PULP

GENETIC CHARACTERIZATION OF AN INDIGENOUS PLASMID IN *Xanthomonas albilineans*

by

KUGENTHIREN PERMAUL

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Deparment of Microbiology in the Faculty of Science at the University of Durban-Westville

Promoter : Prof. B. Pillay

Joint Promoter : Prof. D. Pillay

Date Submitted : February 1999

576.165
PER

This book must be returned on or before the last date stamped below.
Failure to do so, makes borrower liable to a fine of 5 cents per day.

2002 -04- 1 1

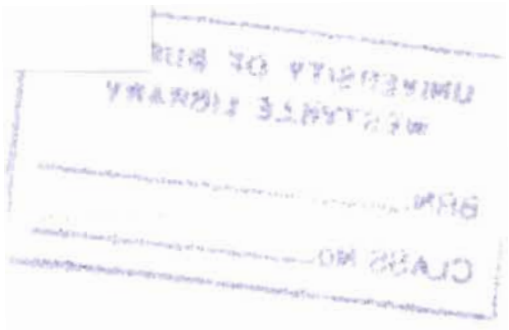
This book is not your property

Please treat it with respect; do not disfigure it by underlining, marking, scribbling, tearing out pages, etc.

is being submitted for the degree of Doctor

, Durban. It has not been submitted before

ty.



K. PERMAUL

February, 1999

A:990279B

UNIVERSITY OF
LIBRARY

TABLE OF CONTENTS

Acknowledgements

The author wishes to thank:

Prof. D. Pillay and Prof. B. Pillay, for their support for the duration of this study;

Prof. D.E. Rawlings and Prof. N.T. Keen, for guidance and provision of strains and vectors;

Mrs. D. James, for carrying out automated DNA sequencing of samples;

Mr. W.F. Rahaman, for jointly carrying out pathogenicity trials;

Dr. G. Govender, for digital capture of images;

Roshini, for her invaluable assistance, patience and support; and

the Foundation for Research Development and the UNESCO/Biotechnology Action Council for financial support and scholarships awarded.

List of Tables

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

Xanthomonas albilineans is a Gram-negative bacterium that causes leaf scald disease in sugarcane (*Saccharum* spp.). The disease is one of the major diseases of sugarcane and it occurs in every important sugar-producing region around the world. Severe losses were reported in the early years of this century but the disease has subsequently been effectively controlled by cultivation of sufficiently resistant varieties. However, the disease has in recent times, spread to regions that have not reported the disease previously (Hoy and Grisham, 1994; Irvine *et al.*, 1993; Isakeit and Irvine, 1995; Ovalle *et al.*, 1995) and re-occurred in countries in which the disease is endemic (Chen *et al.*, 1993; Comstock and Shine, 1992; Rott *et al.*, 1995). The disease therefore still poses a serious threat to the sugarcane industry.

Genetic studies that involve the identification of genes and gene products in the pathogen *X. albilineans*, are rare. Apart from a recent studies identifying gene clusters involved in albicidin biosynthesis and resistance (Wall and Birch, 1997), albicidin production (Rott *et al.*, 1996) and an earlier study identifying certain metabolic enzymes (Liu *et al.*, 1987; Lucas *et al.*, 1987), further genetic characterization of *X. albilineans* has not been reported. Genetic studies involving transfer of a broad-host range plasmid from *Escherichia coli* to *X. albilineans* and between *X. albilineans* strains showed that this plasmid can be stably maintained in the pathogen (Birch and Patil, 1987a). Molecular techniques have also been successfully used in studying genetic variation (Davis *et al.*, 1997) and specific detection (Honeycutt *et al.*, 1995) of the pathogen.

The presence of plasmid DNA in bacteria is common and they have frequently been identified for being of special benefit to their hosts or carrying desirable traits. While this is true for the well-known plasmids, the majority of plasmids in plant pathogenic bacteria have no well-

defined function (Coplin, 1989). It is becoming increasingly apparent that plasmids are stable in their natural hosts but this stability is not due to selective pressure as was previously thought, but due to plasmid-encoded phenomena. This new approach to the study of plasmids has resulted in the ability of plasmids to spread themselves in a bacterial population being regarded as a means of plasmid survival, rather than merely distribution of a particular trait.

In this study, genetic characterization of a plasmid found in a South African strain of *X. albilineans* was carried out in order to identify traits encoded by this plasmid and to determine whether the plasmid was involved in the disease process.

1.1 LEAF SCALD DISEASE

The exact origin of leaf scald disease remains unclear. The disease was probably present well before it was recognised, in the early 1920s, as a separate bacterial vascular disease of sugarcane by Wilbrink in Java, and North in Australia and Fiji (cited by Ricaud and Ryan, 1989). After being reported in Australia, Fiji and Java prior to 1920, the disease spread to the Philippines in 1923 (Lee, 1923), Mauritius in 1928 and Hawaii in 1930 (Martin and Robinson, 1961). It was confined to these Indian Ocean and mid-Pacific Ocean areas, but later spread to several South American countries (1940s - 1960s), to various Caribbean islands (1950s - 1960s), to North America, most of southern Africa and one west African country and to southern Asia in the 1960s (Egan, 1970). The spread of the disease into all major sugarcane-growing regions of the world and its increasing incidence is apparent after considering that in 1950, leaf scald was reported in nine countries (Martin, 1950), by 1961 in 15 countries (Martin and Robinson, 1961), by 1974 in 28 countries (International Society for Sugar-Cane Technologists [ISSCT] Standing Committee on Sugarcane Diseases, 1974), by 1989 in 44 countries and by 1994 in 58 geographic locations around the world (Rott *et al.*, 1994b). The disease now exists or has been recorded in 63 sugar-

producing regions in the world (Table 1.1). The disease was first recorded in KwaZulu Natal, South Africa and in neighbouring Swaziland in 1968 (Thomson, 1969), although it may have been present much earlier.

Two distinct forms of leaf scald disease occurs. These are called the chronic phase and the acute phase. Previously, the two phases were often regarded as separate diseases. The two phases sometimes occur independently, but more often there is a gradual change from the chronic phase to the acute phase (Martin and Robinson, 1961). Infection can also result in a symptomless latent phase. A fourth phase, known as the eclipse phase, has also been proposed (Rott, 1993). The symptoms of leaf scald disease and other aspects of the disease process have been extensively reviewed (Hughes, 1978; Martin and Robinson, 1961; Moffett and Croft, 1983; Ricaud and Ryan, 1989; Rott, 1993).

The chronic phase is characterized by the appearance of white "pencil-lines" that extend along the length of the leaf blades, parallel to the leaf veins (Fig. 1.1). This symptom is the most important diagnostic aid in visual identification of infected plants. The specific name of the pathogen is actually derived from the appearance of these streaks. It is often the only external symptom in resistant varieties. In young newly attacked cane, the long narrow white stripes are the earliest symptoms of infection. The white lines tend to broaden and become more diffuse as the leaves mature, leading to partial or complete chlorosis or whitening of the leaf blade.

Broadening of the streak coincides with withering of the leaf tissue. The withering process commences at the tip of the leaf and proceeds along the line of the streak, giving a scalded appearance from which the name of the disease is derived. Another characteristic feature of the chronic phase, especially on mature stalks, is the development of side shoots. These shoots also display the symptoms found on the main shoots. Internal symptoms of chronically infected stalks are characterized by bright-red colouration of the vascular bundles, especially at the nodal regions.

TABLE 1.1 Incidence of leaf scald disease in sugar-producing regions of the world (Davis *et al.*, 1997; Ricaud and Ryan, 1989; Rott *et al.*,1994b)

Argentina	India	St. Lucia
Australia	Indonesia	St. Vincent
Barbados	Jamaica	South Africa
Belize	Japan	Sri Lanka
Benin	Java	Surinam
Brazil	Kenya	Swaziland
Burkina Faso	Madagascar	Taiwan
Burma	Malawi	Tanzania
Cameroon	Malaysia	Thailand
Chad	Martinique	Trinidad
China	Mauritius	Uruguay
Congo	Mexico	USA: Florida
Cuba	Morocco	Hawaii
Dominican Republic	Mozambique	Louisiana
Fiji	Nigeria	Texas
Ghana	Pakistan	Venezuela
Grenada	Panama	Vietnam
Guadeloupe	Philippines	Zaire
Guatemala	Peurto Rico	Zimbabwe
Guyana	Réunion	
Ivory Coast	St. Kitts	

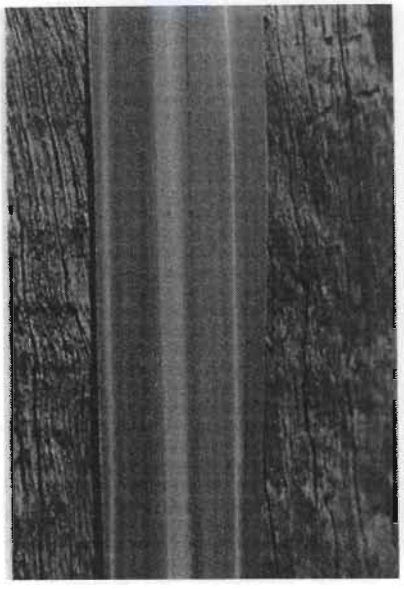


Fig. 1.1 White pencil-line stripes along the leaf blade which is characteristic of the chronic phase of leaf scald disease (Ricaud and Ryan, 1989).

Lysigenous cavities may develop in stalks which are severely diseased.

The acute phase is characterized by sudden wilting and death of mature stalks which often have not shown any previous chronic phase symptoms. It generally occurs after prolonged dry weather or dry weather following a rainy period (Ricaud and Ryan, 1989). This phase of the disease is usually observed only when intolerant, highly susceptible cultivars are grown in areas where the disease is endemic. The latent phase refers to the apparent recovery of plants after initially showing leaf scald symptoms. This usually occurs only in tolerant or resistant varieties. The pathogen is present in the plant during latency and this can be demonstrated by isolation techniques (Comstock and Shine, 1992; Davis *et al.*, 1994; Koike, 1992) and immunoassays (Comstock and Shine, 1992; Leoville and Coleno, 1976 - cited by Ricaud and Ryan, 1989). The mechanism of latency is not well understood but chronic as well as acute phase symptoms can occur once latency is terminated after ratooning or by climatic conditions. Latent infections can

result in undetected spread of the disease within and between sugarcane growing regions. A fourth phase of the disease, the eclipse phase, can occur at the same time as the latent phase. After symptoms have disappeared from the senescent leaves, the plant appears healthy and can be diagnosed as both diseased and healthy, depending on the date of observation (Rott, 1993).

Leaf scald is transmitted mainly by infected cuttings, especially if the disease is latent. It can also spread within a crop by contaminated harvesting implements (Taylor *et al.*, 1988). While other means of transmission are considered unlikely (Ricaud and Ryan, 1989), aerial transmission has been reported in Mauritius (Autrey *et al.*, 1992 - cited by Rott *et al.*, 1995). Klett and Rott (1994) have demonstrated presence of the causal organisms in water droplets on leaf surfaces, on dry leaf surfaces, in infected roots and rhizosphere soil, and provided evidence for aerial dispersal and fuzz transmission. Leaf scald can be effectively controlled by cultivation of resistant varieties. This involves breeding varieties that are resistant to the disease by using resistance genes which have all originated from *Saccharum spontaneum* clones (Egan, 1971). Strict control of the disease implies that all new varieties should undergo testing for resistance and the various factors that need to be considered in these tests were outlined by Koike (1971). Another useful control measure is the elimination of the disease in cuttings using a thermotherapy treatment (Egan and Sturgess, 1980; Steindl, 1971). Other measures of control involve disinfection of harvesting tools and strict quarantine procedures to avoid entry of new strains of the pathogen in a particular area (Ricaud and Ryan, 1989).

1.2 TOXIN PRODUCTION IN LEAF SCALD DISEASE

The production of phytotoxins by bacterial pathogens is common (Strobel, 1977). However, this is rare in the genus *Xanthomonas*. Toxin production has been reported in *X. campestris* pv. *manihotis* (Cherian and Mathew, 1983) and in *X. campestris* pv. *arecae* (Sampath

Kumar, 1984), but in these cases symptoms are due to bacterial exopolysaccharides rather than specific toxins. A few other toxins produced by xanthomonads that have not yet been chemically identified have been reported (Rudolph, 1993). *X. albilineans* produces a family of related toxins called albicidins which in addition to being phytotoxins, are bactericidal to a range of Gram-negative and Gram-positive bacteria (Birch *et al.*, 1990). These toxins, the major purified component called albicidin, are the only antibiotics produced by the genus *Xanthomonas.*

Orian (1942 - cited by Birch and Patil, 1983) noticed that chlorosis always developed on sugarcane leaves before they emerged from the spindle. He proposed that this effect was due to a toxic metabolite acting on plastids and that development of chlorotic zones around white lines were due to slow disorganization of plastids and consequent chlorophyll destruction. Bacterial production of toxins causing chlorosis in infected plants is well documented and examples are: tabtoxin in *Pseudomonas syringae* pv. *tabaci* and *P. syringae* pv. *coronafaciens* (Engst and Shaw, 1992); phaseolotoxin in *P. syringae* pv. *phaseolicola* (Mitchell, 1976); tagetitoxin in *P. syringae* pv. *tagetis* (Jutte and Durbin, 1979); syringomycin in *P. syringae* pv. *syringae* (Surico and De Vay, 1982); rhizobitoxin in *Rhizobium japonicum* (Johnson *et al.*, 1959); and a polysaccharide toxin in *X. campestris* pv. *arecae* (Sampath Kumar, 1984).

In an ultrastructural investigation, Birch and Patil (1983) confirmed that *X. albilineans* was confined to xylem vessels which often became occluded by the bacteria. They also confirmed Orian's finding that chloroplasts were absent from cells surrounding invaded vessels but, more importantly, also in uninvaded emerging white leaves. In these emerging leaves, immature forms of chloroplasts were observed whereas mature or degenerating chloroplasts were absent. This evidence suggested that the pathogen produced a diffusible phytotoxin that blocked chloroplast differentiation which resulted in chlorosis. Another study (Birch and Patil, 1985) showed that chlorosis-inducing isolates produced an antibiotic that inhibited growth of *E. coli*, whereas non-

chlorosis-inducing isolates failed to produce this antibiotic. This antibiotic, albicidin, was the major purified product of a family of structurally related compounds that completely blocked DNA synthesis in *E. coli* and partially blocked RNA and protein synthesis. Inhibition of DNA synthesis without evidence of DNA binding, suggested that albicidin interfered with an essential replication protein. Since chloroplasts employ procaryotic mechanisms of DNA synthesis and protein synthesis, and are thought to have evolved from ingested ancestral procaryotes, it seemed plausible that albicidin or a member of the albicidin family was responsible for leaf scald chlorosis. The antibiotic/phytotoxin was further characterized to reveal that it comprised approximately 38 carbon atoms arranged in a several-ringed aromatic structure (Birch and Patil, 1985). In a study involving mutants in albicidin production and revertants in which production was restored, close correlation was shown between albicidin production and the ability to cause chlorosis in sugarcane (Birch and Patil, 1987a). Mutants which lacked the ability to kill *E. coli* were found to also lack the ability to cause chlorosis, in contrast to antibiotic-producing strains. However, some non-inhibitor-producing strains still retained the ability to cause chlorosis. In addition, mutants still colonized plants and were present in high numbers in symptomless leaves. Further evidence that albicidin causes chlorosis by blocking plastid DNA replication was provided by Birch and Patil (1987b) in a paper describing the effects of purified albicidin on sugarcane plants and isolated plastids. Inhibition of DNA synthesis was found to be the primary mode of action of albicidin in undifferentiated plastids. This was not observed in mature chloroplasts. Permanent chlorosis was observed in callus cultures of sugarcane when albicidin was incorporated in the medium. However, chlorosis did not develop when intact plants were injected with purified albicidin.

1.2.1 **Resistance to Albicidin**

The identification of albicidin as the phytotoxin that causes chlorosis in leaf scald disease led to the search for mechanisms of albicidin resistance. The ultimate aim of such research is to create transgenic plants that would be albicidin-resistant. Resistance mechanisms identified thus far include altered uptake systems in *E. coli* (Birch *et al.*, 1990), albicidin-binding proteins in *Klebsiella oxytoca* (Walker *et al.*, 1988) and *Alcaligenes denitrificans* (Basnayake and Birch, 1995) and enzymatic detoxification in *Pantoea dispersa;* formerly *Erwinia herbicola* (Zhang and Birch, 1996). *E. coli* mutants that were resistant to albicidin were shown to have an altered nucleoside uptake pathway. Albicidin is normally accumulated by *E. coli* via the outer-membrane protein Tsx which actively transports the antibiotic into the cell. This explains why *E. coli* is inhibited by extremely low concentrations of albicidin (Birch *et al.*, 1990). Co-inoculation of *P. dispersa* and *X. albilineans* resulted in almost complete biocontrol against leaf scald disease (Zhang and Birch, 1996; 1997). This was due, in part, to the production of an enzyme, AlbD, that detoxified albicidin, but more importantly to the fact that *P. dispersa* outcompeted *X. albilineans* at the wound site (Zhang and Birch, 1997). The *X. albilineans* gene for albicidin resistance has recently been cloned (Wall and Birch, 1997).

1.3 THE BACTERIUM *X. albilineans*

Prior to the classification of bacteria into the genus *Xanthomonas*, the causal agent of leaf scald disease was known by a variety of generic names, including *Bacterium, Phytomonas, Agrobacterium*, and *Pseudomonas* (Ricaud and Ryan, 1989). The specific name *albilineans* was proposed by Ashby (1929 - cited by Ricaud and Ryan, 1989) and refers to the white lines caused by the pathogen. The bacterium is a short, slender, Gram-negative rod that is rapidly motile. It is the only species of the genus *Xanthomonas* that does not produce the extracellular polysaccharide, xanthan gum. The yellow colour of the colonies are due to the formation of

water-soluble, brominated aryl-polyene pigments called xanthomonadins, which are unique to xanthomonads. These pigments have been shown to be active in protecting the bacteria from photobiological damage (Jenkins and Starr, 1982).

X. albilineans grows optimally between 25 - 30°C in Wilbrink's medium. The original medium was modified to include sodium sulphite and yeast extract (Ricaud and Ryan, 1989). The bacterium also grows well in a selective medium (XAS medium), useful in isolation of the pathogen (Davis *et al.*, 1994). Growth and biochemical characteristics of the bacterium are described in a paper proposing an improved taxonomy of the genus (Van den Mooter and Swings, 1990). The average genome size in *Xanthomonas* is 3.75×10^6 base pairs (Swings *et al.*, 1993) and the G + C content for *X. albilineans* is 64.5% (Vauterin *et al.*, 1993).

X. albilineans has a narrow host range and appears to be restricted to sugarcane and other members of the family *Poaceae*. Alternative hosts include *Imperata cylindrica* var. *major* (blady grass), *Paspalum dilatatum* and *P. conjugatum* (sour grass), *Brachiaria piligera* and *Zea mays* (maize). These hosts could serve as long-term inoculum sources of the pathogen, making eradication of the disease difficult. Artificial hosts in which leaf scald symptoms occur include *Coix lacryma-Jobi* (Job's tears), *Sorghum verticilliflorum* (wild sorghum), *Bambusa vulgaris* (tall bamboo), *Pennisetum purpureum* (elephant grass) and *Zea mays* convar. *saccharata* (sweet corn). Job's tears and sweet corn are particularly susceptible to the bacterium, making them useful for pathogenicity tests (Martin and Robinson, 1961; Moffett and Croft, 1983; Ricaud and Ryan, 1989).

1.3.1 **Differentiation of *X. albilineans* Strains**

The taxonomical status of members within the genus *Xanthomonas* is uncertain and is constantly being reviewed. This stems from the fact that members were classified according to

their pathogenic rather than their inherent characteristics (Vauterin and Swings, 1997). While this is true for the numerous pathovars of *X. campestris*, the status of *X. albilineans* strains is more obvious as demonstrated previously (Van den Mooter and Swings, 1990). Taxonomic studies have therefore tended to examine differences within the species.

Early reports described differences in colony and cell morphology and suggested that this caused the variation seen in severity of the disease within and between countries (Ricaud and Ryan, 1989). This led Egan (1969; 1970) to postulate that strains of the pathogen existed since he found that the sugarcane variety Pindar was highly resistant in Australia, but moderately susceptible in Guyana and highly susceptible in Zimbabwe. The application of modern typing methods has resulted in differentiation of *X. albilineans* and clustering it into taxonomic groups. Serological and lysotypical variability was observed in 28 *X. albilineans* strains tested (Rott *et al.*, 1986) and this led to the classification of these bacteria into three serotypes and six lysovars. A degree of correlation was found between the serovars and lysovars and this could also be related to the geographical origin of the strains. In an extension of this study, two strains which belonged to different serovars and lysovars were shown to behave like pathotypes after infection of *in vitro* plantlets of an apparently resistant variety (Rott and Chagvardieff, 1987). However, symptoms on sensitive varieties were statistically similar.

A study investigating the serological variability of 218 strains from 28 regions around the world revealed that significant antigenic heterogeneity existed among strains of the pathogen (Rott *et al.*, 1994b). Again, three serovars were distinguished. Serovar I was the largest group, consisting of strains from diverse locations such as Australia, USA, Guadeloupe, India, Mauritius, Reunion, South Africa and Taiwan. Serovar II only contained strains from the African countries Burkina Faso, Cameroon, Ivory Coast, Kenya, Malawi, Zaire and Zimbabwe. Serovar III was the smallest group and contained strains from the Caribbean islands, Fiji and Sri Lanka. In separate

studies, five strains of *X. albilineans* were differentiated using analysis of LPS profiles (Pillay *et al.*, 1993) and RAPD analysis (Permaul *et al.*, 1996). Other techniques used to demonstrate heterogeneity at the species level include protein profiles and fatty acid methyl-esters (Yang *et al.*, 1993) as well as molecular techniques such as production of monoclonal antibodies and DNA fingerprinting (Alvarez *et al.*, 1996). A study of 218 strains from 31 different locations using rare-cutting restriction endonuclease digestion and pulsed-field gel electrophoresis (Davis *et al.*, 1997), grouped these strains into eight genetic groups and 54 haplotypes. It was also found that recent outbreaks in Florida, Louisiana and Texas and possibly Guadeloupe and Taiwan could be attributed to introduction of a new haplotype in these areas.

1.4 PLASMIDS IN PLANT PATHOGENIC BACTERIA

After initial genetic investigations in the genus *Xanthomonas* in the 1950s, and observations of virulence acquisition by nonpathogenic agrobacteria in 1969, fertility and resistance plasmid transfer via conjugation by *Erwinia* and *Pseudomonas* was demonstrated in the early 1970s (Panopoulos and Peet, 1985). The subsequent demonstration that plasmids are virulence determinants in *Agrobacterium* (Watson *et al.*, 1975) led to a massive research effort that culminated in a system for genetic engineering of plants. The study of plasmids in plant pathogens has focussed primarily on the following areas: the characterization and detection of pathogens (Civerolo, 1985; Clark and Lawrence, 1986); the identification of plasmid-encoded traits (Chiou and Jones, 1991; Sundin *et al.*, 1989); plasmid host range in various bacteria (Buchholz and Thomashow, 1984); the development of genetic tools (Meletzus and Eichenlaub, 1991; Taylor *et al.*, 1993); and in plant transformation systems (Klee *et al.*, 1987).

Plasmids are widespread in plant pathogenic bacteria and occur in most major genera which contain plant pathogens (Table 1.2). The occurrence of one or more plasmids in these

TABLE 1.2 Occurrence of plasmids in plant pathogenic bacteria

(Garcia-de los Santos and Brom, 1996; Permaul, 1994; Swings *et al.*, 1993; Taylor *et al.*, 1993)

Agrobacterium
- *radiobacter*
- *rhizogenes*
- *tumefaciens*

Clavibacter
- *michiganense* subsp.
  - *insidiosum*
  - *michiganense*
  - *nebraskense*
  - *sepedonicum*
- *rathayi*
- *tritici*
- *xyli* subsp. *cynodontis*

Curtobacter
- *flaccumfaciens* pv.
  - *oorti*
  - *poinsettiae*

Erwinia
- *amylovora*
- *carotovora* subsp.
  - *carotovora*
- *chrysanthemi* pv. *zeae*
- *herbicola*
- *stewartii*
- *uredovora*

Pseudomonas
- *cepacea*
- *glumae*
- *solanacearum*
- *syringae* pv.
  - *angulata*
  - *atropurpurea*
  - *coronafaciens*
  - *glycinea*
  - *lachrymans*
  - *papulans*
  - *phaseolicola*
  - *pisi*
  - *savastanoi*
  - *striafaciens*
  - *syringae*
  - *tabaci*
  - *tomato*

Rhizobium
- *etli*
- *leguminosarum* bv.
  - *trifolii*
  - *viciae*
- *meliloti*
- *phaseoli*

Rhodococcus
- *fascians*

Spiroplasma
- *citri*

Xanthomonas
- *albilineans*
- *campestris* pv.
  - *campestris*
  - *citri*
  - *cyamopsidis*
  - *dieffenbachiae*
  - *glycines*
  - *hederae*
  - *malvacearum*
  - *mangiferaeindicae*
  - *manihotis*
  - *pelargonii*
  - *phaseoli*
  - *pruni*
  - *vesicatoria*
  - *vignicola*
  - *vitians*
- *oryzae* pv. *oryzae*
- *populi*

bacteria appears to be the rule rather than the exception (Panopoulos and Peet, 1985). It is difficult to ascertain which bacteria typically lack plasmids as negative results are seldom reported. The size and number of plasmids in plant pathogens is not unusual although it is well known that they harbour large plasmids. Examples are the Ti and Ri plasmids, which range in size from 200 - 250 kb. Megaplasmids occur in rhizobia and *Ralstonia solanacearum*. In rhizobia these large plasmids range in size from 150 kb to 1 600 kb (García-de los Santos and Brom, 1996) and can constitute up to 40% of the bacterial genomic information (Brom *et al.*, 1996). In *E. stewartii*, virulent strains contain at least eight plasmids and most strains contain between 11 and 13 plasmids ranging in size from 4.1 kb to 320 kb (Coplin, 1989).

Plasmids that are conserved among strains of a pathogen are useful in strain characterization and in detection of these pathogens. Plasmid profiles have been found to be useful in differentiating strains of *X. campestris*, even at the pathovar level (Lazo and Gabriel, 1987; Lazo *et al.*, 1987). Common plasmids occurring in nearly all strains are found in *E. stewartii* and *P. syringae* pathovars (Coplin, 1989). pCS1 is a 50 kb plasmid conserved in strains of *C. michiganense* subsp. *sepedonicum*. Clark and Lawrence (1986) studied 13 isolates from different countries and found the plasmid in 11 isolates. Mogen *et al.* (1988) extended this work to include 49 strains, 23 of which contained the plasmid. Plasmid DNA probes revealed that all but one of the remaining 26 strains harboured the plasmid integrated in the host chromosome. Another conserved plasmid, pEA28, was reported to be present in all strains of *E. amylovora* isolated from various hosts and areas (Falkenstein *et al.*, 1989; Laurent *et al.*, 1989).

The functions of the majority of plasmids discovered in plant pathogens are still unknown. Some plasmids have been shown to confer survival capabilities to their hosts when living as saprophytes . The production of bacteriocins by *E. herbicola*, *E. carotovora* subsp. *carotovora*, *A. radiobacter* and *A. tumefaciens* is plasmid-encoded and gives these bacteria a competitive

advantage in ecological niches (Coplin, 1989). Similarly, plasmids encoding catabolic pathways or other nutritional capabilities offer their hosts alternative metabolic pathways (Coplin, 1989; Laurent *et al.*, 1989; Panopoulos and Peet, 1985). Other plasmid-encoded traits which help bacteria survive in adverse environments include pigmentation (Thiry, 1984) and resistance to heavy metals (Desomer *et al.*, 1988; Stall *et al.*, 1986).

The best characterized plasmids, that are involved in pathogenicity, are the Ti plasmids of *A. tumefaciens*. The involvement of these plasmids in crown gall disease of dicotyledenous plants will be described later. Similar plasmids in *A. rhizogenes*, which can reach sizes of up to 400 kb, cause hairy root disease (Coplin, 1989). Tumerous growths on olives and oleanders are also caused by plasmids in *P. syringae* pv. *savastanoi* (Comai and Kosuge, 1980). These plasmids (pIAA1, pIAA2 and pIAA3) specify the production of auxins and cytokinins that cause tumours. However, plant tissue is not transformed and tumour growth is dependant on the presence of plasmids in the pathogen. The plasmid pPT23A is a 101 kb plasmid in *P. syringae* pv. *tomato* which carries a gene cluster responsible for the synthesis of the chlorosis-inducing phytotoxin coronatine (Bender *et al.*, 1989). This is the only plasmid known to encode a phytotoxin. Plasmids present in *P. syringae* pv. *pisi* were implicated in pathogenicity as pathogenic strains contained two to four plasmids compared to non-pathogenic strains which contained none (Mazarei and Kerr, 1991). Characterization of multiple cryptic plasmids in *E. stewartii* revealed that avirulent strains tended to have fewer plasmids than virulent strains (Coplin *et al.*, 1981). Rhizobia are generally considered to be symbiotic bacteria but are actually highly evolved parasites that share much in common with pathogens in their interactions with plants (Coplin, 1989). The formation of nodules, the process of nitrogen fixation and the spectrum of the host range is conferred by large plasmids called symbiotic plasmids or pSym. Replication of these plasmids is similar to the Ti plasmids of *A. tumefaciens* (Turner *et al.*, 1997). In *R. etli*, six megaplasmids

ranging in size from 180 kb to 600 kb are present (Brom *et al.*, 1997). In addition to pSym, other plasmids were required for nodulation competiveness and growth.

Plasmids are also responsible for host range and race/cultivar specificity of some bacterial pathogens. Since bacterial phytopathogens do not produce host-specific toxins, the reason for susceptibility of plants to some pathogens and resistance to others is not well understood. It is known however, that host range in pathogens and disease resistance in plants depend on the genotype of both the pathogen and the plant (Coplin, 1989). A plant will be resistant to a given pathogen if it carries a dominant resistance gene that interacts with a dominant "avirulence" (*avr*) gene in the pathogen. This type of relationship between plant and microbe is termed a "gene-for-gene" relationship (Coplin, 1989). Avirulence genes are probably involved in the recognition of pathogens by plants and only indirectly with pathogenicity (Coplin, 1989). The *avrD* gene of *P. syringae* pv. *tomato* is present on an 83 kb indigenous plasmid pPT23B and specifies the production of syringolides, which are elicitors of plant defence reactions (Murillo *et al.*, 1994). Six *avr* genes have been reported to be clustered on a 90 kb plasmid in *X. campestris* pv. *malvacearum* (De Feyter and Gabriel, 1991). Avirulence loci are also located on plasmids in *P. syringae* pv. *glycinea* (Coplin, 1989) and on a 193 kb plasmid in *X. campestris* pv. *vesicatoria* (Stall *et al.*, 1986). Other genes, however, appear to be required for pathogenicity on susceptible hosts as well as induction of a hypersensitive response on resistant cultivars or non-hosts. These genes are necessary for expression of an (a)virulence phenotype and have been termed *hrp* (*h*ost *r*ange and *p*athogenicity) genes. *Hrp* genes are normally chromosomal (Bauer and Beer, 1991; Daniels *et al.*, 1988) but are present on a megaplasmid in *R. solanacearum* (Boucher *et al.*, 1986).

1.5 BACTERIAL CONJUGATION

Bacterial conjugation is a mechanism for transfer of a plasmid from a plasmid-containing

donor cell to a plasmid-less recipient cell. Conjugation in *E. coli* was first discovered by Lederberg and Tatum (1946) and initially, it was believed that it was a form of sexual reproduction in bacteria. Although chromosomal genes can be transferred during conjugation, this occurs by accident and at very low frequencies. The conjugative process is therefore a means of spreading a plasmid in a bacterial population and the essential genes required for this process are plasmid-encoded. Conjugation can be divided into four phases (Freifelder, 1987). They are: a) formation of donor-recipient pairs or mating pair formation; b) preparation for DNA transfer or mobilization; c) DNA transfer; and d) formation of a replicative functional plasmid in the recipient. Not all plasmids can carry out all of these processes. Mobilizable plasmids can only prepare their DNA for transfer and depend on conjugative or self transmissible plasmids for transfer into a recipient cell.

Mating pair formation requires the synthesis of a plasmid-encoded pilus. A cluster of genes are involved in pilus biosynthesis. It is generally accepted that the tip of the pilus identifies and binds to receptors in the LPS of the recipient cell (Fig. 1.2) and retracts into the donor cell, bringing the mating pair into cell-to-cell contact (Fig. 1.3). This triggers DNA transfer (Porter, 1991). However, the exact nature of the recognition between the pilus tip and the cell receptors and the nature of the pore that allows DNA transfer through the cell envelope is still unknown. DNA mobilization is initiated by a signal arising from contact of the pilus tip with a recipient cell (Everett and Willets, 1980). This involves nicking of the plasmid at *oriT* (origin of transfer) and synthesis of a single strand by rolling-circle replication of DNA. The single stranded DNA is then transferred to the recipient via the mating pore/bridge. While the proteins that constitute the mating pore in the donor have been identified, it is not known if proteins are involved in entry of the single-stranded DNA in the recipient cell.

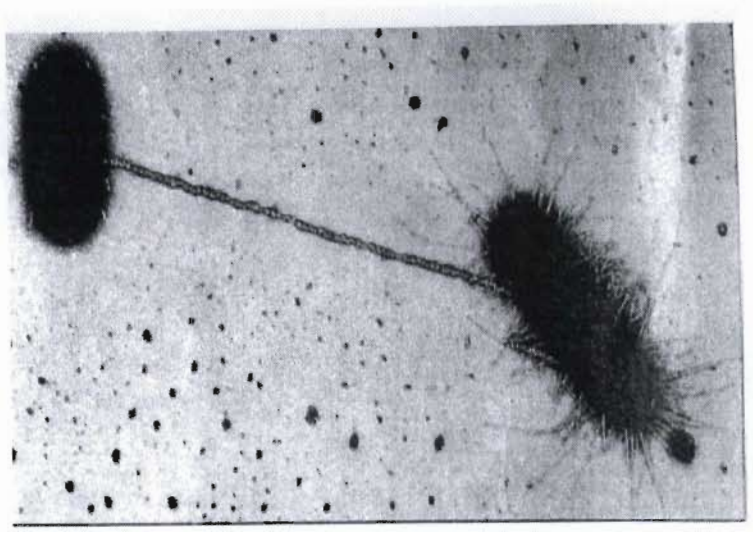


Fig. 1.2 Attachment of *E. coli* cells via conjugative pilus. The donor cell is on the right (Prescott *et al.*, 1996).

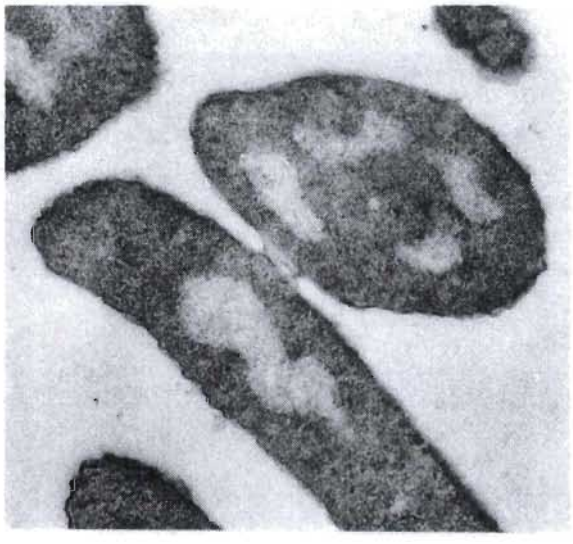


Fig. 1.3 Cell-to-cell contact during conjugation of two *E. coli* cells. The larger cell is the donor (Freifelder, 1987).

In a phenomenon called retrotransfer which can occur at the same frequency as conjugation, the recipient cell becomes a donor cell and is able to transfer its own plasmid into the attached cell (Mergeay *et al.*, 1987). But the proteins forming the mating pore structure inside this donor cell has been shown to be encoded by the conjugative plasmid that was transferred

initially (Sia *et al.*, 1996). Completion of the cycle of bacterial conjugation requires circularization of the transferred plasmid and its replication. Synthesis of complementary DNA is thought to occur concurrently with entry of the plasmid. This implies that a discontinuous mode of replication requiring multiple primers is needed. In some conjugation systems a plasmid-encoded primase is supplied by the donor cell (Wilkins, 1996). Establishment of the immigrant plasmid can also be facilitated by genes present in the leading region of the transferred strand. These include genes encoding suppression of the SOS stress response in recipient cells and genes conferring antirestriction properties (Wilkins, 1996).

1.5.1 **The F Plasmid**

The "fertility factor" or F plasmid was the first plasmid to be discovered by Jacob and Wollman in 1956 (Perlin, 1991). This plasmid is 100 kb in size and has four distinct regions (Fig. 1.4). The *inc,rep* region is responsible for vegetative replication and incompatibility functions. Four transposable elements are located in the second region. A region having only a few distinct genetic functions is called the silent region. The fourth region of the F plasmid is the 33.3 kb *tra* region which is responsible for the conjugative ability of the plasmid. The *tra* region encodes 37 genes, of which 26 have been shown to be associated with conjugation (Skurray and Firth, 1997). The majority of these genes are involved in pilus biosynthesis. The rest of the genes are involved in gene regulation, DNA metabolism, stabilization of mating aggregates and surface exclusion.

The *traA* gene encodes the pilus subunit, pilin. The initial gene product is a 121 amino acid (aa) protein, propilin, which is cleaved to form mature pilin (70 aa). The filament of the pilus was originally thought to consist only of pilin subunits, but a recent study has shown that the F-pilus tip is a specialized adhesive structure comprising other minor protein components (Anthony

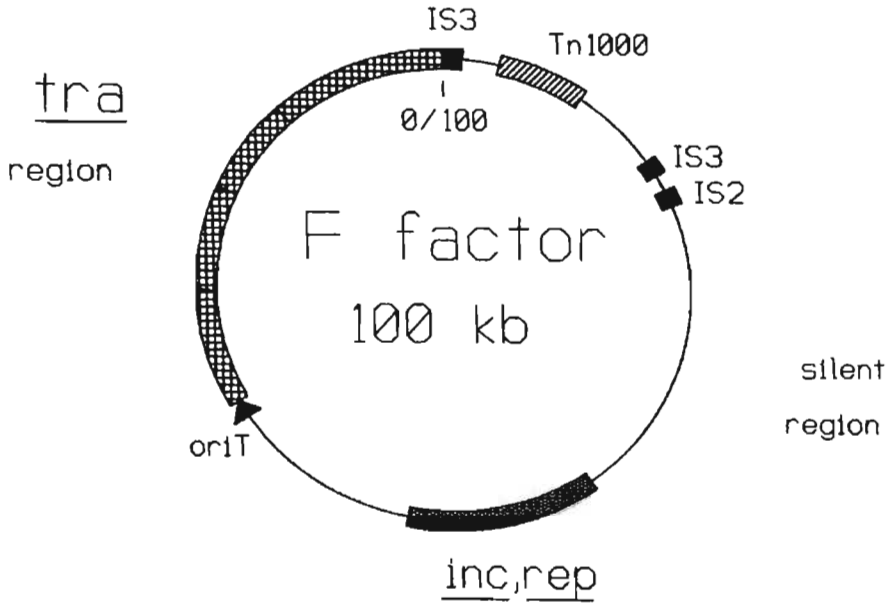


Fig. 1.4 Map of the F plasmid. Four distinct regions have been identified on the plasmid (Porter, 1991).

and Frost, 1996). The filament possesses a central hydrophilic lumen, measuring 2 nm in diameter (Silverman, 1997). The *traQ* and *traX* genes were found to be involved in processing of the propilin into mature pilin. Mutations in *traB, traC, traE, traF, traG, traH, traK, traL, traU, traV, traW* and *trbC* affected piliation and drastically decreased mating efficiency (Frost *et al.*, 1994). These genes are assumed to be involved in pilus assembly. TraD is involved in actively transporting the transferred strand into the recipient cell. The TraG protein also plays a role in mating pair stabilization, together with the TraN protein. *traS* and *traT* encode proteins that are involved in surface exclusion. The roles of the remainder of the gene products identified to be involved in conjugation were summarized recently (Frost *et al.*, 1997). TraJ is the major positive regulator of the *tra* operon and is in turn negatively regulated by the fertility inhibition system

consisting of *finP* and *finO*. The first gene in the *tra* operon, *traY*, appears to be important for *tra* operon expression as well as *traM* transcription. TraY also has an important function in directing the TraI protein to *oriT*. The TraI protein is a relaxase that nicks at the *oriT* site to initiate DNA transfer. TraM is a regulatory protein and its levels are modulated by a variety of mechanisms, in addition to being autoregulated. High levels of TraM are necessary for efficient F plasmid transfer (Frost *et al.*, 1997). The entire process of transfer of the 100 kb plasmid takes approximately five minutes (Frost *et al.*, 1994).

