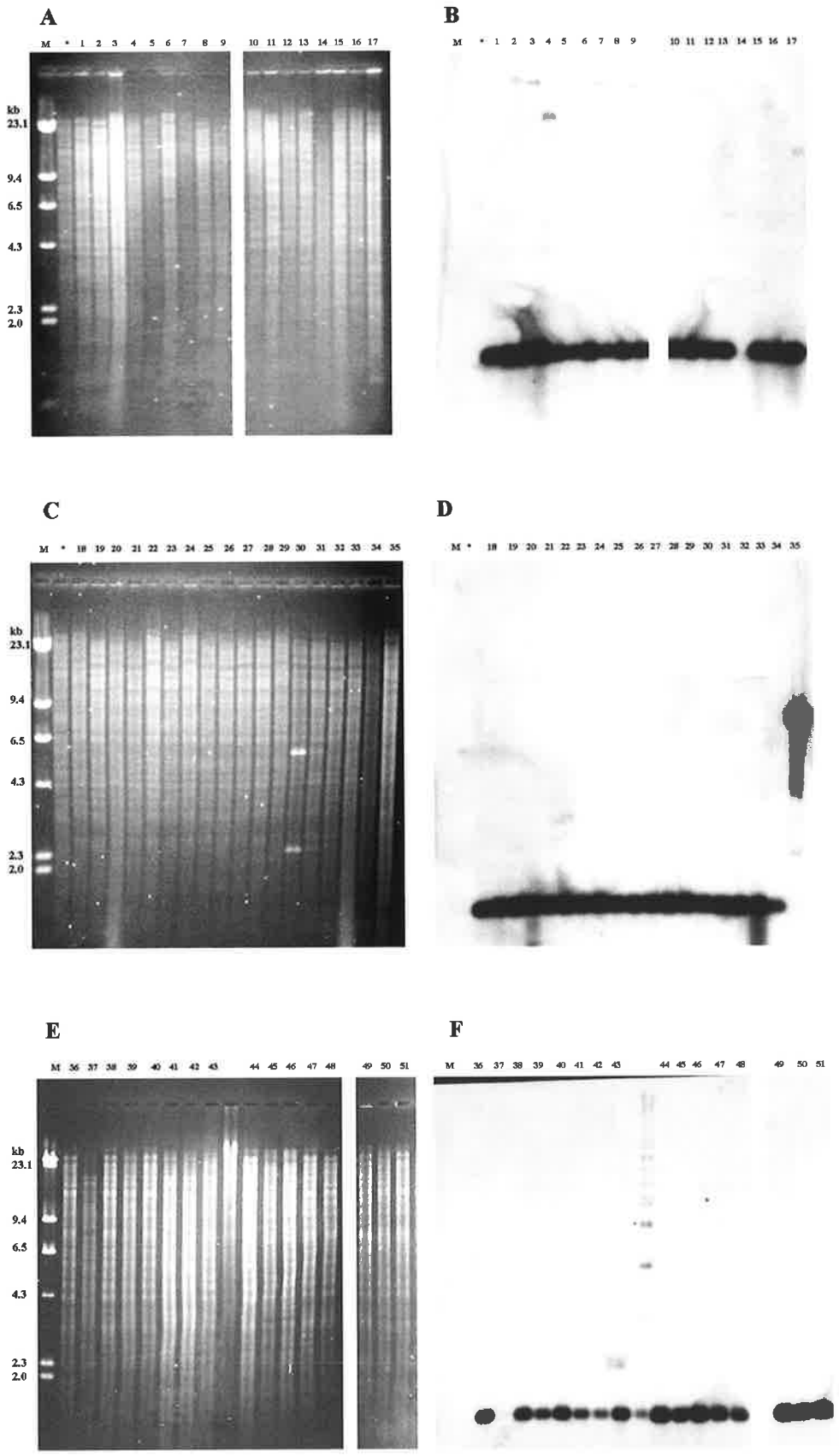
A, C and E. Agarose gel electrophoresis of *Eco*RI-digested genomic DNA of 51 *R. solanacearum* strains.

M. Molecular size marker of λDNA digested with *Hind*III

1. CIP0007 strain
2. CIP0009 strain
3. CIP0066 strain
4. CIP0070 strain
5. CIP0120 strain
6. CIP0218 strain
7. CIP0301 strain
8. CIP0418 strain
9. CIP0419 strain
10. CIP0121 strain
11. CIP0264 strain
12. CIP0283 strain
13. PD1431 strain
14. ACH0262 strain
15. ACH0319 strain
16. ACH1066 strain
17. CIP0277 strain
18. ACH0158 strain
19. CIP0117 strain
20. CIP0310 strain
21. CIP00359 strain
22. CIP0402 strain
23. CIP0403 strain
24. CIP0405 strain
25. South Africa 28 strain
26. South Africa 45 strain
27. South Africa 46 strain
28. South Africa 49 strain
29. South Africa 65 strain
30. PD1427 strain
31. PD1435 strain
32. ACH0127R strain
33. ACH1068 strain
34. CIP0358 strain
35. ACH0732 strain
36. ACH0171 strain
37. ACH0190 strain
38. ACH0369 strain
39. ACH0574 strain
40. ACH0671 strain
41. ACH1024 strain
42. ACH1064 strain
43. ACH1070 strain
44. South Africa 8 strain
45. South Africa 12 strain
46. South Africa 53 strain
47. South Africa 93 strain
48. South Africa 112 strain
49. CIP0065 strain
50. CIP0284 strain
51. PD1437 strain

B, D and F. Southern hybridisation of the gels in A, C and E with the *lpxC* gene probe prepared from pIIIF2.4v containing a part of the gene for encoding of UDP-3-*O*-acyl-GlcNAc deacetylase.

The lane between lane 43 and 44 was disregarded because the DNA sample was partially digested. The asterisk (*) indicates the lane containing ACH0171 DNA sample.


A
M * 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
kb
23.1
9.4
6.5
4.3
2.3
2.0
B
M * 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
C
M * 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
kb
23.1
9.4
6.5
4.3
2.3
2.0
D
M * 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
E
M 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
kb
23.1
9.4
6.5
4.3
2.3
2.0
F
M 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51

lane 37). Nevertheless, the result indicates that the *lpxC* gene is an essential, highly conserved gene encoding UDP-3-*O*-acyl-GlcNAc deacetylase in most of *R. solanacearum* isolates tested.

7.3 Construction of standard plasmid DNAs

The construction of a standard plasmid from the p092-03B5 is illustrated in Figure 7.7. For convenience of manipulation, the insert of p092-03B5 was transferred from pCR™II vector (Opina *et al.*, 1997) into the *Eco*RI site of pBluescript (SK+) to give p092-03B5EL (Table 2.2). Construction of the standard plasmid DNA utilised a unique *Bss*HII site (Fig. 7.4) into which was inserted the polylinker of pBluescript II (SK+), which is contained in a 173 bp fragment flanked by *Bss*HII sites. The DNA region of the multi cloning sites was PCR-amplified by M13 primers (-20 and RSP) and the 173 bp insert fragment was prepared by *Bss*HII digestion of the PCR product.

The standard plasmid, designated as p092-03B5EL+170, contains the same primer (AU759f/AU760r) binding sites as p092-03B5, but yields a 455 bp fragment upon PCR amplification (Fig. 7.8-lane 2). A similar strategy gave the second standard plasmid, designated pBs2.10+170 was constructed from pBs2.10 using an internal *Bss*HII site. It yielded a 480 bp PCR product with the AU630f and AU631r primers (Fig. 7.8-lane 4) compared with the 307 bp fragment from pBs2.10.

7.4 Quantitative estimation of *R. solanacearum* DNA by PCR

For quantitative PCR, the standard plasmids, p092-03B5EL+170 and pBs2.10+170, were diluted to make working solutions at different concentrations containing ranging from 1 copy/µl to 200,000 copies/µl. Genomic DNA of the strain ACH0158, with a genome size of approximately 5 Mb (Holloway *et al.*, 1993), was prepared ranging from 1,000 genome copies/µl to 20,000 genome copies/µl. Initially, PCR reactions using AU759f/AU760r and AU630f/AU631r primer pairs were tested in reactions which contained 20,000 genome copies of ACH0158 spiked with various amounts of standard DNA constructs added (Fig. 7.9). Two similar sets of PCR reactions, either with or without 10 ng of purified potato DNA, were carried out to determine whether host plant DNA interacted with the primers or interfered with the PCR reaction.

The results shown in Figure 7.9 indicated that, as expected, there was competition for PCR amplification between the standard DNA and target DNA, which was most

noticeable in reactions where there was a large discrepancy in copy numbers between the construct and the target DNAs. In general, the greater the number of added constructs, the greater the reduction in the amount of PCR product from the native genome (Fig. 7.9). This effect appears to differ in its extent between the two primer pairs. The AU630f/AU631r primer pair with 2,000 copies of pBs2.10+170 and 20,000 copies of target DNA amplified the standard DNA inefficiently (Fig. 7.9B-lane 3). Whereas, with the same copy numbers, the AU759f/AU760r primer pair was able to efficiently amplify both standard DNA and target DNA (Fig. 7.9A-lane 3). Nevertheless, the PCR assay to quantify 20,000 copies of target DNA with a range of copies of constructs suggested that the amount of PCR product of 20,000 copies of target DNA was highly similar to the amount of PCR product of those of construct DNA (Fig. 7.9A, A', B and B'-lane 2 and 3). This result suggests that the quantitative PCR assay was accomplished with a reasonable accuracy. The addition of potato DNA to the reactions had no observable effect on the results obtained and potato DNA used alone as a template produced no product (Fig. 7.9A' and 7.9B'), suggesting that the PCR method can be applied in practical detection of *R. solanacearum* within host plants.

To evaluate the sensitivity of the method, similar experiments were prepared with 5,000 and 1,000 genome copies of ACH0158 in PCR reactions containing appropriate reduced ranges of copy number of the two standard DNAs (Fig. 7.10). With 5,000 copies of target DNA, both primer pairs gave results that closely corresponded to the expected amount of target (Fig. 7.10A and 7.10B). In the reactions with AU759f/AU760r primers, p092-03B5EL+170 was more efficiently amplified than the target DNA (Fig. 7.10A), while with AU630f/AU631r primers, the two different DNA templates were amplified with similar efficiency (Fig. 7.10B). The smallest amount of the target DNA genomes which were detected only by AU759f and AU760r primers was 1,000 copies (Fig. 7.10C), which equates to approximately 5 pg of bacterial genomic DNA which is much less sensitive than that reported by Opina *et al.* (1997).

To test the practicability of the PCR assay, 20,000 copies of p092-03B5EL+170 construct was added to a reaction containing 5 µl of DNA sample that was extracted from a potato tuber that had been grown in *R. solanacearum* infected soil, so that the sample contains an unknown number of *R. solanacearum* organisms. As shown in Figure 7.11, the PCR products matched at 20,000 copies of construct DNA suggesting that 5 µl of the DNA sample used for the assay contained approximately 20,000 copies

of *R. solanacearum*. Good agreement between the amount of amplified PCR products between 5 µl of the DNA sample and 20,000 copies of construct DNA was also obtained when several PCR reactions were prepared with a range of added copies of construct DNA (Fig. 7.11).

7.5 Summary and discussion

The two primer pairs, AU759f/AU760r and AU630f/AU631r, derived from the p092-03B5 and pBs2.10 clones, were utilised in an investigation of the potential in practical method for specific detection and quantification of *R. solanacearum*. PCR, with appropriate controls, is rapid, sensitive and demonstrably accurate in the quantification of target DNA (Celi *et al.*, 1993; Zachar *et al.*, 1993) and it can be readily automated. It has been argued that standard DNA constructs to be spiked into quantitative assays must be almost identical to target DNA sequences (Pannetier *et al.*, 1993) for the greatest accuracy. However, creating a very similar standard DNA, for example by deletion or insertion of a few base pairs, can make the method difficult to perform and analyse routinely. Technical ease of the assay is particularly important in this case because bacterial wilt disease is most important in the tropical countries, where often only basic laboratory infrastructure is available. The foreign DNA inserted into the standard constructs successfully facilitated the resolution of two different PCR products on agarose gels after co-amplification of native and construct DNA templates.

Quantitative results in which potato DNA was added to PCR reactions had little effect on the results indicating that the method can be applied in the detection of infection in whole plant homogenates. These experiments proved that the PCR method was sensitive to as few as 1,000 cells, equating to approximately 5 pg of ACH0158 genomic DNA, but this level of sensitivity could probably be improved by further optimisation of the procedures. Opina *et al.* (1997) tested the sensitivity of the PCR assay by AU759f and AU760r primers in detection of *R. solanacearum* DNA. The 282 bp fragment was amplified in a genomic DNA sample obtained from as few as 1.4 cells of *R. solanacearum* (Opina *et al.*, 1997). The reason for this difference in the sensitivity is not clear. The result described here was probably affected by the presence of the standard plasmid DNA in the PCR reactions. On the other hand, Ito *et al.* (1998) reported that the AU759f/AU760r primers did not amplify the 282 bp fragment using 10 ng of *R. solanacearum* DNA which was extracted from soil samples and purified by agarose gel and which might carry substances inhibiting *Taq* DNA polymerase during PCR reactions. Further experiments should be designed to optimise and improve the

PCR method in order to quantify the exact number of *R. solanacearum* in real-life samples. Furthermore, the method should be amenable to the use of automated PCR technologies to reduce the amount of post-PCR processing.

PCR-based assays were developed in quantifying *R. solanacearum* DNA. Quantification of 1,000 genome copies of *R. solanacearum* was possible, but optimisation of the methods should allow greater sensitivity. Practicability of the method is also a crucial factor to apply the method in real-life samples. The results indicate that the method can be applied to quantify unknown number of *R. solanacearum* in the presence of host genomic DNA. Technical ease of the assay is important in this case because bacterial wilt disease is most important in several developing countries in the tropics, where often only basic laboratory infrastructure is available.