1.5.2 **Plasmid RK2**

Bacterial plasmids that share replication control or partitioning functions compete for stable inheritance in a cell or its progeny. One of these plasmids is invariably lost from the population and this is termed incompatibility. Plasmids are therefore classified according to incompatibility groups. The F plasmid belongs to the incompatibility group F or is an IncF plasmid. Plasmid RK2, a 60 kb plasmid, belongs to the IncP group of plasmids and its members are characterized by stable maintenance in almost all Gram-negative bacterial species and being self-transmissible at high frequencies (Thomas and Smith, 1987). RK2 is also known by the synonyms RP4, RP1, R68 and R18 as these antibiotic resistance plasmids are indistinguishable (Thomas and Smith, 1987).

The conjugative systems of plasmid RK2 consist of two distinct regions, Tra1 and Tra2 (Fig. 1.5). The Tra1 region contains *oriT* and genes involved mostly with DNA transfer and replication. The Tra2 region originally included a third region, Tra3, but the split in this region is due to insertion of transposable elements. The Tra2 region and TraF from the Tra1 region are involved in the synthesis of a mating bridge during conjugation (Lessl *et al.*, 1992). Twelve genes are present in this region and six of them were found to have similar sequences to the VirB proteins of *A. tumefaciens* (Lessl *et al.*, 1992).

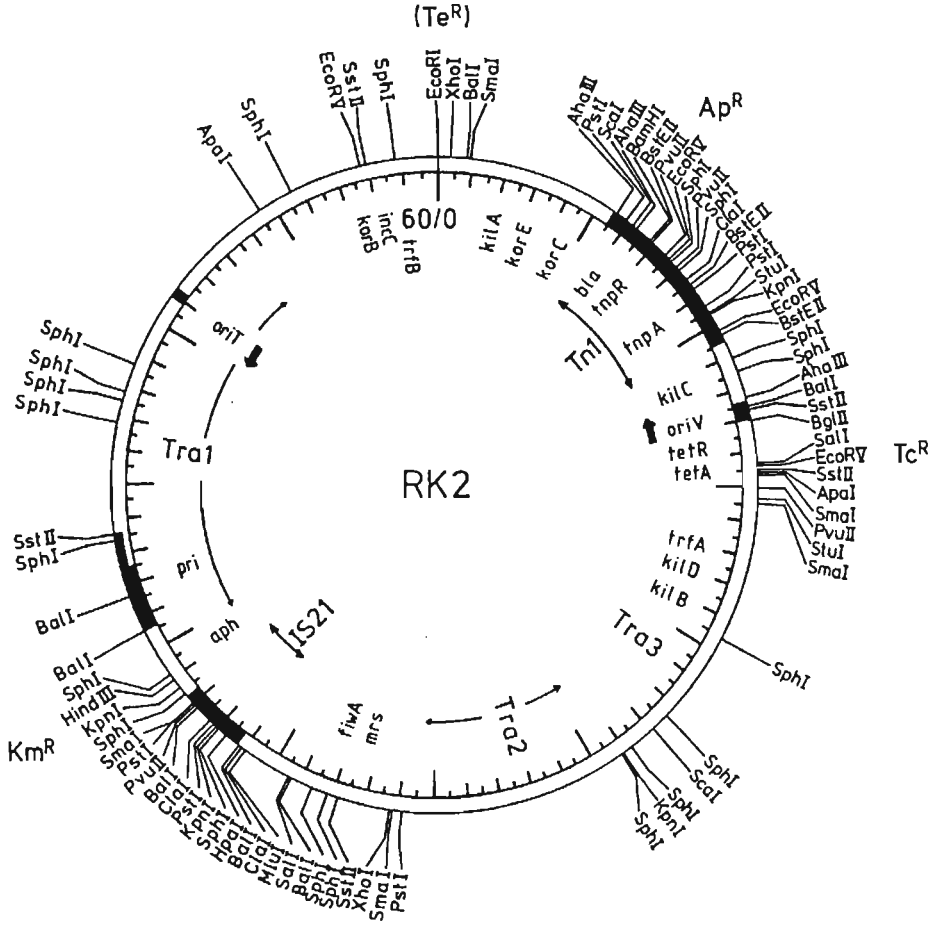


Fig. 1.5 Physical and genetic map of plasmid RK2. Genetic loci: *aph,* aminoglycoside phosphotransferase; *bla*, ß-lactamase; *fiw*, fertility inhibition toward IncW plasmids; *inc*, incompatibility toward IncP plasmids; *kil*, host-lethal function; *kor*, suppression of *kil* effects; *mrs*, multimer resolution system; *oriT*, origin of conjugal DNA transfer; *oriV*, origin of vegetative replication; *pri*, primase; *tet*, tetracycline resistance (Thomas and Smith, 1987).

IncP plasmids are known for their conjugal and replicative promiscuity among diverse bacteria. This is not a common property of plasmids and it implies the existence of novel genetic determinants on these plasmids (Young *et al.*, 1984). RK2 codes for at least 74 genes and displays a remarkable regulatory complexity unmatched by other groups of plasmids (Pansegrau *et al.*, 1994). A unique plasmid regulon, the *kil-kor* regulon was discovered which consists of eight operons (Larsen and Figurski, 1994). *kil* genes are lethal when their expression is unregulated by

kor (kil-override) genes. The *trfA* operon codes for a single-strand DNA-binding protein and the replication protein TrfA, which is essential for the initiation of replication at *oriV*. The *trbB* (Tra2) operon codes for 11 genes that are required for mating pair formation in conjugation. Fives genes are present in the *korAB* operon, including those for the transcriptional repressors KorA and KorB. The other five operons are the *kfrA* operon, the *kilA* operon, the *kilC* operon and two operons in the *kilE* locus (Wilson *et al.*, 1997).

1.5.3 **IncN Plasmids**

Plasmids from IncN, IncP and IncW groups possess morphologically similar conjugal pili and members of these groups, especially IncN plasmids, kill *Klebsiella pneumoniae* cells used as conjugal recipients (Thatte *et al.*, 1985; Winans and Walker, 1985c). The best-studied IncN conjugative system is that of the 35.4 kb plasmid, pKM101 (Fig. 1.6). The *tra* cluster specifying production of the pKM101 pilus consists of 11 genes that are transcribed in the same direction and expressed from two promoters (Winans *et al.*, 1996). Four additional *tra* genes are required for conjugal DNA metabolism, but not for pilus biosynthesis. Therefore, efficient conjugation of pKM101 requires only 15 genes - one of the simplest conjugation systems yet characterized (Winans *et al.*, 1996). The subcellular locations of most of the pilus genes are known and a model has been proposed for the structure of the pKM101 pilus (Fig. 1.7). Other determinants on pKM101 that are common in other conjugative plasmids are an entry exclusion protein (Pohlman *et al.*, 1994b; Winans and Walker, 1985b), at least two sets of *kil/kor* genes (Winans and Walker, 1985c) and a locus for fertility inhibition of co-resident IncP plasmids (Winans and Walker, 1985d).

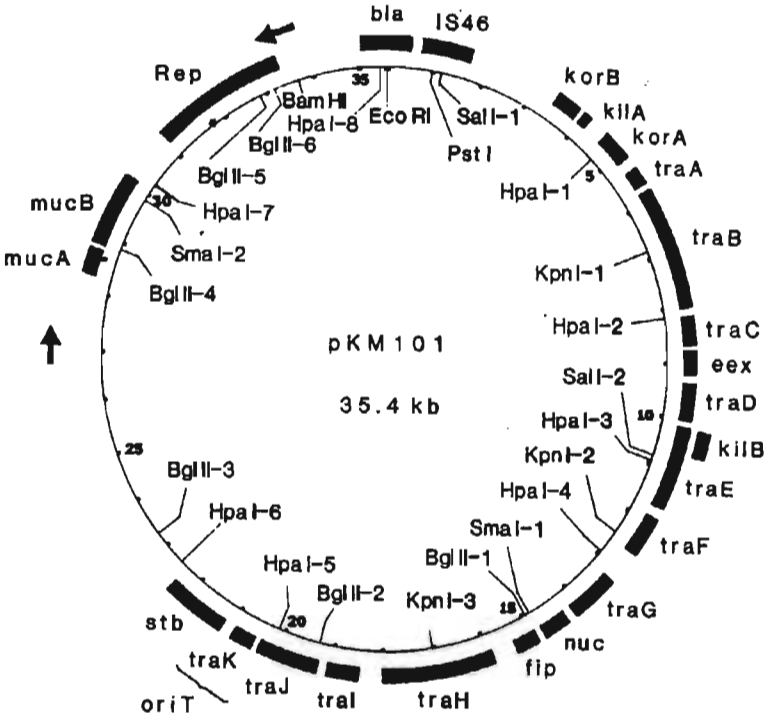


Fig. 1.6 Physical and genetic map of pKM101. Genes *traA* to *traG* are required for pilus synthesis and *traH* to *traK* for conjugal DNA metabolism. *mucA* and *mucB* are mutagenesis sensitive genes; *nuc* codes for a periplasmic endonuclease; *eex* codes for entry exclusion properties; *fip* is responsible for fertility inhibition of IncP plasmids; *IS46* is an insertion sequence; *stb* is required for stability in recombination-proficient hosts; *kilA* and *kilB* are potentially host-lethal; and *korA* and *korB* prevent lethality by *kilA* and *kilB* (Winans and Walker, 1985a).

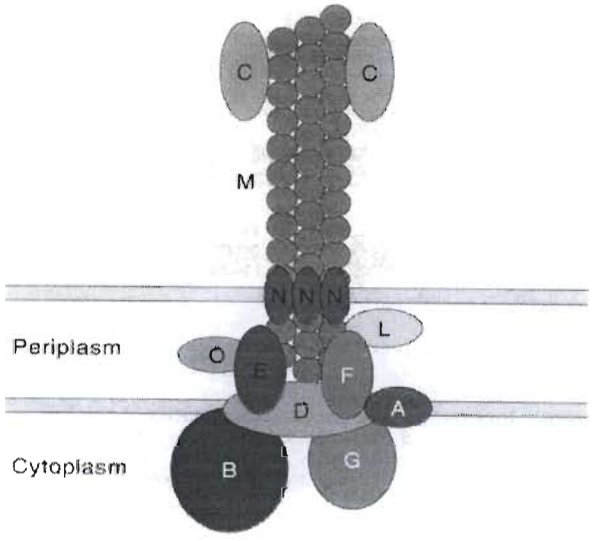


Fig. 1.7 A model describing the possible locations of and interactions between 11 IncN plasmid Tra proteins (Winans *et al.*, 1996).

1.5.4. Conjugative Transfer by the Virulence System of *A. tumefaciens*

The genes required for interkingdom DNA transfer by *A. tumefaciens* are located on a sector of the approximately 200 kb Ti plasmid (Fig. 1.8) close to the transferred DNA (T-DNA). This virulence region is 28 to 35 kb in size and is required for processing the 25 kb T-DNA for delivery into the plant host during infection. The DNA is then directed to the plant cell nucleus and incorporated into the genome. The genes carried on the T-DNA are then expressed and result in the production of growth hormones which cause enlargement and proliferation of the transformed cells, culminating in the formation of tumours. These transformed cells also produce opines which are specifically utilized by *A. tumefaciens* as carbon and nitrogen sources and for

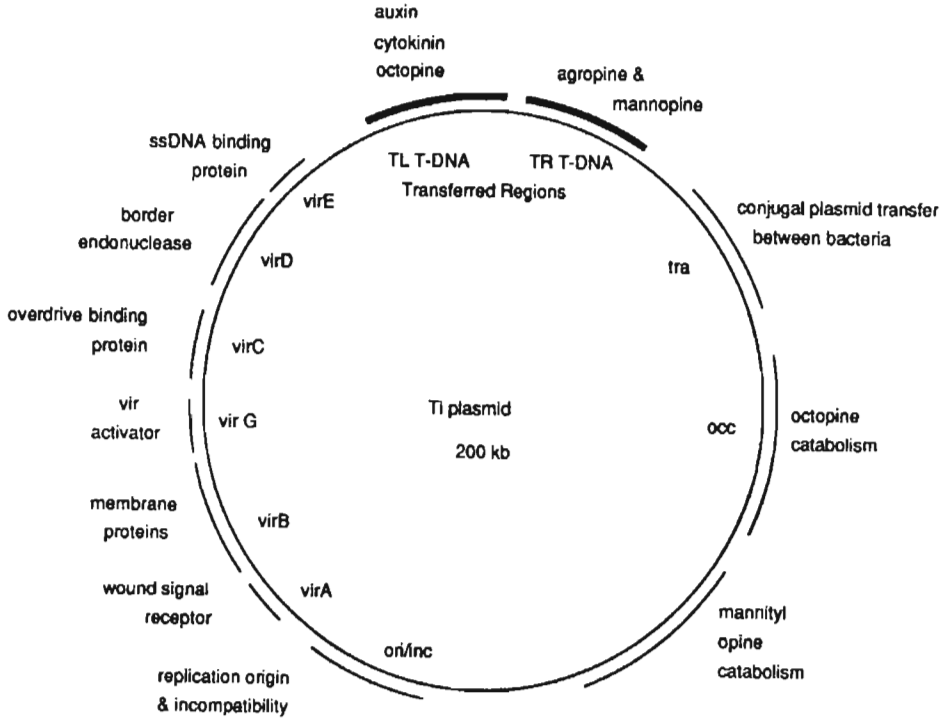


Fig. 1.8 Map of the Ti plasmid of *A. tumefaciens* showing the positions of various coding regions (Ream, 1991).

conjugational activities (Kado, 1994). *vir* genes are contained in six major operons: *virA, virB, virC, virD, virE* and *virF*. Mutations in *virA, virB,* and *virD* eliminates tumour formation in all plant species, whereas mutations in the other operons lead to a restriction in host range of the pathogen or an attenuation in tumorigenicity (Hooykaas and Beijersbergen, 1994).

The virulence system of *A. tumefaciens* has been described as a promiscuous conjugative system (Kado, 1996). T-DNA transfer resembles bacterial conjugation in several respects. Firstly, nicking at the T-DNA borders to release the strand is equivalent to nicking at the conjugal *oriT* region. This was demonstrated when *vir* helper plasmids (lacking T-DNA) were able to transfer IncQ plasmids (containing *oriT*) into tobacco cells (Buchanan-Wollaston *et al.*, 1987). In addition, *vir* genes are also able to transfer plasmids between bacteria (Beijersbergen *et al.*, 1992). Secondly, the T-DNA is similar in structure to the linear, single stranded donor molecule transferred during conjugation. Thirdly, transfer occurs via a similar membrane structure (pilus/pore) during both processes. The 11 genes comprising the *virB* operon and three of the four *virD* genes share remarkable similarities with the transfer genes of conjugative plasmids (Kado, 1994).

This similarity also extends to the protein export system of *B. pertussis* (Weiss *et al.*, 1993). The genes in these systems code for approximately the same number of proteins of similar sizes, which are related (Table 1.3) and the corresponding genes are arranged in a co-linear fashion (Fig. 1.9). It is assumed that these DNA and protein transfer systems share a common ancestry (Christie, 1997; Kado, 1994; Pohlman *et al.*, 1994a; Winans *et al.*, 1996).

TABLE 1.3 Similarity of VirB and VirD proteins to conjugal transfer (*tra)* and pertussis toxin secretion (*ptl)* gene products (Kado, 1994)

Vir	R388	F	RP4	Ptl
VirB1				
VirB2	TrwM (53%)	TraA (48%)	TrbC (50%)	ptlA [31%]
VirB3	TrwL (49%)	TraL (42%)	TrbD (40%)	ptlB [26%]
VirB4	TrwK (54%)	TraC (46%)	TrbE (48%)	ptlC [25%]
VirB5	TrwJ (42%)	TraE (34%)	TrbF (39%)	
VirB6	TrwI (50%)			ptlD [18%]
VirB7	TrwH (49%)			
VirB8	TrwG (45%)			ptlE [32%]
VirB9	TrwF (53%)			ptlF [27%]
VirB10	TrwE (55%)	TraB (42%)	TrbI (53%)	ptlG [32%]
VirB11	TrwD (58%)		TrbB (52%)	ptlH [35%]
VirD1		TraJ (?)		
VirD2	TrwC (59%)	TraI (44%)	TraI (?)	
VirD3				
VirD4	TrwB (44%)	TraD (43%)	TraG [28%]	

Numbers in parentheses indicate percent similarity while numbers in square brackets indicate percent identity. ? = similar size only.

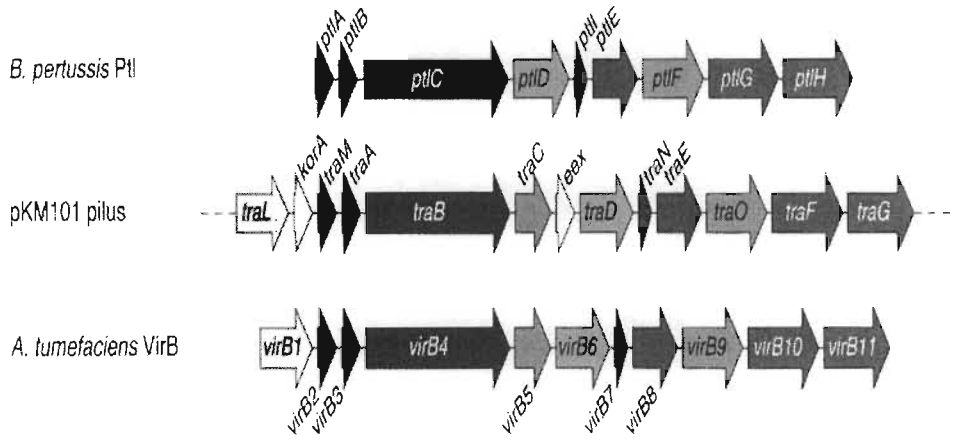


Fig. 1.9 Alignments of the pKM101 pilus cluster genes with the *virB* operon of *A. tumefaciens* and the *ptl* region of *B. pertussis* (Winans *et al.*, 1996)

1.6 SCOPE OF THE PRESENT STUDY

pXA1, the plasmid present in XA86, a South African strain of *X. albilineans*, was the first plasmid to be described in this pathogen (Permaul, 1994). Two previous studies reported absence of plasmids in one (Lazo and Gabriel, 1987) and two (Birch and Patil, 1987a) *X. albilineans* strains. Currently the number of *X. albilineans* strains identified as being hosts to plasmids is 27 out of 68 strains tested (Permaul *et al.*, 1999). In the current study, various strategies were used in an attempt to identify pXA1-encoded phenotypes. As with most bacterial pathogens, it was important to determine whether the plasmid was required for pathogenicity. This was attempted by inoculating a plasmid-cured derivative of XA86, into sugarcane cultivars. Physical and genetic characterization of pXA1 was performed by hybridization with avirulence genes, studying expression of plasmid genes and sequencing of plasmid DNA. It was hoped that these would provide information for explaining the occurrence of pXA1 in XA86 and plasmids in other *X. albilineans* strains.

CHAPTER TWO

THE ROLE OF pXA1 IN PATHOGENICITY

2.1 INTRODUCTION

For most bacterial plant pathogens, the genes encoding pathogenicity determinants such as toxins, degradative enzymes and extracellular polysaccharides are typically located on the bacterial chromosome (Denny and Baek, 1991; Dow *et al.*, 1987; Salch and Shaw, 1988; Turner *et al.*, 1985). Nevertheless, a number of important virulence plasmids have been discovered and it appears that plasmids contribute significantly to the ecological fitness of their hosts (Coplin, 1989).

The discovery of virulence plasmids in *A. tumefaciens* (Van Larebeke *et al.*, 1974; Watson *et al.*, 1975) led to the search for similar plasmids in other pathogens. The indispensable role of the Ti plasmids in pathogenicity has been described in Chapter One. Similar root-inducing (Ri) plasmids are responsible for hairy root diseases caused by *A. rhizogenes* (Coplin, 1989). Analogous megaplasmids have been found in nitrogen-fixing *Rhizobium* and *Bradyrhizobium* species (Banfalvi *et al.*, 1981; Masterson *et al.*, 1982) but these plasmids have evolved to play a role in symbiotic nitrogen fixation rather than pathogenesis. Another plasmid that contributes significantly to virulence is the 101 kb plasmid pPT23A, which is found in *P. syringae* pv. *tomato* (Bender and Cooksey, 1987; Bender *et al.*, 1989). This plasmid carries a gene cluster required for the synthesis of the chlorosis-inducing toxin coronatine. Synthesis of coronatine in *P. syringae* pv. *atropurpurea* and *P. syringae* pv. *glycinea* is also specified by plasmids measuring 88 kb and 92 kb, respectively (Coplin, 1989). The virulence determinant *iaa*M, which causes tumour-like outgrowths, is produced by a 50 kb plasmid, pIAA1, in *P. savastanoi* (Comai and Kosuge, 1981).

Apart from obvious pathogenicity determinants, plasmids can also control host range and race/cultivar specificity in bacterial pathogens. Examples of plasmid-borne *avr* genes occur in *P. syringae* pv. *tomato* and *P. syringae* pv. *glycinea* (Murillo *et al.*, 1994), on a 90 kb plasmid in *X. campestris* pv. *malvacearum* (De Feyter and Gabriel, 1991) and on a 193 kb plasmid in *X. campestris* pv. *vesicatoria* (Stall *et al.*, 1986). *hrp* genes which are normally chromosomal are found on a megaplasmid in *R. solanacearum* (Boucher *et al.*, 1986).

Traditional methods for studying plasmid function involve elimination of plasmids from host cells (curing) or transfer to a suitable host and testing for either loss or acquisition of traits. Non-specific mutations caused by curing agents, transfer of unrelated plasmids in cells carrying multiple plasmids and the lack of selectable markers have impeded these approaches. Newer techniques for studying plasmid genes include cloning of DNA fragments and transposon mutagenesis. Various methods are available for curing of plasmids. Most commonly used are curing agents such as the DNA intercalating dyes acridine orange and ethidium bromide (Comai and Kosuge, 1980; Morales and Sequeira, 1985; Steinberger *et al.*, 1990), the antibiotics novobiocin and coumermycin (Perlin, 1991), mitomycin C (Dittapongpitch and Ritchie, 1993; Stall *et al.*, 1986), nalidixic acid (Thiry, 1984), UV light (Gantotti *et al.*, 1979), growth at elevated temperatures (Thiry, 1984; Ulaganathan and Mahadevan, 1988; Watson *et al.*, 1975) or growth in the presence of the anionic detergent SDS (Thiry, 1984), or simply growth in the absence of selective pressure (Murillo and Keen, 1994). Eviction of plasmids by exploiting plasmid incompatibility has also been used to obtain cured cells (Fu *et al.*, 1998; Laurent *et al.*, 1989; Murillo and Keen, 1994).

This chapter describes the curing of pXA1 using SDS and screening for cured cells by DNA hybridization. Cured cells were tested for their ability to produce albicidin in an agar plate assay and inoculated into sugarcane plants to test their pathogenicity.

2.2 MATERIALS AND METHODS

2.2.1 **Growth and Storage of Bacterial Isolates**

X. albilineans isolates were grown at 28 °C in modified Wilbrink's medium (20 g sucrose, 5 g peptone, 5 g yeast extract, 0.5 g KH_2PO_4, 0.25 g $MgSO_4.7H_2O$ and 0.05 g Na_2SO_3, per litre). Working stocks of these isolates were sub-cultured on agar plates at one month intervals. *E. coli* isolates were grown at 37°C in YT medium (8 g tryptone, 5 g yeast extract and 5 g NaCl, per litre). When required, antibiotics or other supplements were added at the appropriate final concentrations (Table 2.2). For blue-white colony selection during cloning procedures, the chromogenic substrate X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside) was added to YT medium, together with the inducer IPTG (isopropyl-ß-D-thiogalactopyranoside). Alternatively, 50 μl of 2% X-gal and 20 μl of 1M IPTG was spread on YT agar plates, prior to inoculation. Long-term storage of *X. albilineans* isolates involved addition of glycerol (15% final concentration) to log-phase cultures, snap-freezing and storage at -70 °C. *X. albilineans* strains were also successfully preserved using Microbank vials (Prolab Diagnostics) according to the manufacturer's instructions. Long-term storage of *E. coli* cultures involved inoculation of agar deeps in Bijou bottles and air-tight storage at room temperature. *P. syringae* pv. *syringae* was cultured in Nutrient broth/agar . When co-cultures were required, *X. albilineans* and *E.coli* were cultured in Boost medium (23 g Nutrient agar, 20 g sucrose, 10 g yeast extract, per litre).

2.2.2 **Plasmid Curing**

Two methods which involved growth of XA86 in adverse conditions were chosen for curing of XA86. Firstly, the isolate was grown at elevated temperatures, i.e., at 35 °C, 37 °C and 40 °C. Secondly, XA86 was grown in varying concentrations of the anionic detergent SDS. The

TABLE 2.1 Bacterial strains used in this study

Strain	Characteristics/Origin	Source/Reference[†]
X. albilineans		
XA86	Wild type, South Africa	UDW
XA86C	cured derivative of XA86	This study
XA86CS	Sm^R derivative of XA86C	This study
XA86CR	Rif^R dervative of XA86C	This study
XA86CS1	XA86CS harbouring pBX1	This study
GLP7C	cured derivative of GLP7	Ramdeen (1999)
GLP7CS	Sm^R derivative of GLP7C	This study
PDDCC196	Type strain, Fiji	ICMP
TSS	Sm^R derivative of PDDCC196	This study
TSS1	TSS1 harbouring pBX1	This study
P. syringae pv. *syringae*		
3420	Wild type	UP
3420S	Sm^R derivative of 3420	This study
E. coli		
DH5αF′	Nal^R	Sambrook *et al.*(1989)
CSH56	Nal^R	D.E. Rawlings
JC8679	*recA*$^+$, Sm^R	D.E. Rawlings
S17-1	contains RK2 derivative, Sm^R	D.E. Rawlings
S17-11	S17-1 harbouring pBX1	This study
HB101	Sm^R	Sambrook *et al.*(1989)
HBpSa	HB101 harbouring pSa	D.E. Rawlings
Organism X	inhibitor-producing strain	This study

[†] BRL: Bethesda Research Laboratories, USA; ICMP: International Collection of Microorganisms from Plants, Auckland, New Zealand; D.E. Rawlings: Department of Microbiology, University of Stellenbosch, South Africa; UDW: Department of Microbiology, University of Durban-Westville, South Africa; UP: Department of Microbiology and Plant Pathology, University of Pretoria, South Africa.

TABLE 2.2 Antibiotics and other media supplements

Supplement	Stock concentration	Final concentration
Ampicillin (Ap)	100 mg/ml	100-200 μg/ml
IPTG	1 M	1 mM
Nalidixic acid (Nal)	25 mg/ml (pH 11 with NaOH)	25 μg/ml
Rifampicin (Rif)	50 mg/ml (in methanol)	50 μg/ml
Streptomycin (Sm)	50 mg/ml	50-75 μg/ml
X-gal	20 mg/ml (in DMF)	50 μg/ml

latter method was finally chosen as cultures grown at elevated temperatures did not show a substantial increase in turbidity.

SDS curing was accomplished by firstly testing growth of XA86 in Wilbrink's broth containing 0.001, 0.005, 0.01, 0.05 and 0.1% SDS. One hundred μl of a fresh XA86 culture was inoculated into 10 ml of Wilbrink's broth containing SDS and the cells from the flasks growing nearest to the minimum inhibitory concentration (MIC) for SDS were selected. These cells were serially diluted in Wilbrink's broth and plated on Wilbrink's agar plates to obtain single colonies. Individual colonies were subsequently tested for loss of pXA1 by colony hybridization.

2.2.3 **Colony Hybridization**

Since there was no selectable marker available on pXA1, it was not possible to screen for the loss of the plasmid by biochemical tests. Therefore, colony hybridization was used to screen colonies that lacked the plasmid. A plasmid DNA probe that hybridized to total DNA of cells containing the plasmid was used. The DNA probe was an 800 bp *Bam*HI/*Eco*RI fragment of pXA1. This fragment was isolated from agarose gels using the Qiaex II kit (Qiagen, Inc.) and

labelled by random incorporation of DIG-dUTP using the nonradioactive DIG DNA labelling and detection kit (Boehringer Mannheim). Details of genetic manipulation techniques are described in Chapter Three.

A total of 235 colonies were selected for screening and were initially grown on Wilbrink's agar master plates. These colonies together with negative controls comprising five PDDCC196 and *E. coli* DH5α colonies were individually transferred onto two rectangular pieces of nylon hybridization membrane (Hybond-N, Amersham) using the cylindrical end of a glass rod. The cells on the membrane were lysed and the liberated DNA was bound to the membrane using a protocol adapted from Maniatis *et al.* (1982). Three pieces of Whatman 3MM paper were cut to fit neatly into the bottom of plastic boxes. The first piece was saturated with 10% SDS and the excess solution poured off. The nylon membrane was placed, right side up, on the SDS-impregnated 3MM paper for 3 min to lyse the cells, transferred to the second sheet of 3MM paper which had been saturated with denaturing solution (0.5 M NaOH; 1.5 M NaCl) and the DNA denatured for 3 min before the membrane was transferred to the third sheet of 3MM paper which had been saturated with neutralising solution (0.5 M Tris.HCl, pH 8; 1.5 M NaCl). After 5 min, the membrane was removed and transferred to 3MM paper saturated with 2×SSC (300 mM NaCl; 30 mM sodium citrate, pH 7.0) for 15 min. DNA was fixed on the wet membrane by coss-linking with UV light for 3 min and membranes were used immediately for DNA hybridization.

The manufacturer's protocol for DNA hybridization (Boehringer Mannheim) was modified slightly in order to perform colony hybridizations. It was necessary to first remove cellular debris which can result in background signals. The membrane was incubated in 3×SSC/0.1% SDS with shaking for 1-3 h at 68°C. The surface of the membrane was then gently wiped with moistened paper towels. Nylon membranes containing liberated DNA were prehybridized at 68°C in plastic bags containing a solution composed of 5×SSC, 1% blocking reagent (Boehringer Mannheim),

0.1% *N*-lauroyl sarcosine and 0.02% SDS. The solution was replaced with fresh solution containing denatured labelled probe and the membranes were hybridized for at least 15 h at 68 °C. Following hybridization, membranes were washed twice with 2×SSC/0.1% SDS for 5 min at room temperature, followed by two washes with 0.1×SSC/0.1% SDS at 68 °C for 15 min (low and high stringency, respectively). Hybrid DNA was detected by a chemiluminescent reaction using an enzyme-linked immunoassay and recorded by exposure to X-ray film (Hyperfilm-MP, Amersham.) which was developed according to the manufacturer's instructions.

Colonies that appeared to have been cured of pXA1 were identified and re-tested together with negative and positive controls. The negative controls were *E. coli* DH5α and PDDCC196 cells. XA86 cells were used as positive controls. The absence of pXA1 in the strains were further verified by performing miniprep plasmid isolations from cured and wild type strains and analyzing intact DNA and DNA restricted with *Eco*RI on agarose gels. Details of plasmid isolation from XA86 and restriction endonuclease digestions are described in Chapter Three.

2.2.4 **Albicidin Assays**

Cured strains were tested for their ability to produce albicidins using an agar plate method adapted from Birch and Patil (1983). In this method, *E. coli* HB101 was used as an indicator organism as albicidins also inhibit bacterial DNA replication. Isolates to be tested for albicidin production were spotted onto Wilbrink's agar plates and grown for five days to ensure that cells were no longer in the exponential phase of growth. Plates were then overlaid with 2 ml of Mueller-Hinton top agar containing 0.2 ml of an overnight culture of *E. coli*. Plates were incubated at 28 °C and examined for zones of clearing which indicated albicidin production.

2.2.5 **Pathogenicity Trials**

2.2.5.1 Growth of sugarcane plants

Approximately 260 five-week-old sugarcane plantlets were obtained from the South African Sugar Association Experimental Station, Mount Edgecombe, Durban. These plantlets represented five cultivars possessing moderate to high resistance to leaf scald disease, viz., N11, N14, N17, N22 and NCo376. Upon receipt of the plantlets in seedling trays, they were supplemented with nitrogen in the form of a 0.2% solution of NH_4SO_4 and watered daily until they were eight weeks old. The plants were then transferred to pots and grown in potting soil consisting of 50% Gromor potting medium (National Plant Food C.C.) and 50% compost pine bark. Plants were housed in a nursery covered with 20% shadecloth and watered twice a day with an automatic overhead sprinkler system. The potting soil was supplemented with nutrients once a week by addition of a solution containing 1 g/l of Gromor 3.1.3 fertilizer, 50 mg/l of Coastal Blend and calcium nitrate (Evergreen Horticulturalists, Durban).

2.2.5.2 Plant inoculations

Sugarcane plants were inoculated with suspensions of XA86 and a plasmid-cured derivative of XA86 (XA86C). These strains were grown in Wilbrink's broth at 28 °C until the cells were in the exponential phase of growth. The actively growing cells were then harvested in a Beckman JA-14 rotor at 7 000 rpm. The cell pellets were resuspended in Wilbrink's broth and adjusted to an absorbance value of 0.160 at 600 nm with an Ultrospec II spectrophotometer (LKB Biochrom Ltd.). This value corresponded to a cell concentration of 1×10^9 cells/ml (Permaul, 1994). Aliquots of the cells were washed once with sterile distilled water and resuspended in water to yield an identical cell concentration. The cells were spread on Wilbrink's agar plates to ensure that they were viable, prior to inoculation into plants.

Plants were inoculated using the hypodermic syringe method (Martin and Robinson, 1961). Inoculations were performed on healthy 12-week-old sugarcane plants. Each plant was inoculated twice, at the base of the stalks and near the growing point, with 0.5 ml of the inoculum described above. Plants serving as controls were inoculated with sterile distilled water. A total of 252 plants were inoculated. The number of plants inoculated and the number of controls used are listed in Table 2.3. After inoculation, plants were monitored daily for the appearance of leaf scald symptoms.

TABLE 2.3 Inoculation regimen for the pathogenicity trials

	Inoculum			
Cultivar	XA86	XA86C	Controls	Total Plants
N11	15	15	5	35
N14	25	25	10	60
N17	25	25	10	60
N22	22	22	10	54
NCo376	17	17	9	43

2.2.6 **Inhibition of *X. albilineans***

During routine sub-culturing of XA86, an organism that inhibited *X. albilineans* was isolated. This was observed as a zone of clearing around the organism. This isolate was given the tentative name Organism X and further inhibition assays on agar plates were carried out. The procedure used in the assays was essentially the same as that used in the albicidin assays. The inhibitory effect was also tested using *E. coli* DH5α and *P. syringae* pv. *syringae*.

2.3 RESULTS

2.3.1 **Plasmid Curing**

Initial curing experiments involved attempts to evict pXA1 from XA86 by incompatibility. Firstly, *Eco*RI fragments of pXA1, ligated to the pUC-based vector pTZ19R were electroporated into XA86. It was hoped that one of these constructs which contained the pXA1 origin of replication would displace pXA1 in XA86 due to incompatibility. However, transformants were not obtained. The experiment was then modified such that the *Eco*RI fragments were ligated to a kanamycin resistance gene and these constructs were electroporated into XA86. Selection on plates containing kanamycin meant that growth of colonies was due to cells harbouring mini plasmid consisting of the pXA1 replicon and the kanamycin marker gene. However, this technique was also unsuccessful.

The highest concentration of SDS that showed visible increase in turbidity was 0.005%. XA86 grew to saturation at two SDS concentrations only (0.005% and 0.001%). Only a slight increase in turbidity was observed in the other flasks. After serial dilution of cells growing in 0.005% SDS, 235 colonies were selected for screening. Initial screening of the 235 colonies showed that four colonies i.e., H13, J5, J14 and K13 appeared to have failed to hybridize to the plasmid probe (Fig 2.1). This was observed as an absence of a strong hybridization signal after chemiluminescent detection. None of the five negative controls displayed hybridization signals. When this test was repeated with these four colonies, together with appropriate controls, it was observed that the plasmid probe hybridized to colony K13 to give hybridization signals that were as strong as that of XA86 (Fig. 2.2). The probe did not hybridize to the negative controls used (*E. coli* and PDDCC1196 cells), whereas XA86 cells showed strong hybridization signals. Plasmid isolations from the three cultures that appeared to be plasmid-free showed that they had definitely been cured of the plasmid. Colony H13 was selected for further studies and designated XA86C.

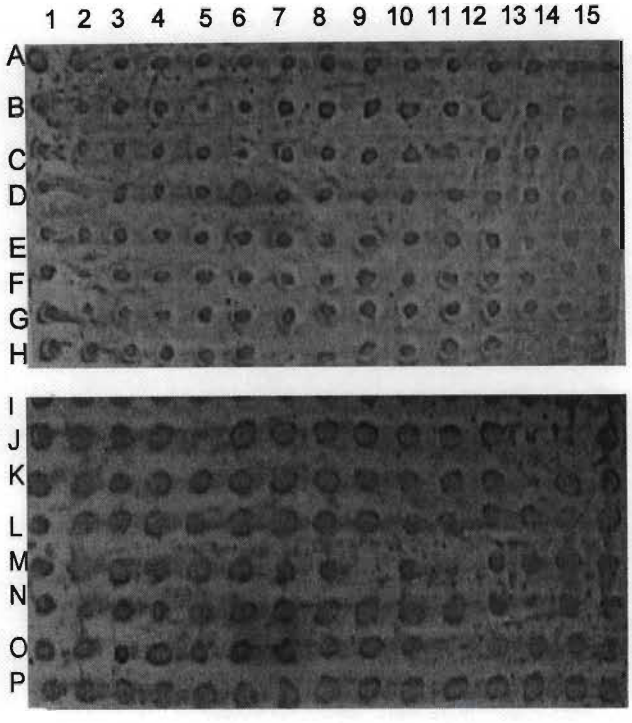


Fig. 2.1 Autoradiographs of colony blots used to screen colonies for loss of pXA1. Cured strains were tentatively identified in positions H13, J5, J14 and K13. PDDCC196 negative controls occur in positions C6, J13 and O11. *E. coli* DH5α negative controls occur in positions H7 and M9.

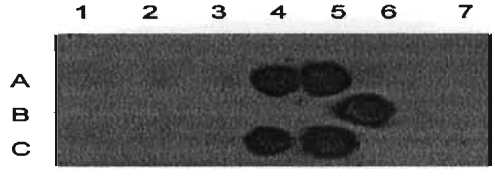


Fig. 2.2 Autoradiograph showing re-testing of suspected cured cells. Colonies H13 (A1 and C1), J5 (A2 and C2) and J14 (A3 and C3) tested negative for pXA1. Colony K13 (A4 and C4), XA86 (A5, B6 and C5) tested positive for pXA1. Negative controls PDDCC196 (A7) and *E. coli* DH5α (C7) did not hybridize to the probe.

Fig. 2.3 shows the results of plasmid isolations from XA86 and XA86C and restriction of their DNA with *Eco*RI. In lane 1, the typical banding pattern after isolation of pXA1 from XA86 is seen, with the CCC form of the plasmid migrating slower than the mixture of linear plasmid and linear chromosomal DNA fragments. The majority of the DNA was isolated as CCC molecules. In lane 2, however, the only band present is the one co-migrating with linear DNA. The characteristic multiple banding pattern associated with pXA1 presence was not observed. Restriction with *Eco*RI showed a typical restriction pattern for pXA1 in lane 3. The six *Eco*RI fragments are not present in lane 4. A 'smear' is present in this lane which indicates that a multitude of *Eco*RI fragments of varying sizes are present, which is characteristic of chromosomal and not plasmid DNA.

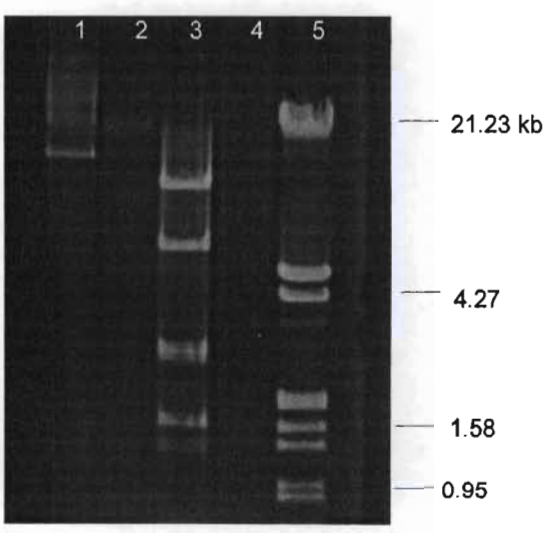


Fig. 2.3 Plasmid isolations from a cured strain (XA86C) and from XA86. Comparison of intact and *Eco*RI-digested DNA showed that pXA1 was absent in the cured strain. Lane 1: uncut XA86 DNA; lane 2: uncut XA86C DNA; lane 3: *Eco*RI-cleaved XA86 DNA; lane 4: *Eco*RI-cleaved XA86C DNA; and lane 5: *Eco*RI/*Hind*III-cleaved λ DNA.

2.3.2 **Albicidin Assays**

Production of albicidin by XA86 was visually demonstrated using the agar plate technique (Fig 2.4a). A zone of clearing was observed indicating that *E. coli* had been inhibited. XA86C also produced a zone of clearing (Fig. 2.4b). Repeated experiments with a similar sized inoculum

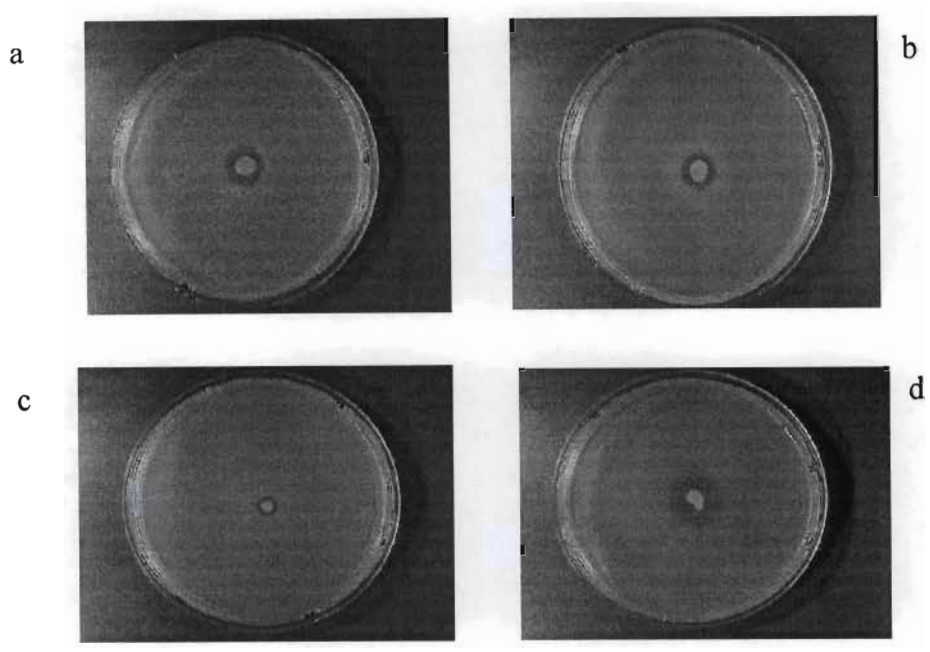


Fig. 2.4 Albicidin production by *X. albilineans* strains: a) XA86; b) XA86C; c) XA86CS1 and d) GLP7C.

at approximately the same stage of growth showed that the zones for XA86 and XA86C were of the same size. Introduction of pBX1, a derivative of pXA1 (Table 4.1), had no effect on albicidin production (Fig. 2.4c). Strain GLP7C, which is a plasmid-cured strain derivative of *X. albilineans* GLP7 also retained the ability to produce albicidin after loss of plasmid pXA2 (Fig. 2.4d), the plasmid indigenous to this strain.

2.3.3 **Pathogenicity Trials**

Two hundred and eight plants were inoculated with the two *X. albilineans* strains. Ten plants died during the course of the experiments, apparently due to physical factors such as dehydration and fungal attack. The 44 control plants representing all five cultivars were healthy and did not display any leaf scald symptoms during the three month period post-inoculation. Of the remaining 198 plants inoculated with *X. albilineans*, 153 plants or 77% displayed leaf scald symptoms. The first symptom observed was the appearance of a chlorotic area at the site of

inoculation on the first leaves to emerge after inoculation. These symptoms which surrounded the needle puncture marks were called "shock" responses (Fig. 2.5). Two weeks after inoculation, the characteristic white pencil lines appeared along the length of the inoculated leaf blades (Fig 2.6). The number of lines on each leaf ranged from one to five in a diffuse band along the middle of the leaves. The majority of the leaves possessed one to three lines only. In some leaves, the central vein also turned white. The majority of the plants in which a shock response was evident also developed the white pencil lines.

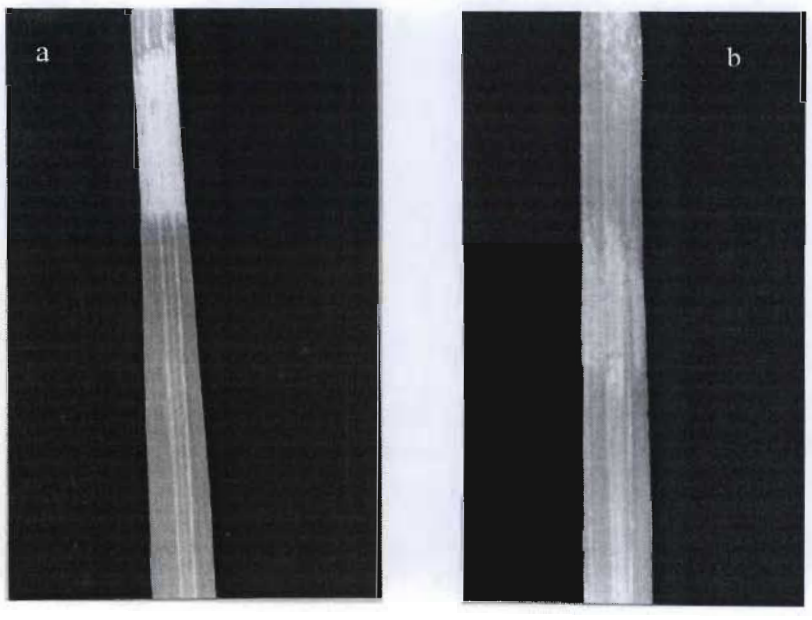


Fig. 2.5 "Shock" responses observed after inoculation of a) N14 and b) N17 cultivars with *X. albilineans* XA86 and XA86C, respectively.

Table 2.4 shows a comparison of the pathogenic ability of XA86 and XA86C. The overall percentage of successful inoculations for both strains were 79% and 75%, respectively. In three cultivars, the percentage of plants exhibiting leaf scald symptoms was higher for XA86 than for XA86C. However, the incidence of leaf scald in these cultivars inoculated with XA86C was still extremely high. In cultivar N22 this pattern was reversed, whereas in cultivar N11 the results were identical. These results clearly showed that XA86C was still pathogenic and since different

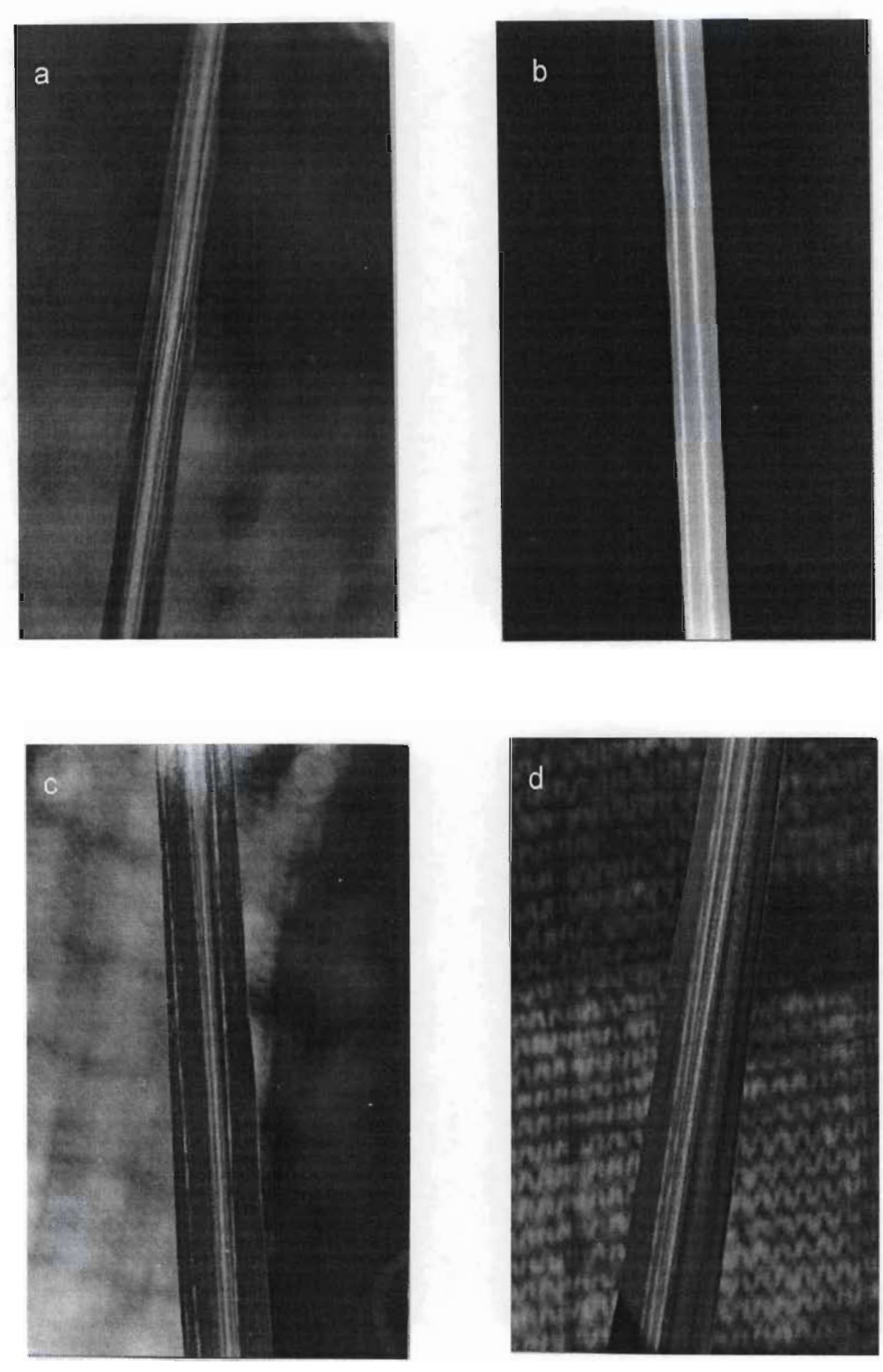


Fig. 2.6 White pencil-lines observed after inoculation of sugarcane cultivars with *X. albilineans* strains. Similar symptoms were produced after inoculation of XA86 into N11 (a) and N14 (b) compared to inoculation of XA86C in N22 (c) and NCo376 (d).

resistance levels of the cultivars with these two strains were not the object of this study, statistical analysis of the results was not performed

Pencil line symptoms were observed on the first leaves to emerge after inoculation and also on leaves that emerged subsequently. However, these symptoms gradually faded from the oldest leaves and newer leaves did not show any pencil lines. After three months, most plants were free of leaf scald symptoms. No other leaf scald symptoms were observed and none of the plants appeared to be seriously affected by the disease.

TABLE 2.4 Incidence of leaf scald symptoms in sugarcane cultivars inoculated with XA86 and XA86C

Cultivar	XA86	XA86C	Total
N11	14/15 = 93%	14/15 = 93%	28/30 = 93%
N14	16/23 = 70%	10/19 = 52%	26/42 = 62%
N17	19/25 = 76%	15/25 = 60%	34/50 = 68%
N22	16/21 = 76%	22/22 = 100%	38/43 = 88%
NCo376	15/17 = 88%	12/16 = 75%	27/33 = 82%
Total	80/101= 79%	73/97 = 75%	153/198 = 77%

2.3.4 **Inhibition of *X. albilineans***

The inhibitory effect first due to contamination of agar plates by Organism X, was repeatedly obtained after inhibition assays. Inhibition of *X. albilineans* cells in an agar overlay was seen as a zone of clearing around Organism X (Fig. 2.7). This effect was not observed when Organism X was grown in Boost medium. However, subsequent transfer to Wilbrink's medium

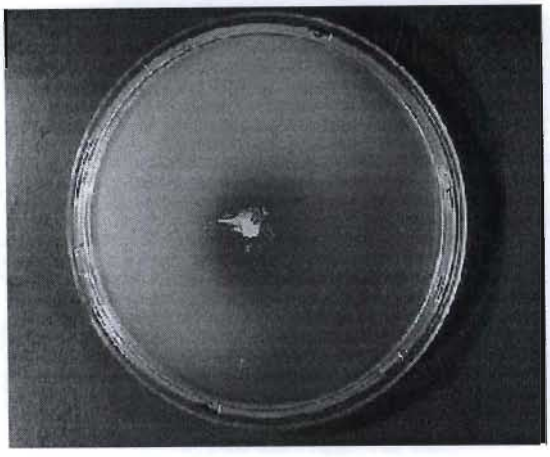


Fig. 2.7 Inhibition of strain 4320 by Organism X. Similar zones of clearing were observed with *X. albilineans*.

restored inhibitor production. Organism X also inhibited growth of *P. syringae* pv. *syringae* 3420 but not *E. coli* DH5α. Due to rapid growth of *P. syringae* pv. *syringae* 3420 compared to *X. albilineans*, this strain was used for the inhibition assays. Inhibitor production was dependant on the growth phase of Organism X. Inhibition zones were not observed when Organism X and *P. syringae* pv. *syringae* 3420 were inoculated onto agar plates at the same time. Zones were only observed after growth of Organism X for a day and subsequent overlay with top agar containing *P. syringae* pv. *syringae* 3420. Due to slow growth of *X. albilineans*, this procedure was unnecessary. Light microscopic examination of Organism X smears showed that the bacterium was a Gram-positive coccus. The majority of the cells occurred singly. Colonies on agar plates were white. Organism X was cultured on either Boost medium or Wilbrink's medium at 28°C.

2.4 DISCUSSION

The majority of plant pathogenic bacteria harbour indigenous plasmids. Even though the overwhelming majority of these plasmids have no known function, plasmids from plant pathogens are well known for conferring pathogenic abilities to their hosts. It was therefore imperative to determine whether pXA1 played a role in pathogenicity of *X. albilineans* in sugarcane.

Since specific loci on the plasmid were not being targeted, but the whole plasmid itself, techniques such as transposon mutagenesis and site-directed mutagenesis could not be used in examining the role of the plasmid in pathogenicity. Therefore, plasmid-curing experiments were performed to obtain plasmid-free derivatives of XA86. Since the plasmid-free strains were going to be subsequently used in pathogenicity trials, it was important to use methods that caused minimal damage to the genetic makeup of the cells. Traditional curing methods using DNA intercalating agents would have resulted in random insertion of ethidium bromide and acridine orange into both chromosomal and plasmid DNA. These mutations can lead to incorrect conclusions regarding plasmid function as was the case with plasmids in *P. syringae* pv. *syringae* and *P. syringae* pv. *phaseolicola* which, after curing experiments, were implicated in syringomycin and phaseolotoxin production (Gantotti *et al.*, 1979). The genes encoding these toxins were subsequently found to be located on the bacterial chromosome (Coplin, 1989). Initial curing experiments involved attempts to evict pXA1 from XA86 by incompatibility. This method of plasmid curing, unlike the two methods mentioned previously, only affects stable maintenance of indigenous plasmids and does not affect the bacterial chromosome. It has been used to evict indigenous plasmids from the plant pathogens *E. amylovora* (Steinberger *et al.*, 1990), *E. stewartii* (Fu *et al.*, 1998) and *P. syringae* pv. *tomato* (Murillo *et al.*, 1994). Similar experiments were performed in XA86. These experiments were not successful, primarily because the *Eco*RI fragment that was suspected to contain the origin of replication could not be cloned (Chapter

Three) and therefore was not used. Failure to evict pXA1 by incompatibility was most likely due to one of the following factors. Firstly, cleavage at the *Eco*RI restriction sites may have disrupted the pXA1 replicon, making it non-functional. Secondly, the electroporation conditions used at that time may not have been optimal for *X. albilineans* as electroporation conditions for this organism have not been published.

The use of supraoptimal growth temperatures and growth in the presence of SDS have been used previously in the curing regimen for the epiphytic bacteria *E. uredovora* (Thiry, 1984) and *E. herbicola* (Gantotti and Beer, 1982). In this study, SDS treatment was successfully used to cure XA86 cells. The efficiency of the treatment was extremely high. Cured cells were detected at a frequency of 0.01 per treated cell. Although XA86 was grown in concentrations of SDS ranging from 0.001% to 0.1%, only cells treated with 0.005% SDS were tested for loss of pXA1. This was the highest concentration of SDS in which the cells grew to saturation. It was assumed that poor growth at SDS concentrations above 0.005% was due to cell lysis. Thiry (1984) found that an increase in SDS concentration from 0.1% to 2% resulted in increased plasmid loss in *E. uredovora*. Plasmid loss was highest at 2% SDS and curing was even reported at an SDS concentration of 5%. These results are surprising since 1% SDS is routinely used to lyse Gram-negative bacteria (Sambrook *et al*, 1989). The exact mechanism of SDS curing is unknown but it is possible that partial disruption of cell membranes by SDS results in leakage of cell constituents including plasmids found in the cytoplasm. Growth at elevated temperatures presumably also creates temporary pores in the cell envelope which results in loss of plasmids. A similar process occurs in the "heat shock" step in transformation protocols, except that in this case temporary pores are created which allow plasmids to enter cells. Both curing techniques are probably restricted to low copy number plasmids only, as all copies of the plasmid have to be lost for curing to be successful. Even if only one copy is left in the cell, the original number in

plasmids in the cell will be restored as plasmid copy number is strictly regulated.

While the curing protocols themselves are simple to perform, successful isolation of plasmid-cured strains depends on a reliable assay for these strains. Most indigenous plasmids are genetically cryptic and it is therefore difficult to assay for loss of these plasmids. The lack of a suitable assay is probably the reason why the majority of these plasmids will remain cryptic. Cells carrying indigenous plasmids which specify recognizable traits such as antibiotic resistance or pigmentation can be easily screened for loss of plasmids. It is possible to "tag" a cryptic plasmid with an antibiotic resistance marker by transposon mutagenesis using a suicide plasmid to deliver the transposon. A successful transposition event into the cryptic plasmid would result in a "tagged" plasmid in the host cell which would subsequently lose the suicide plasmid within a few generations. It would then be possible to evict this plasmid from the cell by following the loss of the antibiotic resistance marker. This strategy was used to evict pCPP60, a self-transmissable plasmid in *E. amylovora* in a study which proved the non-involvement of the plasmid in pathogenicity (Steinberger *et al.*, 1990). pCPP60 was evicted by incompatibility after introduction of the tagged derivative which carried kanamycin resistance. The tagged plasmid was cured using acridine orange which resulted in isolation of a plasmid-free strain.

In this study, after failure of the incompatibility experiments, it was decided to physically screen clones for loss of pXA1 after SDS curing. Plasmid isolations were not attempted as it was considered too labour intensive and it was not expected that the frequency of curing would be as high as 0.01. It was previously shown that an 800 bp pXA1 fragment hybridized strongly to XA86 cells (Permaul, 1994). Furthermore, hybridization was shown to occur only to plasmid DNA and not chromosomal DNA. It was therefore decided to screen colonies for the loss of pXA1 by DNA hybridization of the 800 bp probe to colonies lysed on nylon membranes. In the initial screen, the technique clearly showed that the probe had not hybridized to DNA from four

colonies. Loss of pXA1 was verified by repeating the technique with positive and negative controls, and miniprep plasmid isolations. This protocol was also successful in curing strain GLP7. Hybridization and plasmid isolations revealed that this strain lost the 36.2 kb plasmid pXA2 (Ramdeen, 1999).

The albicidin assays on agar plates proved that pXA1 was definitely not responsible for albicidin production in XA86. GLP7C also retained the ability to produce albicidins after loss of pXA2. This was expected as gene clusters involved in albicidin production were identified on the chromosome of a strain of *X. albilineans* from Florida (Rott *et al.*, 1996). All *X. albilineans* strains do not possess a common plasmid, and it is unlikely that toxin production is specified by plasmids in some strains and chromosomal genes in others. The plate assays also demonstrated that export of the albicidins is not plasmid-encoded in XA86, even though pXA1 appears to have a transfer/export function (Chapter Four). Again this is consistent with the discovery of albicidin-production genes on the chromosome. Genes involved in export of a particular molecule are often found downstream of the genes encoding that molecule (Covacci and Rappuoli, 1993).

The success of pathogenicity trials depends mainly on the number of plants that develop disease symptoms. The preferred method of inoculation of sugarcane with *X. albilineans* is a modification of the decapitation technique described by Koike (1965). Since the sugarcane plants used in this study were housed in a nursery together with plants infected with various fungal pathogens, it was decided that the decapitation technique would not be used. Instead a modification of the hypodermic syringe method (Akhtar *et al*, 1988; Martin and Robinson, 1961; Thompson, 1982) was employed. The original technique involved injection of a suspension of cells at or near the growing point of the plants. In 1994, Rott *et al.* showed that a high concentration of cells were present in the basal nodes of infected cane. The syringe inoculation technique was therefore modifed by inoculating plants at the base of the stalks as well as near the

growing point. This technique resulted in a surprisingly high number of plants showing leaf scald symptoms after inoculation. One hundred and fifty three out of 198 plants (77%) developed white pencil lines after inoculation. All 22 N22 plants inoculated with XA86C displayed disease symptoms, while the lowest percentage of plants showing white lines (52%) was observed in N14 inoculated with XA86C.

Unfortunately, none of the cultivars used were highly susceptible to leaf scald disease. The five commercial cultivars tested possessed moderate to high resistance to the pathogen. Therefore, none of the plants displayed any acute phase symptoms. Furthermore, the white pencil lines gradually disappeared and were not evident three months after infection. These pencil lines that are produced in resistant varieties are due to initial production of toxin by the pathogen. Thereafter, the pathogen is not able to survive in high enough population densities to cause further disease symptoms in these varieties and are unable to colonize the plants effectively. However, it has also been shown that even if symptoms are not being expressed in tolerant varieties, population densities of the pathogen in both tolerant and a susceptible variety were similar (Rott *et al.*, 1994a). The production of white pencil lines by XA86C shows that inoculation was successful and that these strains had reached high densities in the plants. More importantly, it demonstrates that XA86C is capable of albicidin production *in planta*. This together with the plate assays provides conclusive evidence that pXA1 is not required for albicidin production in XA86.

Although it is unlikely that pXA1 may yet play some role in pathogenicity, this possibility cannot be discounted unless susceptible cultivars have been inoculated with XA86C and disease progression is followed for several ratooning stages. Albicidins have been implicated in chlorosis (Birch and Patil, 1987a; b) but they have not been implicated in other leaf scald symptoms. It is not certain whether the toxin is a true virulence or pathogenicity determinant or merely an

antibiotic that provides *X. albilineans* with an advantage in competition with other bacteria. While non-albicidin-producing mutants have been shown to be non-chlorosis inducing (Birch and Patil, 1987a, b), their ability to cause acute phase symptoms in susceptible hosts has not been well documented. Rott *et al.* (1996) have shown that albicidin mutants are present in similar population densities to the toxin producing strain Xa23. Therefore, although pXA1 has been shown to be not required for albicidin production and subsequent development of white lines in plants, its non-involvement in other aspects of pathogenicity has not yet been proved.

The fortuitous isolation of an organism that can inhibit the growth of *X. albilineans* in culture, could ultimately prove useful as a means of biocontrol of the pathogen. At this initial stage it is unclear whether the inhibitor is a true antibiotic active against various bacteria or a metabolic by-product or a modification thereof. It is also possible that this inhibition could be caused by a physiological effect such as acidification of the medium. The inhibitor seems to be produced only in the final stages of the growth cycle of Organism X or in sub-inhibitory concentrations prior to this. It is also unknown whether the effect is an inhibitory or a lethal one. It was noticed that after several days, single colonies appeared in the zone of clearing and the cells eventually grew profusely in this area. It will be interesting to test whether these colonies represent resistant strains. The lack of inhibition of *E. coli* suggests some form of selective inhibition. However, a wide array of test organisms would be required to determine whether the inhibitor is specific for *Pseudomonas* or *Xanthomonas.*

An antibiotic having antimicrobial activity against certain *Xanthomonas* species has been described previously (Shimura *et al.*, 1979). This substance, produced by *Streptomyces zaomyceticus*, was identified as a proline analog that acted as an antimetabolite of proline. Biocontrol of *X. albilineans* has been reported with *P. dispersa* (Zhang and Birch, 1996; 1997). If Organism X does produce an antibiotic and it is capable of growth at wound sites, it is possible

that this organism could be even more effective in biocontrol as it directly inhibits *X. albilineans*. While it is unlikely that a laboratory contaminant would be able to grow on or in sugarcane, growth of the organism on Wilbrink's agar is promising. If Organism X does produce an effective antibiotic, the ultimate form of control would be the creation of transgenic sugarcane plants.

CHAPTER THREE

PHYSICAL CHARACTERIZATION OF pXA1

3.1 INTRODUCTION

Physical characterization of plasmids involves determination of the physical attributes of plasmids. In common with living organisms, plasmids also have a system of classification which groups together plasmids which are evolutionarily linked. Presently, the only reliable system is the grouping of plasmids into incompatibility groups. Plasmids belonging to the same incompatibilty (Inc) group cannot coexist in progeny cells and are therefore termed incompatible. Various Inc groups exist, viz., IncC, IncF, IncH, IncI, IncJ, IncL, IncM, IncN, IncO, IncP, IncQ, IncS, IncT and IncW as well as subgroups within these groups. However, this system of classification is not absolute as plasmids belonging to the same Inc group are not always related. Mutations can also change Inc phenotypes. Other plasmid characteristics that are often determined, although not used in classification, are size, modes of replication, restriction maps, stability functions and copy number.

pXA1 has been partially characterized previously (Permaul, 1994). The plasmid was found in XA86 only, and not in four other *X. albilineans* strains investigated. A restriction map was constructed (Fig. 3.1) for the 24.9 kb plasmid. The structural nature of DNA bands on agarose gels was identified and cloned plasmid fragments were tested for antibiotic resistance and protein production. In another study, the effect of the loss of pXA1 from XA86 was studied using a range of biochemical, antibiotic, metal resistance and physiological tests (Rahaman *et al.*, 1995). However, these tests failed to demonstrate any phenotype encoded by the cryptic plasmid.

Physical characterization has also been performed on two other plasmids found in *X.*

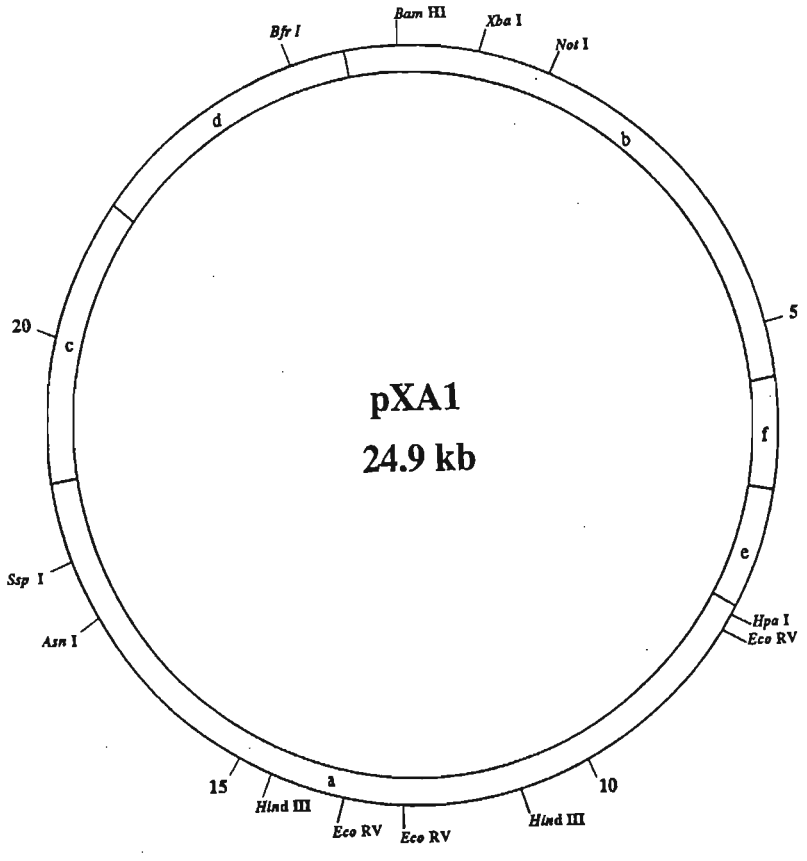


Fig. 3.1 Physical map of restriction endonuclease cleavage sites on pXA1. The *Bam*HI site was arbitrarily chosen as position 0. The inner ring represents six *Eco*RI fragments. The positions of *Eco*RI fragments e and f are interchangeable (Permaul, 1994).

3.2) found in strain GLP7 (Ramdeen, 1999) and the 25.3 kb plasmid pXA3 (Fig. 3.3) found in strain KEN6 (Iyer, 1996). Comparison of the three *X. albilineans* plasmids using restriction analysis data suggests that pXA1 and pXA3 are related (Iyer, 1996; Rahaman, 1995; Ramdeen, 1999).

In this study, physical characterization of pXA1 involved construction of a more detailed restriction map of the plasmid. This map was required for subsequent cloning and sequencing

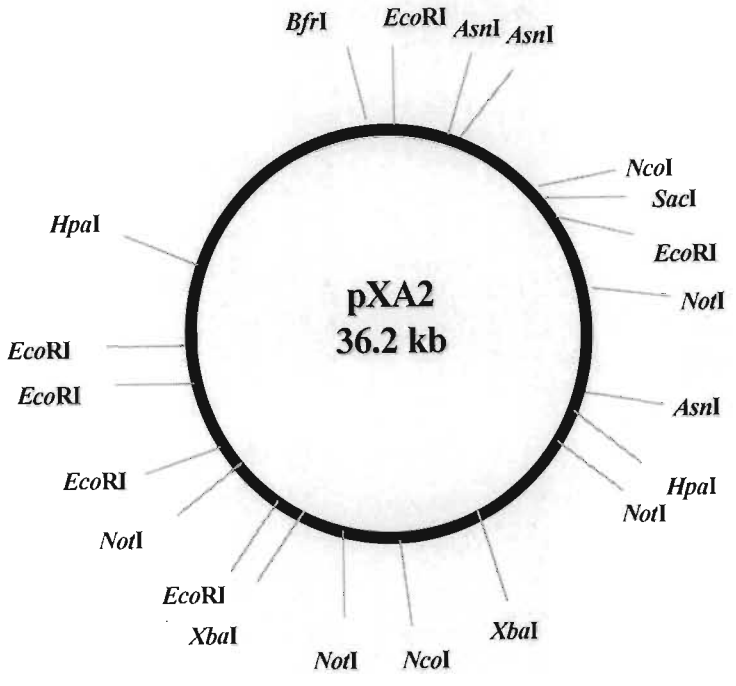


Fig. 3.2 Restriction map of pXA2, the plasmid found in *X. albilineans* GLP7 (Ramdeen, 1999).

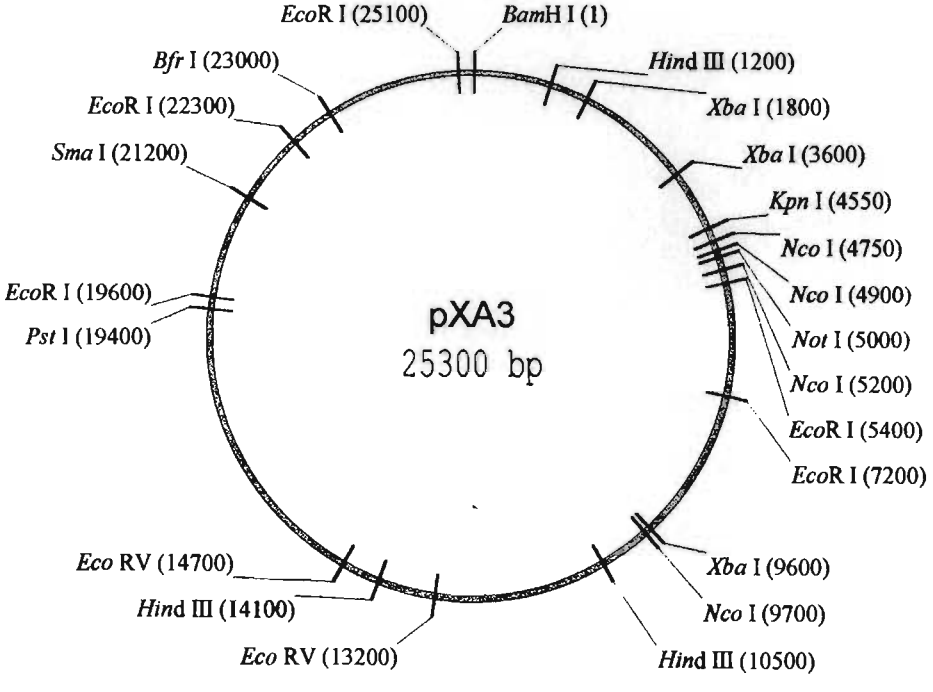


Fig. 3.3 Restriction map of pXA3, the plasmid found in *X. albilineans* KEN6 (Iyer, 1996).

of pXA1 fragments which resulted in identification of a transfer region on the plasmid. Gene probes were also used to determine whether the plasmid harboured any *avr* genes. The presence of *avr* genes in pseudomonad and xanthomonad plant pathogens is common (Hopkins *et al.*, 1992; Kearney and Staskawicz, 1990; Murillo *et al.*, 1994; Osbourn *et al.*, 1990; Swarup *et al.*, 1991) and these genes are often found on conjugative plasmids (Bonas *et al.*, 1989; De Feyter *et al.*, 1993; Murillo *et al.*, 1994; Turner *et al.*, 1985). This chapter describes the isolation and physical characterization of pXA1.

3.2 MATERIALS AND METHODS

3.2.1 **Plasmid DNA Isolation**

Plasmid DNA was isolated from *X. albilineans* and *E. coli* using protocols adapted from the alkaline lysis procedure of Birnboim and Doly (1979). Two types of plasmid preparations were performed: minipreparations (minipreps) and large scale preparations. DNA from minipreps were used to screen cells for plasmids whereas DNA from large scale preparations were used for restriction mapping and other plasmid characterization protocols.

Miniprep isolations were performed on 1.5 ml of *E. coli* broth cultures and from three ml of *X. albilineans* cultures. Cells were pelleted for one min in a microcentrifuge and resuspended in 100 µl of Solution A (25 mM Tris.HCl; 50 mM glucose; 10 mM EDTA, pH 8.0) containing 100 µg/ml RNaseA. The cell suspension was lysed by addition of 200 µl of freshly prepared Solution B (0.2 M NaOH; 1% SDS) and immediate inversion of the microcentrifuge tubes. In this step, cell membranes were disrupted by the anionic detergent SDS. Liberated DNA was denatured with NaOH and cellular RNA degraded by RNase A. As soon as the cell suspension turned clear (indicating cell lysis), plasmid DNA was reannealed and chromosomal DNA and

cellular proteins were precipitated by addition of 150 µl of cold Solution C (3 M sodium acetate, pH 4.8) and incubation on ice for 5 min. Precipitated material was removed by centrifugation at maximum speed (15 800 × *g*) for 15 min. DNA in the supernatant was precipitated by addition of two volumes of cold 100 % ethanol and incubation at -70°C for 15 min. The precipitated DNA was recovered by a 15 min centrifugation step and washed once in 70 % ethanol. The DNA pellet was air-dried, dissolved in 20 µl TE buffer (10 mM Tris.HCl; 1 mM EDTA, pH 8.0) and stored at -20°C. Initially, phenol/chloroform extractions were performed prior to ethanol precipitation of DNA from *X. albilineans* cultures, but this was later found to be unnecessary.

Large scale isolations were essentially scaled-up versions of the miniprep procedure. Five ml of Solution A, 10 ml of Solution B and 7.5 ml of Solution C were used to obtain crude plasmid preparations from 250 ml batches of bacterial broth cultures. After isopropanol precipitation, the plasmid-containing pellets were dissolved in TE buffer. An approximately equivalent volume of cesium chloride (CsCl) crystals was dissolved in the DNA solution, ethidium bromide (EtBr) was added to a final concentration of 200 µg/ml and the refractive index of the solution was adjusted to approximately 1.395. A clearance spin at 8 000 rpm (JA 20.1 rotor, Beckman Instruments Inc.) was used to remove EtBr-protein complexes, which resulted in "cleaner" gradients especially in the upper regions. The solutions were loaded in Quick-Seal centrifuge tubes (Beckman Instruments Inc.), filled with paraffin oil and sealed. The CsCl gradients were formed by centrifugation at 10°C in either a fixed angle rotor (Beckman Ti75) for 24 h at 45 500 rpm or in a vertical rotor (Beckman VTi65) for 7 h at 55 000 rpm. Covalently closed circular (CCC) and linear forms of plasmids were collected and EtBr and CsCl were removed by *N*-butanol extractions and precipitation, respectively (Sambrook *et al.*, 1989). Purified plasmid DNA was resuspended and stored in TE buffer.

3.2.2 Restriction Mapping of pXA1

The restriction map of pXA1 (Fig. 3.1) was extended to include more enzymes and corrected in terms of the position of an *Eco*RI fragment, addition of an *Eco*RI site and revision of the size estimate of the plasmid. Restriction reactions were typically carried out in a total volume of 15 µl, containing an appropriate amount of plasmid DNA, sterile deionised water, 1 × SuRE/Cut buffer (Boehringer Mannheim) and 5 U restriction endonuclease(s) [Boehringer Mannheim]. Reactions were incubated at 37°C, unless otherwise stated by the manufacturer, for 60-90 min. When double or triple digests were performed, the appropriate SuRE/Cut buffer was selected to enable all enzymes to digest at least 75% of the DNA within 1 h. If this was not possible, 2 h digestions were performed and NaCl and Tris.HCl were added after 1 h to ensure that reaction conditions for the second enzyme were optimal.

After restriction endonuclease digestion, 6 × gel loading buffer (40% sucrose; 0.25% bromophenol blue) was added to the plasmid sample and the restriction fragments were analysed by horizontal agarose gel electrophoresis. DNA was loaded in 0.7 - 1.5% agarose gels and electrophoresis was carried out in 1 × TAE electrophoresis buffer (40 mM Tris-acetate; 2 mM EDTA). After electrophoresis, gels were stained in 0.5 µg/ml EtBr for 15 min and fluorescent DNA bands were visualized on a UV transilluminator (UVP Inc.). Alternatively, EtBr was incorporated in the gels at a final concentration of 1 µg/ml and DNA could be viewed during and immediately after electrophoresis. When this procedure was used for longer electrophoretic runs, an appropriate quantity of EtBr was added at the cathode end of the electrophoresis tank to visualize smaller fragments. Gels were photographed with either a Minolta camera using Ilford FP4 Plus film or a Polaroid land camera using Polaroid film.

The sizes of restriction fragments were calculated using the SW5000 Gel Documentation System (UVP Inc.) by comparison with sizes of molecular weight markers (phage λ DNA

cleaved with *Hin*dIII and *Eco*RI/*Hin*dIII). The positions of the restrictions sites on the pXA1 map were deduced using the multiple endonuclease digestion method (Ausubel *et al.*, 1989).

3.2.3 DNA Ligations and Transformation

DNA fragments used in ligation reactions were isolated from agarose gels with the Qiaex II kit (Qiagen). Isolations were performed according to the manufacturer's instructions, except that DNA was eluted from the Qiaex beads at 50°C irrespective of size, and with TE buffer. The concentration and purity of isolated fragments were determined with the Genequant RNA/DNA Calculator (Pharmacia). Ligation of linearised cloning vectors (Table 4.1) and linear pXA1 fragments possessing 5′-overhangs (sticky ends) were performed using T4 DNA ligase (Boehringer Mannheim). Typical ligation reactions contained equimolar amounts of reactants, deionised water, ligation buffer and 1 U of T4 DNA ligase in a total volume ranging from 10 - 20 µl. The DNA reactants were heated at 45°C for 5 min prior to addition of ligation buffer and ligase to ensure that complementary base-pairing between sticky ends were disrupted. Reactions were incubated overnight at room temperature, stopped by the addition of 25 mM EDTA, pH 8.0 (final concentration) and stored at 4°C.

E. coli cells were made competent, i.e., able to take up DNA, using $CaCl_2$. One ml of an overnight culture of *E. coli* DH5αF′ was diluted to 30 ml with YT broth (Section 2.2.1). The culture was incubated at 37°C, in an orbital shaker, until the optical density of the culture reached 0.375 absorbance units at 590 nm. This step was critical as older cells have reduced transformation efficiencies (Ausubel *et al.*, 1989). Cells were immediately placed on ice and kept cold in all subsequent steps, and collected by centrifugation at 10 000 × *g* for 10 min at 4°C. The pellet was gently resuspended in 10 ml cold $CaCl_2$ solution (100 mM $CaCl_2$) and pelleted again. After washing, cells were resuspended gently in 10 ml cold $CaCl_2$ solution and incubated

for 20 min on ice. It was during this step that the cells became competent. Cells were pelleted once more, resuspended gently in 2.5 ml glycerol solution (100 mM $CaCl_2$; 15% glycerol) and stored for 24 h at 4°C to increase competency (Ausubel *et al.*, 1989). One hundred µl aliquots of cells were subsequently snap frozen and stored -70°C.