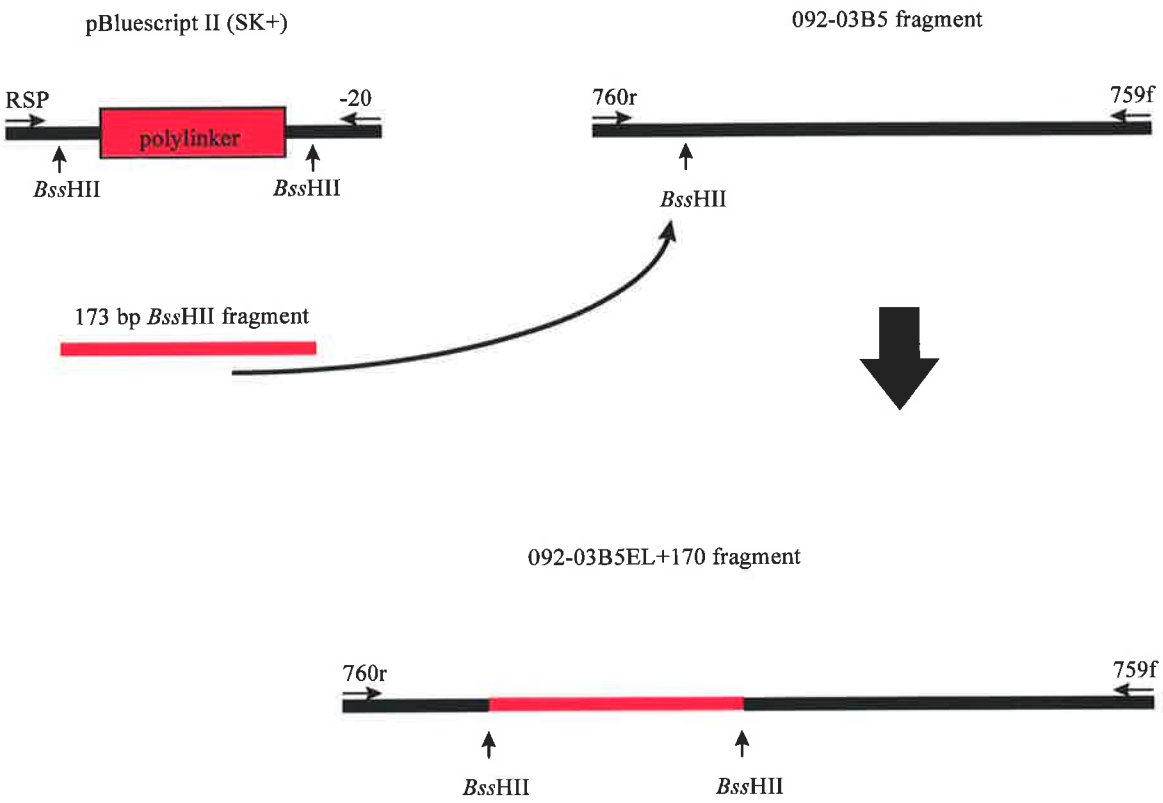


Figure 7.7

A 173 bp *Bss*HII fragment from pBluescript II vector was inserted into a unique *Bss*HII site between AU759f and AU760r primers in an insert fragment of p092-03B5EL, to generate p092-03B5EL+170. RSP and -20 refer to M13 derived sequencing reverse and forward primers, respectively.

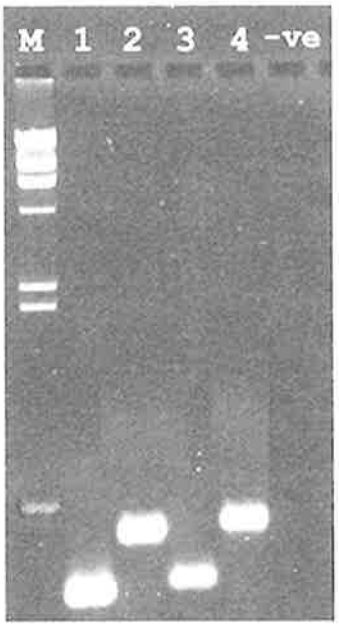


Figure 7.8

A PCR-amplified 282 bp fragment from p092-03B5EL (lane 1) and a 455 bp fragment from p092-03B5EL+170 (lane 2) using AU759f and AU760r primers. A PCR-amplified 307 bp fragment from pBs2.10 (lane 3) and a 480 bp fragment from pBs2.10+170 (lane 4) using AU630f and AU631r primers. Negative DNA control (-ve) was included. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was included as a size marker (M).

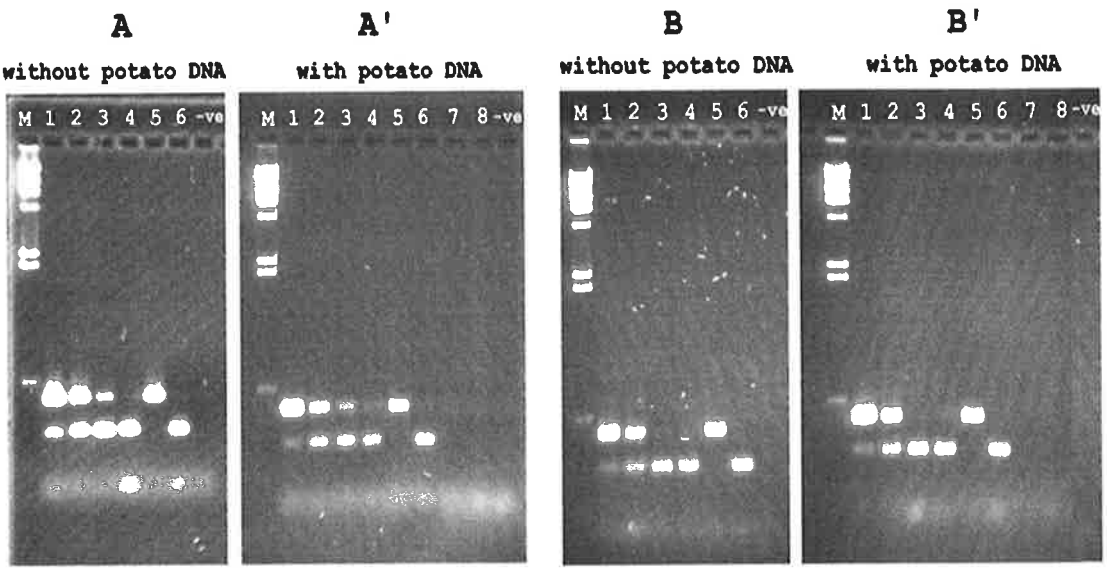


Figure 7.9

Testing the quantitative assay with 20,000 genome copies of *R. solanacearum*.

A and A'. PCR assay using AU759f and AU760r primers to quantify 20,000 copies of *R. solanacearum* target DNA with various copy numbers of standard DNA construct, p092-03B5EL+170. Two sets of experiments prepared, either without (A) or with (A') potato DNA.

B and B'. PCR assay using AU630f and AU631r primers to quantify 20,000 copies of *R. solanacearum* target DNA with various copy numbers of constructed standard DNA, pBs2.10+170. Two sets of experiments prepared, either without (B) or with (B') potato DNA.

The number of spiked copies of construct in all gels is 200,000 (lane 1), 20,000 (lane 2), 2,000 (lane 3) and 200 (lane 4). The reactions in lanes 1-4 also contained 20,000 copies of the ACH0158 genome in all gels. The gels in lane 5 and 6 show separate amplifications from constructed standard and *R. solanacearum* target DNA, respectively. Lane 7 contains 10 ng of potato DNA only. Lane 8 contains 30 ng of potato DNA only. Negative DNA controls are indicated (-ve). λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb used as a size marker (M).

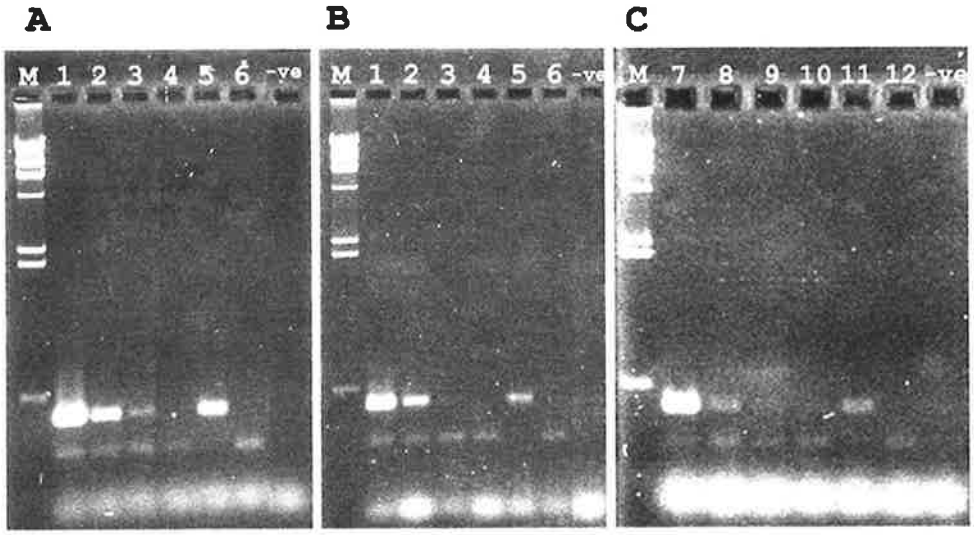


Figure 7.10

Testing the quantitative assay with 5,000 and 1,000 genome copies of *R. solanacearum*.

A. PCR assay using AU759f and AU760r primers to quantify 5,000 copies of *R. solanacearum* target DNA with various copy numbers of p092-03B5EL+170.

B. PCR assay using AU630f and AU631r primers to quantify 5,000 copies of *R. solanacearum* target DNA with various copy numbers of pBs2.10+170. The number of copies of construct added in both gels is 50,000 (lanes 1), 5,000 (lanes 2), 500 (lanes 3) and 50 (lanes 4). The reactions in lanes 1-4 contained 5,000 copies of the ACH0158 genome. In each gel, lane 5-6 show separate amplifications from constructed standard and *R. solanacearum* target DNA, respectively. Negative DNA controls are indicated (-ve). λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was used as a size marker.

C. PCR assay using AU759f and AU760r primers to quantify 1,000 copies of *R. solanacearum* target DNA with various copy numbers of p092-03B5EL+170. The number of copies of construct added in the gel is 10,000 (lanes 7), 1,000 (lanes 8), 100 (lanes 9) and 10 (lanes 10). The reactions in lanes 7-10 contained 1,000 copies of the ACH0158 genome. Lane 11 and 12 show separate amplifications from constructed standard and *R. solanacearum* target DNA, respectively. Negative DNA controls are indicated (-ve). λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was used as a size marker.

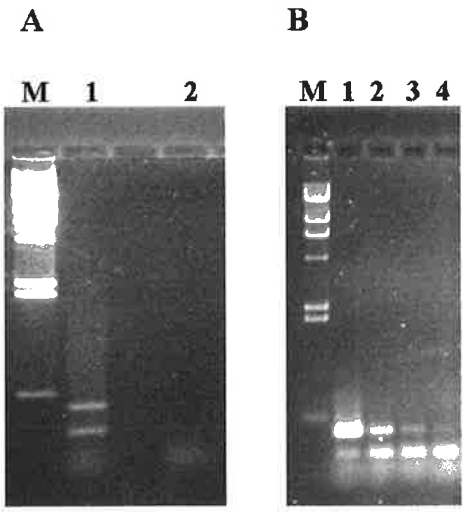


Figure 7.11

The quantitative assay with unknown genome copies of *R. solanacearum*.

A. PCR assay using AU759f and AU760r primers to quantify unknown genome copies of *R. solanacearum* in 5 µl of DNA sample, that was extracted from a potato tuber that have been grown in *R. solanacearum* infected soil, with 20,000 copies of p092-03B5EL+170 (lane 1). Negative DNA controls are indicated (-ve).

B. PCR assay using AU759f and AU760r primers to quantify unknown genome copies of *R. solanacearum* with various copy numbers of p092-03B5EL+170 construct. The number of copies of construct added is 200,000 (lanes 1), 20,000 (lanes 2), 2,000 (lanes 3) and 200 (lanes 4). The reactions in lanes 1-4 contained unknown genome copies of *R. solanacearum* in 5 µl of the DNA sample. λ bacteriophage DNA digested with *Hind*III with fragments of 23.1, 9.4, 6.5, 4.3, 2.3, 2.0 and 0.56 kb was used as a size marker.

APPENDIX

Jeong E-L and Timmis JN. 2000. Novel insertion sequence elements associated with genetic heterogeneity and phenotype conversion in *Ralstonia solanacearum*. Journal of Bacteriology 182(16); 4673-4676.

JOURNAL OF BACTERIOLOGY, Aug. 2000, p. 4673–4676 Vol. 182, No. 16
0021-9193/00/$04.00+0

Novel Insertion Sequence Elements Associated with Genetic Heterogeneity and Phenotype Conversion in *Ralstonia solanacearum*

EUN-LEE JEONG* AND JEREMY N. TIMMIS

Department of Genetics, The University of Adelaide, Adelaide, South Australia 5005, Australia

Received 21 January 2000/Accepted 17 May 2000

Three insertion sequences (IS) elements were isolated from the phytopathogen *Ralstonia solanacearum*. Southern hybridization using these IS elements as probes revealed hybridization profiles that varied greatly between different strains of the pathogen. During a spontaneous phenotype conversion event, the promoter of the *phcA* gene was interrupted by one of these IS elements.

Ralstonia solanacearum is the causative agent of bacterial wilt disease in taxonomically diverse plant hosts, and the biological heterogeneity of the species is reflected in its genetic (4, 6), genomic, and biochemical complexities (12, 13, 16). Genomic variation in many bacterial species is generated by mobile DNA, which is common among bacterial genomes (14). Insertion sequence (IS) elements are the simplest transposable DNA in bacteria and have been major causes of genomic changes, such as insertion, deletion, and inversion (9). In pathogenic bacteria, IS elements modify the expression and cause the transposition of genes controlling interactions of the pathogen with the host organism (8, 10, 15), suggesting an important role for DNA transposition and rearrangement in the evolution of host-pathogen interactions.