Plasmids were introduced into competent *E. coli* DH5αF′ cells by transformation. One hundred µl of frozen competent cells were allowed to thaw slowly on ice. One to three µl of the ligation reaction mixtures was added to the cells and the contents of the tubes mixed gently. Cells were incubated for 30 min on ice to allow the DNA to bind to the competent cells. Cells were subsequently heat-shocked for 2 min in a 42°C water bath, to allow DNA to enter the cells and immediately placed on ice for 15 min. Nine hundred µl of YT broth (Section 2.2.1) was added to the cells in eppendorf tubes and incubated for 1 h at 37°C in a shaking incubator to allow for the replication of recombinant plasmids before a selective medium was used. One hundred µl of the cell suspension was then spread onto X-gal agar plates (Section 2.2.1) which were incubated overnight at 37°C. Recombinant plasmid-containing colonies were selected by blue-white selection (Sambrook *et al.*, 1989).

3.2.4 **Cloning of EIb**

The 6.4 kb *Eco*RI fragment of pXA1 (EIb) could not be cloned in the multifunctional vector pTZ19R even after repeated efforts (Permaul, 1994). Cloning of this fragment was also attempted in this study using various strategies. Cloning of the entire fragment was attempted using the low-copy-number, broad host range vector pRK415 (Table 4.1) and pK194 (Table 4.1) which possesses the origin of replication from pACYC194. Thereafter, the fragment was restricted with *Xba*I and *Bam*HI to pinpoint the exact region that was responsible for this effect. Cloning of this shortened region was then attempted using pTZ19R and pUC118 (Table 4.1): two

cloning vectors with an identical promoter but having their multiple cloning sites arranged in opposite orientations. Cloning was also attempted by growing the transformed cells at a reduced temperature (30°C).

3.2.5 **Hybridization with *avr* Genes**

The presence of *avr* genes in pXA1 was tested by hybridization with *avrBS2* and *avrXa10* gene probes. These genes were present in plasmids pBS2010 and pBSavrXa10, respectively (Table 4.1). A 2.0 kb *Stu*I/*Sac*I fragment of pBS2010 and a 3.1 kb *Bam*HI fragment of pBSavrXa10 were isolated from agarose gels using the Geneclean kit (BIO 101 Inc.) according to the manufacturer's instructions. These fragments were labelled by random-primed incorporation of radioactive dATP or dCTP. Labelling reactions (25 µl) contained template DNA, random hexanucleotide mixture (Boehringer Mannheim), deionised water, 200 µM of dTTP, dGTP, dATP or dCTP, $\alpha[^{32}P]$dATP or $\alpha[^{32}P]$dCTP and 1U of Klenow enzyme (Boehringer Mannheim). Prior to addition of the enzyme and dNTP's, the DNA templates were denatured by boiling for 5 min. Labelling reactions were incubated at 37°C for 45 min and stopped by addition of 0.5 M EDTA, pH 8.0. Labelled DNA was precipitated at -20°C after addition of 0.1 volume of 3 M sodium acetate (NaAc), pH 4.8 and two volumes of 100% ethanol. After centrifugation, the pellet was washed with 70% ethanol and resuspended in 50 µl of TE buffer.

The DNA to be hybridized to the *avr* gene probes, uncleaved pXA1 DNA and pXA1 cleaved with *Eco*RI, were run on agarose gels together with positive controls (pBS2010 and pavrXa10) and DNA molecular weight markers. Following EtBr staining and photography, the DNA in the gels were transferred to positively charged nylon membranes (Genescreen Plus - Du Pont) using an upward alkaline transfer protocol (Ausubel *et al.*, 1989). After capillary transfer,

the membranes were washed with 2 × SSC, air-dried and stored at room temperature between sheets of filter paper.

Membranes were prehybridized for 15 min at 42°C in prehybridization solution (per 10 ml: 5 ml deionized formamide; 2 ml 50% dextran sulphate; 1 ml 10% SDS; 2 ml H_2O and 0.58 g NaCl) with constant agitation in sealed plastic bags. Approximately 100 ng of the labelled DNA probes and 1 mg of salmon sperm DNA were denatured by boiling for 5 min and cooling immediately on ice, and added to the prehybridization solution. The bags were resealed and incubated at 42°C for 15 h with constant agitation. Following hybridization, the membranes were removed and washed twice in the following solutions with constant agitation to prevent high backgrounds: a) 100 ml of 2 × SSC at room temperature for 5 min; b) 200 ml of 2 × SSC/1% SDS for 30 min at 65°C; and c) 100 ml of 0.1 × SSC for 30 min at room temperature. After stringency washes, membranes were air-dried and exposed to X-ray film (Kodak). Autoradiographs were developed using an automatic autoradiograph developer (X-OMAT Film Processor, Kodak).

3.2.6 ***In Vitro* Expression of pXA1 Genes**

In vitro synthesis of radiolabelled proteins encoded by pXA1 was accomplished using the *E. coli* S30 Extract System for Circular DNA (Promega). This kit contains a cell-free *E. coli* extract that allows transcription and translation of naked DNA to yield radiolabelled proteins. In addition to expression of pXA1 genes, expression of genes in pBX1 (Table 4.1) and pXA2 was also studied. Reaction volumes made up to 50 µl with sterile deionized water contained 3 µg plasmid DNA, 5 µl of amino acid mixture minus methionine (Promega), 20 µl of S30 premix without amino acids (Promega), 15 µl of S30 extract (Promega) and 1 µl of [^{35}S]methionine (1 200 Ci/mmole at 15 mCi/ml)[Amersham]. Reactions were incubated at 37°C for 1 h and stopped

by placing on ice. Five µl aliquots were analysed on SDS-polyacrylamide gels after acetone precipitation to remove PEG which causes background staining.

Preparation, running and staining of denaturing polyacrylamide gels in a discontinuous buffer system were performed according to standard protocols and were described previously (Permaul, 1994). After drying of gels, bands were visualised by autoradiography using X-OMAT AR film (Kodak).

3.3 RESULTS

3.3.1 **Plasmid Isolation**

Miniprep and large scale plasmid preparations from *E. coli* and *X. albilineans* routinely yielded DNA of high quality that was suitable for subsequent DNA manipulations. Purity of miniprep DNA was tested by restriction endonuclease digestion whereas DNA from large scale isolations yielded DNA with an OD_{260}/OD_{280} ratio greater than 1.7. In addition to these procedures, plasmid DNA was also reproducibly isolated from bacterial cells using the following kits: Qiagen Plasmid Kit (Qiagen), High Pure Plasmid Isolation Kit (Boehringer Mannheim), Circleprep Spin Kit (BIO 101, Inc.) and the Wizard DNA Purification System (Promega).

3.3.2 **Restriction Mapping**

Another enzyme with a unique recognition site on pXA1, *Stu*I, was found 2.4 kb from the *Ssp*I site (Fig. 3.4). The enzymes *Cla*I and *Sph*I cleave pXA1 at least eight and nine times, respectively (Fig. 3.4) and *Hind*II were useful in restriction mapping. A seventh *Eco*RI fragment, 234 bp in size, was discovered by overloading *Eco*RI-digested DNA on 1.5% agarose gels and viewing the DNA after short electrophoretic runs. The position of this fragment on the pXA1

map was determined by DNA sequencing analysis (Chapter Four). Single sites for *Hin*dII and *Cla*I were found on the 1.3 kb *Eco*RI fragment and the position of this fragment was subsequently changed as a result (Figs 3.1 and 3.5). This change can be seen by comparison of Fig. 3.1 with Fig. 3.5. Fig. 3.5 depicts a restriction map of pXA1, showing selected restriction sites of various enzymes. A detailed list of restriction sites in regions of pXA1 is shown in Appendix 1. The overall size of the plasmid was also increased slightly to 25.054 kb after analysis of DNA sequencing data.

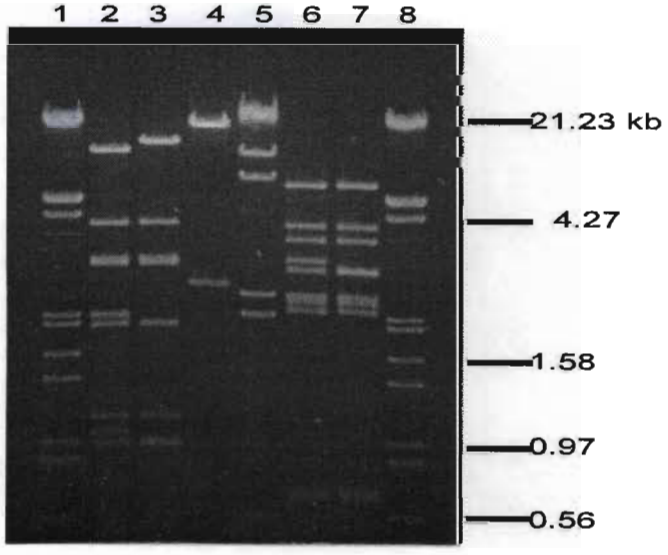


Fig. 3.4. Restriction digests of pXA1 which confirm the positions of *Hpa*I, *Not*I and *Stu*I sites. Lanes 1 and 8: λ DNA cleaved with *Eco*RI/*Hind*III; lane 2: *Cla*I/*Hpa*I; lane 3: *Cla*I/*Not*I; lane 4: *Ssp*I/*Stu*I; lane 5: λ DNA cleaved with *Hind*III; lane 6: *Sph*I; and lane 7: *Sph*I/*Hpa*I.

3.3.3 **Cloning of EIb**

Cloning of this 6.1 kb *Eco*RI fragment was unsuccessful in pTZ19R, even after dephosphorylation of the fragment. The fragment could not be cloned in the low-copy-number vector pRK415 or in pK194 which possesses the p15a origin of replication. Digestion of this

fragment with *Bam*HI yielded a 5.3 kb fragment and an 800 bp fragment. The smaller fragment

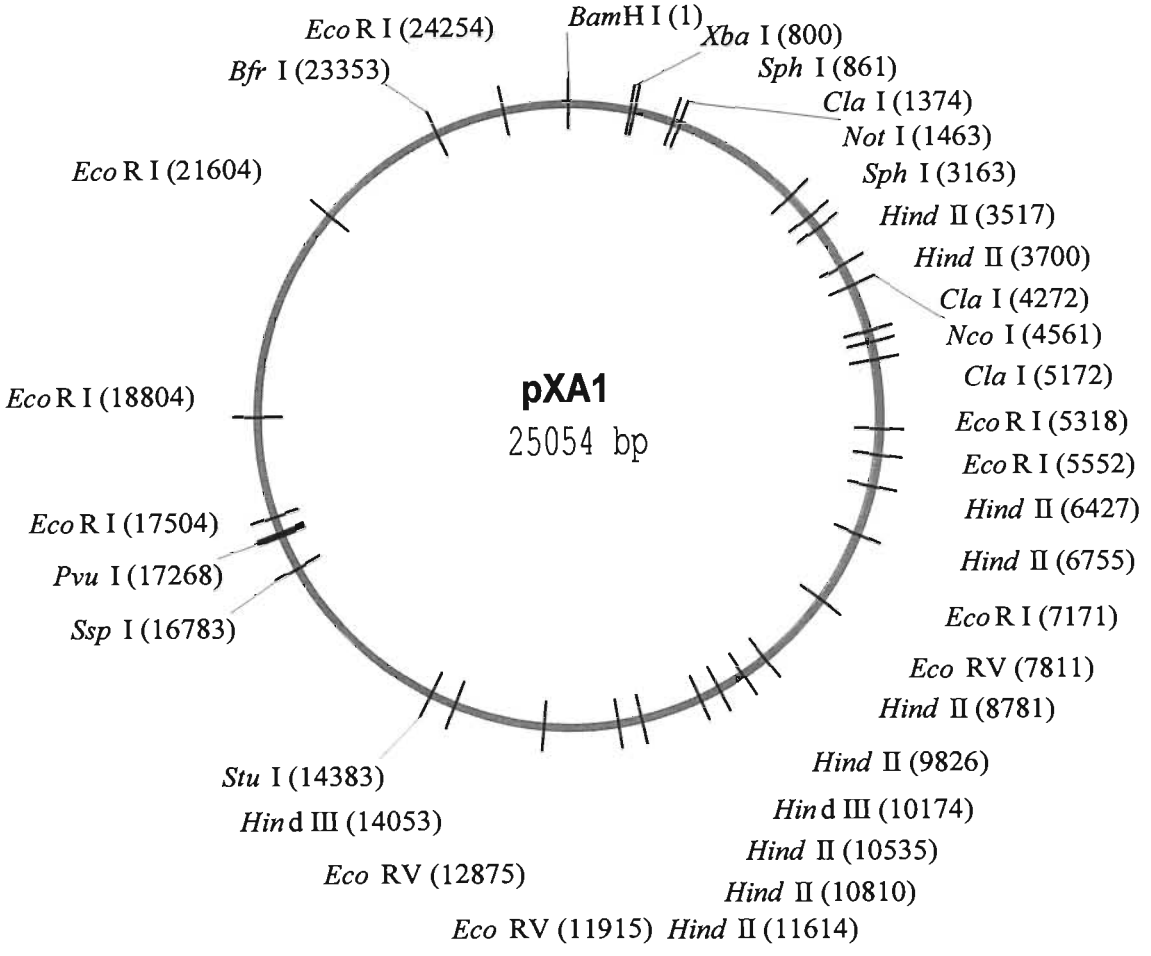


Fig. 3.5 Physical map of pXA1 showing the restriction sites of selected enzymes. The *Bam*HI site was arbitrarily chosen as position 1.

could be cloned easily (Table 4.1) but the larger fragment could still not be cloned. EIb was subsequently digested with *Xba*I to yield a 4.5 kb fragment and a 1.6 kb fragment. In this instance, the larger fragment could be cloned but not the smaller one. This property of lethality or inability to be cloned was therefore located on an 800 bp *Bam*HI/*Xba*I fragment, designated BXb. Attempts to clone BXb using the vectors pTZ19R and pUC118 whose multiple cloning sites are arranged in opposite directions and also by using a lower growth temperature were

unsuccessful.

3.3.4 **Probing with *avr* Genes**

Hybridization of pXA1 DNA to the *avrXa10* gene probe showed that this gene or a homolog was not present on pXA1 (Fig. 3.6). No hybridization bands were observed corresponding to pXA1 *Eco*RI fragments (lane 2) or to uncut DNA (lane 1). The only signal obtained was the positive control which corresponded to the 6.1 kb linear pBSavrXa10 DNA (lane 3). The presence of three additional bands in this lane correspond to partial digestion products which are only faintly visible on the agarose gel. Hybridization with the *avrBs2* probe produced similar results (Fig. 3.7). The probe did not hybridize to pXA1(lanes 1 and 2). The major band on the autoradiograph corresponds to the positive control (lane 3). The probe also hybridized to the 6.1 kb plasmid pBSavrXa10 containing the *avrXa10* gene.

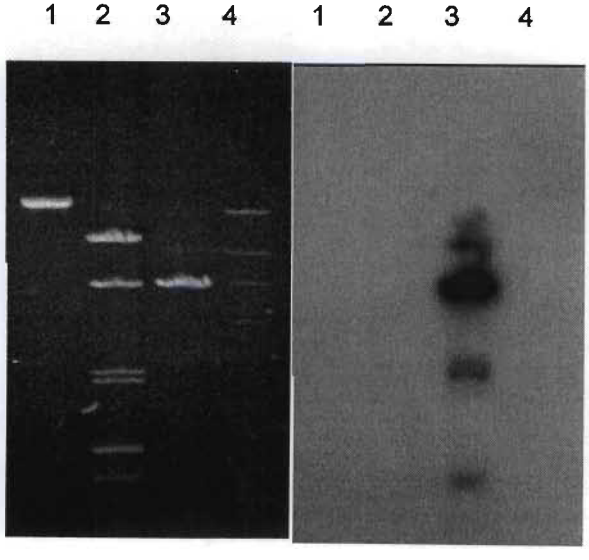


Fig. 3.6 Hybridization of pXA1 DNA to the *avrXa10* gene probe. The agarose gel is shown on the left and the autoradiograph is on the right hand side. Lane 1: uncleaved pXA1; lane 2: *Eco*RI-cleaved pXA1; lane 3: *Eco*RI-cleaved pBSavrXa10; and lane 4: λ DNA cleaved with *Hind*III.

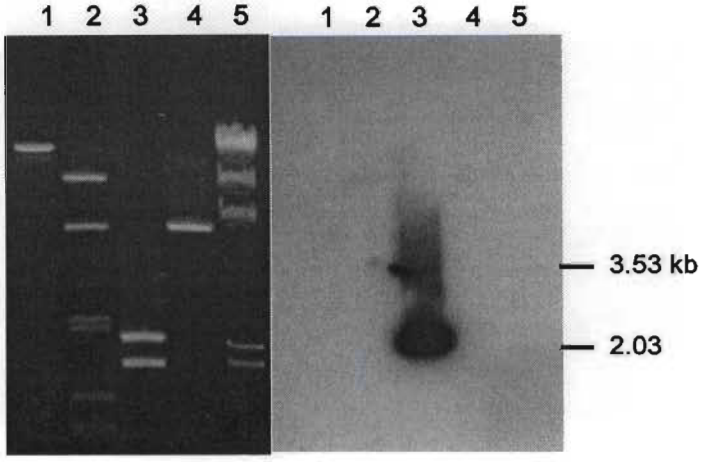


Fig. 3.7 Hybridization of pXA1 DNA to the *avrBs2* gene probe. The agarose gel is shown on the left and the autoradiograph is on the right hand side. Lane 1: uncleaved pXA1; lane 2: *Eco*RI-cleaved pXA1; lane 3: *Sac*I/*Stu*I-cleaved pBS1010; lane 4: *Eco*RI-cleaved pBSavrXa10; and lane 5: λ DNA cleaved with *Hind*III.

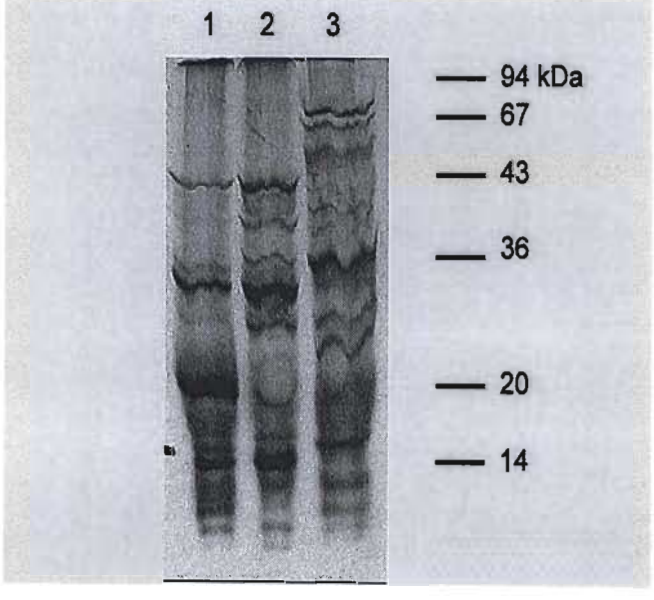


Fig. 3.8 Autoradiograph of an SDS-PAGE gel showing *in vitro* transcription and translation products of *X. albilineans* plasmid genes. Lane 1: pXA1; lane 2: pBX1; and lane 3: pXA2. The sizes of molecular weight standards (LMW Electrophoresis Calibration Kit, Pharmacia) are indicated in kDa according to their positions on the SDS-PAGE gel.

3.3.5 *In Vitro* Gene Expression

Numerous polypeptide bands were observed on SDS-PAGE gels after *in vitro* transcription and translation (Fig. 3.8). As expected, the profiles of pXA1 and pBX1, which consists of a major portion of pXA1 cloned into pBluescript, were similar. However, the intensity of the polypeptide bands differed in these two lanes. High intensity bands corresponded to polypeptides measuring approximately 42 kDa, 32 kDa, 19 kDa and 12 kDa. The profile of the polypeptides synthesized from pXA2 was markedly different except for polypeptides of approximately 36 kDa and 10 kDa (lane 3). A larger number of polypeptides were observed for pXA2. Polypeptides greater than 42 kDa also appeared to be synthesized in larger amounts when pXA2 was used as the DNA template.

3.4 DISCUSSION

This chapter describes the physical characterization of pXA1 and the search for plasmid-encoded phenotypes. Further physical characterization in the form of restriction mapping was necessary as only a rudimentary physical map was available prior to this study (Fig. 3.1). This information was also necessary for the cloning of plasmid fragments and DNA sequencing of these fragments (Chapter Four). Other methods used to search for plasmid genes involved cloning of a putative lethal region on the plasmid, probing with *avr* genes and *in vitro* expression of plasmid genes.

The miniprep method routinely used to isolate pXA1 from *X. albilineans* cells was simplified, without any noticeable decrease in the quality of the DNA. The alkaline lysis protocol of Birnboim and Doly (1979) had been modified previously (Permaul, 1994) when a precipitation step prior to RNase digestion and a second step of phenol/chloroform extractions were eliminated. In this study, RNase was routinely included in the suspension buffer and RNA

digestion occurred immediately after cells were lysed. Degradation of contaminating RNA could be performed at various stages during plasmid isolation, before and even after electrophoresis. It was most convenient to include RNase in the suspension buffer which is a feature of commercial plasmid isolation kits. The enzyme was also added to TE buffer used to dissolve DNA, in restriction reactions, and in gel loading buffers prior to electrophoresis. If RNA was found to obscure low molecular weight DNA fragments after electrophoresis, it was degraded by adding RNase directly onto these regions of the stained agarose gels. The second major modification of the isolation protocol was the complete elimination of the phenol/chloroform and chloroform extractions. These steps remove proteins remaining after NaAc precipitation. Carryover of proteins was minimised when these steps were eliminated, by mixing well several times during the ice incubation step after addition of sodium acetate. The elimination of the organic extraction steps undoubtedly resulted in protein contamination of the plasmid preparations but this did not affect enzymatic restriction of the DNA. The large scale isolation procedure was necessary for isolation of large quantities of plasmid DNA from *X. albilineans*. The miniprep procedure was inadequate for this purpose as isolation of pXA1 from 3 ml of XA86 yielded enough DNA for only a single restriction digest. This indicates that the copy number of pXA1 in XA86 is low.

The visualization of low molecular weight DNA fragments in agarose gels was improved by incorporation of EtBr in the gels. The major disadvantage of this technique is that it results in contamination of electrophoresis equipment with the mutagen. This was minimised by adding EtBr to the gel only and not to the electrophoresis buffer. However, for long electrophoretic runs it was necessary to add EtBr at the cathode end of the tank to ensure this end of the gel contained EtBr.

Amendment of the pXA1 restriction map with regard to *Eco*RI fragments was possible after analysis of DNA sequencing data. A seventh *Eco*RI fragment was predicted to occur at

position 5318 (Fig. 3.5) as an open reading frame (ORF) in this region, corresponding to a *VirB4*-like gene, was interrupted by a gap (Chapter Four). Overloading of *Eco*RI-digested pXA1 revealed that this fragment was indeed present but was not normally visible due to its small size (234 bp). In Figure 3.1, the positions of the 1.6 kb and the 1.3 kb *Eco*RI fragments were uncertain since restriction sites within these fragments were not found. In this study, the discovery of *Hind*II and *Cla*I sites on the 1.3 kb fragment together with DNA sequence data led to a change in the position of this fragment on the map. The overall size of the plasmid was increased slightly to 25 054 kb. It is possible that other *Eco*RI fragments are present at positions 17 504, 18 804, 21 604 and 24 254. However, these fragments would have to be less than 100 bp and would therefore not affect the size of the plasmid substantially.

Cloning of the entire 6.4 kb *Eco*RI fragment EIb was again unsuccessful. The region that could not be cloned was located within an 800 bp *Bam*HI/*Xba*I fragment. Two possible reasons for the inability to clone this fragment are that it contains either a gene that is lethal to *E. coli* host cells or that the pXA1 origin of replication is located on this fragment. However, a lethal gene would have to encode a small protein (<250 aa) and the BXb fragment would have to also include a promoter for this gene as cloning of this fragment was attempted in opposite orientations in the vectors pTZ19R and pUC118. It is more likely that BXb contains the pXA1 origin of replication and inability to clone the fragment stems from incompatibility with *ori* genes on the cloning vectors. This possibility can be tested by ligating BXb to cloning vectors lacking a functional origin of replication. Alternatively, plasmid vectors which cannot replicate in certain hosts can be used in cloning experiments, e.g., pK194 which cannot replicate in pseudomonads (Murillo and Keen, 1994). Another method of locating the pXA1 replicon would be hybridization with *rep* probes (Couturier *et al.*, 1988). An advantage of this method is that in addition to finding the location of the *ori* gene, the plasmid can then be classified into one of the incompatibility groups, depending on which *rep* probe binds to pXA1.

Hybridization of *avr* gene probes to pXA1 showed that pXA1 does not appear to possess *avr* genes. These genes have been found in bacterial and fungal plant pathogens and are often located on plasmids in bacteria. The two genes used in this study are representative of the *avrBs2* and *avrBs3* gene families which are conserved and widespread among members of the genus *Xanthomonas* (Hopkins *et al.*, 1992; Kearney and Staskawicz, 1990). *avr* genes are characterized by a repeat region that determines race specificity (Herbers *et al.*, 1992) but functions for these genes are unknown except for *pthA* (Yang and Gabriel, 1995). It is believed that these genes are targeted by plants containing corresponding resistance genes (Kearney and Staskawicz, 1990). It is possible that *X. albilineans* may possess *avr* genes that are located on the chromosome. The development of a hypersensitive response after inoculation into tobacco plants (Akhtar *et al.*, 1988) is probably due to a gene-for-gene interaction. It is unlikely that *avr* genes other than those from the *avrBs2* and *avrBs3* gene families are present on pXA1 as cross-hybridization was observed between the two genes used in this study, which indicates that xanthomonad *avr* genes are related.

The first evidence that pXA1 contains functional genes that encode specific proteins was provided by *in vitro* transcription and translation assays. These polypeptides most likely correspond to the ORFs identified in Chapter Four which are arranged in an operon that comprises a major portion of pXA1. Identification of specific polypeptide bands corresponding to particular genes would be easier using clones consisting of single genes only. The profiles of pXA1 and pBX1 were similar except for additional bands in the 20 - 14 kDa region. pBX1 consists of the entire pXA1 molecule except for BXb, cloned into pBluescript. It is therefore interesting that additional bands are present in this region since a protein encoded by the 800 bp BXb fragment would not be larger than approximately 21 kDa. Apart from the approximately 30 kDa ß-lactamase band, no other polypeptide bands are due to pBluescript DNA (D.E. Rawlings, pers. comm.).

Other physical characteristics of pXA1 that are yet to be determined include mode of replication and copy number of the plasmid. These characteristics of the plasmid will become important once the origin of replication or replication region of the plasmid have been identified and expression of plasmid-borne genes are studied.

CHAPTER FOUR

IDENTIFICATION OF A TRANSFER REGION ON pXA1

4.1 INTRODUCTION

Transport of most proteins out of the cytoplasm of bacteria occurs via the general secretion pathway (GSP). Exported proteins are retained partially or wholly within the cell envelope whereas secreted proteins are released extracellularly. The term excretion is reserved for the transport of non-proteinaceous compounds (Pugsley, 1993). Bacterial protein secretion occurs via four secretion systems, type I, II, III and IV. Recently, it has become apparent that conjugation systems and protein export systems share sequence similarity and structural features (Christie, 1997). This finding is reinforced by the discovery that some bacterial pathogens of plants and humans use a type III secretion system to directly "inject" proteins into eucaryotes in a process resembling conjugation (He, 1997; Hensel *et al.*, 1998; Lee, 1997). The conjugal and agrobacterial transport systems have tentatively been classified as type IV secretion systems (Christie, 1997).

The presence of a conjugal transfer region on a plasmid is largely dependant on the size of the plasmid. The most well-characterized transfer systems in Gram-negative bacteria are those of fairly large plasmids, e.g., RK2 (60 kb), F plasmid (100 kb) and the Ti plasmid (± 200 kb). In contrast, the IncN plasmids pKM101 (35.4 kb) and pCU1 (39 kb) are much smaller. These plasmids contain transfer regions that are approximately 15 kb in size for pCU1 (Thatte *et al.*, 1985) ands 12.4 kb in pKM101 (Pohlman *et al.*, 1994a). However, conjugative regions of Gram-positive bacteria are even smaller and less complex with apparent involvement of only a single protein in *Streptomyces ghanaensis* (Maas *et al.*, 1998).

The organization and relatedness of DNA sequences in these conjugation systems is

linked to the evolutionary history of the plasmids. It is therefore apparent that transfer regions on plasmids belonging to the same Inc group are similar. However, similarities also extend to plasmids belonging to different Inc groups. This is especially true for the IncP, IncN and IncW plasmids. The conjugal pili in these systems are well known for having a similar shape (pointed ends) and being brittle (Kado, 1994; Thatte *et al.*, 1985; Winans *et al.*, 1996). It is therefore not surprising that the pilus biosynthesis genes and DNA mobilization genes in these systems are co-linear and show sequence similarity (Kado, 1994, Winans *et al.*, 1996). The relatedness of the agrobacterial T-DNA transport genes and a protein export system in *B. pertussis* to these DNA transfer regions is surprising and points to an evolutionary link between these operons.

The similarity between DNA transfer and protein export systems suggests that both these processes may occur via a pilus-like structure. While this has been shown to be true for plasmid transfer, the structure of the pilus and interactions beween the various proteins involved has yet to be determined. A model proposing possible structure of the IncN pilus has been constructed (Fig. 1.7). The majority of the genes in conjugal transfer regions are responsible for biosynthesis of the conjugative pilus. In the 33 kb F-plasmid transfer region, more than half of the 36 genes are involved in construction of the F-pilus (Frost *et al.*, 1994). In the pKM101 system, only four of the 15 *tra* genes are not involved in pilus biosynthesis (Winans *et al.*, 1996). Since the pilus is attached to the cell envelope, many of the pilus proteins have signal sequences for transport across the cytoplasmic membrane. Hydrophobic membrane-spanning regions. suggest that they are located in the cytoplasmic membrane. A model showing possible location of the VirB proteins in *A. tumefaciens* has been constructed after analysis of protein sequence data (Fig. 4.1).

In this Chapter, sequencing of pXA1 is described, together with analysis of the sequencing data. The functioning of a putative transfer region was tested by mating studies.

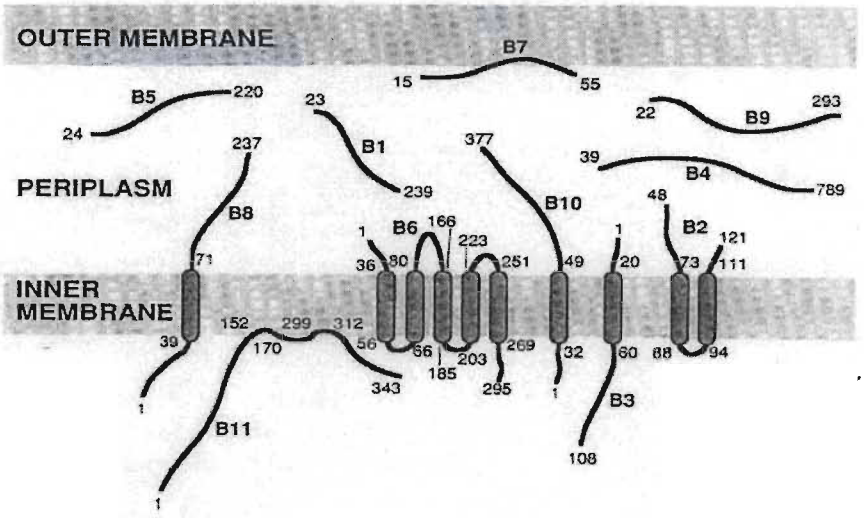


Fig. 4.1 Hypothetical model of the location and topology of the VirB proteins in the agrobacterial membrane. The location of VirB4 is uncertain (Beijersbergen *et al.*, 1994).

4.2 MATERIALS AND METHODS

4.2.1 **Plasmid Isolation**

Various plasmid isolation protocols were used depending on the intended use of the DNA and the host strains harbouring the plasmid to be isolated. DNA for manual dideoxy sequencing reactions were prepared using either the modified alkaline lysis method of Birnboim and Doly (1979) including phenol/chloroform extractions or the Wizard DNA Purification System (Promega). DNA used for automated DNA sequencing was isolated using the High Pure DNA Isolation Kit (Boehringer Mannheim). Verification of cloned DNA and conjugal transfer of plasmids were performed using the shortened alkaline lysis protocol described in Chapter Three.

4.2.2. **Construction of Recombinant Plasmids**

Recombinant plasmids containing pXA1 restriction fragments ligated to plasmid cloning vectors were constructed (Table 4.1) using five different cloning protocols. DNA molecules

having compatible sticky ends (5′ overhangs) were ligated according to the protocol described in Chapter Three. Ligation of DNA molecules having either one or two blunt ends were performed with the Rapid DNA Ligation Kit (Boehringer Mannheim) according to the manufacturer's instructions except that reactants were DNA molecules in TE buffer. Blunt-end fragments were cloned into either the *Eco*RV or *Hind*II sites of pBluescript. The plasmid pBX1 was constructed by ligation of the 24.3 kb *Bam*HI/*Xba*I fragment of pXA1 to pBluescript. Since isolation of DNA fragments larger than 10 kb using silica/glass beads results in breakage of the molecules, gel ligations were used. Slices of low melting temperature agarose gels containing the 24.3 kb fragment were melted at 65°C and 20 µl of these molten mixtures were used in ligation reactions having a total volume of 200 µl. This reduced the agarose concentration to less than 0.1% (Crouse *et al.*, 1983). Purified *Bam*HI/*Xba*I-cleaved pBluescript molecules were added to the molten agarose together with T4 ligation buffer and 1 U of T4 DNA ligase (Boehringer Mannheim). Ligation reactions were incubated overnight at room temperature. Prior to transformation, the ligation reactions were heated briefly and 50 µl was used in transformations (Section 3.2.3). Construction of pBX3 which involved cloning of the 13.583 kb *Stu*I/*Xba*I fragment of pXA1 into pBluescript required yet another cloning strategy. pBX1 was cleaved with *Sma*I and *Stu*I which resulted in the *Bam*HI/*Stu*I fragment of pXA1 being "dropped out" of this plasmid. The larger fragment from this restriction reaction was isolated from agarose gels and its blunt-ended termini were ligated to yield pBX3. When required, restriction fragments were also "polished" to produce blunt ends by incubation with 0.05 mM of each dNTP, 1 U of Klenow enzyme (Boehringer Mannheim) in 1 × SuRE/Cut Buffer H (Boehringer

TABLE 4.1. Plasmids used or constructed in this study

Plasmid	Characteristics	Source/Reference
Indigenous		
pXA1	25.133 kb plasmid in XA86	This study
pXA2	36.2 kb plasmid in GLP7	Ramdeen, 1999
Vectors		
pBluescript	2.98 kb cloning vector, Ap^R	Stratagene
pTZ19R	2.87 kb cloning vector, Ap^R	Pharmacia
pUCBM20	2.722kb cloning vector, Ap^R	BM
pUC118	3.20 kb cloning vector, Ap^R	N.T. Keen
pRK415	10.5 kb bhr vector, Ap^R	N.T. Keen
pK194	2.432 kb vector, p15a origin, Ap^R	N. T. Keen
Constructs		
pBSavrXa10	*avrXa10* gene on a 3.1 kb *Bam*HI fragment in pBS	J.E Leach
pBS1010	*avrBs2* gene on a 1.9 kb *Sac*I/*Stu*I fragment in pBS	N.T. Keen
pBX1	24.3 kb *Bam*HI/*Xba*I fragment of pXA1 in pBS	This study
pBX3	13.583 kb *Stu*I/XbaI fragment of pBX1	This study
pTX1	10.412 kb *Eco*RI fragment of pXA1 in pTZ19R	Permaul, 1994
pTX11	3.879 kb *Hind*III fragment of pTX1	This study
pBX111	1.741 kb *Eco*RV/*Hind*III fragment of pTX11	This study
pBX1111	804 bp *Hind*II fragment of pBX111	This study
pBX1112	361 bp *Hind*II/*Hind*III fragment of pBX111	This study
pBX1113	275 bp *Hind*II fragment of pBX111	This study
pBX112	1.178 *Eco*RV/*Hind*III fragment of pTX11	This study
pBX1121	638 bp *Eco*RV/*Hind*II fragment of pBX112	This study
pBX1122	540 bp *Hind*II/*Hind*III fragment of pBX112	This study
pBX113	960 bp *Eco*RV fragment of pTX11	This study
pTX12	3.5 kb *Eco*RI/*Hind*III fragment of pTX1	This study
pBX121	2.4 kb *Ssp*I/*Stu*I fragment of pTX12	This study
pBX122	729 bp *Eco*RI/*Ssp*I fragment of pTX12	This study
pBX123	1.0 kb *Hind*II fragment of pBX121	This study
pTX13	3.003 kb *Eco*RI/*Hind*III fragment of pTX1	This study

Plasmid	Characteristics	Source/Reference
pBX131	2.363 kb *Eco*RV/*Hind*III fragment of pTX13	This study
pBX132	1.598 kb *Hind*II/*Hpa*I fragment of pTX13	This study
pBX133	1.045 kb *Hind*II fragment of pTX13	This study
pBX135	348 bp *Hind*II fragment of pTX13	This study
pBX136	640 bp *Eco*RI/*Eco*RV fragment of pTX13	This study
pTX21	4.518 kb *Eco*RI/*Xba*I fragment of pXA1	This study
pTX22	800 bp *Bam*HI/*Eco*RI fragment of pXA1	Permaul, 1994
pTX23	2.155 kb *Eco*RI/*Sph*I fragment of pTX21	This study
pBX231	1.109 *Cla*I/*Sph*I fragment of pTX23	This study
pBX232	900 bp *Cla*I fragment of pTX23	This study
pBX234	1.618 kb *Eco*RI/*Hind*II fragment of pTX23	This study
pBX235	354 bp *Hind*II/*Sph*I fragment of pTX23	This study
pBX236	183 bp *Hind*II fragment of pTX23	This study
pBM25	3.098 kb *Nco*I/*Not*I fragment of pTX21 in pUCBM20	This study
pBM26	757 bp *Eco*RI/*Nco*I fragment of pTX23	This study
pBM27	1.700 *Not*I/*Sph*I fragment of pTX21	This study
pBX28	954 bp *Hin*dII fragment of pTX21	This study
pTX3	2.85 kb *Eco*RI fragment of pXA1	Permaul, 1994
pTX31	1.85 kb *Eco*RI/*Hin*dII fragment of pTX3	This study
pTX32	1.0 kb *Eco*RI/*Hin*dII fragment of pTX3	This study
pTX4	2.65 kb *Eco*RI fragment of pXA1	Permaul, 1994
pTX42	800 bp *Eco*RI fragment of pTX4	This study
pTX5	1.566 kb *Eco*RI fragment of pXA1	Permaul, 1994
pTX51	1.409 kb *Eco*RI/*Sph*I fragment of pTX5	This study
pBX53	875 bp *Eco*RI/*Hind*II fragment of pTX5	This study
pBX54	416 bp *Eco*RI/*Hind*II fragment of pTX5	This study
pBX55	328 bp *Hind*II/*Hind*II fragment of pTX5	This study
pTX6	1.27 kb *Eco*RI fragment of pXA1	Permaul, 1994
pTX61	650 bp *Eco*RI/*Hind*II fragment of pTX6	This study
pTX62	621 bp *Eco*RI/*Hind*II fragment of pTX6	This study
pBX7	234 bp *Eco*RI fragment of pBX1	This study

Mannheim) for 30 min at 37°C. Klenow polymerase was denatured at 65°C for 20 min prior to ligations into blunt-end vectors.

4.2.3 **DNA Sequencing**

pXA1 fragments cloned into plasmid vectors were sequenced to identify genes present on the plasmid. Two DNA sequencing methods were used. Initially, sequencing of terminal ends of *Eco*RI restriction fragments were performed manually in order to identify regions of interest. Once these regions were found, automated sequencing was performed to obtain the entire sequence of both DNA strands.

4.2.3.1 Manual DNA sequencing

The T7 Sequencing Kit (Pharmacia Biotech Inc.) was used for dideoxy sequencing of plasmid templates. Annealing of sequencing primers to the templates and labelling and termination reactions were performed according to the manufacturer's instructions. DNA was radioactively labelled with [α^{35}S]dATP (> 1000 Ci/mmol at 10mCi/ml - Amersham International). Chain-terminated reaction products were loaded in 6% polyacrylamide sequencing gels (per 100 ml: 42 g urea; 10 ml of 10 × TBE buffer; 15 ml of 40% acrylamide solution [19:1 acrylamide:bisacrylamide]; 42.5 ml H_2O; 60 µl TEMED; 800 µl of 10% $(NH_4)_2SO_4$). 0.35 mm gels were formed between 40 cm × 45 cm Sigmacote (Sigma)-treated glass plates. Electrophoresis was carried out at 60 W constant power in a Fisher Biotech Adjustable Sequencing Electrophoresis System (Fisher Scientific). After electrophoresis, the gel was fixed by spraying with 10% acetic acid/10% methanol. The gel was then transferred to a supporting sheet of filter paper (Whatman number 1) and dried for 45 min under a vacuum at 80°C in a gel dryer (Model 583 Gel Dryer, Bio-Rad Laboratories Inc.). Dried gels were exposed

to X-ray film (Hyperfilm-MP, Amersham International) overnight at room temperature and the films were developed according to the manufacturer's instructions.

4.2.3.2 Automated DNA sequencing

Automated sequencing of cloned pXA1 DNA was performed by Di James, Department of Microbiology, University of Cape Town using an ALFexpress Automated DNA Sequencer (Amersham Pharmacia Biotech). Chain terminating sequencing reactions (Sanger *et al.*, 1977) were performed using the Thermo Sequenase Fluorescent Labelled Primer Cycle Sequencing Kit (Amersham Pharmacia Biotech) and end-labelled fluorescent primers. All reactions were performed according to the manufacturer's instructions.

4.2.3.3 Analysis of DNA sequences

DNA sequencing data was initially processed at the DNA Sequencing Laboratory (University of Cape Town) using ALFwin version 1.1 software (Amersham Pharmacia Biotech). Raw sequencing data was then manually edited to remove DNA sequences adjacent to the sequencing primers that corresponded to the cloning vectors. Assembly of contiguous DNA sequences, identification of ORFs and translations into corresponding protein sequences were performed using Vector NTI 5.0 (InforMax Inc.) and DNAmend 1.01 (Piet Jonas and Björn Maul, University of Greifswald, Germany). Similarity searches to DNA and protein databases were performed using the BLAST 2.0 algorithm (Altschul *et al.*, 1997). The mass and hydrophobicity profiles of predicted proteins were computed using the PROTEIN.EXE Protein Sequence Editor. Alignments of protein sequences to homologs were performed using the CLUSTALW program on the GenomeNet server. These alignments were edited and analysed using Genedoc version 2.4 (Nicholas and Nicholas, 1997).

4.2.4 Isolation of Antibiotic-Resistant Strains

Bacterial strains resistant to selected antibiotics were isolated in order to study the transfer of pBX1 via conjugation. *X. albilineans* strains XA86C, TS and GLP7C were grown in the presence of 75 µg/ml streptomycin and 50 µg/ml rifampicin until colonies resistant to these antibiotics formed. These mutants were designated XA86CS, XA86CR, TSS, and GLP7CS (Table 2.1). A *P. syringae* pv. *syringae* mutant, resistant to 50 µg/ml streptomycin was also isolated.

4.2.5 Creation of Donor Strains

E. coli and *X. albilineans* strains containing pBX1 were created to act as donor strains in mating experiments. Transformation of *E. coli* cells with pBX1 was accomplished using the procedure described in Section 3.2.3. pBX1 was delivered into donor strains by electroporation using a Gene Pulser (BioRad). Electrocompetent *X. albilineans* cells were prepared in 10% glycerol according to the BioRad Users Manual. 40 µl of these cells were used for electroporation after addition of approximately 50 ng of pBX1. A voltage of 1.35 kV was delivered to cells in 0.1 cm cuvettes (BioRad) using a resistance of 200 Ω and capacitance of 25 µF. Survivors were grown in Wilbrink's broth for 2 h before dilutions were plated on plates containing 200 µg/ml ampicillin. Verification of pBX1-containing cells was performed by plasmid isolation.

4.2.6 Bacterial Conjugation

Bacterial matings were conducted to study the transfer of pBX1, a derivative of pXA1, between *E. coli* strains and between *X. albilineans* strains. Transfer of the plasmid between *E. coli* strains was tested using bi-parental and tri-parental mating procedures. Self-transmissibility

of pBX1 was tested using DH5α (Nal^R) as the donor strain and *E. coli* S17-1 (Sm^R) as the recipient strain. Mobilization of pXA1 was tested using S17-1 as the donor strain as it contains a derivative of RK2 (IncP) integrated into its chromosome, and *E. coli* CSH56 (Nal^R) as the recipient strain. An IncW helper plasmid, pSa, was also used to test mobilization of pBX1. In this case tri-parental matings were performed, using a 1:1:1 ratio, with DH5α (pBX1) as the donor strain, HB101(pSa) providing helper functions and CSH56 as the recipient

Self-transmissibility of pBX1 in *X. albilineans* strains was tested using XA86CS1 (Sm^R) as the donor strain and XA86CR (Rif^R) as the recipient strain. All mating experiments were performed using broth cultures of participating strains grown in the presence of the appropriate antibiotics. The cultures were mixed together in a 1:5 (donor:recipient) ratio. The cultures were pelleted and resuspended in 50 µl of Wilbrink's broth or YT broth. The suspensions were spotted onto non-selective agar plates and incubated overnight at the appropriate temperature. Agar plugs containing the mating aggregates were excised from the plates and placed in McCartney bottles containing 5 ml of 0.85% saline. The bottles were vortexed vigorously for 1 min to wash the cells off the agar plugs and disrupt the mating pairs. One ml of the cell suspensions were washed three times in saline to remove extracellular ß-lactamases. Serial dilutions of the washed cells were plated on agar plates containing antibiotics appropriate for selection of donors and transconjugants. When selecting for transconjugant growth, the ampicillin concentration was doubled in order to prevent growth of satellite colonies. Plasmid isolations from 10 random *X. albilineans* transconjugants were tested for pBX1 presence using the shortened alkaline lysis protocol described in Section 3.2.1. Frequency of transfer of pBX1 was calculated by dividing the number of transconjugants/ml by the number of donors/ml to obtain the frequency of transfer/donor cell.

4.3 RESULTS

4.3.1 Plasmid Isolations and DNA cloning

The High Pure Plasmid DNA Isolation Kit (Boehringer Mannheim) yielded a satisfactory amount of plasmid DNA that was adequate for automated DNA sequencing. The majority of DNA isolations yielded DNA having an OD_{260}/OD_{280} ratio greater than 1.8, which indicates minimal protein contamination. A total of 48 clones carrying pXA1 restriction fragments were constructed (Table 4.1), 42 in this study and six clones in a previous study. These fragments ranged in size from 24.3 kb to approximately 125 bp.

4.3.2 DNA Sequencing

Manual DNA sequencing routinely produced approximately 300 bp of useful sequencing data after analysis of the three separate runs on the gel. In contrast, automated sequencing often yielded more than a 1000 bp of sequencing data. Approximately 600 - 800 bp of this data could be reliably used, as specified by the ALFwin software.

4.3.2.1 Identification of a transfer region

Analysis of the DNA sequence between coordinates 2218 and 17545 revealed the presence of a transfer region on pXA1. Sixteen ORFs have been identified in this region and the majority of these ORFs share sequence similarity with genes identified in similar transfer regions in other bacteria. The first 12 ORFs in this region share sequence similarity as well as gene order and gene size with the *tra* operon of pKM101, the *pilW* operon of R388, the *virB* operon of the Ti plasmids in *A. tumefaciens* and the *ptl* operon in *B. pertussis*. ORF1 to ORF11 were named such that they correspond to the *virB* genes *virB1* to *virB11* as they have the same linear order

and equivalent sizes. Eight of these ORFs show sequence similarity to their corresponding homologs *viz., virB2, virB4, virB5, virB6, virb8, virB9, virB10* and *virB11*. An additional ORF occurs between ORF5 and ORF6 which is not present in the *virB* operon. This ORF which corresponds in size and position to the *eex* gene in pKM101 was designated ORF12. ORF13 is a homolog of the *virD4* gene of the Ti plasmids. ORF14 shares sequence similarity with bacterial transglycosylases. ORFklcA and ORFkorC are homologs of the *klcA* and *korC* genes, respectively, of plasmid RK2. The DNA sequence of this transfer region is shown in Appendix One. A putative promoter region was identified immediately upstream of ORF1 which corresponded exactly with the -35 consensus sequence TTGACA. Multiple alignment of ORFs to protein sequences of homologs identified in other bacteria are shown in Appendix Two while alignment of each ORF to its most closely related homolog is shown in the following section.

4.3.3 **Analysis of ORFs**

4.3.3.1 ORF1

ORF1 is a 375 bp gene that lies immediately downstream of a putative promoter region. The predicted protein product has 124 aa and a mass of 14.362 kDa. The translated DNA sequence did not show significant homology to any protein sequence in the databases searched using the BLAST algorithm. Alignment of the protein sequence with TraL (Fig. 4.2) and VirB1 (Appendix Two) showed very few identical residues. The alignment score (CLUSTALW 1.7 program) between ORF1 and TraL was higher than that beween ORF1 and VirB1 (Appendix Two). ORF1 is approximately half the size of VirB1 and TraL and the weak similarity to ORF1 exists at the N-terminal half of these proteins.

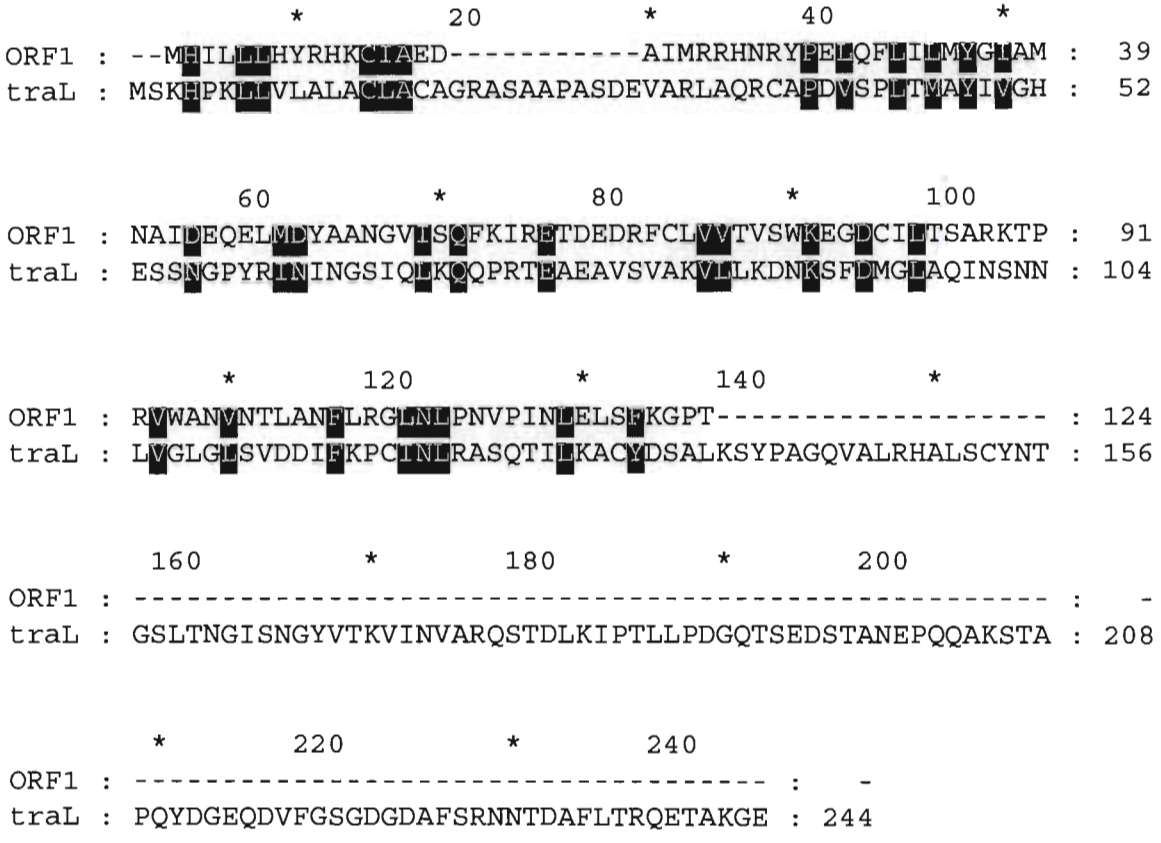


Fig. 4.2 Alignment of ORF1 to the TraL protein (244 aa) of pKM101. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

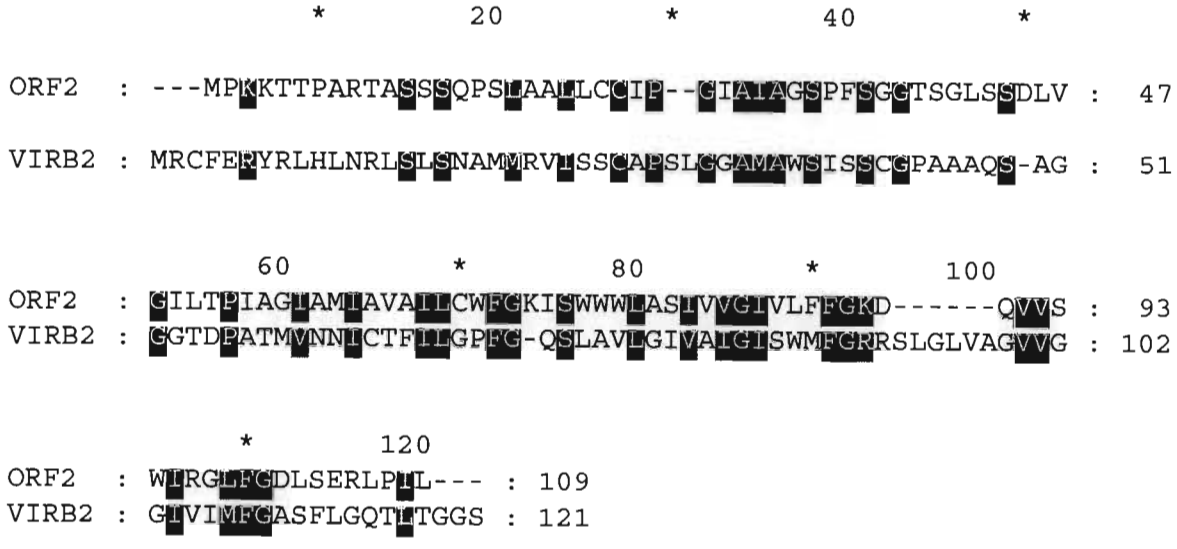


Fig. 4.3 Alignment of ORF2 to the VirB2 protein (121 aa) of *A. tumefaciens*. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.2 ORF2

The proposed gene is 330 bp in size and the translated protein has 109 aa (13.046 kDa). The results of BLAST similarity searches show that homology to VirB2 was detected, although the quality score of 31 was low. BLAST searches also revealed that ORF2 was similar to YggA in *Aeromomas salmonicida* and a putative coding region in *Helicobacter pylori*. Lower similarity was found to a transport protein in *Mycobacterium tuberculosis* and a predicted transport protein in *Salmonella typhimurium*. Alignment of ORF2 to VirB2 is shown in Fig. 4.3.

4.3.3.3 ORF3

BLAST searches did not identify proteins homologous to ORF3. This ORF encodes a 121 aa (13.932 kDa) protein which is slightly larger than proteins in the identical position of other transfer operons. Alignment of ORF3 to TrwL, PtlB, VirB3 and TraA (Appendix 2) showed that ORF3 is most similar to VirB3. The alignment between these two sequences is shown in Fig. 4.4.

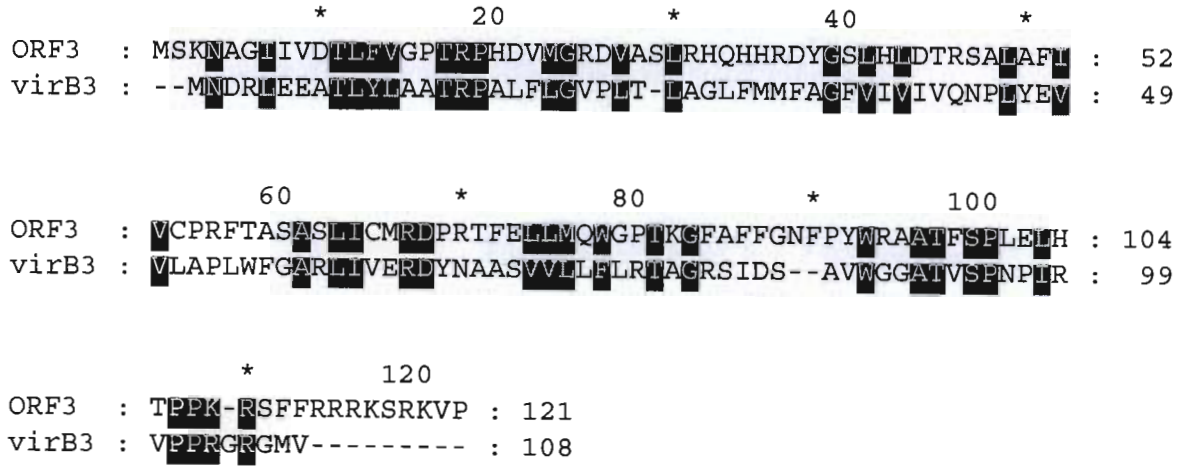


Fig. 4.4 Alignment of ORF3 to the VirB3 protein (108 aa) of *A. tumefaciens*. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

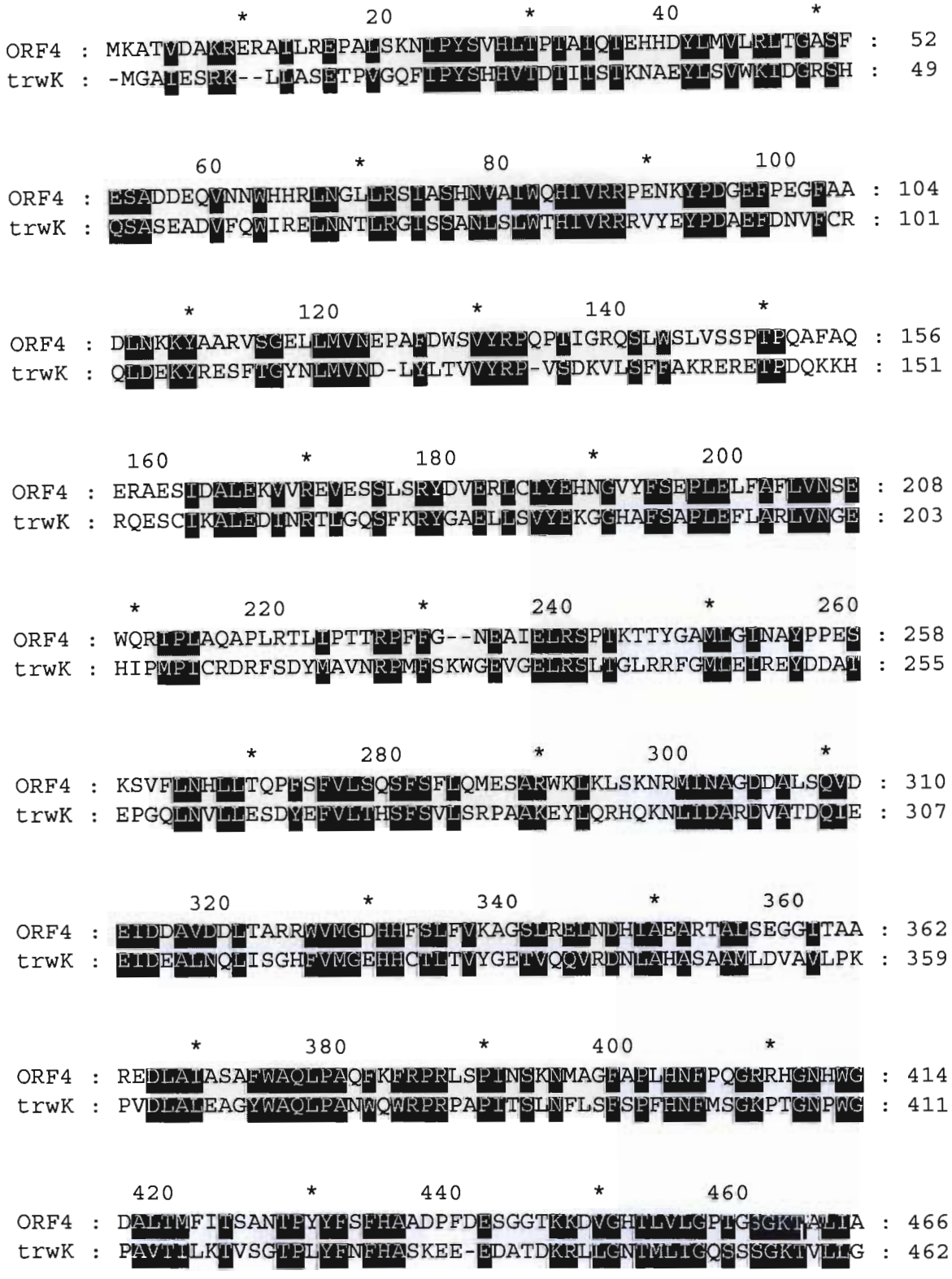
4.3.3.4 ORF4

ORF4 is the largest ORF in the pXA1 transfer region. The predicted gene is 2442 bp in size and is located on three of the seven *Eco*RI fragments of the plasmid. The corresponding protein comprises 813 aa and has a molecular weight of 92.158 kDa. BLAST searches revealed that ORF4 has appreciable homology to various proteins found in bacterial transfer operons. Highest sequence similarity occurred beween ORF4 and TrwK and the alignment between these two sequences is shown in Fig. 4.5. The BLAST quality score of 528 was the highest for all of the pXA1 ORFs. Besides TrwK, homology to VirB4, TraB and PtlC was detected and quality scores ranged from 474 to 376. Among these homologs, the alignment score between ORF4 and TrwK was equal to that of TraB and TrwK (36) and only slightly less than the highest score (37) between PtlC and TrwK. Other homologs were identified in various bacteria including CagE/PicB in *H. pylori*, TrbE conjugal proteins or RK2, *A. tumefaciens R. leguminosarum* and *Enterobacter aerogenes* and a sex pilus assembly and synthesis protein in the Archaeon *Synechocystis* sp. A possible function for ORF4 was identified after a Walker box (nucleotide-binding motif) was found at position 456. This sequence GPTGSGKT conformed exactly with the consensus sequence GXXGXGKT (X = any aa) of ATPases (Shirasu *et al.*, 1994; Lessl *et al.*, 1992). Multiple alignments reveal that this motif is conserved in all homologs (Appendix Two), except for a slight modification in TrwK.

4.3.3.5 ORF5

BLAST searches showed that the 236 aa product of this ORF was homologous to TrwJ (R388), TraC(pKM101) and the EexA/TrbJ protein of the IncP plasmids R18/RK2. In addition, similarity to a yeast nuclear pore protein and the HlyD secretion proteins of *E. coli* was also discovered. Multiple alignment to three homologs showed that the highest alignment score (23)

was between ORF5 and TrwJ. Alignment of ORF4 to TrwJ (Fig. 4.6) indicates that identical and similar aa residues occur throughout the protein sequences.



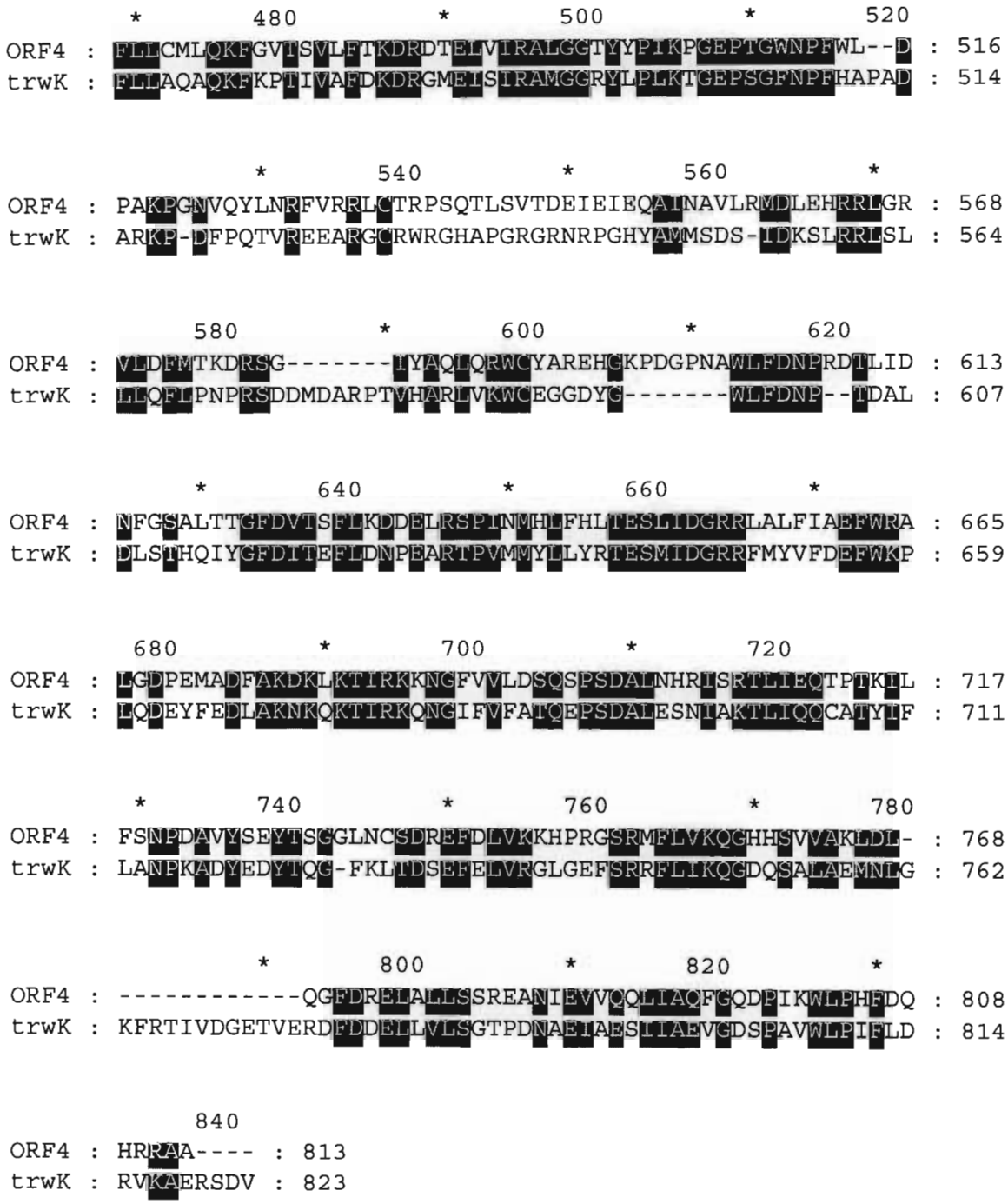


Fig. 4.5 Alignment of ORF4 to the TrwK protein of the plasmid R388. The Walker Box occurs at positions 458 - 462 in the ORF4 sequence. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

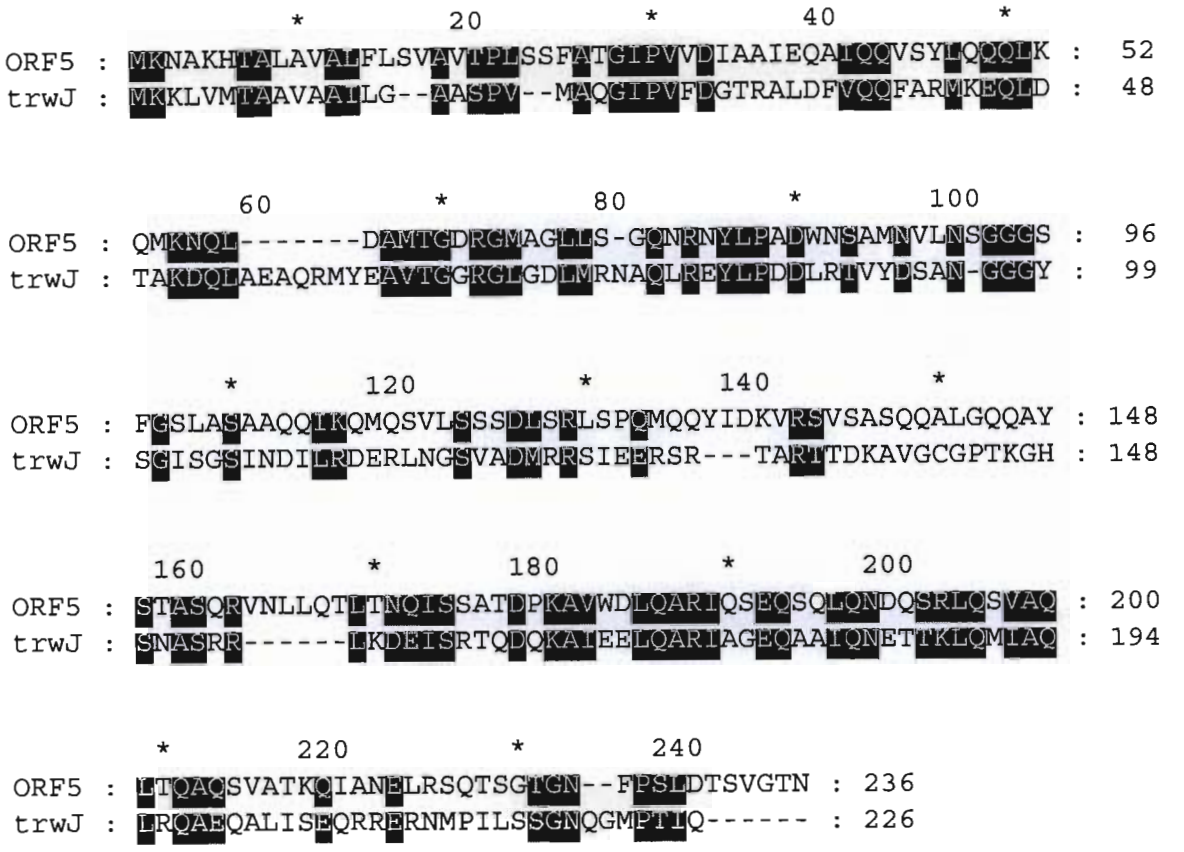


Fig. 4.6 Alignment of ORF5 to the TrwK protein of plasmid R388. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.6 ORF6

A high quality score of 129 for BLAST searches corresponded to similarity with the TrwI protein of R388. These two protein sequences differed in length by just two aa residues (340 aa vs 342 aa). Alignment of these sequences is shown in Fig. 4.7. The similarity searches also detected homology to VirB6 proteins in *A. tumefaciens* and *Rickettsia prowazekii*, TraD and PtlD. The size of the homolog in *A. tumefaciens* is smaller (295 aa) whereas PtlD is larger (463 aa) than the 36.917 kDa ORF6-encoded protein. Similarity was also detected with TrbL of RK2, a TrbL homolog in *E. aerogenes* and a putative gene in *H. pylori*.

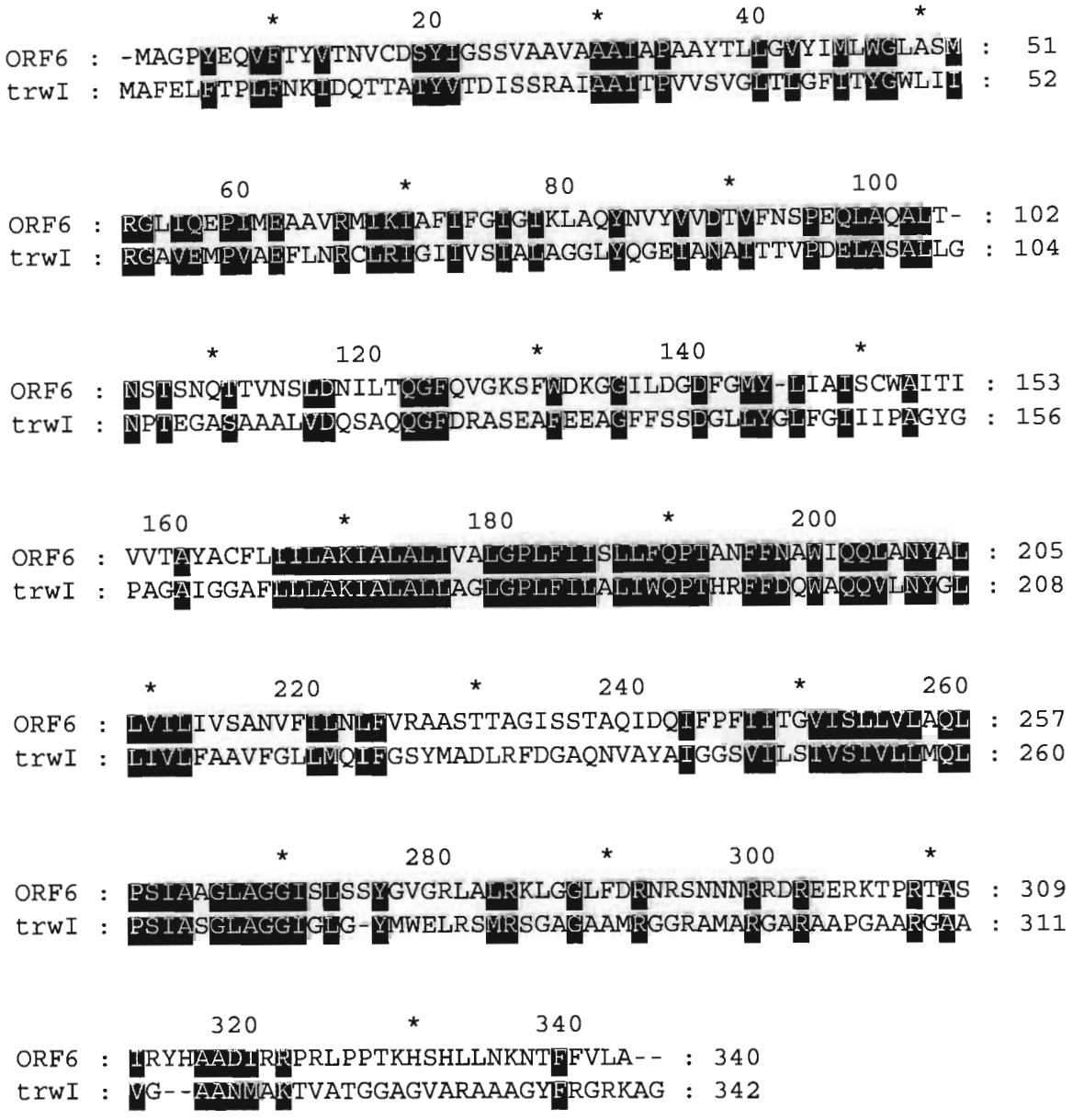


Fig. 4.7 Alignment of ORF6 to the TrwI protein of plasmid R388. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.7 ORF7

BLAST searches failed to identify homologs to this small 52 aa (5.241 kDa) predicted protein. Very low homology (scores between 25 to 27) was found to a transport protein in *M. tuberculosis*, an integral membrane protein, a Na^+/H^+ antiporter, a general secretion pathway lipoprotein in *X. campestris* pv. *campestris* and lipoproteins in *Neisseria meningitidis* and *Sinorhizobium meliloti*. This transport function suggested that ORF7 may be a

VirB7/TrwH/TraN homolog. This was confirmed when a multiple alignment with these homologs was performed and became obvious when ORF7 was aligned with its most closely homolog, TrwH (Fig. 4.8). ORF7 was found to be more closely related to TrwH and TraN than VirB7 as the alignment scores between ORF7 and TrwH/TraN were higher than that between VirB7 and TrwH/TraN.

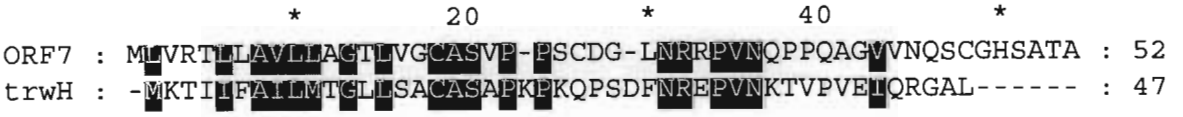
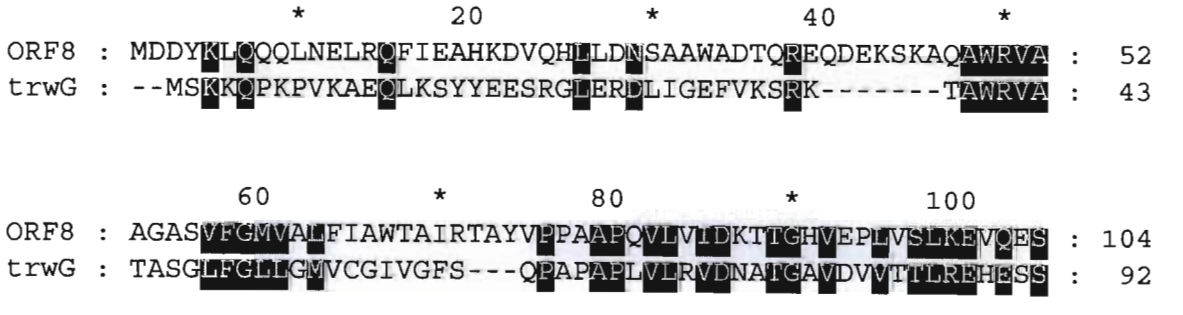


Fig. 4.8 Alignment of ORF7 to the TrwH protein of plasmid R388. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.8 ORF8

The homolog showing the highest quality score after ORF8 BLAST searches, was the R388 protein TrwG. Homology was detected to TraE and VirB8 which are similar in size to TrwG. Surprisingly, homology was not detected to the largest related protein PtlE (276 aa) which is 8 aa larger than the ORF8 protein (29.820 kDa). However, CLUSTALW alignments confirmed that ORF8 and PtlE are the most distantly related homologs (alignment score = 14). Alignment with TrwG (alignment score = 19) is shown in Fig. 4.9.



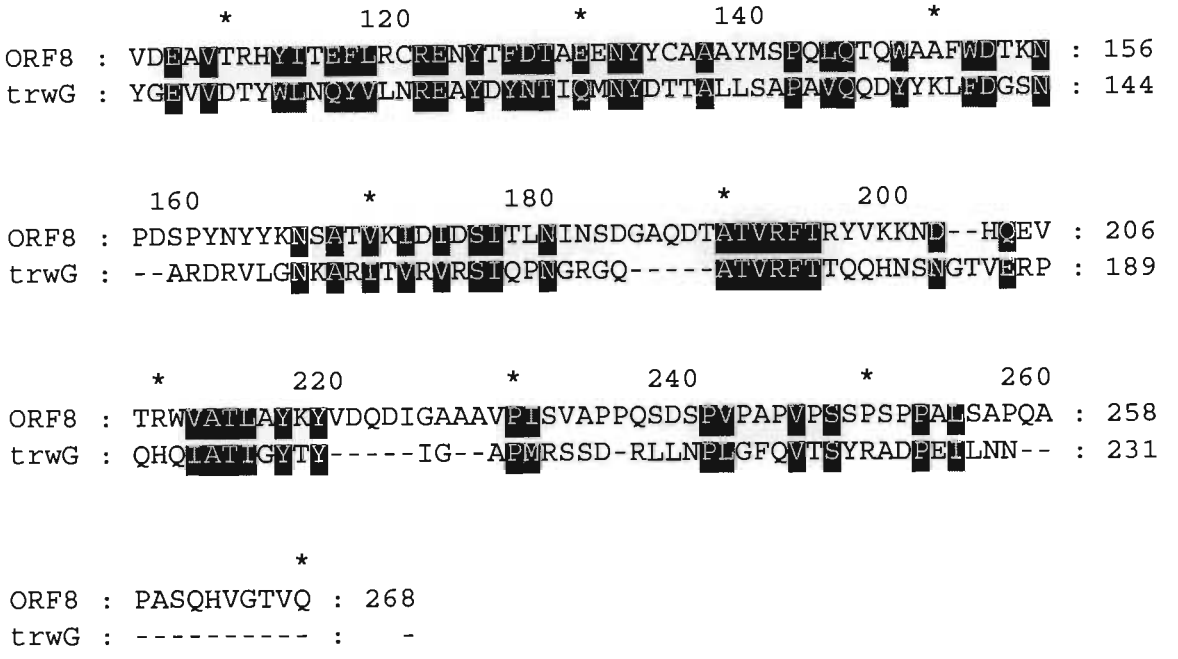


Fig. 4.9 Alignment of ORF8 to the TrwG protein of plasmid R388. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.9 ORF9

BLAST similarity searches revealed that the 266 aa ORF9 protein (29.528 kDa) is homologous to a family of similar proteins. The highest quality score of 132 was achieved for the match between ORF9 and TrwF. These two proteins differ in size by just a single amino acid. The alignment between these two proteins shows a high proportion of identical or similar amino acids (Fig. 4.10). BLAST searches also revealed homology of ORF9 to TraO, VirB9 and PtlF. Multiple alignment of these proteins showed that highest relatedness (alignment score = 29) existed between ORF9 and TrwF and between PtlF and TrwF. Homology was also detected to TrbG proteins of IncP plasmids. Homology was greater to TrbG in conjugal transfer regions of Ti and Ri plasmids in *Agrobacterium* and in a pSym of *R. leguminosarum*, compared to TrbG of plasmid RK2. ORF9 is also homologous to the Cag8 protein located in a *H. pylori*.