During a screening for repetitious DNA that was unique to biovar 2 strain ACH0158 of *R. solanacearum* (Table 1), we isolated and characterized three novel IS elements, designated IS*Rso4*, IS*Rso3*, and IS*Rso2*. Phenotype conversion (PC) (2) was also investigated for potential association with the movement of an IS element. PC is a phenomenon in which wild-type strains of *R. solanacearum* spontaneously lose their ability to produce large amounts of extracellular polysaccharide and a range of secreted proteins and become more motile, resulting in mutant strains with greatly reduced virulence (2). In several cases, PC has been reported to result from mutations in *phcA* (1, 2), a global regulatory gene controlling the transcription of genes associated with the virulence of the pathogen (18). In this study, ACH0158-M81C (Table 1) was identified as a PC mutant resulting from the insertion of IS*Rso4* within *phcA* (1, 2).

Isolation of IS elements. Plasmid clones of *Sau*3AI- and *Sal*I-digested genomic DNA of the wild-type *R. solanacearum* biovar 2 strain ACH0158 were prepared with pBluescript or pUC19 (Table 1), and duplicate DNA blots from colonies were hybridized with [α-^{32}P]dATP-labeled total genomic DNA from ACH0158 and ACH0171 (a taxonomically distant biovar 3 strain). Colonies showing strong hybridization to the ACH0158 probe and little or no hybridization to the ACH0171 probe were selected for further characterization. Several such clones were sequenced, and those that showed similarity to known IS elements were used as probes for a plasmid library of *Eco*RI-digested genomic DNA from ACH0158 to yield clones containing full-length IS elements. A 2.7-kb *Eco*RI fragment (pSV102; Table 1) was sequenced and found to contain two adjacent IS elements, designated IS*Rso4* and IS*Rso3*. Similarly, sequencing of a 5.2-kb *Eco*RI fragment (pISBE; Table 1) revealed a third IS element, designated IS*Rso2*.

Characterization of IS elements. IS*Rso4* was readily identified because of its similarity to known IS elements, such as IS*1031* (7). It is 855 bp long and has perfectly matched 17-bp inverted repeats (IRs) flanked by 4-bp CTAG direct repeats (DRs). The deduced amino acid sequence for a putative transposase within IS*Rso4* showed similarity to those of several IS elements belonging to the IS*5* family from various bacterial species (17), including IS*1031* from *Acetobacter xylinum* (35.2% identical, 57.2% similar). Amino acid sequences for the putative transposase of IS*Rso4* and those of the homologous IS elements had two conserved domains (17), N3 with D1 I(G/A)(Y/F) and C1 with R1 3E as invariant motifs that are typical features of the IS*5* family.

The second element, IS*Rso3*, is 1,209 bp long and has imperfect 18-bp IRs with 4 nucleotide mismatches and 4-bp CTAG DRs. The 3′ DR of IS*Rso3* is shared with the 5′ DR of IS*Rso4*. For IS*Rso3*, the amino acid sequence of the predicted transposase showed about 50% identity and 70% similarity to those of IS elements from distantly related *Pseudomonas* species and *Xanthomonas campestris* pv. *vesicatoria*. This second group also shared the typical features of the IS*5* family, but the core sequences of two domains were further apart than in IS*Rso4* (compare Fig. 1A and 1C). Therefore, IS*Rso3* is also a member of the IS*5* family but has significant differences from IS*Rso4* in sequence and size.

IS*Rso2* is 864 bp long and has 15-bp IRs with a base-pair mismatch. Its putative open reading frame (ORF) (encoding 134 amino acids) showed sequence similarity to several IS elements from diverse bacteria, including *Neisseria meningitidis* (with only one of the two ORFs of IS*1301* showing 28% identity and 49% similarity) (11). This group of IS elements has conserved amino acid sequences but as yet has not been classified into a particular family. Some bacterial species in which elements similar to IS*Rso2* have been found are taxonomically distant from *R. solanacearum*, such as *N. meningitidis*, a human pathogen causing meningococcal meningitis.

Each of the three IS elements from *R. solanacearum* showed conserved domains and strong sequence similarities, both nucleotide and derived amino acid, to different families of IS elements from a diverse range of bacteria.

* Corresponding author. Mailing address: Department of Genetics, The University of Adelaide, Adelaide, South Australia 5005, Australia. Phone: (61) 8 8303-3013. Fax: (61) 8 8303-4399. E-mail: eljeong@genetics.adelaide.edu.au.

TABLE 1. Bacterial strains and plasmids

R. solanacearum strain or plasmid	Relevant characteristic(s)[a]	Source or reference[b]
Strains		
ACH0158	Wild-type strain isolated from potato	A. C. Hayward
ACH0158-M81C	Spontaneous PC mutant of ACH0158	This study
ACH0158-M3	Spontaneous PC mutant of ACH0158	This study
ACH0158-M8	Spontaneous PC mutant of ACH0158	This study
ACH1061	Wild-type strain isolated from potato	A. C. Hayward
ACH1068S	Wild-type strain isolated from potato	A. C. Hayward
CIP418	Wild-type strain isolated from peanut	CIP
CIP419	Wild-type strain isolated from peanut	CIP
Plasmids		
pKS(1), pKS(2), pSK(1), pSK(2)	Bluescript vectors; Ap^r	Stratagene
pUC19	Cloning vector; Ap^r	New England Biolabs
pGEM-T Easy	Cloning vector; Ap^r	Promega
pISBE	5.2-kb *Eco*RI fragment containing IS*Rso2* from wild-type ACH0158 cloned into pKS; Ap^r	This study
pSV102	2.7-kb *Eco*RI fragment containing IS*Rso4* and IS*Rso3* from wild-type ACH0158 cloned pKS; Ap^r	This study
pWT	1.46-kb *Eco*RI fragment containing *phcA* from wild-type ACH0158 cloned into pGEM-T Easy; Ap^r	This study
pM81C	1.64-kb *Eco*RI fragment containing *phcA81C* from PC mutant ACH0158-M81C cloned into pGEM-T Easy; Ap^r	This study
pM3	1.33-kb *Eco*RI fragment containing *phcA3* from PC mutant ACH0158-M3 cloned into pGEM-T Easy; Ap^r	This study
pM8	1.46-kb *Eco*RI fragment containing *phcA8* from PC mutant ACH0158-M8 cloned into pGEM-T Easy; Ap^r	This study

[a] Ap^r, resistance to ampicillin.
[b] CIP, International Potato Centre, Lima, Peru.

Distribution of IS elements. To ascertain the copy numbers and sequence distribution of the IS elements within the *R. solanacearum* genome and to look for differences between genetically different isolates, the three IS elements were used individually as probes in Southern analyses. IS*Rso4* hybridized to six *Eco*RI fragments in three biovar 2 strains (Fig. 2A) and was concluded to be present in at least six copies because of the lack of internal *Eco*RI sites in its sequence. In these biovar 2 strains, IS*Rso3* hybridized to numerous (. 40) *Eco*RI fragments (Fig. 2B) and IS*Rso2* hybridized to approximately 12 fragments (Fig. 2C). For IS*Rso3* and IS*Rso2*, these copy numbers may be underestimated because of the possibility of multiple copies of the elements in the more strongly hybridizing fragments. The differences in hybridization intensity may reflect the presence of truncated fragments or relics that have diverged sufficiently in sequence to reduce the level of hybridization at high stringency. As we have sequenced only single examples of IS*Rso3* and IS*Rso2*, we cannot distinguish between these possibilities.

Southern hybridization of DNA from three biovar 2 strains revealed similar hybridization patterns for a particular probe. Fourteen additional biovar 2 race 3 strains were also tested

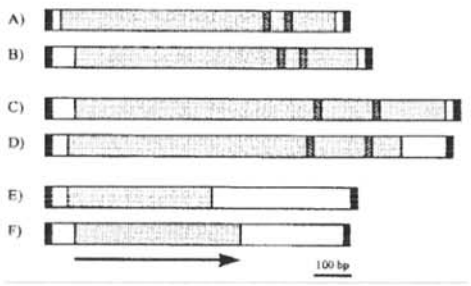


FIG. 1. Structural organization of IS*Rso4*, IS*Rso3*, and IS*Rso2* compared with that of known IS elements. (A) IS*Rso4*. (B) IS*1031* from *A. xylinum* (GenPept accession number AAA25029) (7). (C) IS*Rso3*. (D) IS*1384* from *P. putida* plasmid pPGH1 (GenPept accession number AAC98743). (E) IS*Rso2*. (F) IS*1301* from *N. meningitidis* (GenPept accession number CAA88914). Only one of the two ORFs in IS*1301* was significantly similar to the IS*Rso2* sequence (11). Symbols: n , IRs of each IS element; u , ORFs of each IS element; p , locations of the signatures for N3 [D1 I(G/A)(Y/F)]; o , C1 (R1 3E) motif. The arrow indicates the direction of transcription.

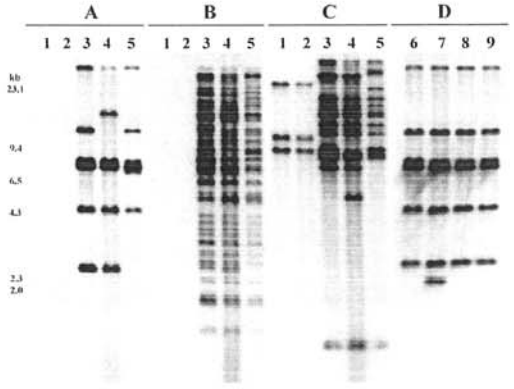


FIG. 2. Southern hybridization of genomic DNA from five strains of *R. solanacearum* and three PC mutants derived from wild-type ACH0158. Lane 1, CIP418 (biovar 1); lane 2, CIP419 (biovar 1); lanes 3 and 6, ACH0158 (biovar 2); lane 4, ACH1061 (biovar 2); lane 5, ACH1068S (biovar 2); lane 7, ACH0158-M81C; lane 8, ACH0158-M3; lane 9, ACH0158-M8. Southern membranes were hybridized with IS*Rso4* (A and D), IS*Rso3* (B), and IS*Rso2* (C) probes. Molecular sizes are shown on the left. The figure was produced by use of Adobe Photoshop 5.0.

(results not shown), and hybridization profiles very similar to those shown in Fig. 2 were obtained. The overall similarity of the Southern hybridization patterns in biovar 2 strains is consistent with genetic uniformity in biovar 2 strains. Based on restriction fragment length polymorphism analysis of several isolates of the species, Cook and others (4, 6) suggested that biovar 2 (race 3) strains have almost identical genomes, commensurate with the uniformity of their metabolic and other biological characteristics regardless of their geographical origins. Taken together, these results support the suggestion that all biovar 2 strains have a clonal origin in South America (6). Cook and Sequeira (5, 6), however, also observed that a race 3-specific 2-kb DNA fragment isolated by subtractive hybridization was present in a minority (less than 6%) of non-race 3 strains and suggested the possibility of lateral gene transfer between distantly related strains within the species. Our results support this suggestion as we observed IS*Rso2*-like sequences in two biovar 1 (non-race 3) strains (Fig. 2C, lanes 1 and 2), although a small number of strains was tested.

IS*Rso4* disrupts the *phcA* gene in a PC mutant. To investigate the possibility of active transposition of these IS elements, the behavior of IS*Rso4* during PC was examined (2). Three spontaneous PC mutants were isolated from wild-type ACH0158, and their *Eco*RI-digested genomic DNA was probed with IS*Rso4* in Southern analyses. The hybridization pattern of ACH0158-M81C showed an additional 2.1-kb band (Fig. 2D, lane 7) compared with the patterns of the wild type (Fig. 2D, lane 6) and the other two PC mutants (Fig. 2D, lanes 8 and 9).

The regions flanking the new insertion in ACH0158-M81C were obtained by inverse PCR and cloned as pM81C (Table 1). Sequence comparisons of this clone and DNA databases revealed greater than 99% amino acid sequence identity to the product of the *phcA* gene of *R. solanacearum* strain AW1 (2). The *phcA* gene has been characterized as a global regulator of virulence genes and as being central to the mechanism of PC (1, 3). Brumbley et al. (1) characterized the PhcA protein as a member of the LysR family of transcriptional regulators that control the expression of genes encoding virulence factors (3, 18). Brumbley and Denny (2) observed hybridization of the sequences flanking *phcA* to repetitious DNA in the genomes of AW1 and related strains.

The new copy of IS*Rso4*, integrated between nucleotide positions 120 and 121 within the *phcA* gene of ACH0158-M81C, created 5-bp (CTGAG) DRs (results not shown), and the sequence differed from the CTAG DRs in pSV102 and the CTAG DRs in a third, independently isolated copy of IS*Rso4* (results not shown). Variations in target duplication length and sequence at the point of insertion are not unusual (9) and may be influenced by different helix conformations of the nucleotide sequence of the *phcA* gene during the initial cleavage by IS*Rso4*. However, additional variations in the length and sequence of the DR motif between different genomic copies of IS*Rso4* cannot be ruled out, as we have characterized only three of seven possible insertion sites. These results suggest that IS elements contribute to genomic variation and heterogeneity among strains and also to significant phenotypic changes related to virulence in *R. solanacearum*.