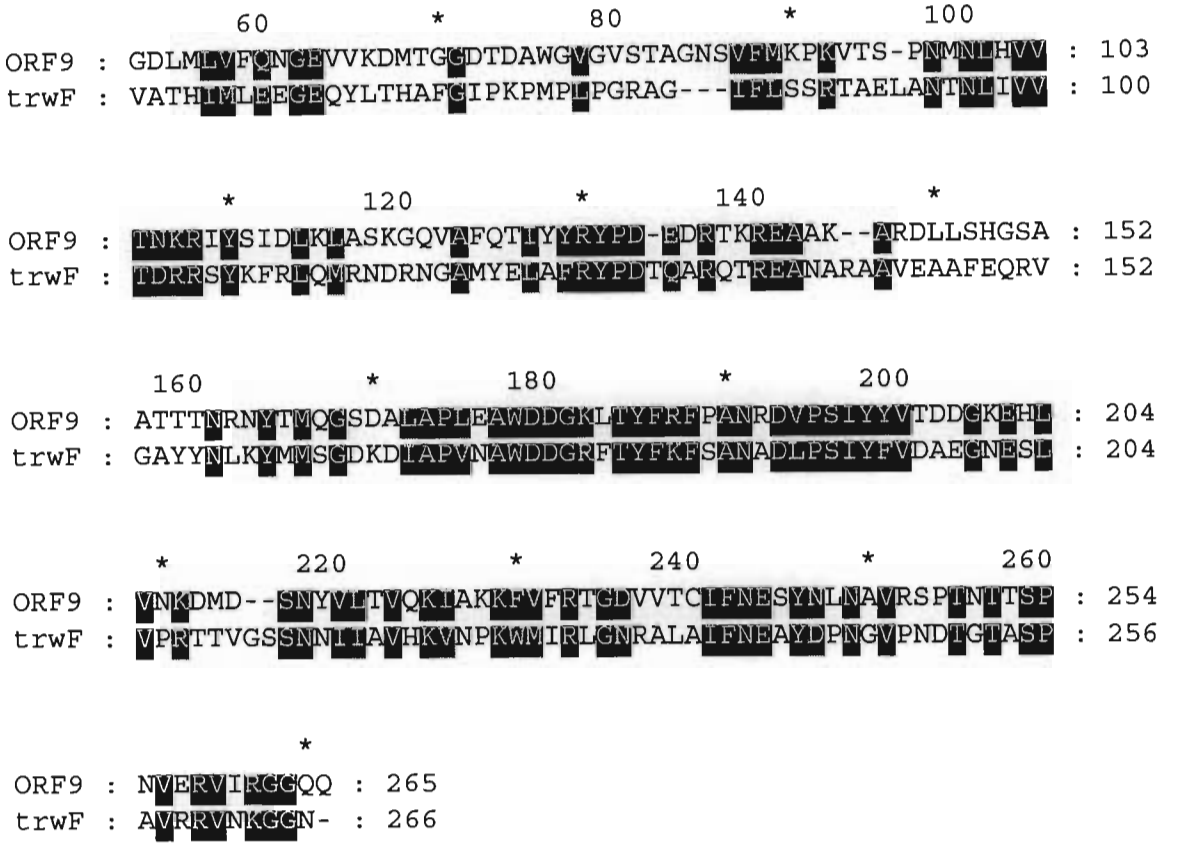


Fig. 4.10 Alignment of ORF9 to the TrwF protein of plasmid R388. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.10 ORF10

BLAST searches with the ORF10 protein sequence (393 aa) produced results similar to that with ORF9. Highest similarity occurred with the TrwE protein (395 aa). The quality score for this match was 187. Matches with quality scores ranging from 120 to 158 were achieved for the homologs PtlG, TraF and VirB10. Multiple alignments of these proteins showed that identical residues were prevalent especially at the carboxyl ends of the proteins. Alignment of ORF9 to TrwE also showed the same pattern but more identical residues were apparent at the amino terminals of the proteins (Fig. 4.11). Homology to TrbI proteins of the Ti and Ri plasmid conjugal regions and plasmid RK2 were also evident. A homologous protein in *H. pylori*, Cag7, was also detected in BLAST searches.

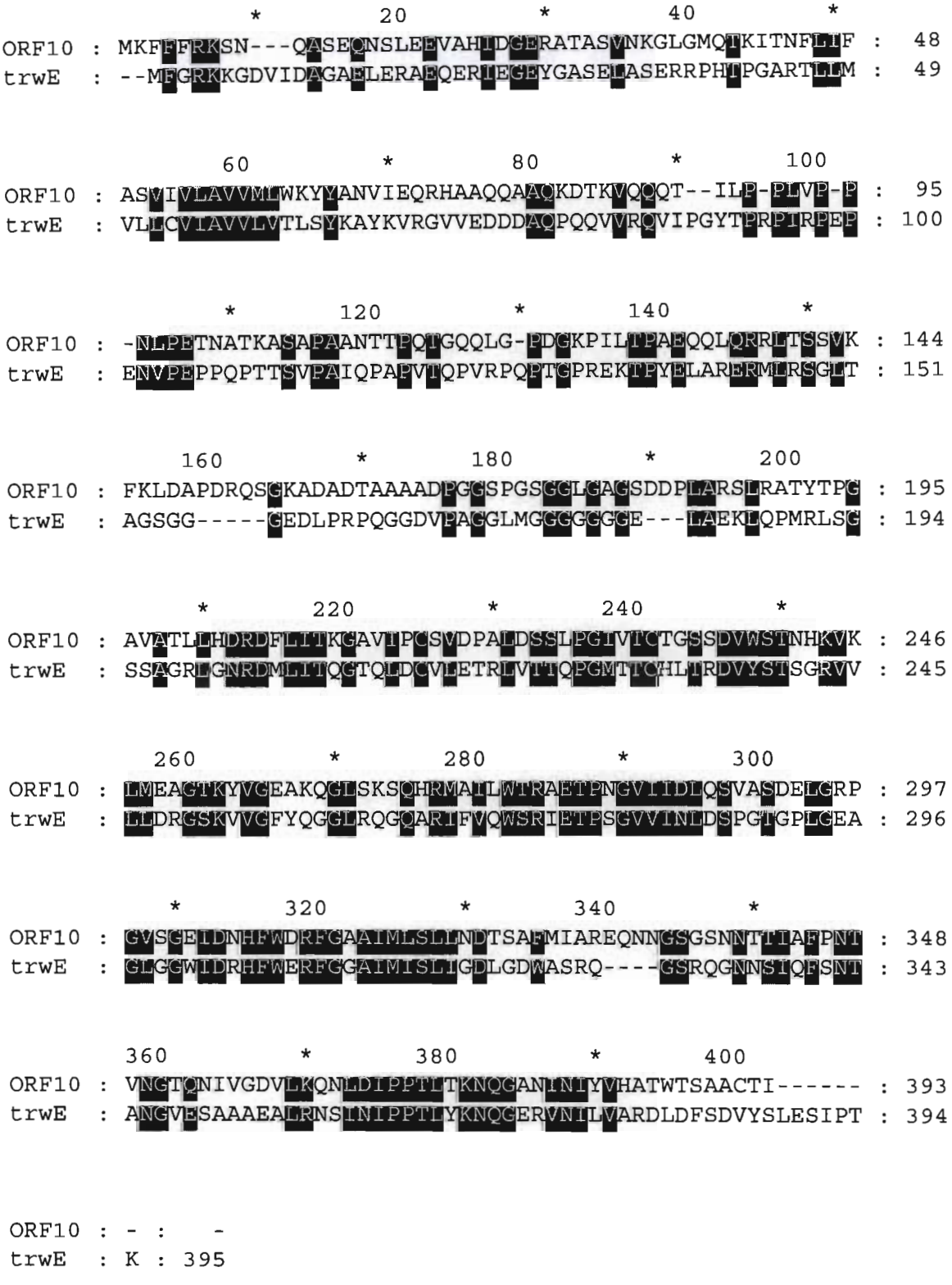


Fig. 4.11 Alignment of ORF10 to the TrwE protein of plasmid R388. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.11 ORF11

In contrast to the previous ORFs, the protein displaying highest homology to ORF11 was not from the *pilW, virB, tra* or *ptl* operons described previously. A virB11 protein from *R. prowazekii* matched to ORF11 with a quality score of 216 after BLAST searches. These searches revealed that ORF11 is a member of a family of homologous proteins that is widespread among both Gram-negative and Gram-positive bacteria as well as in Archeae. ORF11 was most closely related to VirB11, PtlH, TrwD and TraG (quality scores between 216 and 142). Alignment of ORF11 to the VirB11 protein of *R. prowazekii* is shown in Fig. 4.12. In contrast with other ORFs in the pXA1 transfer region, ORF11 had appreciable homology with proteins in the DNA transfer regions of *H. pylori*, the TrbB conjugal transfer proteins of *A. tumefaciens*, *R. leguminosarum* and plasmid RK2 (quality scores between 142 and 99). BLAST searches also revealed that ORF11 was homologous to secretory protein kinases in *Chlorobium* and *Methanobacterium*, type II and type III secretion proteins in *E. coli* and various Archeae, type IV pilus assembly proteins in eubacteria and Archeae, fimbrial and flagella assembly proteins and a VirB homolog in *Sulfolobus solfataricus*. An interesting homolog to ORF11 was the competence protein in *Streptococcus pneumoniae*. A common motif present in ORF11 and all homologs identified, was the Walker Box motif which also occurred in ORF4.

```
                      *         20         *         40
ORF11     : MTGEKSTLEIHLEPLRPFLDDPANNEIVINNPLVVWTESRGQWVTHDV :  48
rpVIRB11  : MNEEFAALETFLLPFKNLFAEEGINEIMVNKPGEAWVEKRGDIYSKQI :  48

            *         60         *         80         *
ORF11     : PSITSEWCDELSKLVANFSDQKIDVEHPMIGSTLPTRERIQIVIPPVV :  96
rpVIRB11  : PELDSDHLLALGRLVAQSTEQMLSEEKPLLSATLPNGYRIQIVFPPAC :  96
             100         *        120         *        140
ORF11     : KT--VSVTIRRPSADVMTFDEIYERGTFDDIRCEQSSRLDAEEREAIE : 142
rpVIRB11  : EIGQIIYSIRKPSGMNLTLDEYAQMGAFDNIATESLVDEDAD------ : 138
```

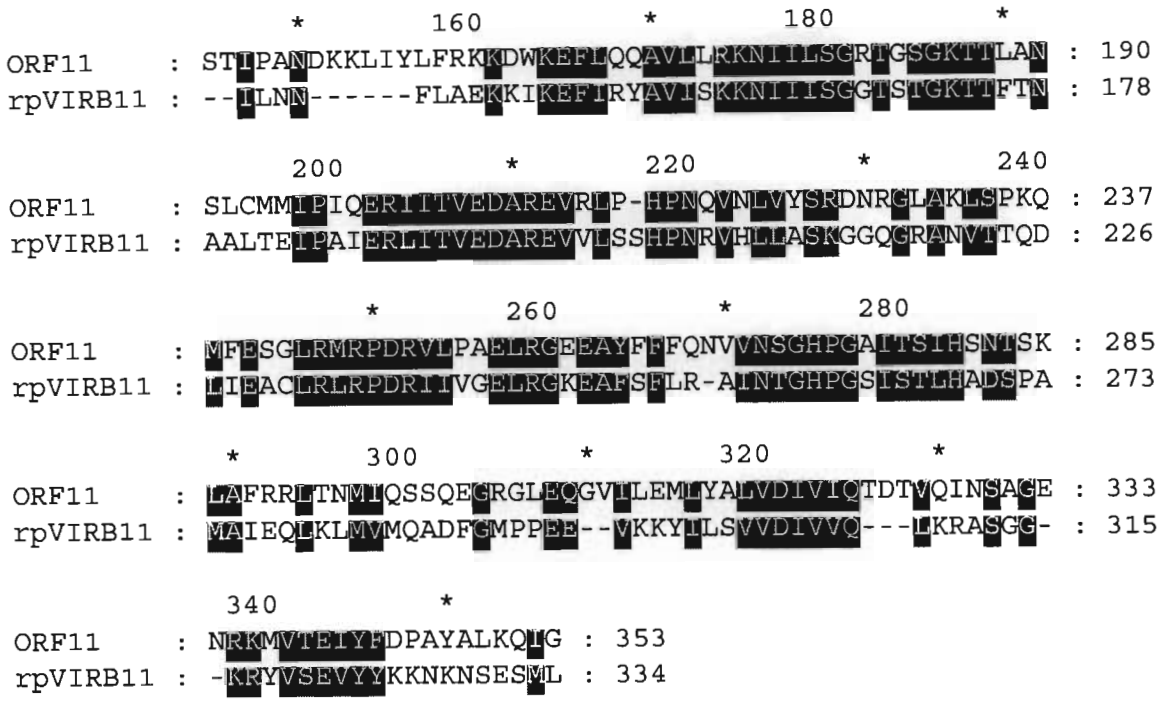


Fig. 4.12 Alignment of ORF11 to the VirB11 protein of *R. prowazekii*. The Walker Box motif occurs in positions 179 - 186 of the ORF11 sequence. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.12 ORF12

ORF12 corresponds to a 99 aa (10.473 kDa) protein that lies between ORF5 and ORF6. BLAST searches showed that this ORF is related to the *eex* gene of the R388 *pilW* operon (*eexW*). The quality score for this match (39) is not high but alignment of these two sequences (Fig. 4.13) confirms that these proteins are related (alignment score = 27). The alignment score between EexW and the pKM101 Eex protein is only slightly higher (score = 29).

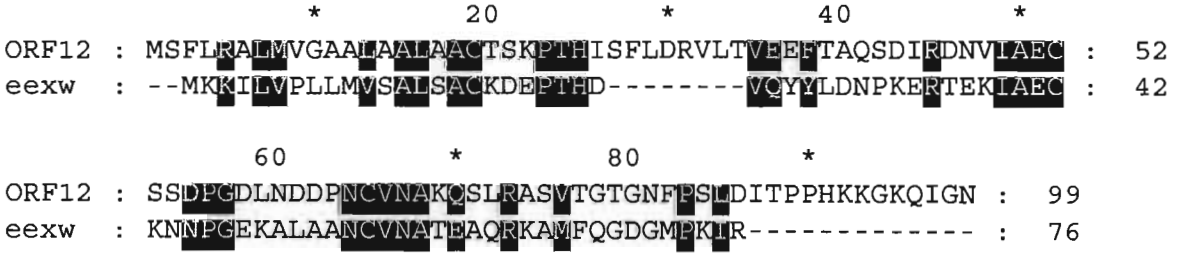


Fig. 4.13 Alignment of ORF12 to the EexW protein of plasmid R388. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.13 ORF13

ORF13 is the second largest ORF in the pXA1 transfer region. The predicted protein is 639 aa in size and has a molecular mass of 72.116 kDa. BLAST searches revealed that the proteins most closely related to ORF13 is a VirD4 protein in *R. prowazekii* (quality score = 279) and a plasmid encoded TaxB in *E. coli* (quality score = 196). ORF13 is homologous to various VirD4 proteins in *A. tumefaciens* and *A. rhizogenes*. Even more TraG homologs were identified. These conjugal DNA proteins are present in the IncP plasmids RK2 and R751, in Ti and Ri plasmids of *Agrobacterium*, in the megaplasmid pSym of *R. leguminosarum* and several *E. coli* plasmids. ORF13 is also homologous to a conjugal protein, TrsK, in the Gram-positive bacteria *Lactococcus lactis* and *Staphylococcus aureus* (quality score = 106 and 103, respectively) and Cag5 in the pathogenicity island of *H. pylori* (quality score = 145). A nucleotide-binding motif common to most homologs is present at position 139 on the ORF9 protein sequence. The consensus sequence is AXXRXGKG. Alignment of ORF13 to VirD4 is shown in Fig. 4.14.

```
                  *         20          *         40          *
ORF13  :  MILTDIQSKRVRSVVISGLLVCGLILGLYLAGYFFLWKVKLKPYSATPLTI  :  51
VIRD4  :  -----MNSSKTTPQRLAVSIVCSLAAGFCAASLYVTFRHGFNGEAMMTFSV  :  46

             60          *         80          *         100
ORF13  :  IDYWAYYSDNPQLRKLMKYCLGFGFILSYGLVAMLFMPVR----RALHGDA  :  98
VIRD4  :  FAFWYETPLYMGHATPVFYC-GLAIVVSTSIVVLLSQLIISFRNHEHHGTA  :  96

             *         120         *         140         *
ORF13  :  RFAKDKEVRDADLLGEHGLILG---------KWGDRFIMLAGQLGAICAAP  : 140
VIRD4  :  RWAGFGEMRHAGYLQRYNRIKGPIFGKTCGPRWFGSYLTNGEQPHSLVVAP  : 147

          160         *         180         *         200
ORF13  :  PRTGKGAGLVQPNMLNWLQSVVLLDVRQESYRLTSGFRKMFSD-VFLFNPV  : 190
VIRD4  :  TRAGKGVGVVIPTLLTFKGSVIALDVKGELFELTSRARKAGRDAVFKFSPL  : 198
```

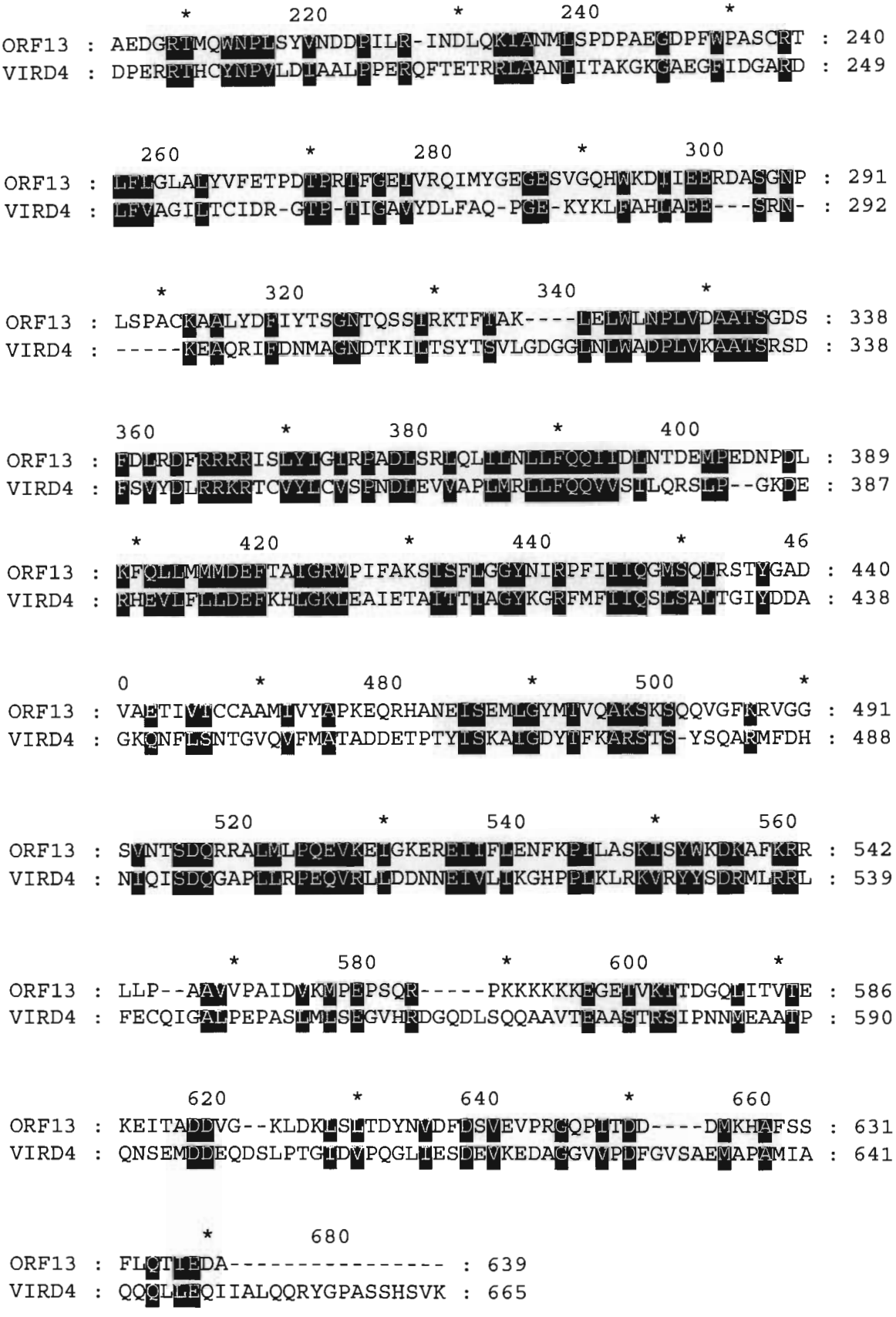


Fig. 4.14 Alignment of ORF13 to the VirD4 protein of *A. tumefaciens*. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.14 ORF14

ORF14 encodes a 209 aa protein that was found to be homologous to lytic transglycosylases in bacteria and phages. The most closely related homolog was an 181 aa protein in *Bacillus subtilis* (quality score = 61). Homology to a murein-degrading transglycosylase in *Sphingomonas aromaticivorans* was also observed. In. *E. coli*, homology towards both soluble and membrane-bound transglycosylases was detected. In addition, homologs in phage PRD1 and phage SPBc2 were also identified. Alignment of ORF14 to the *B. subtilis* transglycosylase is shown in Fig. 2.15. Alignment of ORF14 to TraL and VirB1 (Appendix Two) showed that ORF14 was related to these proteins and not ORF1.

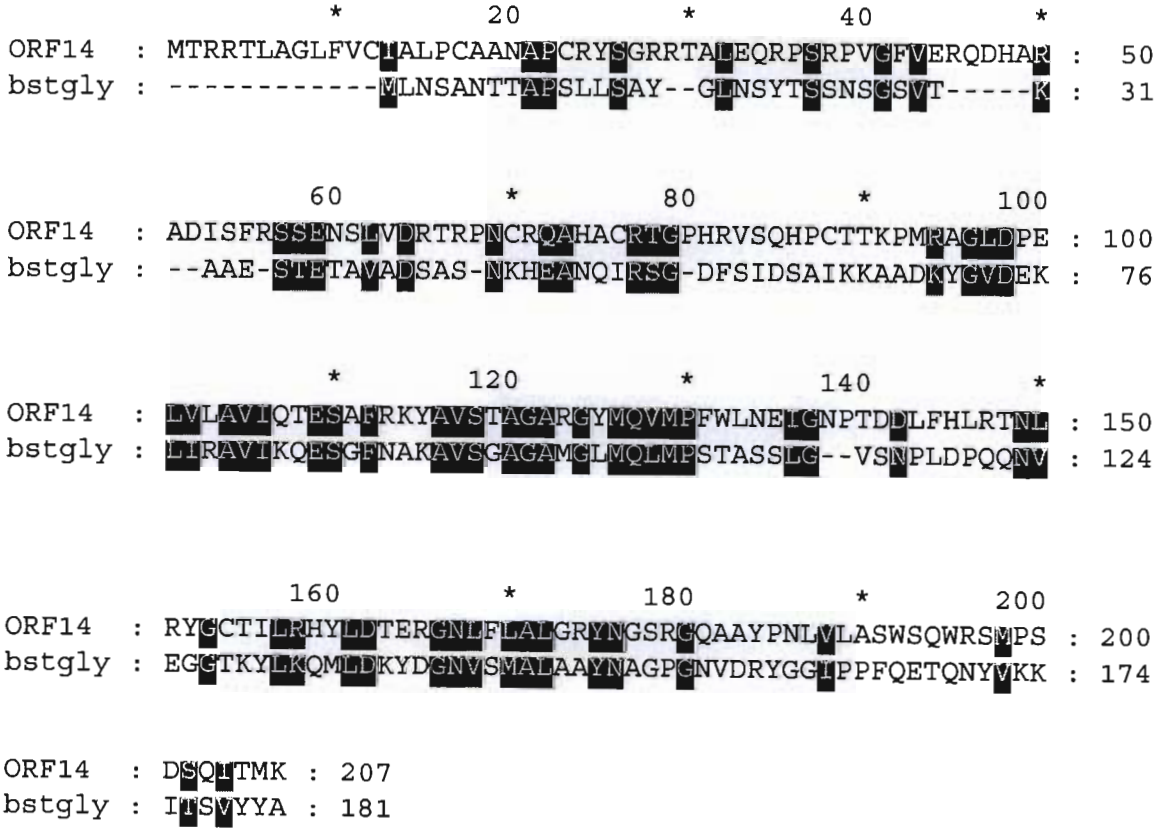


Fig. 4.15 Alignment of ORF14 to a transglycosylase enzyme in *Bacillus subtilis*. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.15 ORFklca

The ORFklca protein product (146 aa) is homologous to the KlcA antirestriction proteins in *E. coli* and *E. aerogenes*. The size of this protein product is similar to the homologs (142- 148 aa). The plasmid RK2 Klca protein from the *kil/kor* regulon is 146 aa in length. Quality scores for these BLAST matches range from 89 - 96. Alignment of ORFklca to the KlcA protein from *E. coli* and *E. aerogenes* shows conserved region in the protein sequences (Fig. 2.16).

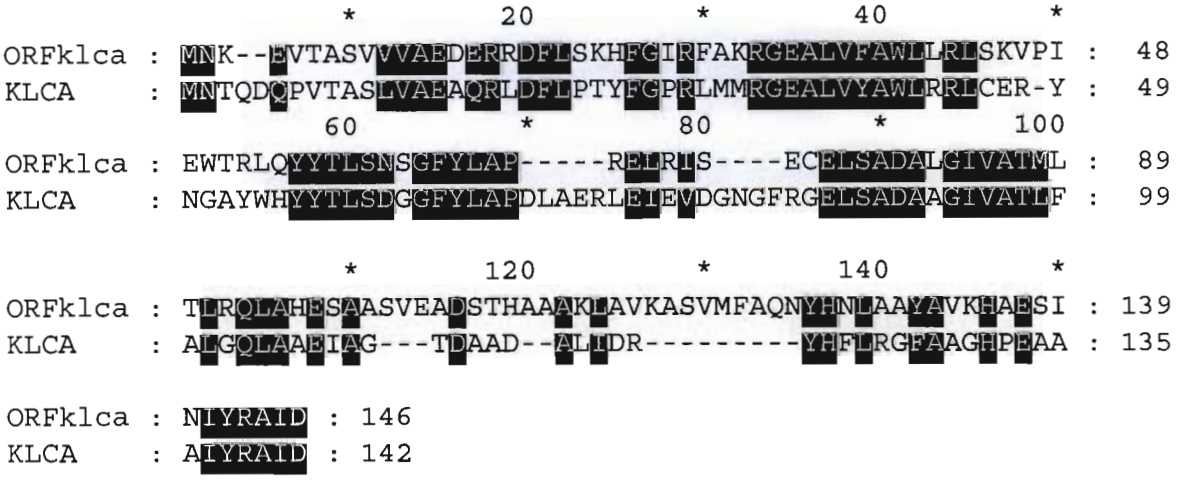


Fig. 4.16 Alignment of ORFklca to the KlcA protein of *E. coli* and *E. aerogenes*. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.3.16 ORFkorC

The product product predicted from this ORF is homologous to the transcriptional repressor protein, KorC, from plasmid RK2 (quality score = 91) and a KorC protein in *E. aerogenes* (quality score = 82). The predicted protein is 10 aa shorter than the KorC proteins. BLAST searches also showed that ORFkorC is related to a DNA cytosine methyltransferase in *Salmonella enteritidis*. This relatedness occurs at the N-terminus of the much larger (379 aa) methyltransferase. Alignment of ORFkorC to the RK2 KorC protein is shown in Fig. 4.17. The

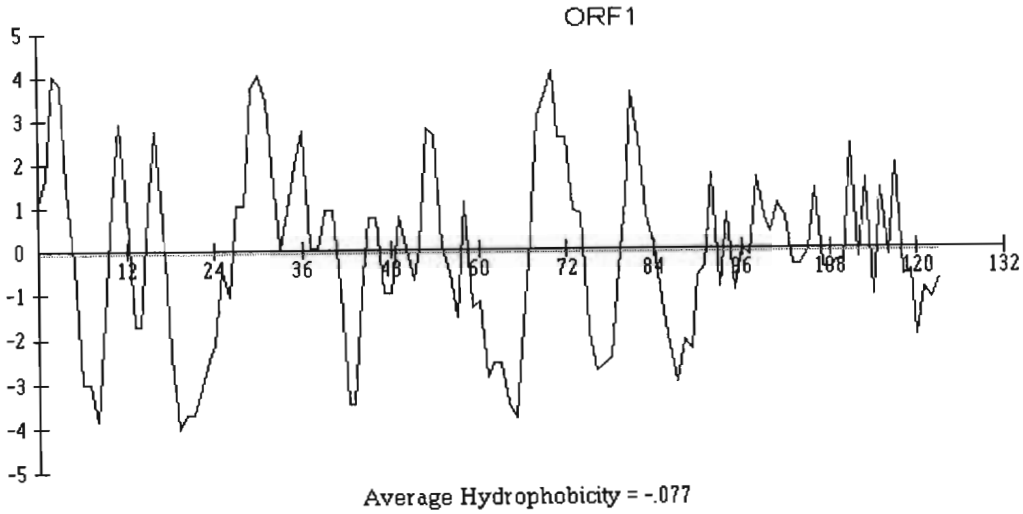
positions 28 to 47 on the KorC protein was reported to be a helix-turn-helix motif (Kornacki *et al.*, 1990) containing highly conserved residues at positions 5 (A), 9 (G) and 15 (I). The corresponding region (27 to 46) in ORFkorc contains A, G and the similar residue L at these positions. Analysis of the structure of both these regions showed that neither was in the H-T-H configuration.

```
                  *         20          *         40          *
ORFkorC  : -MHNEIRPQCLRPAECWTQPTANEIRAVLRMAGMTGGAAAKLLGLGAKGDR : 50
KorC     : MSDVNIRLECLRPAERWVQPTGAEIREVLHLAGLTGGQAARILGLGAKGDR : 51

                60          *         80
ORFkorC  : TIRRWIGGDS---RILWQRQIHDGPAHA------ : 75
KorC     : TVRRWVGEDSPIPYAAWAILCDLAGIGAIWKGQG : 85
```

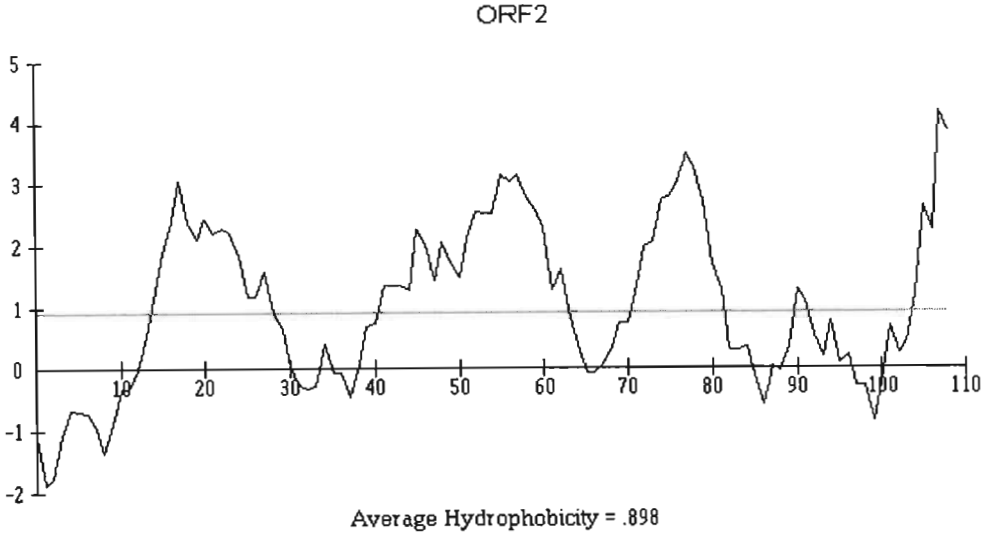
Fig. 4.17 Alignment of ORFkorC to the KorC protein of plasmid RK2. Darkly shaded boxes represent identical amino acid residues and lightly shaded boxes represent conserved amino acids.

4.3.4 **Hydropathy Profiles**

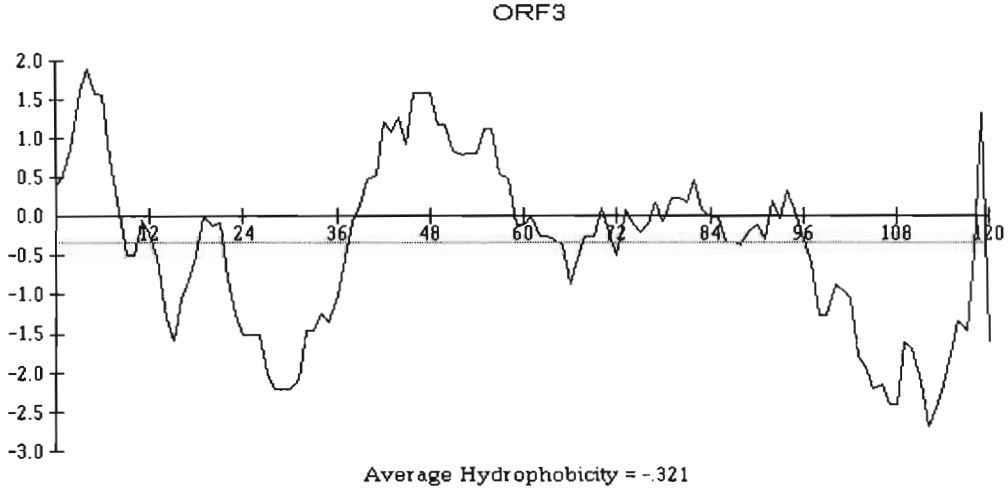
The hydrophobicity of predicted pXA1 proteins from ORF1 to ORF13 were examined to determine whether they were membrane associated. The results of hydropathy plots using the Kyte and Doolittle (1982) algorithm indicated that proteins encoded by ORF2, ORF3, ORF6, ORF8 and ORF10 have membrane-spanning regions. These regions are characterized by stretches of 16 - 35 hydrophobic amino acids, flanked on either side by charged amino acid residues (Lessl *et al.*, 1992). Examination of the profiles of the pXA1 ORFs (Fig. 4.18) suggest that ORF2 has two such hydrophobic regions while ORF3, ORF8 and ORF10 possess a single membrane-spanning region. ORF6 which has an average hydrophobicity of 0.582 appears to possess five such regions. These domains were identified by their size and extent of hydrophobicity.