The *phcA* genes were also cloned from wild-type ACH0158 and two additional PC mutants (ACH0158-M3 and ACH0158-M8) in clones pWT, pM3, and pM8, respectively (Table 1). The two PC mutants showed sequence lesions within *phcA*, confirming the importance of this gene in PC. However, IS*Rso4* was not involved, and neither were IS*Rso3* and IS*Rso2*. ACH0158-M3 (Table 1) had between nucleotides 23 and 155 a 132-bp deletion that removed the start codon of *phcA* (results not shown). ACH0158-M8 (Table 1) contained at nucleotide position 648 in the ORF of *phcA* a 2-bp (TG) insertion causing a frameshift that resulted in an early stop codon and a truncated protein similar to that previously described for the *phcA1* allele from a spontaneous PC mutant, AW1-PC (1). Therefore, the movement of IS*Rso4* contributes to a range of different mutations that cause PC. However, the transposition of DNA into *phcA* is not a common cause of PC. Southern analysis indicated that genomic DNA of only 1 in 10 PC mutants probed with *phcA* showed a band shift consistent with IS element insertion (results not shown). The mutation in ACH0158-M3 is clearly irreversible, whereas the reversibility of PC may be important in vivo. The latter is possible in ACH0158-M81C by perfect IS element excision. However, so far we have been unable to demonstrate the restoration of wild-type ACH0158 from either ACH0158-M81C or ACH0158-M8. It is possible that many insertions are merely "dead ends"—giving rise to cells which not only are incapable of infection but also have lost the capacity to regain these functions.

Concluding remarks. Southern hybridization and sequence analyses suggested that the IS elements isolated in *R. solanacearum* may have been horizontally transferred because of the presence of highly homologous IS elements in very distantly related bacterial taxa and the lack of a clear lineage for the IS elements within diverse strains of *R. solanacearum*.

Only a single member of each element was fully characterized by sequencing, and evidence for the presence of other genomic copies was obtained from Southern hybridization. We describe one event where the transposition of IS*Rso4* to a new genomic location causes a specific gene mutation and a consequent change in phenotype. It remains to be fully explored whether these IS elements are all actively involved in genomic rearrangements that contribute to the extensive genetic heterogeneity among different isolates of this species. The availability of mechanisms, in the form of multiple IS elements, that enable fast and extensive genome changes may be essential in *R. solanacearum* and in other bacterial species that constantly modify the expression of genes responsible for host-pathogen interactions (8), pathogenicity (10), and metabolic pathways (15), either by mutation or regulation.

Nucleotide sequence accession numbers. The nucleotide sequences of IS*Rso4*, IS*Rso3*, and IS*Rso2* have been deposited in GenBank under accession numbers AF079849, AF183890, and AF186082, respectively. The GenBank accession number for the *phcA* gene from *R. solanacearum* strain ACH0158 is AF184046.

We thank The University of Adelaide for an international postgraduate research scholarship to E.-L. Jeong. This study was supported in part by grant PN9452 from the Australian Centre for International Agricultural Research.

We thank Tim Denny, Viji Krishnapillai, Chris Hayward, and Mark Fegan for comments on the manuscript.

REFERENCES

1. **Brumbley, S. M., B. F. Carney, and T. P. Denny.** 1993. Phenotype conversion in *Pseudomonas solanacearum* due to spontaneous inactivation of PhcA, a putative LysR transcriptional regulator. J. Bacteriol. **175:**5477–5487.
2. **Brumbley, S. M., and T. P. Denny.** 1990. Cloning of wild-type *Pseudomonas solanacearum phcA*, a gene that when mutated alters expression of multiple traits that contribute to virulence. J. Bacteriol. **172:**5677–5685.
3. **Clough, S. J., M. A. Schell, and T. P. Denny.** 1994. Evidence for involvement of a volatile extracellular factor in *Pseudomonas solanacearum* virulence gene expression. Mol. Plant Microbe Interact. **7:**621–630.
4. **Cook, D., E. Barlow, and L. Sequeira.** 1989. Genetic diversity of *Pseudomonas solanacearum*: detection of restriction fragment length polymorphisms with DNA probes that specify virulence and the hypersensitive response. Mol. Plant Microbe Interact. **2:**113–121.
5. **Cook, D., and L. Sequeira.** 1991. The use of subtractive hybridization to

obtain a DNA probe specific for *Pseudomonas solanacearum* race 3. Mol. Gen. Genet. **227:**401–410.
6. **Cook, D., and L. Sequeira.** 1994. Strain differentiation of *Pseudomonas solanacearum* by molecular genetic methods, p. 77–93. *In* A. C. Hayward and G. L. Hartman (ed.), Bacterial wilt: the disease and its causative agent, *Pseudomonas solanacearum*. CAB International, Wallingford, United Kingdom.
7. **Coucheron, D. H.** 1993. A family of IS*1031* elements in the genome of *Acetobacter xylinum*: nucleotide sequences and strain distribution. Mol. Microbiol. **9:**211–218.
8. **Galán, J. E., and A. Collmer.** 1999. Type III secretion machines: bacterial devices for protein delivery into host cells. Science **284:**1322–1328.
9. **Galas, D. J., and M. Chandler.** 1989. Bacterial insertion sequences, p. 109–162. *In* D. Berg and M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
10. **Hacker, J., G. Blum-Oehler, I. Mühldorfer, and H. Tschäpe.** 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol. Microbiol. **23:**1089–1097.
11. **Hammerschmidt, S., R. Hilse, J. P. van Putten, R. Gerardy-Schahn, A. Unkmeir, and M. Frosch.** 1996. Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element. EMBO J. **15:**192–198.
12. **Hayward, A. C.** 1964. Characteristics of *Pseudomonas solanacearum*. J. Appl. Bacteriol. **27:**265–277.
13. **Hayward, A. C.** 1994. The hosts of *Pseudomonas solanacearum*, p. 9–24. *In* A. C. Hayward and G. L. Hartman (ed.), Bacterial wilt: the disease and its causative agent, *Pseudomonas solanacearum*. CAB International, Wallingford, United Kingdom.
14. **Healy, F. G., R. A. Bukhalid, and R. Loria.** 1999. Characterization of an insertion sequence element associated with genetically diverse plant pathogenic *Streptomyces* spp. J. Bacteriol. **181:**1562–1568.
15. **Kallastu, A., R. Hõrak, and M. Kivisaar.** 1998. Identification and characterization of IS*1411*, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida*. J. Bacteriol. **180:**5306–5312.
16. **Opina, N., F. Tavner, G. Hollway, J.-F. Wang, T.-H. Li, R. Maghirang, M. Fegan, A. C. Hayward, V. Krishnapillai, W. F. Hong, B. W. Holloway, and J. N. Timmis.** 1997. A novel method for development of species- and strain-specific DNA probes and PCR primers for identifying *Burkholderia solanacearum* (formerly *Pseudomonas solanacearum*). Asia Pacific J. Mol. Biol. Biotechnol. **5:**19–30.
17. **Rezsöhazy, R., B. Hallet, J. Delcour, and J. Mahillon.** 1993. The IS*4* family of insertion sequences: evidence for a conserved transposase motif. Mol. Microbiol. **9:**1283–1295.
18. **Schell, M. A.** 1996. To be or not to be: how *Pseudomonas solanacearum* decides whether or not to express virulence genes. Eur. J. Plant Pathol. **102:**459–469.

REFERENCES

Abdullah H. 1982. Resistance of winged bean (*Psophocarpus tetragonolobus*) to *Pseudomonas solanacearum*. Malays. Appl. Biol. 11; 35-39.

Akiew E, Trevorrow PR and Tonello PE. 1993. Management of bacterial wilt of tobacco. In *Bacterial wilt proceedings of an international conference held at Kaohsiung, Taiwan, 28-31 October 1992 ACIAR Proceedings No 45.* Ed. Hayward AC and Hartman GL. pp 270-275.

Allen C and Gay J. 1995. Structure of a polygalacturonase regulatory operon in *Pseudomonas solanacearum*. Phytopathology 85; 1159 (abstract).

Allen C, Gay J and Simon-Buela L. 1997. A regulatory locus, *pehSR*, controls polygalacturonase production and other virulence functions in *Ralstonia solanacearum*. Mol. Plant-Microbe Interact. 10(9); 1054-1064.

Allen C, Huang Y and Sequeira L. 1991. Cloning of genes affecting polygalacturonase production in *Pseudomonas solanacearum*. Mol. Plant-Microbe Interact. 4; 147-154.

Allen C, Simon L, Atkinson M and Sequeira L. 1993. Analysis of polygalacturonase as a component of bacterial wilt disease. In *Bacterial wilt*. Ed. Hartman GL and Hayward AC. ACIAR press, Canberra, Australia. pp 238-244.

Altshul SF. 1991. Amino acid substitution matrices from an information theoretic perspective. J. Mol. Biol. 219; 555-565.

Altshul SF, Gish W, Miller W, Myers EW and Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215; 403-410.

Anonymous. 1992. Quarantine procedure no. 26. *Pseudomonas solanacearum*. Bull. OEPP 20; 255-262.

Anonymous. 1993. Validation of the publication of new names and new combinations previously effectively published outside the IJSB-list no. 45. Int. J. Syst. Bacteriol. 43(2); 398-399.

Anwar RA and Vlaovic M. 1978. Purification of UDP-*N*-acetylenolpyruvoylglucosamine reductase from *Escherichia coli* by affinity chromatography, its subunit structure and the absence of flavin as the prosthetic group. Can. J. Biochem. 57: 188-196.

Arlat M and Boucher CA. 1991. Identification of a *dsp* DNA region controlling aggressiveness of *Pseudomonas solanacearum*. Mol. Plant-Microbe Interact. 4(2); 211-213.

Arlat M, Gough CL, Zischek C, Barberis PA, Trigalet A and Boucher C. 1992. Transcriptional organisation and expression of the large *Hrp* gene cluster of *Pseudomonas solanacearum*. Mol. Plant-Microbe Interact. 5; 187-193.

Arlat M, van Gijsegem F, Genin S, Gough CL, Zischek C, Barberis PA and Boucher C. 1993. Studies of the *hrp* pathogenicity genes from *Pseudomonas solanacearum* GMI1000. In *Bacterial wilt proceedings of an international conference held at Kaohsiung, Taiwan, 28-31 October 1992 ACIAR Proceedings No 45.* Ed. Hayward AC and Hartman GL. pp 232-237.

Arlat A, van Gijsegem F, Huet JC, Pernollet JC and Boucher CA. 1994. PopA1, a protein which induces a hypersensive-like response on specific *Petunia* genotypes, is secreted via the Hrp pathway of *Pseudomonas solanacearum*. EMBO J. 13(3); 543-553.

Beall B and Lutkenhaus J. 1987. Sequence analysis, transcriptional organization, and insertional mutagenesis of the *envA* gene of *Escherichia coli*. J. Bacteriol. 169; 5408-5415.

Bennett and Leitch. 1995. Nuclear DNA amounts in angiosperms. Ann. Bot. 76; 113-176.

Bertolla F, Frostegård Å, Brito B, Nesme X and Simonet P. 1999. During infection of its host, the plant pathogen *Ralstonia solanacearum* naturally develops a state of competence and exchanges genetic material. Mol. Plant-Microbe Interact. 12(5); 467-472.

Bertolla F, van Gijsegem F, Nesme X and Simonet P. 1997. Conditions for natural transformation of *Ralstonia solanacearum*. Appl. Environ. Microbiol. 63(12); 4965-4968.

Blanc-Potard A-B, Solomon F, Kayser J and Groisman EA. 1999. The SPI-3 pathogenicity island of *Salmonella enterica*. J. Bacteriol. 181(3); 998-1004.

Bolland S and Kleckner N. 1996. The three chemical steps of Tn*10*/IS*10* transposition involve repeated utilization of a single active site. Cell 84; 223-33.

Boucher CA, Barberis PA, Trigalet APh and Demery DA. 1985. Transposon mutagenesis of *Pseudomonas solanacearum*: isolation of Tn*5*-induced avirulent mutants. J. General Microbiology 131; 2449-2457.

Boucher CA, Martinel A, Barberis PA, Alloing G and Zischek C. 1986. Virulence genes are carried by a megaplasmid of the plant pathogen *Pseudomonas solanacearum*. Mol. Gen. Genet. 205; 270-275.

Boucher CA, van Gijsegem F, Barberis PA, Arlat M and Zischek C. 1987. *Pseudomonas solanacearum* genes controlling both pathogenicity on tomato and hypersensitivity on tobacco are clustered. J. Bacteriol. 169(12); 5626-5632.

Boyd EF and Hartl DL. 1999. Analysis of the type 1 pilin gene cluster *fim* in *Salmonella*: its distinct evolutionary histories in the 5' and 3' regions. J. Bacteriol. 181(4); 1301-1308.

Brachet P, Eisen H and Rambach A. 1970. Mutations of coliphage lambda affecting the expression of replicative functions O and P. Mol. Gen. Genet. 108; 266-276.

Brito B, Marenda M, Barberis P, Boucher C and Genin S. 1999. *prhJ* and *hrpG*, two new components of the plant signal-dependent regulatory cascade controlled by PrhA in *Ralstonia solanacearum*. Mol. Microbiol. 31(1); 237-251.

Brumbley SM, Carney BF and Denny TP. 1993. Phenotype conversion in *Pseudomonas solanacearum* due to spontaneous inactivation of PhcA, a putative LysR transcriptional regulator. J. Bacteriol. 175(17); 5477-5487.

Brumbley SM and Denny TP. 1990. Cloning of wild-type *Pseudomonas solanacearum phcA*, a gene that when mutated alters expression of multiple traits that contribute to virulence. J. Bacteriol. 172(10); 5677-5685.

Buddenhagen IW. 1986. Bacterial wilt revisited. In *Bacterial Wilt Disease in Asia and the South Pacific No. 13*. Ed. Persley GJ. ACIAR, Canberra, Australia. pp 126-143.

Buddenhagen IW and Kelman A. 1964. Biological and physiological aspects of bacterial wilt caused by *Pseudomonas solanacearum*. Ann. Rev. Phytopathol. 2; 203-230.

Buddenhagen IW, Sequeira L and Kelman A. 1962. Designation of races of *Pseudomonas solanacearum*. Phytopathol. 52; 726.

Bukhalid RA, Chung SY and Loria R. 1998. *nec1*, a gene conferring a necrogenic phenotype, is conserved in plant pathogenic *Streptomyces* species and linked to a transposase pseudogene. Mol. Plant-Microbe Interact. 11; 960-967.

Busse J and Auling G. 1988. Polyamine pattern as a chemotaxonomic marker within the Proteobacteria. System. Appl. Microbiol. 11; 1-8.

Burrows LL, Charter DF and Lam JS. 1996. Molecular characterization of the *Pseudomonas aeruginosa* serotype O5 (PAO1) B-band lipopolysaccharide gene cluster. Mol. Microbiol. 22(3); 481-495.

Byrne ME. 1989. Nucleotide sequence analysis of IS*256* from the *Staphylococcus aureus* gentamicin-tobramycin-kanamycin-resistance transposon Tn*4001*. Gene 81; 361-367.

Byrne AM and Lessie TG. 1994. Characteristics of IS*401*, a new member of the IS*3* family implicated in plasmid rearrangements in *Pseudomonas cepacia*. Plasmid 31; 138-147.

Calos MP, Johnsrud L and Miller JH. 1978. DNA sequence at the integration sites of the insertion element IS*1*. Cell 13; 411-418.

Carney BF and Denny TP. 1990. A cloned avirulence gene from *Pseudomonas solanacearum* determines incompatibility on *Nictiana tobacum* at the host species level. J. Bacteriol. 172; 4836-4843.

Caruso P, Llop P, Palomo JL, Garcia P, Morente C and Lopez MM. 1998. Evaluation of methods for detection of potato seed contamination by *Ralstonia solanacearum*. In *Bacterial wilt disease-molecular and ecological aspects*. Ed. Prior P, Allen C and Elphinstone J. Springer. 128-132.

Cassier-Chauvat C, Poncelet M and Chauvat F. 1997. Three insertion sequences from the cyanobacterium *Synechocystis* PCC6803 support that occurrence of horizontal DNA transfer among bacteria. Gene 195; 257-266.

Celi FS, Zenilman ME and Shuldiner AR. 1993. A rapid and versatile method to synthesize internal standards for competitive PCR. Nucleic Acids Research 21(4); 1047.

Chandler M and Fayet O. 1993. Translational frameshifting in the control of transposition in bacteria. Mol. Microbiol. 7(4); 497-503.

Chapman M and Kao CC. 1998. EpsR modulates production of extracellular polysaccharides in the bacterial wilt pathogen *Ralstonia (Pseudomonas) solanacearum*. J. Bacteriol. 180(1); 27-34.

Chen J-H, Hsieh Y-Y, Hsiau S-L, Lo T-C and Shau C-C. 1999. Characterization of insertions of IS*476* and two newly identified insertion sequences, IS*1478* and IS*1479*, in *Xanthomonas campestris* pv. *campestris*. J. Bacteriol. 181(4); 1220-1228.

Chen CW, Yu T-W, Chung H-M and Chou C-F. 1992. Discovery and characterization of a new transposable element, Tn*4811*, in *Streptomyces lividans* 66. J. Bacteriol. 174(23); 7762-7769.

Ciampi L and Sequeira L. 1980. Multiplication of *Pseudomonas solanacearum* in resistant potato plants and the establishment of latent infections. Am. Potato J. 57; 319-329.

Ciampi L, Sequeira L and French ER. 1980. Latent infection of potato tubers by *Pseudomonas solanacearum*. Am. Potato J. 57; 377-386.

Clough SJ, Flavier AB, Schell MA and Denny TP. 1997. Differential expression of virulence genes and motility in *Ralstonia (Pseudomonas) solanacearum* during exponential growth. Appl. Environ. Microbiol. 63(3); 844-850.

Clough SJ, Lee K-E, Schell MA and Denny TP. 1997. A two-component system in *Ralstonia (Pseudomonas) solanacearum* modulates production of PhcA-regulated virulence factors in response to 3-hydroxypalmitic acid methyl ester. J. Bacteriol. 179(11); 3639-3648.

Clough SJ, Schell MA and Denny TP. 1994. Evidence for involvement of a volatile extracellular factor in *Pseudomonas solanacearum* virulence gene expression. Mol. Plant-Microbe Interact. 7; 621-630.

Clough SJ, Schell MA and Denny TP. 1995. Involvement of a two-component system and 3-hydroxypalmitic acid methyl ester in the regulation of virulence in *Pseudomonas solanacearum*. Phytopathology 85; 1159 (abstract).

Cook D, Barlow E and Sequeira L. 1989. Genetic diversity of *Pseudomonas solanacearum*: detection of restriction fragment length polymorphisms with DNA probes that specify virulence and the hypersensitive response. Mol. Plant-Microbe Interact. 2; 113-121.

Cook D and Sequeira L. 1991a. The use of subtractive hybridization to obtain a DNA probe specific for *Pseudomonas solanacearum* race 3. Mol. Gen. Genet. 227 401-410.

Cook D and Sequeira L. 1991b. Genetic and biochemical characterization of a *Pseudomonas solanacearum* gene cluster required for extracellular polysaccharide production and for virulence. J. Bacteriol. 173; 1654-1662.

Cook D and Sequeira L. 1994. Strain differentiation of *Pseudomonas solanacearum* by molecular genetic methods. In *Bacterial Wilt: The disease and its causative agent, Pseudomonas solanacearum*. Ed. Hayward AC and Hartman GL. CAB International, UK. pp. 77-93.

Coplin DL and Cook D. 1990. Molecular genetics of extracellular polysaccharide biosynthesis in vascular phytopathogenic bacteria. Mol. Plant-Microbe Interact. 3(5); 271-279.

Costerton JW, Cheng K-J, Geesey GG, Ladd TI, Nickel JC, Dasgupta M and Marrie TJ. 1987. Bacterial biofilms in nature and disease. Annu. Rev. Microbiol. 41; 435-464.

Coucheron DH. 1993. A family of IS*1031* elements in the genome of *Acetobacter xylinum*: nucleotide sequences and strain distribution. Mol. Microbiol. 9; 211-218.

Cowan SW, Schirmer T, Rummel G, Steiert M, Ghosh R, Pauptit RA, Jansonius JN and Rosenbusch JP. 1992. Crystal structures explain functional properties of two *E. coli* porins. Nature 358; 727-733.

Davis DH, Doudoroff M, Stanier RY and Mandel M. 1969. Proposal to reject the genus *Hydrogenomonas*: taxonomic implications. Int. J. Syst. Bacteriol. 19; 375-390.

De Meirsman C, van Soom C, Verreth C, van Gool A and Vanderleyden J. 1990. Nucleotide sequence analysis of IS*427* and its target sites in *Agrobacterium tumefacience* T37. Plasmid 24; 227-234.

Denny TP and Baek S-R. 1991. Genetic evidence that extracellular polysaccharide is a virulence factor of *Pseudomonas solanacearum*. Mol. Plant-Microbe Interact. 4(2); 198-206.

Denny TP, Brumbley SM, Carney BF, Clough SJ and Schell MA. 1993. Regulation of virulence in *Pseudomonas solanacearum*. In *Bacterial wilt proceedings of an international conference held at Kaohsiung, Taiwan, 28-31 October 1992 ACIAR Proceedings No 45.* Ed. Hayward AC and Hartman GL. pp 252-256.

Denny TP, Carney BF and Schell MA. 1990. Inactivation of multiple virulence genes reduces the ability of *Pseudomonas solanacearum* to cause wilt symptoms. Mol. Plant-Microbe Interact. 3; 293-300.

Denny TP, Makini FW and Brumbley SM. 1988. Characterization of *Pseudomonas solanacearum*Tn*5* mutants deficient in extracellular polysaccharide. Mol. Plant-Microbe Interact. 1; 215-223.

Denny TP and Schell MA. 1994. Virulence and pathogenicity of *Pseudomonas solanacearum*: genetic and biochemical perspectives. In *Biotechnology and plant protection-bacterial pathogenesis and disease resistance.* Ed. Bills DD and Kung S-D. World scientific publishing Co. Pty. Ltd. pp 127-148.

Dianese JC and Dristig MCG. 1994. Strain characterization of *Pseudomonas solanacearum* based on membrane protein patterns. In *Bacterial Wilt: The disease and its causative agent, Pseudomonas solanacearum.* Ed. Hayward AC and Hartman GL. CAB International, UK. pp 113-121.

Dodd HM, Horn N and Gasson MJ. 1990. Analysis of the genetic determinant for production of the peptide antibiotic nisin. J. General Microbiol. 136; 555-566.

Duggleby RG, Kinns H and Rood JI. 1981. A computer program for determining the size of DNA restriction fragment. Anal. Biochem. 110; 49-55.

Elphinstone JG, Stanford HM and Stead DE. 1998. Detection of *Ralstonia solanacearum* in potato tubers, *Solanum dulcamara* and associated irrigation water. In *Bacterial wilt disease-molecular and ecological aspects.* Ed. Prior P, Allen C and Elphinstone J. Springer. pp 133-139.

Engler JA and van Bree MP. 1981. The nucleotide sequence and protein-coding capability of the transposable element IS*5*. Gene 14; 155-163.

Falkow S. 1988. Molecular koch`s postulates applied to microbial pathogenicity. Rev. Infect. Dis. 10 (suppl. 2); 274-276.

Fang Z, Morrison N, Watt B, Doig C and Forbes KJ. 1998. IS*6110* transposition and evolutionary scenario of the direct repeat locus in a group of closely related *Mycobacterium tuberculosis* strains. J. Bacteriol. 180(8); 2102-2109.

Farabaugh PJ, Schmeissner U, Hofer M and Miller JH. 1978. Genetic studies of the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *laci* gene of *E. coli*. J. Mol. Biol. 126; 847-863.

Farkas T, Kiss J and Olasz F. 1996. The construction and characterization of an effective transpositional system based on IS*30*. FEBS Letters 390; 53-58.

Fegan M, Holoway G, Hayward AC, Timmis JN. 1998. Development of a diagnostic test based on the polymerase chain reaction (PCR) to identify strains of *R. solanacearum* exhibiting the Biovar 2 genotype. *In Bacterial wilt disease-molecular and ecological aspects*. Ed. Prior P, Allen C and Elphinstone J. Springer. pp 34-43.

Feinberg AP and Vogelstein B. 1983. A technique for radiolabeling DNA restriction fragments to high specific activity. Anal. Biochem. 132: 6-13.

Feng X, Ou L-T and Ogram A. 1997. Cloning and sequence analysis of a novel insertion element from plasmids harbored by the carbofuran-degrading bacterium, *Sphingomonas* sp. CFO6. Plasmid 37; 169-179.

Ferrante AA and Lessie TG. 1991. Nucleotide sequence of IS*402* from *Pseudomonas cepacia*. Gene 102; 143-144.

Fiandt M, Szybalski W and Malamy MH. 1972. Polar mutations in *lac, gal*, and phage λ consist of a few IS-DNA sequences inserted with either orientation. Mol. Gen. Genet. 119; 223-231.

Flavier AB, Clough SJ, Schell MA and Denny TP. 1997. Identification of 3-hydroxypalmitic acid methyl ester as a novel autoregulator controlling virulence in *Ralstonia solanacearum*. Mol. Microbiol. 26(2); 251-259.

Flavier AB, Ganova-Raeva LM, Schell MA and Denny TP. 1997. Hierarchical autoinduction in *Ralstonia solanacearum*: control of acyl-homoserine lactone production by a novel autoregulatory system responsive to 3-hydroxypalmitic acid methyl ester. J. Bacteriol. 179(22); 7089-7097.

Flavier AB, Schell MA and Denny TP. 1998. An RpoS(σ^s) homologue regulates acylhomoserine lactone-dependent autoinduction in *Ralstonia solanacearum*. Mol. Microbiol. 28(3); 475-486.

French ER. 1986. Interaction between strains of *Pseudomonas solanacearum*, its host and the environment. In *Bacterial wilt disease in Asia and the South Pacific. Proc. Int. Workshop, PCARRD, Los Banos, Philippines, 8-10 Octobe. 1985 ACIAR Proceedings No 13.* Ed. Persley GJ. pp 99-104.

French ER. 1988. Strategies for bacterial wilt control. In *Bacterial disease of the potato. Report of the planning conference on bacterial diseases of the potato, 1987* Ed. International Potato Center (CIP). pp 133-141.

French ER, Aley P, Torres, E and Nydegger U. 1993. Diversity of *Pseudomonas solanacearum* in peru and Brazil. In *Bacterial wilt proceedings of an international conference held at Kaohsiung, Taiwan, 28-31 October 1992 ACIAR Proceedings No 45.* Ed. Hayward AC and Hartman GL. pp 70-77.

French ER, Gutarra L, Aley P and Elphinstone J. 1995. Culture media for *Pseudomonas solanacearum* isolation, identification and maintenance. Fitopatologia 30(3); 126-130.

Fudyk TC, Maclean IW, Simonsen JN, Njagi EN, Kimani J, Brunham RC and Plummer FA. 1999. Genetic diversity and mosaicism at the *por* locus of *Neisseria gonorrhoeae*. J. Bacteriol. 181(18); 5591-5599.

Fuqua C, Winans SC and Greenberg EP. 1996. Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum-sensing transcriptional regulators. Annu. Rev. Microbiol. 50; 727-751.

Galán JE and Collmer A. 1999. Type III secretion machines: bacterial devices for protein delivery into host cells. Science 284; 1322-1328.

Galas DJ and Chandler M. 1989. Bacterial insertion sequences. In *Mobile DNA*. Ed. Berg D and Howe M. American Society for Microbiology. Washington, DC. pp. 109-162.

Garcia MI, Labigne A and Bouguence C. 1994. Nucleotide sequence of the afimbrial-adhesin-encoding *afa-3* gene cluster and its translocation via IS*1* insertion sequences. J. Bacteriol. 176(24); 7601-7613.

Genin S, Gough CL, Zischek C and Boucher CA. 1992. Evidence that the *hrpB* encodes a positive regulator of *hrp* genes from *Pseudomonas solanacearum*. Mol. Microbiol. 6; 3065-3076.