Average Hydrophobicity = -.077

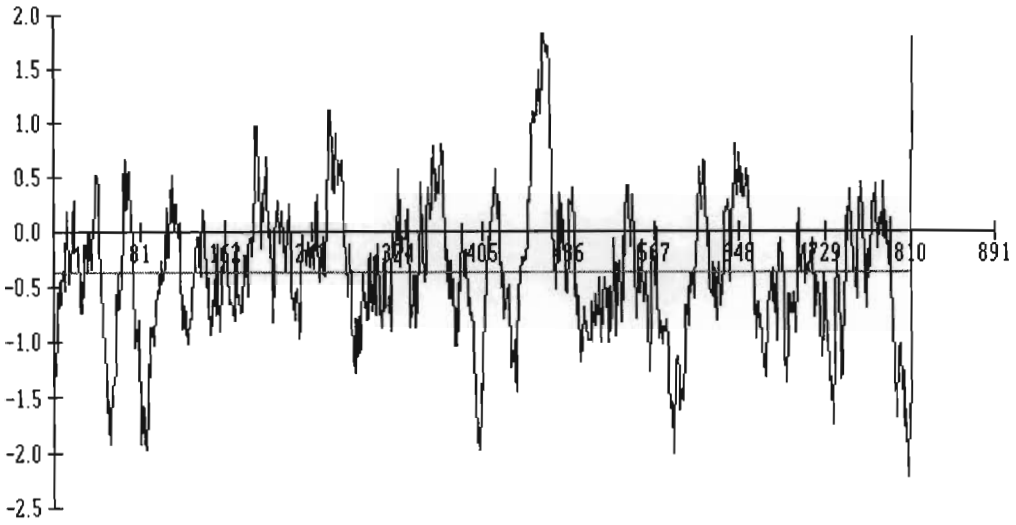


Average Hydrophobicity = .898



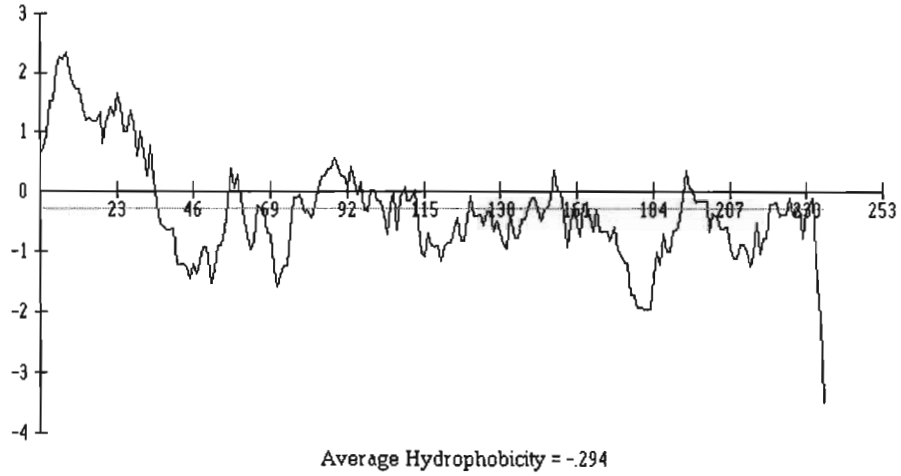
Average Hydrophobicity = -.321

ORF4



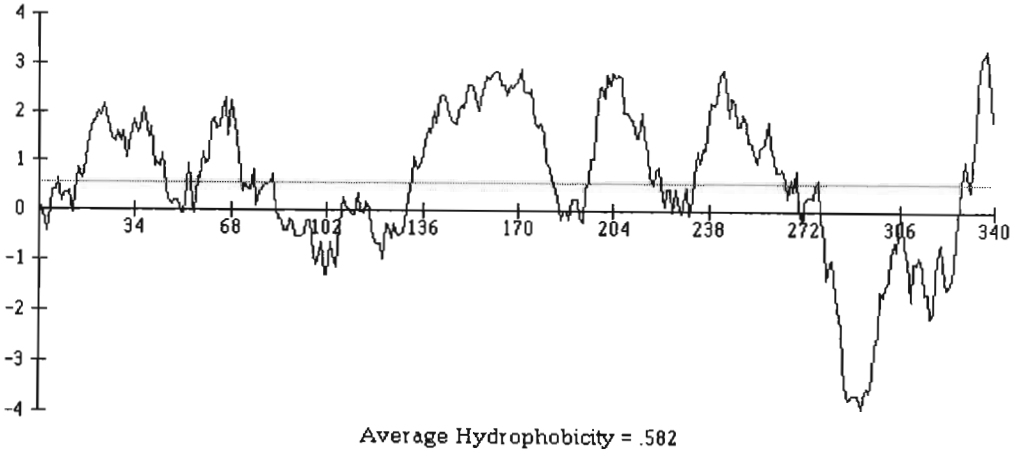
Average Hydrophobicity = -.367

ORF5



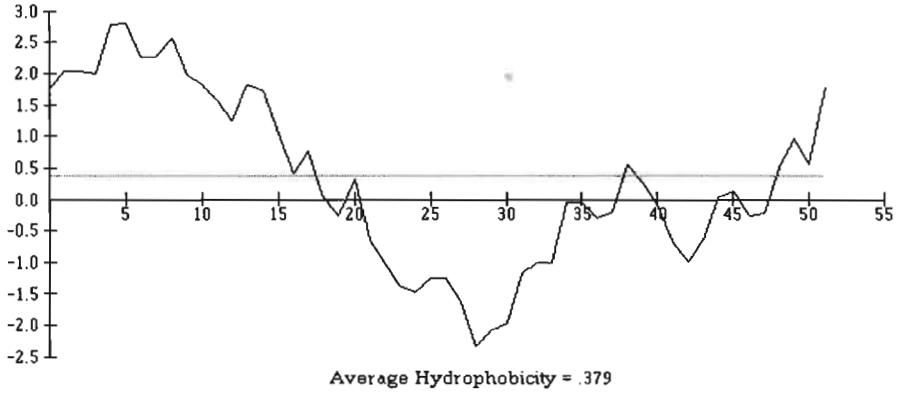
Average Hydrophobicity = -.294

ORF6



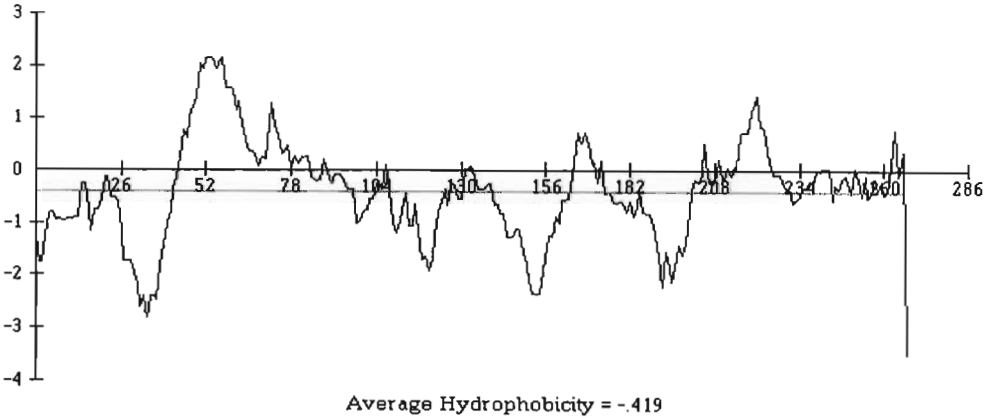
Average Hydrophobicity = .582

ORF7



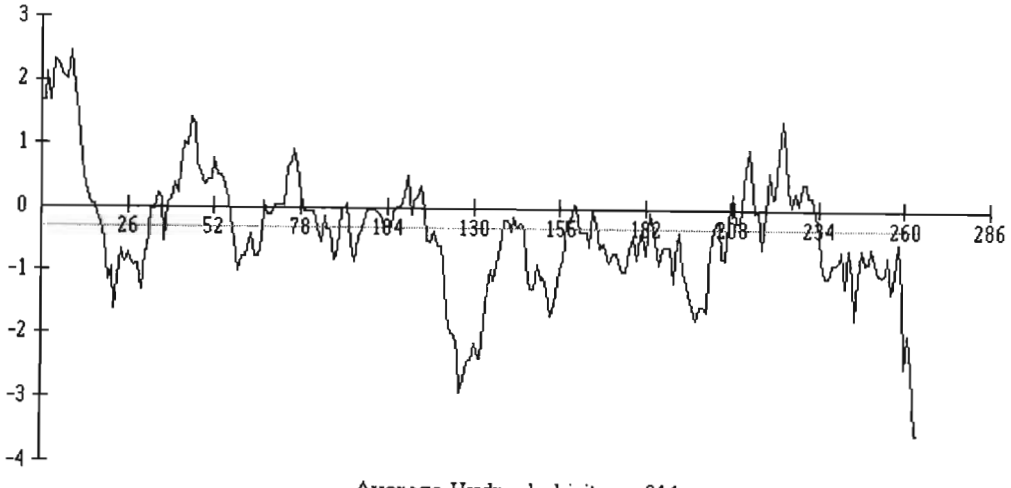
Average Hydrophobicity = .379

ORF8



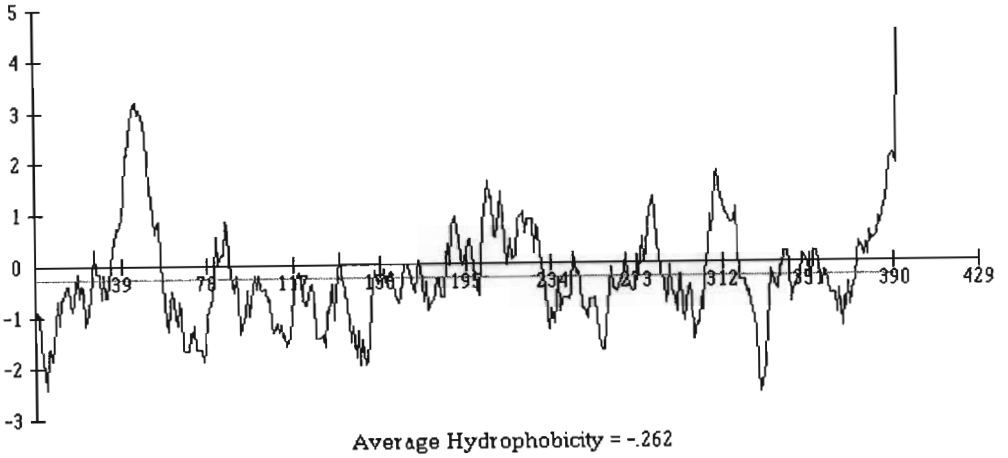
Average Hydrophobicity = -.419

ORF9



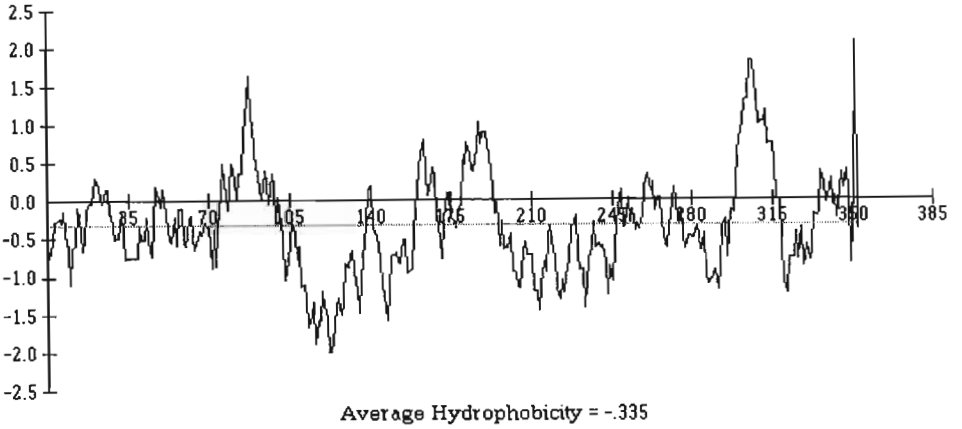
Average Hydrophobicity = -.314

ORF10



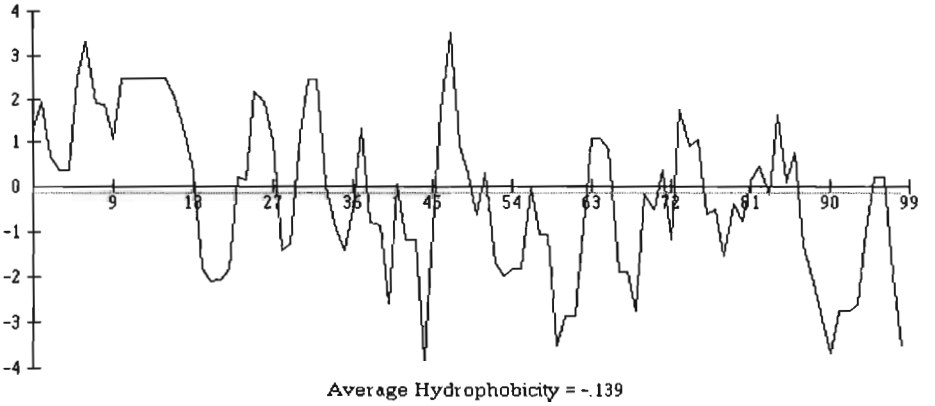
Average Hydrophobicity = -.262

ORF11



Average Hydrophobicity = -.335

ORF12



Average Hydrophobicity = -.139

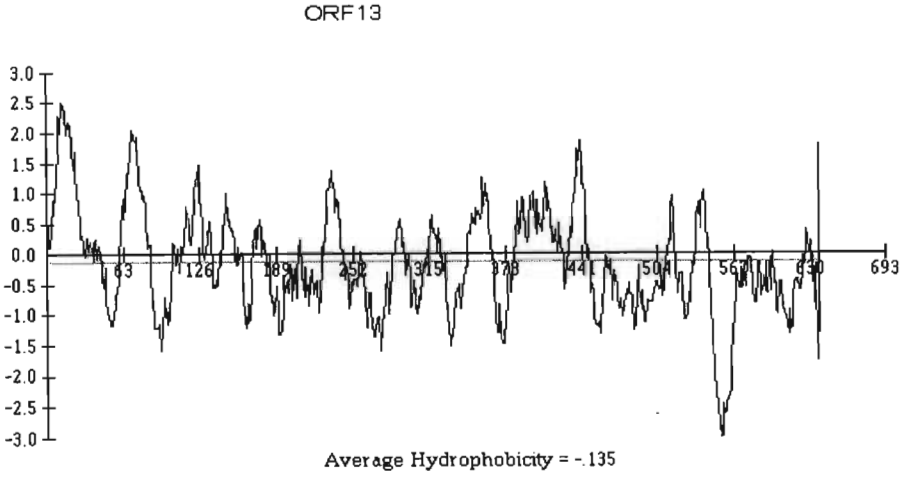


Fig. 4.18 Hydropathy profiles of the thirteen membrane-associated pXA1 ORFs. The plots were generated using the algorithm of Kyte and Doolittle (1982). The average hydrophobicity of the protein is shown as a single dotted line. Positive values represent hydrophobic regions whereas negative values represent hydrophilic regions

ORF2, ORF5, ORF7, ORF9, ORF12 and ORF14 all possess strongly hydrophobic regions near their N-terminals. Analysis of the protein sequences revealed that these regions were not membrane-spanning domains but signal sequences for transport across the cytoplasmic membrane. These signal peptides contain three regions: a) an "N" domain (2 - 15 residues) which is polar and carries a net positive charge due to the presence of 1 or more K or R residues; a hydrophobic "H" domain (> 8 residues) which is rich in A and L but devoid of P, K, R, D, E and H residues; and a "C" domain that contains a P or a G at the -6 position relative the cleavage site and A, G, or S at the -3 and -1 sites (Pugsley, 1993). The signal peptide of ORF14 exhibiting these characteristics is shown in Fig. 4.19.

Fig. 4.19 N-terminal region of the ORF14 protein precursor showing the signal sequence and signal peptidase cleavage site. Charged residues in the "N" domain are underlined and aa positions of signals recognised by the peptidase (bold) are relative to the cleavage site.

The signal sequences of ORF7 and ORF12 differ slightly from the standard protein signals and they correspond to lipoprotein signal peptides. These peptides have similar "N" domains, but the "H" domain is shorter and extremely hydrophobic. The turn residue (P or G) in the "C" domain is usually absent and C, A/G and L residues are present at +1, -1 and -3 positions, respectively, relative to the cleavage site (Pugsley, 1993). The lipoprotein signal sequences of ORF7 and ORF12 are shown in Fig. 4.20.

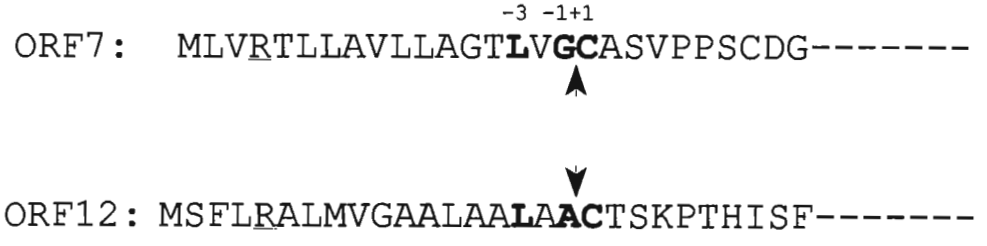


Fig. 4.20 Lipoprotein signal sequences of ORF7 and ORF12. Charged residues in the "N" domain are underlined. The "H" domain contains a high proportion of A and L. Conserved residues before and after the peptidase cleavage site are indicated in bold type.

4.3.5 **Location of additional pXA1 genes**

In addition to the 16 ORFs, two other regions in the plasmid showed similarity to gene sequences stored in DNA databases. Sequencing of the 1.3 kb *Eco*RI fragment showed that it contained a gene homologous to the *tlpA* gene located on a virulence plasmid in *S. typhimurium*. This gene product is a member of a family of α-helical coiled coil proteins in bacteria (Koski *et* al., 1992). BLAST similarity searches revealed that the pXA1 homolog was smaller than TlpA (371 aa) and the major portion of the putative gene was present on the *Eco*RI/*Hin*dII fragment of pTX61. Comparison of a stretch of protein sequence showed 79/148 (53%) identical amino acids and 99/148 (67%) similar amino acids.

Sequencing of the *Xba*I-end of pTX21 showed that a gene homologous to a gene on a plasmid, pZM2, in *Zymomonas mobilis* was present in this region. The hypothetical protein in *Z. mobilis* is 584 aa in length.

The regions containing these genes were not sequenced completely and ORF analysis was therefore not performed. A map showing the location of the pXA1 transfer region and the position of these two regions was constructed (Fig. 4.21).

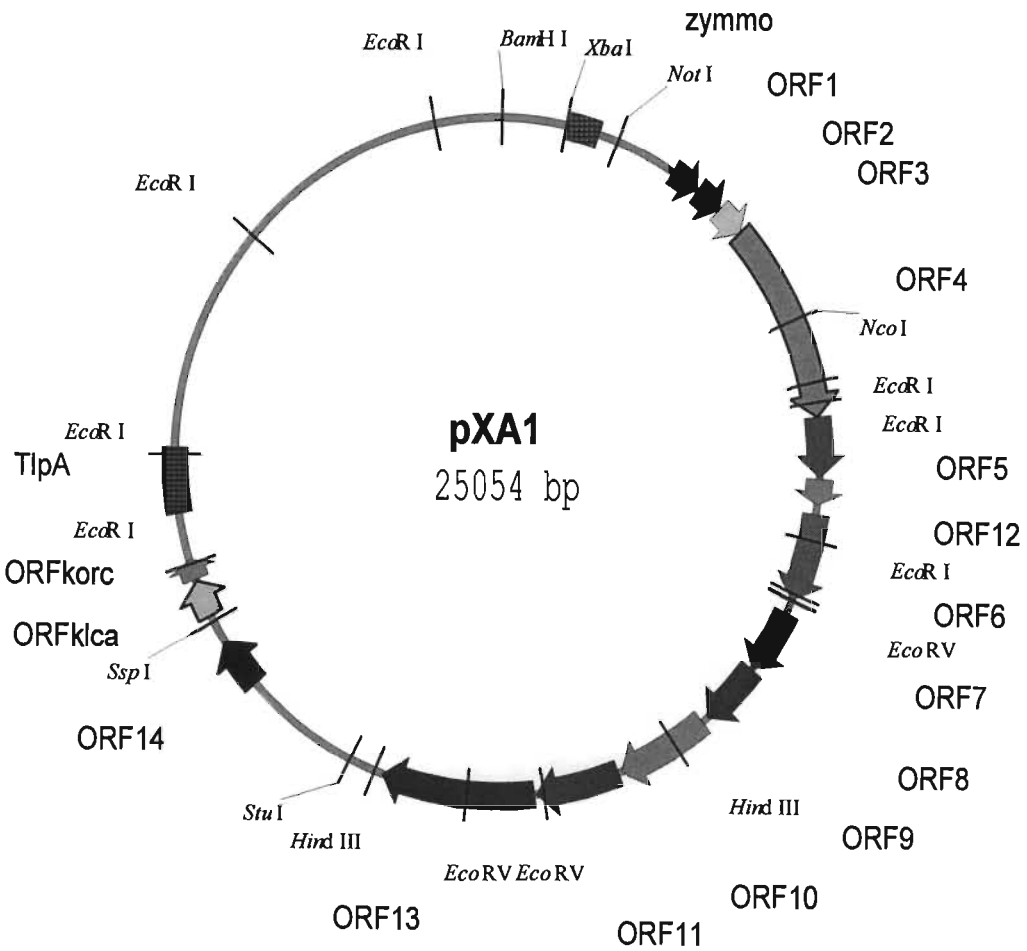


Fig. 4.21 Map of pXA1 showing the location of coding regions. There are 16 ORFs in the transfer region. pXA1 DNA sequences homologous to *tlpA* and a hypothetical protein in *Z. mobilis* (zymmo) are shown.

4.3.6 Creation of Donor Strains

After construction of pBX1 which consists of the 24.3 kb *Bam*HI/*Xba*I pXA1 fragment cloned into pBluescript, mating experiments were performed to determine whether the transfer region on this construct was operative. Electroporation of pBX1 into the *X. albilineans* donor strains XA86CS and TSS was successful. Plasmids isolations from these strains revealed the presence of autonomous pBX1 molecules (Fig. 4.22). The characteristic banding pattern of *Eco*RI-digested pXA1 was still observed, except that the second-largest fragment migrated slower (contains pBluescript) and an additional 800 bp *Eco*RI fragment (cleavage in the multiple cloning region) was present. In strain GLP7CS and the *E. coli* strain JC8679, Ap^R-colonies were isolated but plasmid isolations failed to demonstrate presence of the plasmid. Electroporation of the *P. syringae* pv. *syringae* strain 3420S was unsuccessful using the protocol for *X. albilineans*. Attempts to transform XA86C with a deletion derivative of pBX1, pBX3, were unsuccessful as Ap^R colonies were not produced after electroporation.

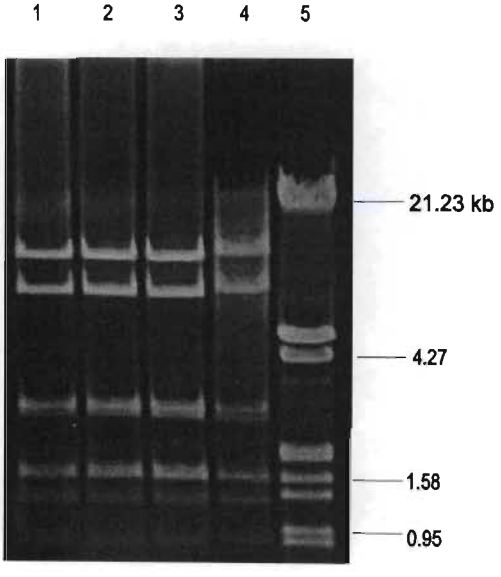


Fig. 4.22 Successful electroporation of *X. albilineans* XA86C with pBX1. Plasmid DNA from three randomly-selected colonies (lanes 1-3) was restricted with *Eco*RI. Lane 4: pBX1; and lane 5: λ DNA digested with *Eco*RI/*Hin*dIII.

4.3.7 **Bacterial Mating**

pBX1 was successfully transferred from XA86CS1into XA86CR by conjugation. Appearance of Ap^R/Rif^R cells indicated that XA86CR cells had acquired Ap^R. Spontaneous mutation of the donor and recipient cells could not account for the high number of colonies observed. Verification of pBX1-transfer was accomplished by plasmid isolation from ten randomly-selected colonies. The results of *Eco*RI cleavage of the isolated DNA proved that pXA1 was present in all ten transconjugants tested (Fig. 4.23). The frequency of plasmid transfer was calculated to be 2×10^{-3} per donor cell.

Attempts to transfer pBX1 out of *E. coli* donor cells were unsuccessful. Conjugation was not detected in bi-parental matings (testing self-transmissibility) or tri-parental matings with helper functions being provided by the IncP plasmid RK2 and the IncW pSa. Although TSS1 contained autonomous copies of pBX1, transfer of the plasmid to XA86CR was not detected.

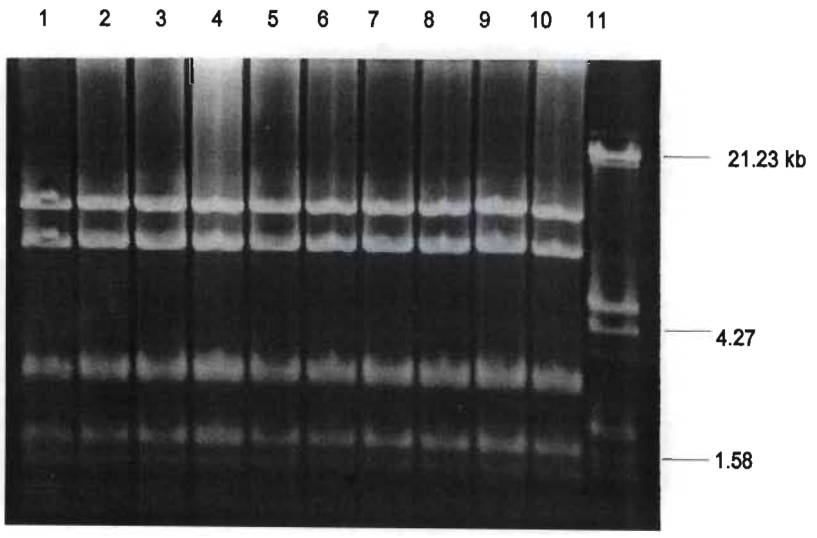


Fig. 4.23 Isolation of pBX1 from *X. albilineans* transconjugants. Plasmid DNA from ten randomly-selected colonies (lanes 1-10) cleaved with *Eco*RI showed identical banding patterns. Lane 11 contains λ DNA cleaved with *Eco*RI/*Hin*dIII.

4.4 DISCUSSION

Analysis of DNA sequencing data has clearly shown that pXA1 contains a transfer region. There are many overall similarities with conjugal transfer systems in other bacteria, in addition to the obvious relatedness in the DNA sequence of these genes. The order of the ORFs in the putative pXA1 transfer operon is similar to the order of the genes in the *tra, pilW, virB* and *ptlC* operons. The corresponding homologs are in the same positions and are of similar size, with a few exceptions. The operons are tightly packed with overlapping stop and start codons. They appear to be transcribed on a single polycistronic mRNA transcript, using a promoter upstream of the first gene. The genes in the operons are involved in pilus biosynthesis, DNA processing, murein degradation, antirestriction activity and gene regulation. It is significant that the pXA1 operon has homologs that appear to possess all of these activities. While DNA sequencing data has shown that these homologs are present on the plasmid, mating experiments have demonstrated that these genes are functional, as pXA1 is capable of independant transfer.

Only the *tra* operon of pKM101 appears to have an additional promoter between the *TraD* and *traN* genes. The putative pXA1 promoter region is tentative since a -10 RNA polymerase-binding region has not been identified. It is possible that additional ORFs are present upstream of ORF1 as this region has not yet been sequenced. The ORF1 protein is not homologous to any known protein sequences according to BLAST searches and it has very weak homology to the VirB1/TraL proteins when CLUSTALW alignments were performed. However, ORF14 has much stronger homology to VirB1 and TraL and is also homologous to various lytic transglycolsylases. The ORF14 protein sequence shows conservation of important amino acid residues as described by Lehnherr *et al.* (1998). It is therefore apparent that ORF14 and not ORF1 encodes homologs to TraL and VirB1. These murein-degrading enzymes have been found

to be present in certain bacteria, conjugal systems and in certain bacteriophages (Lehnherr *et al.*, 1998; Koraimann *et al.*, 1996). It is believed that they introduce small localized gaps in the peptidoglycan molecule which allows passage of viral and plasmid genomes. It is therefore apparent that the position of this gene in the pXA1 operon is different to that of the *tra* and *virB* operons as VirB1 and TraL are also similar to bacterial lytic transglycosylases. These genes are not essential for conjugation and mutations in *virB1* result in decreased transfer efficiencies (Winans *et al.*, 1996). Interestingly, the *pilW* operon of R388 does not possess a similar homolog. It is possible that conjugation systems can also use bacterial autolysins to break down peptidoglycan although it is preferable to have a plasmid-encoded enzyme. It remains to be proved whether conjugative plasmids have sequestered these autolysins in order to increase their transfer capabilities.

Homology of the ORF2 protein to VirB2, TraL and TrWL suggests that it may be the pilin subunit for the pXA1 conjugative pilus. While it is known that the TraA protein of the F plasmid is the pilin subunit this has not been proved for the other systems. Hydropathy profiles of VirB2 and ORF2 show that two membrane-spanning domains exist. While this is not clear in the proposed model of the pKM101 pilus (Fig. 1.7), these domains could be the responsible for attachment of the pilin molecules to the cytoplasmic membrane as seen in Fig. 4.1. If this is so, folding of the pilin molecules forming the filament of the pilus is necessary, to ensure that these hydrophobic domains are not exposed to the environment.

BLAST comparisons did not reveal homology between the ORF3 and ORF7 proteins and their corresponding proteins in other transfer operons. This was not surprising as the searches typically display low scores when comparing small proteins (Deveraux *et al.*, 1984). However, alignment of ORF3 to VirB3 showed that these two proteins were homologs.

The ORF4 and ORF11 proteins are predicted to be ATPases as they possess nucleotide-

binding motifs. The VirB4 and VirB11proteins have been shown to have ATPase activity *in vitro* and the Walker Boxes are crucial for T-DNA transfer (Shirasu *et al.*, 1994; Stephens *et al.*, 1995). It is believed that these proteins and their homologs in conjugal systems are required for activation of proteins comprising the transport machinery in the cell envelope (Christie, 1997). The VirB4 protein is also responsible for accumulation of VirB3 in the outer membrane (Jones *et al.*, 1994).

ORF5 homologs are thought to encode a pilus-associated protein as mutants can be complemented extracellularly (Winans *et al.*, 1996), presumably by recycling pilus components. The presence of a signal sequence and a largely hydrophilic hydropathy profile suggest that ORF5 may also be a pilus-associated protein occurring outside of the cell. The characteristics of the ORF6, ORF8 and ORF10 proteins are similar to their counterparts in terms of size and hydropathy profiles.

ORF7 and ORF12 were found to be lipoproteins. It has been observed that ORF7 homologs interact with ORF9 homologs via disulphide links (Christie 1997). The obvious similarity of ORF12 to other lipoprotein homologs involved in entry exclusion and homology of ORF5 to TrbJ of RK2 suggest that ORF5 and ORF12 are involved in entry exclusion functions as are the homologs TrbJ and TrbK in RK2 and TraT and TraS in the F plasmid (Lessl *et al.*, 1992). This property of conjugative plasmids prevents entry of similar plasmids into the cell and also prevents the cell from mating with itself if the end of the pilus attaches to receptors on the donor cell.

The relatedness of ORF11 to conjugal and various other transport systems including type II and type III protein secretion systems and competence proteins indicates that conjugation systems are evolutionarily linked to other transport systems and possibly have common mechanisms of transport. This link was also reinforced by homology of ORF7 to the type I HlyD

secretion systems responsible for hemolysin export in enteropathogenic *E. coli*.

The presence of ORF13, a VirD4 homolog, in pXA1 was interesting as this gene is present in the *virD* operon in the Ti plasmid. However, it occurs immediately after ORF11 in pXA1 and is still part of the polycistronic mRNA transcript. This arrangement also occurs in R388. It was even more surprising that ORF13 showed highest homology to a VirD4 protein in *R. prowazekii*, an obligate human pathogen. The nucleotide binding motif in these proteins is required for DNA transfer (Balzar *et al.*, 1992). It is believed that VirD4 is the coupler between the DNA processing and the DNA transfer systems. Homologs to this protein are present in the F plasmid, R388 and RP4. However, a homolog in pKM101 has not been identified (Christie, 1997). This is the only DNA processing protein that has been identified in the pXA1 operon. It is a requirement of conjugal transfer systems that a nickase is present to cleave at the *oriT* site to begin the transfer process. In R388 the nickase (TrwA) occurs after the VirD4 homolog, TrwB. The pXA1 operon could be arranged similarly. ORFs have not yet been identified on an approximately 2 kb region between ORF14 and ORFklca. This region could contain the pXA1 nickase as well as other DNA-processing proteins. The pKM101 tra operon has four such genes occurring immediately after the pilus biosynthesis genes (Winans *et al.*, 1996).

The presence of ORFklca at the far end of the operon, opposite to the end containing the promoter, suggests that this end is the first to enter the recipient cell. ORFKlca is a homolog of the antirestriction protein KlcA. KlcA is related to the pKM101 antirestriction protein, Ardb, but both these proteins do not share sequence similarity with other known proteins (Larsen and Figurski, 1994). Antirestriction proteins are known to be present in the leading region of the incoming plasmid strand during conjugation and they protect the plasmid from type I restriction endonucleases (Chilley and Wilkins, 1996). ORFklca shares extensive homology with KlcA including several possible motifs. Both these proteins and Ardb possess the sequence IYRAID

at their C-terminal ends. KlcA has also been shown to be lethal to *E. coli* cells unless regulated by KorC (Larsen and Figurski. 1994). This relationship probably also occurs in pXA1 as ORFkorc is found immediately after ORFklca. The helix-turn-helix motif (H-T-H) in KorC which is the proposed DNA binding region (Kornacki *et al.*, 1990) shared conserved residues with pXA1. However, analysis of the *E. coli* KorC protein using the program PROTEIN.EXE showed that this region is not a true H-T-H region as most of the residues on the latter half of the motif are in the extended configuration. In ORFkorc, the entire region is in the helical configuration. Thomas *et al.* (1995) have found that this region is not well conserved.

The homology of ORFkorC to the N-terminal region of a cytosine methyltransferase is intriguing. This homology has not been reported previously. The relatedness is significant at 40% identity and 48% similarity for a 45 aa stretch. This region includes a sequence GDRTIRRW which is also found on the *E. coli* KorC proteins. It is also interesting that relatedness to the methyltransferase begins immediately after the proposed DNA-binding region. DNA methylation could therefore be the mode of action of these transcriptional repressors. Since DNA methylation is also a mechanism of evading restriction enzymes, this activity could also implicate KorC in the antirestriction effect.

Other possible functions for the *kil/kor* homologs in pXA1 are in regulation of the transfer operon and in plasmid stability. The *kil/kor* gene pairs are analogous to the poison/antidote pairs in proteic plasmid stability systems (Jensen and Gerdes, 1995). It is likely that the lethal effect of ORFklca is functional in pXA1 as a plasmid (pTX12) carrying these genes was consistently isolated at much lower yields compared to other plasmids. This was found to be due to a lower concentration of cells in the broth cultures used for plasmid isolations. After DNA sequencing and ORF analysis, it was found that this 3.5 kb *Eco*RI/*Hin*dIII fragment contained a full length ORFklca but ORFkorC was truncated at the C-terminal. It is therefore possible that the lethal *kil*

kil effect was being only partially overridden by the KorC homolog. Slow-growing strains have also been observed for pKM101 derivatives, which is believed to be due to the *kilA* gene (Winans and Walker, 1985c). In *Streptomyces*, the *kil* genes have been shown to result in "pocking" or localized lysis of the cell envelope (Perlin, 1991).

The construction of pBX1 was necessary to test whether the conjugal transfer genes on pXA1 were functional. Cloning of this 24.3 kb fragment into pBluescript is near the upper limit for cloning into plasmid vectors and was only possible using gel-ligation techniques. The use of an Ap^R marker was not ideal since *X. albilineans* is resistant to 100 μg/ml of Ap. However, no growth occurred at 200 μg/ml Ap. The isolation of transconjugants after transfer of pBX1 at an acceptable frequency of 2×10^{-3} per donor cell, showed that the plasmid is self-transmissible and that the transfer genes are functional. It also demonstrates that pXA1 does have the necessary DNA processing genes to initiate transfer. Since pBX1 has only been mated out of its original host cell thus far, other possible scenarios are that either a conjugative transposon or a megaplasmid in XA86 are responsible for mobilization of pXA1. Restriction analysis of plasmids rescued from transconjugants have clearly shown that such a transposon insertion did not occur. While it is still possible that a megaplasmid has provided helper functions for pXA1, the discovery of the pilus biosynthesis operon has made this possibility seem remote.

The inability of pBX1 to transfer itself out of *E. coli* donor cells could be due to several possible reasons. The most likely reason is that these transfer genes are not expressed in *E. coli* or that an appropriate signal for construction of the pilus is not present. It was noticed that the copy number of pBX1 was higher in *E. coli* compared to *X. albilineans* hosts. This was apparent from differences in yield of the plasmid, even when a greater number of *X. albilineans* cells were used in plasmid isolations. This raises the possibility that pBX1 still harbours an intact replication region and that the plasmid uses the pBluescript and pXA1 replication regions in *E.*

albilineans hosts, respectively. It is therefore also possible that the promoter for the pXA1 transfer region is poorly expressed in *E. coli*. Secondly, conjugative transfer regions are not normally expressed constitutively (Frost *et al.*, 1994). The formation of the *virB* pilus is dependant on a lower temperature and the presence of acetylsyringone (Kado, 1996). The pXA1 conjugative pilus may require such a signal for formation of the pilus and transfer of the plasmid. Conjugation was also found to be dependant on a $RecA^+$ phenotype. Most cloning strains such as DH5α are $RecA^-$. However, attempts to introduce pBX1 into JC8679 cells in mating experiments resulted in inability to recover autonomous forms of the plasmid. This was also apparent in GLP7C. It is possible that the plasmid integrates into the chromosome of these strains which would suggest that they share common regions.

The pXA1 operon showed highest overall sequence relatedness to the *pilW* operon. The arrangement of the genes in the pXA1 operon also matched this operon exactly except the position of the first gene. This suggests that pXA1 may be an IncW plasmid.

CHAPTER FIVE

GENERAL DISCUSSION

The presence of plasmids in bacteria raises the question whether these plasmids constitute a stable characteristic of the host cell or whether the plasmids are genetic nomads whose stay in any bacterial population is transient. In the past, plasmids have been associated with certain bacteria only. While this is still largely true for most bacteria, the existence of broad host range conjugative plasmids lifts this restriction on the lifestyle of plasmids. It has become apparent from analysis of DNA sequencing of plasmids that they are comprised of DNA segments from various sources.

We have shown in our laboratory that 28 out of 69 *X. albilineans* strains tested contained plasmids. This incidence is not uncommon in other bacteria and illustrates the point that these plasmids cannot be regarded as being a characteristic of the bacterium. In certain well known cases, such as the Ti plasmids in *Agrobacterium* or the symbiotic plasmids of *Rhizobium,* the relationship between the host and plasmid is characteristic of the species since the bacteria rely on the functions encoded by the plasmids. However, this type of relationship is rare and does not occur for the overwhelming majority of plasmids. Unfortunately, the interest created by such infrequent cases has led to extensive efforts to find functions for all plasmids isolated.

A similar effort was made to find a function for the plasmid found in a South African strain of *X. albilineans*. The creation of plasmid-cured strains established that the plasmid was not responsible for production of a novel antibiotic produced by *X. albilineans*. Pathogenicity trials with a plasmid-free strain showed that the plasmid was not responsible for the symptoms associated with this antibiotic. The plasmid did not contain sequences homologous to *avr* genes conserved in other *Xanthomonas* strains. Another study showed that the plasmid did not

contribute towards a range of biochemical and physiological phenotypes tested (Rahaman, 1995). These results question the need for the plasmid by the host organism. Considering that the majority of plasmids in plant pathogenic bacteria have no known function (Coplin, 1989), this question can be extrapolated to query the existence of plasmids in all bacteria. Are plasmids required to contribute towards the coding potential of the host organism or do they represent selfish DNA strands that are parasitic in bacteria?

Sequencing of pXA1 DNA has revealed the presence of valid ORFs. *In vitro* transcription and translation assays demonstrated that pXA1 genes are expressed to produce polypeptides. It is therefore unlikely that the cryptic plasmid genes are dormant in bacteria. It is more likely that success in searching for roles of plasmids will be achieved if such searches are targeted towards investigation of plasmid maintenance functions and not phenotypes that are of benefit to the host.

Conjugation experiments together with DNA sequencing data showed that the majority of DNA in pXA1 is involved in transfer of the plasmid via conjugation. If plasmids are considered to be parasitic, then this ability to transfer itself between bacteria is obviously advantageous to the plasmid. In contrast to non-transferable or mobilizable plasmids, this ability allows conjugative plasmids to spread among a population. However, this spread does not occur indiscriminately. Conjugation systems are well regulated and are normally repressed for transfer (Frost *et al.*, 1994). An exception is the F plasmid in which one of these regulatory genes is inactive. It therefore seems as if plasmids are in control of their transfer or do so upon receipt of some signal. In *Agrobacterium*, *virB*-encoded pili are not normally present on the cell surface but only occur upon a decrease in temperature (Kado, 1996). It is possible that transfer of plasmids occurs only during periods of stress experienced by the host. This would mean that conjugation is simply an escape mechanism for perpetuation of a plasmid.

Until recently, conjugation was considered to be a sexual process between bacteria. The conjugative pili were referred to as "sex pili." This was due to concentration of research on the F plasmid only. Unfortunately, F is not a typical plasmid as its transfer region is expressed constitutively, as described previously. The ability of this plasmid to integrate into the bacterial chromosome resulted in erroneous transfer of chromosomal genes and even entire chromosomes into the recipient cell. However, study of other conjugal systems has shown that conjugation is a mechanism for transfer of plasmids only, and transfer of chromosomal DNA is accidental.

DNA sequencing of pXA1 showed that in addition to the pilus biosynthesis genes, the plasmid contained genes similar to those from different bacteria. This included a DNA-processing gene most similar to VirD4, a transglycosylase having greater similarity to a *B. subtilis* homolog compared to other plasmid homologs, kil/kor regulatory genes very similar to those from RK2, a rare prokaryotic α-helical coiled coil protein similar to that found on a virulence plasmid in *S. typhimurium* and a region similar to an unknown protein found on a plasmid in *Z. mobilis*. In another *X. albilineans* plasmid, pXA2, preliminary sequencing studies have shown that regions of the plasmid are similar to various genes including those coding for a mutidrug efflux protein, an auxin-induced protein, a proteinase and a cell filamentation protein (Ramdeen, 1999). A third *X. albilineans* plasmid, pXA3, codes for a cold-shock protein (Iyer, 1996). This curious assortment of proteins from various bacteria suggests that plasmids may pick up these proteins after transfer into different bacteria. This would also explain why strains of a pathogen possess unrelated plasmids. Alternatively, conservation of plasmids in strains of a bacterium suggests that the resident plasmid encodes a phenotype crucial to the host bacterium.

There are no known sequences that are exclusive to plasmids only. This suggests that most of the plasmid sequences originated from the bacterial chromosomes. Many of the functions associated with plasmids such as stability systems are also found on the bacterial

chromosome (Jensen and Gerdes, 1995). It is believed that conjugal transfer systems actually recognise only protein sequences and that DNA transfer is effected since the single strand is coated with DNA-binding proteins (Christie, 1997). If this is so, then plasmids have adopted a protein transport apparatus to aid in their dissemination. This is not surprising if one considers the relatedness of the *ptl* operon to the DNA transfer operons. Proteins from conjugation systems also have counterparts in type I, II and III protein secretion systems. ORF11 was found to be homologous to a competence protein in *S. pneumoniae*. This raises the possibility that these two systems have common mechanisms for passage of DNA through the cell membranes. It is still unclear how the transferred strand enters the recipient cell. Considering that it requires a pilus comprising of 10 proteins to transverse the donor cell envelope, it needs analogous structures in the recipient cell to reach the cytoplasm. Competence proteins in the recipient cell may play a role in accepting the DNA strand. It has been shown that natural competence in bacteria involves migration of proteins to the cell membrane and autolysin activity to degrade peptidoglycan and allow for DNA entry (Steips, 1991).

Excision of a functional replication region on the chromosome is the only requirement for creation of a new plasmid. There are numerous examples of small plasmids which possess only a replication region and no coding regions. An example is the 3.4 kb plasmid, pGT5, in *Pyrococcus abyssi* (Marsin *et al.*, 1996). The sole function of this plasmid is replication. Plasmids can acquire genes mainly by transposon insertions. Once a plasmid possesses an *oriT* site, it can be mobilized to other bacteria by conjugative plasmids. Insertions, deletions and rearrangements and high rates of mutation can render such plasmids indistinguishable from the parent molecules. Merging with other plasmids or acquisition of a conjugation system makes a plasmid transfer-proficient. The Ti plasmid was presumably created in such a way by joining of an IncP plasmid with an IncN plasmid as its conjugation region resembles IncP pili while its

virB region resembles IncN or IncW pili. It follows from this progression that megaplasmids have acquired genes from a variety of sources and are minichromosomes. The symbiotic plasmids of Rhizobium carry such a large quantity of DNA that they are intimately involved in the biochemical processes of the cell. This is in keeping with the definition of plasmids as selfish parasites seeking to perpetuate their lineage, as megaplasmids are so valuable to the cell that they will not be lost.

Conjugative broad host range plasmids show some of the characteristics that we normally associate with living organisms. They are capable of replication and can produce multiple copies of themselves. Joining of plasmid molecules can result in larger plasmids. They are capable of independent transfer from one host to another. During transfer, plasmids can carry their own primases as DNA-binding proteins so that it is not totally dependant on host enzymes. Plasmids also encode their own "virulence factors" in the form of transglycosylases which facilitates exit from and possible entry into cells. They can also force their hosts to maintain them by encoding stability systems which kills the host cell if the plasmid is lost. Although they cannot exist outside of a host cell, this property is shared by all obligate parasites. Viruses are currently considered to be in the "grey area" between living and non-living organisms, but plasmids can quite easily be included in this definition. At the very least, they fit the description of a "living molecule" perfectly.

REFERENCES

Akhtar, M.A., K.B. Malik and M. Aslam. 1988. Sugarcane leaf scald (*Xanthomonas albilineans*) in Pakistan. *Tropic. Agric. (Trinidad)* **65**:281-282.

Altschul, S.F., T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389-3402.

Alvarez, A.M., S. Schenck and A.A. Benedict. 1996. Differentiation of *Xanthomonas albilineans* strains with monoclonal antibody reaction patterns and DNA fingerprints. *Plant Pathol.* **45**:358-366.

Anthony, K.G. and L.S. Frost. 1996. Probing the nature of the F pilus tip. *Proc. Plasmid Biol. '96*, p 2.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl. 1989. Current protocols in molecular biology, Vol. 2. Greene Publishing Associates and Wiley-Interscience, New York.

Balzar, Z., G. Ziegelin, W. Pansegrau, V. Kruft and E. Lanka. 1992. KorB protein of promiscuous plasmid RP4 recognizes inverted sequence repetitions in regions essential for conjugative transfer. *Nucleic Acids Res.* **20**:1851-1858.

Banfalvi, Z., V. Sakanyan, C. Koncz, A. Kiss, I. Dusha and A. Kondorosi. 1981. Location of nodulation and nitrogen fixation genes on a high molecular weight plasmid in *R. meliloti*. *Mol. Gen. Genet.* **184**:318-325.