Gerischer U, D'Argenio DA and Ornston LN. 1996. IS*1236*, a newly discovered member of the IS*3* family, exhibits varied patterns of insertion into the *Acinetobacter calcoaceticus* chromosome. Microbiol. 142; 1825-1831.

Gillings M, Fahy P and Davies C. 1993. Restriction analysis of an amplified polygalacturonase gene fragment differentiates strains of the phytopathogenic bacterium *Pseudomonas solanacearum*. Letters in Applied Microbiology 17; 44-48.

Gish W and States DJ. 1993. Identification of protein coding regions by database similarity search. Nat. Genet. 3; 266-272.

Gough CL, Genin S, Zischek C and Boucher CA. 1992. *hrp* genes of *Pseudomonas solanacearum* are homologous to pathogenicity determinants of animal pathogenic bacteria and are conserved among plant pathogenic bacteria. Mol. Plant-Microbe Interact. 5; 384-389.

Grindley ND and Leschziner AE. 1995. DNA transposition: from a black box to a color monitor. Cell 83; 1063-1066.

Grunstein M and Hogness D. 1975. Colony hybridisation: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA 72; 3961.

Hacker J, Blum-Oehler G, Mühldorfer I and Tschäpe H. 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol. Microbiol. 23; 1089-1097.

Hallet B and Sherratt DJ. 1997. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. FEMS Microbiol. Rev. 21; 157-178.

Hammerschmidt S, Hilse R, van Putten JP, Gerardy-Schahn R, Unkmeir A and Frosch M. 1996. Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element. EMBO J. 15; 192-198.

Hammond SM, Lambert PA and Rycroft AN. 1984. The bacterial cell surface. Croom Helm.

Harris DC. 1972. Intraspecific variation in *Pseudomonas solanacearum. In Proceedings of the third international conference of plant pathogenic bacteria.* Wageningen, The Netherlands. pp289-292.

Hay B and Short JM. 1992. ExAssist™helper phage and SOLR™ cells for Lambda ZAP®II excisions. Strategies 5; 16-18.

Hayward AC. 1964. Characteristics of *Pseudomonas solanacearum*. J. Appl. Bacteriol. 27; 265-277.

Hayward AC. 1976. Systematics and relationships of *Pseudomonas solanacearum*. In *Proceedings of the 1st planning conference and workshop on the ecology and control of bacterial wilt caused by Pseudomonas solanacearum. Raleigh, North Carolina State University.* Ed. Sequeira L and Kelman A. pp 6-21.

Hayward AC. 1991. Biology and epidemiology of bacterial wilt caused by *Pseudomonas solanacearum*. Ann. Rev. Phytopathol. 29; 65-87.

Hayward AC. 1994a. The hosts of *Pseudomonas solanacearum*. In *Bacterial Wilt: The disease and its causative agent, Pseudomonas solanacearum.* Ed. Hayward AC and Hartman GL. CAB International, UK. pp 9-24.

Hayward AC. 1994b. Systematics and phylogeny of *Pseudomonas solanacearum* and related bacteria. In *Bacterial Wilt: The disease and its causative agent, Pseudomonas solanacearum.* Ed. Hayward AC and Hartman GL. CAB International, UK. pp. 123-135.

Hayward AC, El-Nashaar HM, Nydegger U and de Lindo L. 1990. Variation in nitrate metabolism in biovars of *Pseudomonas solanacearum*. J. Appl. Bacteriol. 69; 269-280.

Hayward AC, Sequeira L, French ER, El-Nashaar H and Nydegger U. 1991. Tropical strain of biovar 2 of *Pseudomonas solanacearum*. In *Proceedings of the XV meeting of the asociation latinoamericana de la Papa, March, 1991, Lima, peru (Abstract).*

He LY. 1990. Control of bacterial wilt of groundnut in China with emphasis on cultural and biological properties. In *Bacterial wilt of Groundnuts. Proc. ACIAR/ICRISAT collaborative res. Plan. Meet., Genting Highlands, Malaysia, 18-19 March. ACIAR Proceedings No 31.* Ed. Middleton KJ and Hayward AC. pp 22-25.

He LY, Sequeira L and Kelman A. 1983. Characteristics of strains of *Pseudomonas solanacearum* from China. Plant Disease 67; 1357-1361.

Healy FG, Bukhalid RA and Loria R. 1999. Characterization of an insertion sequence element associated with genetically diverse plant pathogenic *Streptomyces* spp.. J. Bacteriol. 181; 1562-1568.

Hirsch H-J, Starlinger P and Brachet P. 1972. Two kinds of insertions in bacterial genes. Mol. Gen. Genet. 119; 191-206.

Hodgson CP and Fisk RZ. 1987. Hybridisation probe size control: optimized oligolabelling. Nucleic Acids Research 15; 6259.

Holloway BW, Bowen AR St G and Escuadra MD. 1993. Whole genome analysis of *Pseudomonas*. In *Bacterial wilt proceedings of an international conference held at Kaohsiung, Taiwan, 28-31 October 1992 ACIAR Proceedings No 45.* Ed. Hayward AC and Hartman GL.

Holmes DS and Quigley M. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 14: 193-197.

Höltje J-V and Schwarz U. 1985. Biosynthesis and growth of the murein sacculus. In *Molecular cytology of Escherichia coli.* Ed. Nanninga N. Academic Press, Inc. (London), Ltd., London.

Huang Q and Allen C. 1997. An exo-poly-α-D-galacturonase, PehB, is required for wild-type virulence of *Rslatonia solanacearum*. J. Bacteriol. 179(23); 7369-7378.

Huang J, Carney BF, Denny TP, Weissinger AK and Schell MA. 1995. A complex network regulates expression of *eps* and other virulence genes of *Pseudomonas solanacearum*. J. Bacteriol. 177(5); 1259-1267.

Huang J, Denny TP and Schell MA. 1993. VsrB, a regulator of virulence genes of *Pseudomonas solanacearum*, is homologous to sensors of the two-component regulatory family. J. Bacteriol. 175(19); 6169-6178.

Huang J and Schell MA. 1990a. Evidence that extracellular export of the endoglucanase encoded by *egl* of *Pseudomonas solanacearum* occurs by a two-step process involving a lipoprotein intermediate. J. Biol. Chem. 265; 11628-11632.

Huang J and Schell MA. 1990b. DNA sequence analysis of *pglA* and mechanism of export of its polygalacturonase product from *Pseudomonas solanacearum*. J. Bacteriol. 172; 3879-3887.

Huang J and Schell MA. 1992. Role of the two-component leader sequence and mature amino acid sequences in extracellular export of endoglucanase EGL from *Pseudomonas solanacearum*. J. Bacteriol. 174(4); 1314-1323.

Huang J and Schell MA. 1995. Molecular characterization of the *eps* gene cluster of *Pseudomonas solanacearum* and its transcriptional regulation at a single promoter. Mol. Microbiol. 16(5); 977-989.

Huang J, Sukordhaman M and Schell MA. 1989. Excretion of the *egl* gene product of *Pseudomonas solanacearum*. J. bacteriol. 171; 3767-3774.

Husain A and Kelman A. 1958. Relation of slime production to mechanism of wilting and pathogenicity of *Pseudomonas solanacearum*. Phytopathology 48; 155-165.

Hyland SA, Eveland SS and Anderson MS. 1997. Cloning, expression, and purification of UDP-3-*O*-acyl-GlcNAc deacetylase from *Pseudomonas aeruginosa*: a metalloamidase of the lipid A biosynthesis pathway. J. Bacteriol. 179(6); 2029-2037.

Iida S, Meyer J and Arber W. 1983. Prokaryotic IS elements. In *Mobile Genetic Elements*. Ed. Shapiro JA. New York. Academic Press, pp 159-221.

Ito S, Ushijima Y, Fujii T, Tanaka S, Kameya-Iwaki M, Yoshiwara S and Kishi F. 1998. Detection of viable cells of *Ralstonia solanacearum* in soil using a semiselective medium and a PCR technique. J. Phytopathology 146; 379-384.

Iversen T-G, Standal R, Pedersen T and Coucheron DH. 1994. IS*1032* from *Acetobacter xylinum*, a new mobile insertion sequence. Plasmid 32; 46-54.

Janse JD. 1991. Infra and intraspecific classification of *Pseudomonas solanacearum* strains using whole cell fatty acid analysis. Syst. Appl. Microbiol. 14; 335-345.

Janse JD. 1996. Potato brown rot in western Europe-history, present occurrence and some remarks on possible origin, epidemiology and control strategies. Bull. OEPP 26; 679-695.

Jap BK, Walian PJ and Gehring K. 1991. Structural architecture of an outer membrane channel as determined by electron crystallography. Nature 350; 167-170.

Jeanteur D, Lakey JH and Pattus F. 1991. The bacterial porin superfamily: sequence alignment and structure prediction. Mol. Microbiol. 5(9); 2153-2164.

Jeong E-L and Timmis JN. 2000. Novel insertion sequence elements associated with genetic heterogeneity and phenotype conversion in *Ralstonia solanacearum*. J. Bacteriology 182(16); 4673-4676.

Johnson JM and Church GM. 1999. Alignment and structure prediction of divergent protein families: periplasmic and outer membrane proteins of bacterial efflux pumps. J. Mol. Biol. 287; 695-715.

Jordan E, Saedler H and Starlinger P. 1968. 0-zero and strong polar mutations in the gal operon are insertions. Mol. Gen. Genet. 102; 353-363.

Kaiser D. 1996. Bacteria also vote. Science 272; 1598.

Kallastu A, Hõrak R and Kivisaar M. 1998. Identification and characterization of IS*1411*, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida*. J. Bacteriol. 180; 5306-5312.

Kang Y, Saile E, Schell MA and Denny TP. 1999. Quantitative immunofluorescence of regulated *eps* gene expression in single cells of *Ralstonia solanacearum*. Appl. Environ. Microbiol. 65(6); 2356-2362.

Kao CC, Barlow E and Sequeira L. 1992. Extracellular polysaccharide is required for wild-type virulence of *Pseudomonas solanacearum*. J. Bacteriol. 174(3); 1068-1071.

Kao CC, Gosti F, Huang Y and Sequeira L. 1994. Characterization of a negative regulator of exopolysaccharide production by the plant-pathogenic bacterium *Pseudomonas solanacearum*. Mol. Plant-Microbe Interact. 7(1); 121-130.

Kao CC and Sequeira L. 1991. A gene cluster required for coordinated biosynthesis of lipopolysaccharide and extracellular polysaccharide also affects virulence of *Pseudomonas solanacearum*. J. Bacteriol. 173(24); 7841-7847.

Kao CC and Sequeira L. 1994. The function and regulation of genes required for extracellular polysaccharide synthesis and virulence in *Pseudomonas solanacearum*. In *Molecular mechanisms of bacterial virulence.* Kluwer Academic, Dordrecht, The Netherlands. Ed Kado CI and Crosa JH. pp 93-108.

Kelman A. 1953. The bacterial wilt caused by *Pseudomonas solanacearum*. North Carolina Agricultural Experiment Station Technical Bulletin. 99; 194.

Kelman A. 1954. The relationship of pathogenicity in *Pseudomonas solanacearum* to colony appearance on a tetrazolium medium. Phytopathology 44; 693-695.

Kelman A, Hartman GL and Hayward AC. 1994: Introduction. In *Bacterial Wilt: The disease and its causative agent, Pseudomonas solanacearum* Ed. Hayward AC and Hartman GL. CAB International, UK. 1-8.

Kelman A and Sequeira L. 1965. Root-to-root spread of *Pseudomonas solanacearum*. Phytopathology 55; 304-309.

Kennedy AK and Haniford DB. 1996. Isolation and characterization of IS*10* transposase separation of function mutants: identification of amino acid residues in transposase that are important for active site function and the stability of transposition intermediates. J. Mol. Biol. 256; 533-547.

Kersters K, Ludwig W, Vancanneyt M, De Vos P, Gillis M and Schleifer K-H. 1996. Recent changes in the classification of *Pseudomonas*: an overview. Syst. Appl. Microbiol. 19; 465-477.

Khan E, Mack JPG, Katz RA, Kulkosky J and Skalka AM. 1991. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. Nucleis Acids research 19; 851-860.

Kim JF, Charkowski AO, Alfano JR, Collmer A and Beer SV. 1998. Sequences related to transposable elements and bacteriophages flank avirulence genes of *Pseudomonas syringae*. Mol. Plant-Microbe Interact. 11; 1247-1252.

Kivisaar M, Kasak L and Nurk A. 1991. Sequence of the plasmid-encoded catechol 1,2-dioxygenase-expressing gene, *pheB*, of phenol-degrading *Pseudomonas* sp. strain EST1001. Gene 98; 15-20.

Klaer R, Kühn S, Tillman E, Fritz H-J and Starlinger P. 1981. The sequence of IS*4*. Mol. Gen. Genet. 181; 169-175.

Kondo K and Horinouchi S. 1997a. Characterization of an insertion sequence, IS*12528*, from *Gluconobacter suboxydans*. Appl. Environ. Microbiol. 63; 1139-1142.

Kondo K and Horinouchi S. 1997b. A new insertion sequence IS*1452* from *Acetobacter pasteurianus*. Microbiol. 143; 539-546.

Krishnapillai V. 1996. Horizontal gene transfer. J. Genet. 75(2); 219-232.

Kröger M and Hobom G. 1982. Structural analysis of insertion sequence IS*5*. Nature 297; 159-162.

Kyrpides NC and Olsen GJ. 1999. Archaeal and bacterial hyperthermophiles: horizontal gene exchange or common ancestry? Trends in Genetics 15; 298-299.

Labrador M and Corces VG. 1997. Transposable element-host interactions: regulation of insertion and excision. Annu. Rev. Genet. 31; 381-404.

Laine MJ, Zhang Y-P and Metzler MC. 1994. IS*1237*, a repetitive chromosomal element from *Clavibacter xyli* subsp. *cynodontis*, is related to insertion sequences from gram-negative and gram-positive bacteria. Plasmid 32; 270-279.

Lawrence JG, Ochman H and Hartl DL. 1992. The evolution of insertion sequences within enteric bacteria. Genetics 131; 9-20.

Leveau JHJ and van der Meer JR. 1997. Genetic characterization of insertion sequence ISJP4 on plasmid pJP4 from *Ralstonia eutropha* JMP134. Gene 202; 103-114.

Lewin B. 1994. Genes V. Oxford University Press, Oxford.

Lewis JP and Macrina FL. 1998. IS*195*, an insertion sequence-like element associated with protease genes in *Porphyromonas gingivalis*. Infection and Immunity 66(7); 3035-3042.

Liao BS, Shan ZH, Duan NX, Tan YJ, Lei Y, Li D and Mehan VK. 1998. Relationship between latent infection and groundnut bacterial wilt resistance. In *Bacterial wilt disease-molecular and ecological aspects.* Ed. Prior P, Allen C and Elphinstone J. Springer. 294-299.

Lieb M. IS*5* increases recombination in adjacent regions as shown for the repressor gene of coliphage lambda. Gene 12; 277-280.

Lindgren PB, Peet RC and Panopoulos NJ. 1986. Gene cluster of *Pseudomonas syringae* pv. *phaseolicola* controls pathogenicity of bean plants and hypersensitivity of non-host plants. J. Bacteriol. 168; 512-522.

Ludwig W, Rosselló-Mora R, Aznar R, Klugbauer S, Spring S, Reetz K, Beimfohr C, Brockmann E, Kirchhof G, Dorn S, Bachleither M, Klugbauer N, Springer N, Lane D, Nietupsky R, weizenegger M, Schleifer K-H. 1995. Comparative sequence analysis of 23S rRNA from Proteobacteria. Syst. Appl. Microbiol. 18; 164-188.

Lugtenberg B. 1985. Structure and function of outer membrane proteins. In *Enterobacterial surface antigens-methods for molecular characterisation.* Ed. Korhonen TK, Dawes EA and Mäkelä PH. Elsevier-FEMS symposium No. 25. 3-16.

Lum KY. 1973. Cross inoculation studies of *Pseudomonas solanacearum* from ginger. MARDI Res. Bull. 1; 15-21.

Ma QS, Chang MF, Tang JL, Feng JX, Fan MJ, Han B and Liu T. 1988. Identification of DNA sequences involved in host specificity in the pathogenesis of *Pseudomonas solanacearum* strain T2005. Mol. Plant-Microbe Interact. 1; 169.

Maekawa T and Ohtsubo E. 1994. Identification of the region that determines the specificity of binding of the transposases encoded by Tn*3* and γδ to the terminal inverted repeat sequences. Jpn. J. Genet. 69; 269-285.

Mahillon J and Chandler M. 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. 62; 725-774.

Maizel and Lenk. 1981. Enhanced Graphic Matrix Analysis of Nucleic Acid and Protein Sequences. Proc. Natl. Acad. Sci. USA. 78; 7665-7669.

Malamy MH. 1966. Frameshift mutations in the lactose operon of *E. coli*. Cold Spring Harb. Symp. Quant. Biol. 31; 189-201.

Marenda M, Brito B, Callard D, Genin S, Barberis P, Boucher CA and Arlat M. 1998. PrhA controls a novel regulatory pathway required for the specific induction of *Ralstonia solanacearum hrp* genes in the presence of plant cells. Mol. Microbiol. 27(2); 437-453.

McWilliams R, Chapman M, Kowalczuk KM, Hersberger D, Sun J-H and Kao CC. 1995. Complementation analyses of *Pseudomonas solanacearum* extracellular polysaccharide mutants and identification of genes responsive to EpsR. Mol. Plant-Microbe Interact. 8(6); 837-844.

Meneghetti F, Alberghini S, Tola E, Giacomini A, Ollero FJ, Squartini A and Nuti MP. 1996. Presence of unique repeated insertion sequences in nodulation genes of *Rhizobium 'hedysari'*. 186; 113-120.

Menéndez C, Igloi GL and Brandsch R. 1997. IS*1473*, a putative insertion sequence identified in the plasmid pAO1 from *Arthrobacter nicotinovorans*: isolation, characterization, and distribution among *Arthrobacter* species. Plasmid 37; 35-41.

Meyer TF. 1991. Evasion mechanisms of pathogenic *Neisseriae*. Behring Inst. Mitt. 88; 194-199.

Napiere CM. 1980. Varying inoculum levels of bacteria-nematodes and the severity of tomato bacterial wilt. Annals of Tropical Research 2; 129-134.

Needleman SB and Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48; 443-453.

Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Nelson WC, Ketchum KA, McDonald L, Utterback TR, Malek JA, Linher KD, Garrett MM, Stewart AM, Cotton MD, Pratt MS, Phillips CA, Richardson D, Heidelberg J, Sutton GG, Fleischmann RD, Eisen JA, White

O, Salzberg SL, Smith HO, Venter JC and Fraser CM. 1999. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. Nature 399; 323-329.

Nikaido H. 1992. Porins and specific channels of bacterial outer membranes. Mol. Microbiol. 6(4); 435-442.

Nikaido H. 1993. Transport across the bacterial outer membrane. J. Bioenergetics and Biomembranes 25(6); 581-589.

Nurk A, Kasak L and Kivisaar M. 1991. Sequence of the gene (*pheA*) encoding phenol monooxygenase from *Pseudomonas* sp. EST1001: expression in *Escherichia coli* and *Pseudomonas putida*. Gene 102; 13-18.

Nyangeri JB, Gathuru EM and Mukunya DM. 1984. Effect of latent infection on the spread of bacterial wilt of potatoes in Kenya. Trop. Pest Manage 30; 163-165.

Ohtsubo E and Sekine Y. 1996. Bacterial insertion sequences. Current Topic in Microbiology and Immunology (vol. 204). Ed. Saedler H and Gierl A. Springer-Verleg. 1996.

Oka A, Sugisaki H and Takanami M. 1981. Nucleotide sequence of the kanamycin resistance transposon Tn*903*. J. Mol. Biol. 147; 217-226.

Olasz F, Farkas T, Kiss J, Arini A and Arber W. 1997. Terminal inverted repeats of insertion sequence IS*30* serve as targets for transposition. J. Bacteriol. 179(23); 7551-7558.

Olasz F, Kiss J, König P, Buzás Z, Stalder R and Arber W. 1998. Target specificity of insertion element IS*30*. Mol. Microbiol. 28(4); 691-704.

Opina N, Tavner F, Hollway G, Wang J-F, Li T-H, Maghirang R, Fegan M, Hayward AC, Krishnapillai V, Hong WF, Holloway BW and Timmis JN. 1997. A novel method for development of species and strain-specific DNA probes and PCR primers for identifying *Burkholderia solanacearum* (formerly *Pseudomonas solanacearum*). Asia Pacific J. Molecular Biology and Biotechnology 5(1); 19-30.

Orgambide G, montrozier H, Servin P, Roussel J, Trigalet-Demery D and Trigalet A. 1991. High heterogeneity of the exopolysaccharides of *Pseudomonas solanacearum* strain GMI1000 and the complete structure of the major polysaccharide. J. Biol. Chem. 266; 8312-8321.

Palleroni NJ. 1975. General properties and taxonomy of the genus *Pseudomonas*. In Genetics and biochemistry of *Pseudomonas*. Ed. Clarke PH and Richmond MH. John Wiley and Sons. London. pp 1-37.

Palleroni NJ, Kunisawa R, Contopoulou R and Doudoroff M. 1973. Nucleic acid homologies in the genus *Pseudomonas*. Int. J. Syst. Bacteriol. 23; 333-339.

Pannetier C, Delassus S, Darche S, Saucier C and Kourilsky P. 1993. Quantitative titration of nucleic acids by enzymatic amplification reactions run to saturation. Nucleic Acids Research 21(3); 577-583.

Paulsen IT, Park JH, Choi PS and Saier Jr MH. 1997. A family of Gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals from Gram-negative bacteria. FEMS Microbiology letters 156; 1-8.

Pegg KG and Moffett ML. 1971. Host range of the ginger strain of *Pseudomonas solanacearum* in Queensland. Australian Journal of Experimental Agriculture and Animal Husbandry 11; 696-698.

Polard P, Prère MF, Chandler M and Fayet O. 1991. Programmed translational frameshifting and initiation at an AUU codon in gene expression of bacterial insertion sequence IS*911*. J. Mol. Biol. 222; 465-477.

Polard P, Prère MF, Fayet O and Chandler M. 1992. Transposase-induced excision and circularization of the bacterial insertion sequence IS*911*. EMBO J. 11; 5079-5090.

Poussier S, Vandewalle P and Luisetti J. 1999. Genetic diversity of African and worldwide strains of *Ralstonia solanacearum* as determined by PCR-restriction fragment length polymorphism analysis of the *hrp* gene region. Appl. Environ. Microbiol. 65(5); 2184-2194.

Quintiliani RJ and Courvalin P. 1996. Characterization of Tn*1547*, a composite transposon flanked by the IS*16* and IS*256*-like elements, that confers vancomycin resistance in *Enterococcus faecalis* BM4281. Gene 172; 1-8.

Ralston E, Palleroni NJ and Doudoroff M. 1973. *Pseudomonas pickettii*, a new species of clinical origin related to *Pseudomonas solanacearum*. Int. J. Syst. Bacteriol. 23; 15-19.

Ratnaningsih E, Dharmsthiti S, Krishnapillai V, Morgan A, Sinclair M and Holloway BW. 1990. A combined physical and genetic map of *Pseudomonas aeruginosa* PAO. Journal of General Microbiology 136; 2351-2357.

Reinitz DM, Inverso JA and Mansfield JM. 1989. Complete nucleotide sequence of an *E. coli* IS*5* insertion element containing an internal 88 base pair direct repeat (IS*5*-D). Nucleic Acids Research 17(10); 3990.

Rezsöhazy R, Hallet B and Delcour J. 1992. IS*231*D, E and F, three new insertion sequences in *Bacillus thuringiensis*: extension of the IS*231* family. Mol. Microbiol. 6(14); 1959-1967.

Rezsöhazy R, Hallet B, Delcour J and Mahillon J. 1993. The IS*4* family of insertion sequences: evidence for a conserved transposase motif. Mol. Microbiol. 9; 1283-1295.

Rice P, Craigie R and Davies DR. 1996. Retroviral integrases and their cousins. Curr. Opin. Struct. Biol. 6; 76-83.

Roberts DP, Denny TP and Schell MA. 1988. Cloning of the *egl* gene of *Pseudomonas solanacearum* and analysis of its role in phytopathogenicity. J. Bacteriol. 170(4); 1445-1451.

Robles J and Doers M. 1994. pGEM®-T vector systems troubleshooting guide. Promega Notes 45; 19-20.

Roth JR. 1974. Frameshift mutations. Annu. Rev. Genet. 8; 319-346.

Saedler H, Reif HJ, Hu S and Davidson N. 1974. IS*2*, a genetic element for turn-off and turn-on of gene activity in *E. coli*. Mol. Gen. Genet. 132; 265-289.

Sambrook J, Fritsch EF and Maniatis T. 1989. Molecular cloning : A laboratory manual. (2nd Ed.). Cold spring harbour laboratory press, Cold spring harbour, USA.

Sánchez-Beato A, García E, López R and García JL. 1997. Identification and characterization of IS*1381*, a new insertion sequence in *Streptococcus pneumoniae*. J. Bacteriol. 179(7); 2459-2463.

Sanger F, Nicklen S and Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74; 5463-5467.

Saunders JR, O'Sullivan H, Wakeman J, Sims G, Hart CA, Virji M, Heckels JE, Winstanley C, Morgan JA and Pickup RW. 1993. Flagella and pili as antigenically variable structures on the bacterial surface. J. Appl. Bacteriol. 74 (suppl.); 33S-42S.

Sawadogo M and van Dyke MW. 1991. A rapid method for the purification of deprotected oligodeoxynucleotides. Nucleic Acids Research 19; 674.

Schell MA. 1987. Purification and characterization of an endoglucanase from *Pseudomonas solanacearum*. Appl. Environ. Microbiol. 53; 2237-2241.

Schell MA. 1996. To be or not to be: how *Pseudomonas solanacearum* decides whether or not to express virulence genes. European J. Plant Pathology 102; 459-469.

Schell MA, Denny TP, Clough SJ and Huang J. 1993. Further characterization of genes encoding extracellular polysaccharide of *Pseudomonas solanacearum* and their regulation. In *Advances in molecular genetics of plant-microbe interactions vol. 2.* Ed Nester EW and Verma DPS. Kluwer Academic, Dordrecht, The Netherlands. pp 231-239.

Schell MA, Denny TP and Huang J. 1994a. Extracellular virulence factors of *Pseudomonas solanacearum*: role in disease and their regulation. In *Molecular mechanisms of bacterial virulence* Ed Kado CI and Crosa JH. Kluwer Academic, Dordrecht, The Netherlands. pp 311-324.

Schell MA, Denny TP and Huang J. 1994b. VsrA, a second two-component sensor regulating virulence genes of *Pseudomonas solanacearum*. Mol. Microbiol. 11; 489-500.

Schell MA, Roberts DP and Denny TP. 1988. Cloning of the *pglA* gene of *Pseudomonas solanacearum* and its involvement in phytopathogenicity. J. Bacteriol. 170; 4501-4508.

Schmiediche P. 1986. Breeding potatoes for resistance to bacterial wilt caused by *Pseudomonas solanacearum*. In *Bacterial wilt disease in Asia and the South Pacific. Proc. Int. Workshop, PCARRD, Los Banos, Philippines, 8-10 Octobe. 1985 ACIAR Proceedings No 13.* Ed. Persley GJ. pp 105-111.

Schmiediche P. 1988. Breeding for resistance to *Pseudomonas solanacearum*. In *Bacterial disease of the potato. Report of the planning conference on bacterial diseases of the potato, 1987* Ed. International Potato Center (CIP) pp 19-27.

Schmit J. 1978. Microscopic study of early stages of infection by *Pseudomonas solanacearum* E.F.S. on "*in vitro*" grown tomato seedlings. In *Proceedings of the forth international conference on plant pathogenic bacteria. Station de Pathologie Vegetale et Phytobacteriologie, Anger, France, 27 August-2 September, 1978.* INRA, Angers, France pp 841.

Schoner B and Kahn M. 1981. The nucleotide sequence of IS*5* from *Escherichia coli*. Gene 14; 165-174.

Seal SE, Jackson LA, Young JPW and Daniels MJ. 1993. Differentiation of *Pseudomonas solanacearum, pseudomonas syzygii, pseudomonas pickettii* and the blood disease bacterium by partial 16S rRNA sequencing: construction of oligonucleotide primers for sensitive detection by polymerase chain reaction. J. General Microbiology 139; 1587-1594.

Seifert HS and So M. 1988. Genetic mechanisms of bacterial antigen variation. Microbiol. Rev. 52; 327-336.

Sekine Y, Eisaki N and Ohtsubo E. 1994. Translational control in production of transposase and in transposition of insertion sequence IS*3*. J. Mol. Biol. 235; 1406-1420.

Sekine Y and Ohtsubo E. 1989. Frameshifting is required for expression of IS*1* transposase. Proc. Natl. Acad. Sci. USA 86; 4609-4613.

Shapiro JA. 1969. Mutations caused by insertion of genetic material into the galactose operon of *Escherichia coli*. J. Mol. Biol. 40; 93-105.

Sharp PM and Li W-H. 1987. The codon adaptation index- a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Research 15(3); 1281-1295.

Shekhawat GS, Gadewar AV, Bahal VK and Verma RK. 1988. Cultural practices for managing bacterial wilt of potatoes. In *Bacterial disease of the potato. Report of the planning conference on bacterial diseases of the potato, 1987* Ed. International Potato Center (CIP). pp 65-84.

Sherratt D. 1989. Tn3 and related transposable elements; site-specific recombination and transposition. In *Mobile DNA.* Ed. Berg D and Howe M. American Society for Microbiology, Washington, DC. pp. 163-184.

Short JM, Fernandez JM, Sorge JA and Huse WD. 1988. λZAP: a bacteriophage λ expression vector with in vivo excision properties. Nucleic Acids Research 16; 7583-7600.

Simons RW and Kleckner N. 1983. Translational control of IS*10* transposition. Cell 34; 683-691.

Smith EF. 1896. A bacterial disease of tomato, pepper, eggplant and irish potato (*Bacillus solanacearum* nov. sp.). United States Department of Agriculture, Division of Vegetable Physiology and Pathology Bulletin 12; 1-28.

Smith D and Waterman R. 1980. Adv. Appl. Math.2; 482-489.

Solinas F, Marconi AM, Ruzzi M and Zennaro E. 1995. Characterization and sequence of a novel insertion sequence, IS*1162*, from *Pseudomonas fluorescens*. Gene 155; 77-82.

Spök A, Stubenrauch G, Schorgendorfer K and Schwab H. 1991. Molecular cloning of a pectin esterase gene from *Pseudomonas solanacearum*. J. Gen. Microbiol. 137; 131-140.

Stackebrandt E, Murray RGE and Trüper HG. 1988. Proteobacteria classis nov., a name for the phylogenetic taxon including the "purple bacteria and their relatives". Int. J. syst. Bacteriol. 38; 321-325.

Staskawicz BJ, Dahlbeck D, Miller J and Damm D. 1983. Molecular analysis of virulence genes in *Pseudomonas solanacearum*. In *Molecular genetics of the bacteria-plant interaction.* Ed Pühler A. Springer-Verlag, Heidelberg. pp 345-352.

Stead DE. 1993. Classification and identification of *Pseudomonas solanacearum* and other pseudomonads by fatty acid profiling. ACIAR Proc. 45; 49-53.

Streisinger G and Owen JE. 1985. Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. Genetics 109; 633-659.

Stroeher UH and Manning PA. 1997. *Vibrio cholerae* serotype O139: swapping genes for surface polysaccharide biosynthesis. Trends in Microbiology 5(5); 178-180.

Sunaina V, Kishore V and Shekhawat GS. 1989. Latent survival of *Pseudomonas solanacearum* in potato tubers and weeds. Pflanzenkr. Pflanzenschutz 96; 361-364.

Taghavi M, Hayward AC, Sly LI and Fegan M. 1996. Analysis of the phylogenetic relationships of strains of *Burkholderia solanacearum*, *Pseudomonas syzygii*, and the blood disease bacterium of banana based on 16S rRNA gene sequences. Int. J. Syst. Bacteriol. 46; 10-15.

Tans-Kersten J, Guan Y and Allen C. 1998. *Ralstonia solanacearum* pectin methylesterase is required for growth on mathylated pectin but not for bacterial wilt virulence. Appl. Environ. Microbiol. 64(12); 4918-4923.

Tauch A, Zheng Z, Pühler A and Kalinowski J. 1998. *Corynebacterium striatum* chloramphenicol resistance transposon Tn*5564*: genetic organization and transposition in *Corynebacterium glutamicum*. Plasmid 40; 126-139.

Tavakoli NP, DeVost J and Derbyshire KM. 1997. Defining functional regions of the IS*903* transposase. J. Mol. Biol. 274; 491-504.

Tavner, F. 1993. Honours thesis. Department of Genetics, The University of Adelaide, South Australia, Australia.

Tenzen T, Matsutani S and Ohtsubo E. 1990. Site-specific transposition of insertion sequence IS*630*. J. Bacteriol. 172; 3830-3836.

Thomas CE, Carbonetti NH and Sparling PF. 1996. Pseudo-transposition of a Tn*5* derivative in *Neisseria gonorrhoeae*. FEMS Microbiology Letters 145; 371-376.

Thompson JD, Higgins DG and Gibson TJ. 1994. "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Research 22; 4673-4680.

Thuring RWJ, Sanders JPM and Borst P. 1975. A freeze-squeeze method for recovering long DNA from agarose gels. Anal. Biochem. 66; 213-220.

Thurston HD. 1976. Resistance to bacterial wilt (*Pseudomonas solanacearum*). In *Proc. Int. Plann. Conf. Workshop Ecol. Control. Bact. Wilt, 1st Raleigh. NC State Univ.* Ed. Sequeira L and Kelman A. pp 166.

Triglia T, Peterson MG and Kemp DJ. 1988. A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Research 16; 8186.

Tsujimoto H, Gotoh N, Yamagishi J-I, Oyamada Y and Nishino T. 1997. Cloning and expression of the major porin protein gene *opcP* of *Burkholderia* (formerly *Pseudomonas*) *cepacia* in *Escherichia coli*. Gene 186; 113-118.

Tümmler R, Römling U, Ratnaningsih E, Morgan AF, Krishnapillai V and Holloway BW. 1992. A common system of nomenclature for the physical map of the chromosome of *Pseudomonas aeruginosa* PAO. In *Pseudomonas, molecular biology and biotechnology*. Ed. Galli E, Silver S and Witholt B. Washington: American Society for Microbiology.

Tyler SD, Rozee KR and Johnson WM. 1996. Identification of IS*1356*, a new insertion sequence, and its association with IS*402* in epidemic strains of *Burkholderia cepacia* infecting cystic fibrosis patients. J. Clinical Microbiol. 34(7); 1610-1616.

Umeda M and Ohtsubo E. 1990. Mapping of insertion element IS*5* in the *Escherichia coli* K-12 chromoome: chromosomal rearrangements mediated by IS*5*. J. Mol. Biol. 213; 229-237.

van Alfen NK. 1982. Wilts: concepts and mechanisms. In *Phytopathogenic prokaryotes*. Ed. Mount MS and Lacy GH. Academic Press, New York pp 459-474.

van Alfen NK. 1989. Reassessment of plant wilt toxins. Annu. Rev. Phytopathol. 27; 533-550.

van Alfen NK, McMillan BD, Turner V and Hess WM. 1983. Role of pit membranes in macromolecular-induced wilt of plants. Plant physiology 73; 1020-1-23.

van Gijsegem F, Gough C, Zischek C, Niqueux E, Arlat M, Genin S, Barberis P, German S, Castello P and Boucher CA. 1995. The *hrp* gene cluster of *Pseudomonas solanacearum*, which controls the production of a type III secretion system, encodes eight prooteins related to components of the bacterial flagella biogenesis complex. Mol. Microbiol. 15; 1095-1114.

van Spanning RJM, De Boer APN, Slotboom D-J, Reijnders WNM and Stouthamer AH. 1995. Isolation and characterization of a novel insertion sequence element, IS*1248*, in *Paracoccus denitrificans*. Plasmid 34; 11-21.

Varanid L, Tattoos RL, Wolf YI, Walker DR and Koonin EV. 1998. Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. Trends in Genetics 14; 442-444.

Vögele K, Schwartz E, Welz C, Schiltz E and Rak B. 1991. High-level ribisomal frameshifting directs the synthesis of IS*150* gene products. Nucleic Acids Res. 19; 4377-4385.

Wall S, Ghanekar K, McFadden J and Dale JW. 1999. Context-sensitive transposition of IS*6110* in mycobacteria. Microbiology 145; 3169-3176.

Wallis FM and Truter SJ. 1978. Histopathology of tomato plants infected with *Pseudomonas solanacearum*, with emphasis on ultrastructure. Physiol. Plant Pathol. 13; 307.

Ward LJH, Brown JCS and Davey GP. 1996. Identification and sequence analysis of IS*1297*, an IS*S1*-like insertion sequence in a *Leuconostoc* strain. Gene 174; 259-263.

Weiss MS and Schulz GE. 1992. Structure of porin refined at 1.8 A resolution. J. Mol. Biol. 227; 493-509.

Williams MG and Rogers M. 1987. Expression of the *arg* genes of *Escherichia coli* during arginine limitation dependent upon stringent control of translation. J. Bacteriol. 169; 1644-1650.

Wood MS, Byrne A and Lessie TG. 1991. IS*406* and IS*407*, two gene-activating insertion sequences from *Pseudomonas cepacia*. Gene 105; 101-105.

Wullings BA, van Beuningen AR, Janse JD and Akkermans ADL. 1998. Detection of *Ralstonia solanacearum*, which causes brown rot of potato, by fluorescent in situ hybridisation with 23S rRNA-targeted probes. Appl. Environ. Microbiol. 64(11); 4546-4554.

Yabuuchi E, Kosako Y, Oyaizu H, Yano I, Hotta H, Hashimoto Y, Ezaki T and Arakawa M. 1992. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov.. Microbiology and immunology 36(12); 1251-1275.

Yabuuchi E, Kosako Y, Yano I, Hotta H and Nishiuchi Y. 1995. Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. nov.: Proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) comb. nov., *Ralstonia solanacearum* (Smith 1896) comb. nov. and *Ralstonia eutropha* (Davis 1969) comb.nov.. Microbiology and Immunology 39(11); 897-904.

Yamada T, Lee P-D and Kosuge T. 1986. Insertion sequence elements of *Pseudomonas savatanoi*: nucleotide sequence and homology with *Agrobacterium tumefaciens* transfer DNA. Proc. Natl. Acad. Sci.USA. 83; 8263-8267.

Yao W, Yang Y and Chiao J. 1994. IS*204*: an insertion sequence from *Nocardia asteroides* (*mexicana*) YP21. Plasmid 32; 262-269.

Yeo CC and Poh CL. 1996. IS*1394* from *Pseudomonas alcaligenes* N.C.I.B. 9867: identification and charaxterization of a member of the IS*30* family of insertion elements. Gene 175; 109-113.

Yeo CC and Poh CL. 1997. Characterization of IS*1474*, an insertion sequence of the IS*21* family isolated from *Pseudomonas alcaligenes* NCIB 9867. FEMS Microbiology Letters 149; 257-263.

Young K, Silver LL, Bramhill D, Cameron P, Eveland SS, Raetz CRH, Hyland SA and Anderson MS. 1997. The *envA* permeability/cell division gene of *Escherichia coli* encodes the second enzyme of lipid A biosynthesis UDP-3-*O*-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase. J. Biol. Chem. 270; 30384-30391.

Zachar V, Thomas RA and Goustin AS. 1993. Absolute quantification of target DNA: a simple competitive PCR for efficient analysis of multiple samples. Nucleic Acids Research 21(8); 2017-2018.

Zekri S and Toro N. 1996. Identification and nucleotide sequence of *Rhizobium meliloti* insertion sequence IS*Rm6*, a small transposable element that belongs to the IS*3* family. Gene 175; 43-48.

Ziebuhr W, Krimmer V, Rachid S, Löβner I, Götz F and Hacker J. 1999. A novel mechanism of phase variation of virulence in *Staphylococcus epidermidis*: evidence for control of the polysaccharide intercellular adhesin synthesis by alternating insertion and excision of the insertion sequence element IS*256*. Mol. Microbiol. 32(2); 345-356.

Errata

Eun-Lee Jeong, PhD thesis 2000 Insertion sequence elements in *Ralstonia solanacearum*: roles in genomic heterogeneity.

1. Page 72: the last to sentences in paragraph two should be substituted with the following text:

   The result suggested that IS*Rso*3 might be present throughout the genome of the strains tested, on the basis of the observation of several *Spe*I fragments hybridised with IS*Rso*3 probe. However, some of these *Spe*I fragments showed increased levels of hybridisation with the probe, implying that these genomic regions may have contained multiple copies of the element. The possibility of the presence of clusters of IS*Rso*3 in those regions could not be ruled out.

2. Figure 5.5: the following sentence should be added to the legend:

   These diagrams depict scenarios that are intended to explain the evolutionary origin of each of the IS elements discovered within *R.solanacearum*. They are not based on a formal phylogenetic analysis of the data.