Basnayake, W.V.S. and R.G. Birch. 1995. A gene from *Alcaligenes denitrificans* that confers albicidin resistance by reversible anitibiotic binding. *Microbiology* **141**:551-560.

Bauer, D.W. and S.V. Beer. 1991. Further characterization of an *hrp* gene cluster in *Erwinia amylovora*. *Mol. Plant-Microbe Interact.* **4**:493-499.

Beijersbergen A., A.D. Dulk-Ras, R.A. Schilperoot and P.J.J. Hooykaas. 1992. Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256**:1324-1327.

Beijersbergen, A., S.J. Smith and P.J.J. Hooykaas. 1994. Localization and topology of VirB proteins of *Agrobacterium tumefaciens*. *Plasmid* **32**:212-218.

Bender, C.L. and D.A. Cooksey. 1987. Molecular cloning of copper resistance genes from *Pseudomonas syringae* pv. *tomato*. *J. Bacteriol.* **169**:470-474.

Bender, C.L., D.K. Malvick and R.E. Mitchell. 1989. Plasmid-mediated production of the phytotoxin coronatine in *Pseudomonas syringae* pv. *tomato*. *J. Bacteriol.* **171**:807-812.

Birch, R.G. and S.S. Patil. 1983. The relation of blocked chloroplast differentiation to sugarcane leaf scald disease. *Phytopathology* **73**:1368-1374.

Birch, R.G. and S.S. Patil. 1985. Preliminary characterization of an antibiotic produced by *Xanthomonas albilineans* which inhibits DNA synthesis in *Escherichia coli*. *J. Gen. Microbiol.* **131**:1069-1075.

Birch, R.G. and S.S. Patil. 1987a. Correlation between albicidin production and chlorosis induction by *Xanthomonas albilineans*, the sugarcane leaf scald pathogen. *Physiol. Mol. Plant Pathol.* **30**:199-206.

Birch, R.G. and S.S. Patil. 1987b. Evidence that an albicidin-like phytotoxin induces chlorosis in sugarcane leaf scald disease by blocking plastid DNA replication. *Physiol. Mol. Plant Pathol.* **30**:207-214.

Birch, R.G., J.M. Pemberton and W.V.S. Basnayake. 1990. Stable albicidin resistance in *Escherichia coli* involves an altered outer-membrane nucleoside uptake system. *J. Gen. Microbiol.* **136:**51-58.

Birnboim, H.C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1523.

Bonas, U, R.E. Stall and B. Staskawicz. 1989. Genetic and structural characterization of the avirulence gene *avrBs*3 from *Xanthomonas campestris* pv. *vesicatoria. Mol. Gen. Genet.* **218**:127-136.

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

Brom, S., A. Garcia-de los Santos, L. Girard and D. Romero. 1996. *Rhizobium etli* plasmids: participation in symbiosis, transference and free-living functions. *Proc. Plasmid Biol. '96,* p. 100.

Buchanan-Wollaston, V., J.E. Passatore and F. Cannon. 1987. The *mob* and *oriT* mobilization functions of a bacterial plasmid promotes its transfer to plants. *Nature* **328**:172-175.

Buchholz, W.G. and M.F. Thomashow. 1984. Comparison of the T-DNA oncogene complements of limited and wide host range tumour-inducing plasmids of *Agrobacterium tumefaciens. J. Bacteriol.* **160**:319-326.

Chen, C.T., P.H. Juang, Y.G. Liang and P.T. Hsieh. 1993. Recurrence of sugarcane leaf scald in Taiwan. (Abstr.) *Rev. Plant Pathol.* **72**:634.

Cherian, M.T. and J. Mathew. 1983. Toxin production by *Xanthomonas campestris* pv. *manihotis* - the cassava bacterial blight pathogen. *Indian Phytopathol.* **36**:156-157.

Chilley, P.M. and B.M. Wilkins. 1996. Transcriptional induction of leading region genes during bacterial conjugation. *Proc. Plasmid Biol. '96*, p. 10.

Chiou, C.-S. and A.L. Jones. 1991. The analysis of plasmid-mediated streptomycin resistance in *Erwinia amylovora. Phytopathology* **81**:710-714.

Christie, P.J. 1997. *Agrobacterium tumefaciens* T-complex transport apparatus: a paradigm for a new family of multifunctional transporters in eubacteria. *J. Bacteriol.* **179**:3085-3094.

Civerolo, E.L. 1985. Indigenous plasmids in *Xanthomonas campestris* pv. *citri. Phytopathology* **75**:524-528.

Clark, M.C. and C.H. Lawrence. 1986. Characterization of a plasmid in isolates of *Corynebacterium sepedonicum. Can. J. Microbiol.* **32**:617-622.

Comai, L. and T. Kosuge. 1980. Involvement of plasmid DNA in IAA synthesis in *Pseudomonas savastanoi. J. Bacteriol.* **143**:950-957.

Comai, L. and T. Kosuge. 1981. Cloning and characterization of *iaa*M, a virulence determinant of *Pseudomonas savastanoi. J. Bacteriol.* **149**:40-46.

Comstock, J.C. and J.M. Shine Jnr. 1992. Outbreak of leaf scald of sugarcane caused by *Xanthomonas albilineans*, in Florida. *Plant Dis.* **76**:426.

Coplin, D.L. 1989. Plasmids and their role in the evolution of plant pathogenic bacteria. *Ann. Rev. Phytopathol.* **27**:187-212.

Coplin, D.L., R.G. Rowan, D.A. Chisholm and R.E. Whitmoyer. 1981. Characterization of plasmids in *Erwinia stewartii. Appl. Environ. Microbiol.* **42**:599-604.

Couturier, M., F. Bex, P.L. Bergquist and K. Maas. 1988. Identification and classification of bacterial plasmids. *Microbiol. Rev.* **52**:357-395.

Covacci, A. and R. Rappuoli. 1993. Pertussis toxin export requires accessory genes located downstream from the pertussis toxin operon. *Mol. Microbiol.* **8**:429-434.

Crouse, G.F., A. Frischauf and H. Lehrach. 1983. An integrated and simplified approach for cloning into plasmids and single-stranded phages. *Meth. Enzymol.* **101**:78-89.

Daniels, M.J., J.M. Dow and A.E. Osbourn. 1988. Molecular genetics of pathogenicity in phytopathogenic bacteria. *Ann. Rev. Phytopathol.* **26**:285-312.

Davis, M.J., P. Rott, P. Baudin and J.L. Dean. 1994. Evaluation of selective media and immunoassays for detection of *Xanthomonas albilineans*, causal agent of sugarcane leaf scald disease. *Plant Dis.* **78**:78-82.

Davis, M.J., P. Rott, C.J. Warmuth, M. Chatenet and P. Baudin. 1997. Intraspecific genomic variation within *Xanthomonas albilineans*, the sugarcane leaf scald pathogen. *Phytopathology* **87**:316-324.

De Feyter, R. and D.W. Gabriel. 1991. At least six avirulence genes are clustered on a 90-kilobase plasmid in *Xanthomonas campestris* pv. *malvacearum. Mol. Plant-Microbe Interact.* **4**:423-432.

De Feyter, R., Y. Yang and D.W. Gabriel. 1993. Gene-for-gene interactions between cotton R genes and *Xanthomonas campestrs* pv. *malvacearum avr* genes. *Mol. Plant-Microbe Interact.* **6**:225-237.

Denny, T.P. and S.-R. Baek. 1991. Genetic evidence that extracellular polysaccharide is a virulence factor in *Pseudomonas solanacearum. Mol. Plant-Microbe Interact.* **4**:198-206.

Desomer, J., P. Dhaese and M. van Montague. 1988. Conjugative transfer of cadmium resistance plasmids in *Rhodococcus fascians* strains. *J. Bacteriol.* **170**:2401-2405.

Deveraux, J., P. Haeberli and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:11114-11120.

Dittapongpitch, V. and D.F. Ritchie. 1993. Indigenous miniplasmids in strains of *Xanthomonas campestris* pv. *vesicatoria. Phytopathology* **83**:959-964.

Dow, J.M., G. Scofield, K. Trafford, P.C. Turner and M.J. Daniels. 1987. A gene cluster in *Xanthomonas campestris* pv. *campestris* required for pathogenicity controls the excretion of polygalacturonate lyase and other enzymes. *Physiol. Mol. Plant Pathol.* **31**:261-271.

Egan, B.T. 1969. Leaf scald strains - a good subject for international co-operation. *Sugarcane Pathol. Newsl.* **3**:12.

Egan, B.T. 1970. Leaf scald disease of sugarcane. *Sugarcane Pathol. Newsl.* **5**:28-29.

Egan, B.T. 1971. Breeding for resistance to leaf scald disease. *Proc. Int. Soc. Sugar Cane Technol.* **14**:920-924.

Egan, B.T. and O.W. Sturgess. 1980. Commercial control of leaf scald disease by thermotherapy and a clean seed programme. *Proc. Int. Soc. Sugar Cane Technol.* **17**: 1602-1606.

Engst, K. and P.D. Shaw. 1992. Identification of a *lysA*-like gene required for tabtoxin biosynthesis and pathogenicity in *Pseudomonas syringae* pv. *tabaci* strain PTBR2.024. *Mol. Plant-Microbe Interact.* **5**:322-329.

Everett, R. and N. Willetts. 1980. Characterization of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor. *J. Mol. Biol.* **136**:129-150.

Falkenstein, H., W. Zeller and K. Geider. 1989. The 29 kb plasmid, common in strains of *Erwinia amylovora*, modulates development of fireblight symptoms. *J. Gen. Microbiol.* **135**:2643-2650.

Frost, L.S., R. Fekete, S. Penfold and J. Manchak. 1997. Factors that control TraM expression: The level of TraM defines the F plasmid transfer efficiency. *Proc. Plasmid Biology '96*, p. 4.

Frost, L.S., K. Ippen-Ihler and R.A. Skurray. 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.* **58**:162-210.

Freifelder, D.M. 1987. Essentials of Molecular Biology. Jones and Bartlett Publishers, Portola Valley, Boston.

Fu, J-F., J-M. Hu, Y-S. Chang and S-T. Liu. 1998. Isolation and characterization of plasmid pSW200 from *Erwinia stewartii. Plasmid* **40**:100-112.

Gantotti, B.V. and S.V. Beer. 1982. Plasmid borne determinants of pigmentation and thiamine prototrophy in *Erwinia herbicola. J. Bacteriol.* **151**:1620-1627.

Gantotti, B.V., S.S. Patil and M. Mandel. 1979. Apparent involvement of a plasmid in phaseotoxin production by *Pseudomonas phaseolicola. Appl. Environ. Microbiol.* **37**: 511-516.

Garcia-de los Santos, A. and S. Brom. 1996. Plasmids from diverse *Rhizobium* species share structural and functional homology. *Proc. Plasmid Biol. '96*, p. 86.

He, S.Y. 1997. Hrp-controlled interkingdom protein transport: learning from flagellar assembly? *Trends Microbiol.* **5**:489-495.

Hensel, M., J.E. Shea, S.R. Waterman, R. Mundy, T. Niklaus, G. Banks, A. Vazquez-Torres, C. Gleeson, F.C. Fang and D.W. Holden. 1998. Genes encoding putative effector proteins of the type III secretion system of *Salmonella* pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. *Mol. Microbiol.* **30**:163-174.

Herbers, K., J. Conrad-Strauch and U. Bonas. 1992. Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. *Nature* **356**:172-174.

Honeycutt, R.J., B.W.S. Sobral and M. McClelland. 1995. tRNA intergenic spacers reveal polymorphisms diagnostic for *Xanthomonas albilineans. Microbiology* **141**:3229-3239.

Hooykaas, P.J.J. and A.G.M. Beijersbergen. 1994. The virulence system of *Agrobacterium tumefaciens*. *Ann. Rev. Phytopathol.* **32**:157-179.

Hopkins, C.M., F.F. White, S.-H. Choi, A. Guo and J.E. Leach. 1992. Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. *Mol. Plant-Microbe Interact.* **5**:451-459.

Hoy, J.W. and M.P. Grisham. 1994. Sugarcane leaf scald distribution, symptomatology, and the effect on yield in Louisuana. *Plant Dis.* **78**:1083-1087.

Hughes, C.G. 1978. Diseases of sugarcane - a review. *PANS* **24**:143-159.

Irvine, J.E., J.M. Amador, R. Gallo, M.I. Riess and J.C. Comstock. 1993. First report of leaf scald, caused by *Xanthomonas albilineans*, of sugarcane in Mexico. *Plant Dis.* **77**: 46.

Isakeit, T. and J.E. Irvine. 1995. First report of leaf scald, caused by *Xanthomonas albilineans*, of sugarcane in Texas. *Plant Dis.* **79**:860.

International Society for Sugar-Cane Technologists Standing Committee on Sugarcane Diseases. 1974. The Diseases of Sugarcane, their Causal Agents and Distribution.

Iyer, R. 1996. Characterization of the plasmid found in a Kenyan (KEN6) isolate of *Xanthomonas albilineans*. B.Sc.(Hons) Dissertation, University of Durban-Westville, Durban, South Africa.

Jenkins, C.L. and M.P. Starr. 1982. The brominated aryl-polyene (xanthomonadin) pigments of *Xanthomonas juglandis* protect against photobiological damage. *Curr. Microbiol.* **7**:323-326.

Jensen, R.B. and K. Gerdes. 1995. Programmed cell death in bacteria: proteic plasmid stabilization systems. *Mol. Microbiol.* **17**:205-210.

Johnson, H.W., U.M. Means and F.E. Clark. 1959. Responses of seedlings to extracts from soybean nodules bearing selected strains of *Rhizobium japonicum*. *Nature* **183**:308-309.

Jones, A.L., K. Shirasu and C.I. Kado. 1994. The product of the *vir*B4 gene of *Agrobacterium tumefaciens* promotes accumulation of VirB3 protein. *J. Bacteriol.* **176**:5255-5261.

Jutte, S.M. and R.I. Durbin. 1979. Ultrastructural effects in zinnia leaves of a chlorosis-inducing toxin from *P. tagetis*. *Phytopathology* **69**:839-842.

Kado, C.I. 1994. Promiscuous DNA transfer system of *Agrobacterium tumefaciens*: role of the *vir*B operon in sex pilus assembly and synthesis. *Mol. Microbiol.* **12**:17-22.

Kado, C.I. 1996. Evolution of plasmids cuminates in microbial specialization for optimized survival in nature. *Proc. Plasmid Biol. '96.* p. 90.

Kearney, B. and B. Staskawicz. 1990. Widespread distribution and fitness contribution of *Xanthomonas campestris* avirulence gene *avrBs2*. *Nature.* **346**:385-386.

Klee, H.J., R.B. Horsch, M.A. Hinchee, M.B. Hein and N.L. Hoffman. 1987. *Agrobacterium*-mediated plant transformation and its further applications to plant biology. *Ann. Rev. Plant Physiol.* **38**:467-486.

Klett, P. and P. Rott. 1994. Inoculum sources for the spread of leaf scald disease of sugarcane caused by *Xanthomonas albilineans*. *J. Phytopathol.* **142**:283-291.

Koike, H. 1965. The aluminium-cap method for testing sugarcane varieties against leaf scald disease. *Phytopathology* **55**:317-319.

Koike, H. 1971. Testing sugarcane varieties for leaf scald disease resistance. *Proc. Int. Soc. Sugar Cane Technol.* **14**:909-919.

Koike, H. 1992. Leaf scald disease symposium held at Belle Glade, Florida. *Sugar J.* **52**:26-28.

Koraimann, G., G. Hammerle, E. Rassi, K. Bischof and G. Hogenauer. 1996. Lytic transglycosylases in macromolecular transport systems - protein 19 of plasmid R1 and its role in conjugal DNA transport. *Proc. Plasmid Biol. '96*, pp.18-19.

Kornacki, J.A., R.S. Burlage and D.H. Figurski. 1990. The *kil-kor* regulon of broad-host-range plasmid RK2: nucleotide sequence, polypeptide product and expression of regulatory gene *kor*C. *J. Bacteriol.* **172**:3040-3050.

Koski, P., H. Saarilahti, S. Sukupolvi, S. Taira, P. Riikonen, K. Osterlund, R. Hurme and M. Rhen. 1992. A new α-helical coil protein encoded by the *Salmonella typhimurium* virulence plasmid. *J. Biol. Chem.* **267**:12258-12265.

Kyte, J. and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**:105-132.

Larsen, M.H. and D.H. Figurski. 1994. Structure, expression and regulation of the *kil*C operon of promiscuous IncPα plasmids. *J. Bacteriol.* **176**:5022-5032.

Laurent, J., M.-A. Barny, A. Kotoujansky, P. Dufriche and J.L. Vanneste. 1989. Characterization of a ubiquitous plasmid in *Erwinia amylovora*. *Mol. Plant-Microbe Interact.* **2**:160-164.

Lazo, G.R. and D.W. Gabriel. 1987. Conservation of plasmid DNA sequences and pathovar identification of strains of *Xanthomonas campestris*. *Phytopathology* **77**:448-453.

Lazo, G.R., R. Roffey and D.W. Gabriel. 1987. Pathovars of *Xanthomonas campestris* are distinguishable by restriction fragment length polymorphisms. *Int. J. Syst. Bacteriol.* **37**:214-221.

Lederberg, J. and E.L. Tatum. 1946. Gene recombination in *Escherichia coli. Nature* **158**: 558.

Lee, C.A. 1997. Type III secretion systems: machines to deliver bacterial proteins into eukaryotic cells? *Trends Microbiol.* **5**:148-156.

Lee, H.A. 1923. Gum diseases of sugarcane in the Philippines. *Phytopathology* **13**:504.

Lehnherr, H., A.M. Hansen and T. Ilyina. 1998. Penetration of the bacterial cell wall: a family of lytic transglycosylases in bacteriophages and conjugative plasmids. *Mol. Microbiol.* **30**:453-457.

Lessl, M., D. Balzer, W. Pansegrau and E. Lanka. 1992. Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. *J. Biol. Chem.* **267**:20471-20480.

Liu, C.-Q., R.J. Lucas, J.M. Daly, P.A.D. Rickard and N.W. Dunn. 1987. Cloning of *Xanthomonas* DNA that expresses D-xylose catabolic enzymes. *J. Biotechnol.* **6**:159-165.

Lucas, R.J., R.A. Austen and N.W. Dunn. 1987. Cloning of genes encoding endoglucanases from a cellulolytic xanthomonad. *J. Biotechnol.* **6**:83-90.

Maas, R.M., J. Gotz, W. Wohlleben and G. Muth. 1998. The conjugative plasmid pSG5 from *Streptomyces ghanaensis* DSM 2932 differs in its transfer functions from other *Streptomyces* rolling-circle-type plasmids. *Microbiology* **144**:2809-2817.

Maniatis, T., E.F. Fritsch and J. Sambrook. 1982. Molecular cloning a laboratory manual, Cold Spring Harbor Laboratory, Cold Spring, NY.

Marsin, S., Y. Zivanovic, N. Benbouzid-Rollet, D. Prieur and P. Forterre. 1996. Plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidences for rolling circle replication in a hyperthermophile. *Proc. Plasmid Biol. '96*, p. 54.

Martin, J.P. 1950. Sugar-cane diseases and their world distribution. *Proc. Int. Soc. Sugar Cane Technol.* 7:435-452.

Martin, J.P. and P.E. Robinson. 1961. Leaf scald, p. 78-107. *In:* J. P. Martin, E. V. Abbot and C. G. Hughes (eds.), Sugar-cane diseases of the world. Vol. I. Elsevier Publishing Co., Amsterdam.

Masterson, R.V., P.R. Russell and A.G. Atherly. 1982. Nitrogen fixation (*nif*) genes and large plasmids of *Rhizobium japonicum. J. Bacteriol.* **152**:928-931.

Mazarei, M. and A. Kerr. 1991. Plasmids in *Pseudomonas syringae* pv. *pisi* carry genes for pathogenicity. *Plant Pathol.* **40**:408-414.

Meletzus, D. and R. Eichenlaub. 1991. Transformation of the phytopathogenic bacterium *Clavibacter michiganense* subsp. *michiganense* by electroporation and development of a cloning vector. *J. Bacteriol.* **173**:184-190.

Mergeay, M., P. Lejeune, A. Sadouk, J. Gerits and L. Fabry. 1987. Shuttle transfer (or retrotransfer) of chromosomal markers mediated by plasmid pULB113. *Mol. Gen. Genet.* **209**:61-70.

Mitchell, R.E. 1976. Bean halo-blight toxin. *Nature* **260**:75-76.

Moffett, M.L. and B.J. Croft. 1983. *Xanthomonas*, p. 217-220. *In*: P.C. Fahy and G.J. Persley (eds), Plant Bacterial Diseases: a diagnostic guide. Academic Press, New York.

Mogen, B.D., A.E. Oleson, R.B. Sparks, N.C. Gudmestad and G.A. Secor. 1988. Distribution and partial characterization of pCS1, a highly conserved plasmid present in *Clavibacter michiganense* subsp. *sepedonicum. Phytopathology* **78**:1381-1386.

Morales, V.M. and L. Sequeira. 1985. Indigenous plasmids in *Pseudomonas solanacearum. Phytopathology* **75**:767-771.

Murillo, J. and N.T. Keen. 1994. Two native plasmids of *Pseudomonas syringae* pathovar tomato strain PT23 share a large amount of repeated DNA, including replication sequences. *Mol. Microbiol.* **12**:941-950.

Murillo, J., H. Shen, D. Gerholh, A. Sharma, D.A. Cooksey and N.T. Keen. 1994. Characterization of pT23B, the plasmid involved in syringolide production by *Pseudomonas syringae* pv.*tomato* PT23. *Plasmid* **31**:275-287.

Nicholas, K.B. and H.B. Nicholas Jnr. 1997. Genedoc: a tool for editing and annotating multiple sequence alignments. Distribution by the author. www.cris.com/~ketchup/genedoc.shtml

Osbourn, A.E., B.R. Clarke and M.J. Daniels. 1990. Identification and DNA sequence of a pathogenicity gene of *Xanthomonas campestris* pv. *campestris. Mol. Plant-Microbe Interact.* **3**:280-285.

Ovalle, W., J. Comstock, J. Huarez and G. Soto. 1995. First report of leaf scald of sugarcane (*Xanthomonas albilineans*) in Guatemala. *Plant Dis.* **79**:212.

Panopoulos, N.J. and R.C. Peet. 1985. The molecular genetics of plant pathogenic bacteria and their plasmids. *Ann. Rev. Phytopathol.* **23**:381-419.

Pansegrau, W., E. Lanka, P.T. Barth, D.H. Figurski, D.G. Giuney, D. Haas, D.R. Helinski, H, Schwab, V.A. Stanisich and C.M. Thomas. 1994. Complete nucleotide sequence of Birmingham IncPα plasmids: compilation and comparative analysis. *J. Mol. Biol.* **239**:623-663.

Perlin, M.H. 1991. Plasmids other than F, p. 123-155. *In*: U.N Steips and R.E. Yasbin (eds), Modern microbial genetics. Wiley-Liss, New York.

Permaul, K. 1994. Molecular characterisation and detection of *Xanthomonas albilineans*, the sugarcane leaf scald pathogen. M.Sc. Dissertation, University of Durban-Westville, Durban, South Africa.

Permaul, K., D. Pillay and B. Pillay. 1996. Random-amplified polymorphic DNA (RAPD) analysis shows intraspecies differences among *Xanthomonas albilineans* strains. *Lett. Appl. Microbiol.* **23**:307-311.

Permaul, K., S. Ramdeen and D. Pillay. 1999. Conjugative plasmids in *Xanthomonas albilineans*. *Proc. 37th Congress South African Soc. Plant Pathol.* p. 76.

Pillay, D., B. Pillay, E.A. Wachters and L. Korsten. 1993. Electrophoretic and immunological analysis of lipopolysaccharides of *Xanthomonas albilineans* from three geographical regions. *Lett. Appl. Microbiol.* **21**:210-214.

Pohlman, R.F., H.D. Genetti and S.C. Winans. 1994a. Common ancestry between IncN conjugal transfer genes and macromolecular export systems of plant and animal pathogens. *Mol. Microbiol.* **14**:655-668.

Pohlman, R.F., H.D. Genetti and S.C. Winans. 1994b. Entry exclusion of the IncN plasmid pKM101 is mediated by a single hydrophilic protein containing a lipid attachment motif. *Plasmid* **31**:158-165.

Porter, R.D. 1991. Conjugation, p. 157-189. *In*: U.N Streips and R.E. Yasbin (eds), Modern microbial genetics. Wiley-Liss, New York.

Prescott, L.M., J.P. Harley and D.A. Klein. 1996. Microbiology, third edition. Wm. C. Brown Publishers, Dubuque.

Pugsley, A.P. 1993. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **57**:50-108.

Rahaman, W.F. 1995. Virulence and biochemical properties of a plasmid in a South African isolate of *Xanthomonas albilineans*. B. Sc.(Hons) Dissertation, University of Durban-Westville, Durban, South Africa.

Rahaman, W.F., K. Permaul, D. Pillay and B. Pillay. 1995. Virulence and biochemical properties in a South African isolate of *Xanthomonas albilineans*. *Proc. 34th Congress South African Soc. Microbiol.* p. 9.

Ramdeen, S. 1999. Molecular analysis of plasmids from Guadeloupean strains of *Xanthomonas albilineans*. M. Sc. Dissertation, University of Durban-Westville, Durban, South Africa.

Ream, W. 1991. The essentials of *Agrobacterium* genetics, p. 431-453. *In*: U.N Steips and R.E. Yasbin (eds), Modern Microbial Genetics. Wiley-Liss, New York.

Ricaud, C. and C.C. Ryan. 1989. Leaf scald, p. 39-53. *In*: C. Ricaud, B.T. Egan, A.G. Gillaspie Jr., and C.G. Hughes (eds.), Diseases of sugarcane - major diseases. Elsevier Science Publishers B. V., Amsterdam.

Rott, P. 1993. *Xanthomonas albilineans*: cause of leaf scald of sugarcane, p. 70-75. *In:* J.G. Swings and E.L. Civerolo (eds), *Xanthomonas*, Chapman and Hall, London.

Rott, P.C., M. Abel, D. Soupa, P. Feldmann and P. Letourmy. 1994a. Population dynamics of *Xanthomonas albilineans* in sugarcane plants as determined with an antibiotic-resistant mutant. *Plant Dis.* **78**:241-247.

Rott, P., M. Arnaud and P. Baudin. 1986. Serological and lysotypical variability of *Xanthomonas albilineans* (Ashby) Dowson, causal agent of sugarcane leaf scald disease. *J. Phytopathol.* **116**:201-211.

Rott, P. and P. Chagvardieff. 1987. In vitro reactions of sugarcane (*Saccharum* sp.) plantlets inoculated with 2 strains of *Xanthomonas albilineans* (Ashby) Dowson. *Plant Cell, Tissue and Organ Culture* **9**:9-18.

Rott, P.C., L. Costet, M.J. Davis, R. Frutos and D.W. Gabriel. 1996. At least two separate gene clusters are involved in albicidin production by *Xanthomonas albilineans. J. Bacteriol.* **178**:4590-4596.

Rott, P.C., M.J. Davis and P. Baudin. 1994b. Serological variability in *Xanthomonas albilineans*, causal agent of leaf scald disease of sugarcane. *Plant Pathol.* **43**:344-349.

Rott, P.C., D. Soupa, Y. Brunet, P. Feldmann and P. Letourmy. 1995. Leaf scald (*Xanthomonas albilineans*) incidence and its effect on yield in seven sugarcane cultivars in Guadeloupe. *Plant Pathol.* **44**:1075-1084.

Rudolph, K. 1993. Infection of plants by *Xanthomonas*, p. 193-264. *In:* J.G. Swings and E.L. Civerolo (eds), *Xanthomonas*, Chapman and Hall, London.

Salch, Y.P. and P.D. Shaw. 1988. Isolation and characterization of pathogenicity genes of *Pseudomonas syringae* pv. *tabaci. J. Bacteriol.* **170**:2584-2591.

Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. Molecular cloning: A laboratory manual. Second edition, Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.

Sampath Kumar, S.N. 1984. A study in toxin production by *Xanthomonas campestris* pv. *arecae* incitant of bacterial leaf stripe of arecanut. *Indian Phytopathol.* **37**:516-519.

Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**:5463-5467.

Shimura, M., M. Iwata, S. Omoto and Y. Sekizawa. 1979. The structure of SF-185 substance: a new antibiotic against *Xanthomonas* plant pathogens. *Agri. Biol. Chem.* **11**:2279-2281.

Shirasu, K., Z. Koulkolikova-Nicola, B. Hohn and C.I. Kado. 1994. An inner-membrane associated virulence protein essential for T-DNA transfer from *Agrobacterium tumefaciens* to plants exhibits ATPase activity and similarities to conjugative transfer genes. *Mol. Microbiol.* **11**:581-588.

Sia, E.A., D.M. Kuehner and D.H. Figurski. 1996. Mechanism of retrotransfer in conjugation: prior transfer of the conjugative plasmid is required. *J. Bacteriol.* **178**:1457-1464.

Silverman, P.M. 1997. Towards a structural biology of bacterial conjugation. *Mol. Microbiol.* **23**:423-429.

Skurray, R.A. and N.A. Firth. 1997. Conjugation: Macromolecular transport mechanisms in Gram-negative and Gram-positive bacteria. *Plasmid* **37**:230.

Stall, R.E., D.C. Loschke and J.B. Jones. 1986. Linkage of copper resistance and avirulence loci on a self-transmissible plasmid in *Xanthomonas campestris* pv. *vesicatoria*. *Phytopathology* **76**:240-243.

Steinberger, E.M., G-Y. Cheng and S.V. Beer. 1990. Characterization of a 56-kb plasmid from *Erwinia amylovora* Ea322: its noninvolvement in pathogenicity. *Plasmid* **24**:12-24.

Steindl, D.R.L. 1971. The elimination of leaf scald from infected plant material. *Proc. Int. Soc. Sugar Cane Technol.* **14**:925-929.

Steips, U.N. 1991. Transformation, p. 191-216. *In* U.N. Steips and R.E. Yasbin (eds), Modern Microbial Genetics. Wiley-Liss, New York.

Stephens, K.M., C. Roush and E. Nester. 1995. *Agrobacterium tumefaciens* VirB11 proetin requires a consensus nucleotide-binding site for function in virulence. *J. Bacteriol.* **177**: 27-36.

Strobel, G.A. 1977. Bacterial phytotoxins. *Ann. Rev. Microbiol.* **31**:205-224.

Sundin, G.W., A.L. Jones and D.W. Fulbright. 1989. Copper resistance in *Pseudomonas syringae* pv. *syringae* from cherry orchards and its associated transfer in vitro and in planta with a plasmid. *Phytopathology* **79**:861-865.

Surico, G. and J.E. De Vay. 1982. Effect of syringomycin and syringotoxin produced by *Pseudomonas syringae* pv. *syringae* on structure and function of mitochondria isolated from holcus spot resistant and susceptible maize lines. *Physiol. Plant Pathol.* **21**:39-53.

Swarup, S., R. De Feyter, R.H. Brlansky and D.W. Gabriel. 1991. A pathogenicity locus from *Xanthomonas campestris* to elicit cankerlike lesions on citrus. *Phytopathology* **81**: 802-809.

Swings, J., L. Vauterin and K. Kersters. 1993. The bacterium *Xanthomonas*, p. 121-156. *In:* J.G. Swings and E.L. Civerolo (eds), *Xanthomonas*, Chapman and Hall, London.

Taylor, J., R.S. Stearman and B.B. Uratini. 1993. Development of a native plasmid as a cloning vector in *Clavibacter xyli* subsp. *cynodontis*. *Plasmid* **29**:241-244.

Taylor, P.W.J., C.C. Ryan and R.G. Birch. 1988. Harvester transmission of leaf scald and ratoon stunting disease. *Sugar Cane* **4**:11-14.

Thatte, V., D.E. Bradley and V.N. Iyer. 1985. N conjugative transfer system of plasmid pCU1. *J. Bacteriol.* **163**:1229-1236.

Thiry, G. 1984. Plasmids of the epiphytic bacterium *Erwinia uredovora*. *J. Gen. Microbiol.* **130**:623-631.

Thomas, C.M. and C.A. Smith. 1987. Incompatibility group P plasmids: genetics, evolution, and use in genetic manipulation. *Ann. Rev. Microbiol.* **41**:77-101.

Thomas, C.M., C.A. Smith, J.P. Ibbotson, L. Johnston and N. Wang. 1995. Evolution of the *korA-oriV* segment of promiscuous IncP plasmids. *Microbiology* **141**:1201-1210.

Thomson, G.M. 1969. Leaf scald in South Africa and Swaziland. *Sugarcane Pathol. Newsl.* **3**:4.

Thompson, V. 1982. A rapid method of testing the resistance of sugarcane to leaf scald disease. *Sugarcane Pathol. Newsl.* **29**:34:35.

Turner, P., C. Barber and M. Daniels. 1985. Evidence for clustered pathogenicity genes in *X. campestris* pv. *campestris*. *Mol. Gen. Genet.* **199**:338-343.

Turner, S.L., L. Rigottier-Gois, N. Amarger and J.P.W. Young. 1997. Multiple related *repC* plasmid replication sequences in *Rhizobium leguminosarum*. *Plasmid* **37**:233.

Ulaganathan, K. and A. Mahadevan. 1988. Isolation of a plasmid from *Xanthomonas campestris* pv. *vignicola*. *J. Phytopathol.* **123**:92-96.

Van den Mooter, M. and J. Swings. 1990. Numerical analysis of 295 phenotypic features of 266 *Xanthomonas* strains and related strains and an improved taxonomy of the genus. *Int. J. Syst. Bacteriol.* **40**:348-369.

Van Larebeke, N., G. Engler, M. Holsters, S. Van den Elsacker, I. Zaenen, R.A. Schilperoort and J. Schell. 1974. Large plasmid in *Agrobacterium tumefaciens* essential for crown gall inducing ability. *Nature* **252**:169-171.

Vauterin, L., B. Hoste, P. Yang, A. Alvarez, K. Kersters and J. Swings. 1993. Taxonomy of the genus *Xanthomonas*, p. 157-192. *In*: J.G. Swings and E.L. Civerolo (eds), *Xanthomonas*, Chapman and Hall, London.

Vauterin, L. and J. Swings. 1997. Are classification and phytological diversity compatible in *Xanthomonas*? *J. Ind. Microbiol. Biotechnol.* **19**:77-82.

Walker, M.J., R.G. Birch and J.M. Pemberton. 1988. Cloning and characterization of an albicidin resistance gene from *Klebsiella oxytoca. Mol. Microbiol.* **2**:443-454.

Wall, M.K. and R.G. Birch. 1997. Genes for albicidin biosynthesis and resistance span at least 69 kb in the genome of *Xanthomonas albilineans. Lett. Appl. Microbiol.* **24**:256-260.

Watson, B., T.C. Currier, M.P. Gordan, M.D. Chilton and E.W. Nester. 1975. Plasmid required for virulence in *Agrobacterium tumefaciens. J. Bacteriol.* **123**:255-264.

Weiss, A.A., F.D. Johnson and D.I. Burns. 1993. Molecular characterization of an operon required for pertussis toxin secretion. *Proc. Natl. Acad. Sci. USA* **90**:2970-2974.

Wilkins, B.M. 1996. Establishment of the transferred plasmid in the new cell. *Proc. Plasmid Biol. '96*, p. 15-16.

Wilson, J.W., E.A. Sia and D.H. Figurski. 1997. The *kilE* locus of promiscuous plasmid RK2 is required for stable maintenance in *Pseudomonas aeruginosa. Plasmid* **37**:234-235.

Winans, S.C., D.L. Burns and P.J. Christie. 1996. Adaptation of a conjugal transfer system for the export of pathogenic macromolecules. *Trends Microbiol.* **4**:64-68.

Winans, S.C. and G.C. Walker. 1985a. Conjugal transfer system of the IncN plasmid pKM101. *J. Bacteriol.* **161**:402-410.

Winans, S.C. and G.C. Walker. 1985b. Entry exclusion determinant(s) of IncN plasmid pKM101. *J. Bacteriol.* **161**:411-416.

Winans, S.C. and G.C. Walker. 1985c. Fertility inhibition of RP1 by IncN plasmid pKM101. *J. Bacteriol.* **161**:425-427.

Winans, S.C. and G.C. Walker. 1985d. Identification of pKM101-encoded loci specifying potentially lethal gene products. *J. Bacteriol.* **161**:417-424.

Yang, Y. and D.W. Gabriel. 1995. Intragenic recombination of a single plant pathogen gene provides a mechanism for the evolution of new host specificities. *J. Bacteriol.* **177**:4963-4968.

Yang, P., P. Rott, L. Vauterin, B. Hoste, P. Baudin, K. Kersters and J. Swings. 1993. Intraspecific variability of *Xanthomonas albilineans. Syst. Appl. Microbiol.* **16**:420-426.

Young, C., D.H. Bechhofer and D.H. Figurski. 1984. Gene regulation in plasmid RK2: positive control by *kor*A in the expression of *kor*C. *J. Bacteriol.* **157**:247-252.

Zhang, L. and R.G. Birch. 1996. Biocontrol of sugar cane leaf scald disease by an isolate of *Pantoea dispersa* which detoxifies albicidin phytotoxins. *Lett. Appl. Microbiol.* **22**:132-138.

Zhang, L. and R.G. Birch. 1997. Mechanisms of biocontrol by *Pantoea dispersa* of sugar cane leaf scald disease caused by *Xanthomonas albilineans*. *J. Appl. Microbiol.* **82**:448-454.

APPENDIX ONE

a) ORF1

```
  1  ATGCACATTT TGTTGCTACA CTATCGCCAC AAATGTATTG CAGAAGATGC
     TACGTGTAAA ACAACGATGT GATAGCGGTG TTTACATAAC GTCTTCTACG

 51  TATCATGCGG CGACATAACC GGTATCCCGA GCTACAATTT TTGATTTTGA
     ATAGTACGCC GCTGTATTGG CCATAGGGCT CGATGTTAAA AACTAAAACT

                               ClaI
                              ~~~~~~
101  TGTACGGGAT AGCCATGAAC GCAATCGATG AGCAAGAACT GATGGACTAC
     ACATGCCCTA TCGGTACTTG CGTTAGCTAC TCGTTCTTGA CTACCTGATG

151  GCGGCAAACG GGGTCATTTC GCAGTTCAAG ATTCGCGAGA CTGACGAAGA
     CGCCGTTTGC CCCAGTAAAG CGTCAAGTTC TAAGCGCTCT GACTGCTTCT

201  CCGGTTTTGC CTAGTTGTAA CCGTATCGTG GAAAGAGGGG GACTGCATCC
     GGCCAAAACG GATCAACATT GGCATAGCAC CTTTCTCCCC CTGACGTAGG

251  TAACAAGCGC ACGCAAGACA CCTCGCGTAT GGGCCAACGT GAATACGCTG
     ATTGTTCGCG TGCGTTCTGT GGAGCGCATA CCCGGTTGCA CTTATGCGAC

301  GCGAATTTCC TTCGGGGCTT AAACCTTCCT AACGTCCCGA TCAATCTGGA
     CGCTTAAAGG AAGCCCCGAA TTTGGAAGGA TTGCAGGGCT AGTTAGACCT

351  ATTGTCATTC AAGGGACCAA CATGA
     TAACAGTAAG TTCCCTGGTT GTACT
```

b) ORF2

```
  1  ATGCCCAAGA AAACAACTCC CGCCCGCACC GCCTCATCTT CACAGCCTTC
     TACGGGTTCT TTTGTTGAGG GCGGGCGTGG CGGAGTAGAA GTGTCGGAAG

 51  TCTAGCCGCC CTCCTCTGCT GCATTCCTGG CATCGCCATC GCTGGCTCCC
     AGATCGGCGG GAGGAGACGA CGTAAGGACC GTAGCGGTAG CGACCGAGGG

101  CGTTCTCGGG TGGCACGTCT GGCTTGTCGA GCGATCTTGT AGGCATCCTG
     GCAAGAGCCC ACCGTGCAGA CCGAACAGCT CGCTAGAACA TCCGTAGGAC

151   ACGCCAATTG CTGGCATCGC CATGATCGCA GTCGCCATTC TTTGCTGGTT
     TGCGGTTAAC GACCGTAGCG GTACTAGCGT CAGCGGTAAG AAACGACCAA

                               MluNI
                              ~~~~~~
201  CGGCAAAATC TCATGGTGGT GGTTGGCCAG CATCGTAGTT GGCATTGTCC
     GCCGTTTTAG AGTACCACCA CCAACCGGTC GTAGCATCAA CCGTAACAGG

251  TGTTCTTCGG CAAGGATCAA GTCGTGAGCT GGATTCGCGG GCTGTTCGGC
     ACAAGAAGCC GTTCCTAGTT CAGCACTCGA CCTAAGCGCC CGACAAGCCG

301  GATTTAAGCG AGAGACTCCC AATACTATGA
     CTAAATTCGC TCTCTGAGGG TTATGATACT
```

c) ORF3

```
                                  HindII
                                   ~~~~~~
  1  ATGAGCAAGA ACGCCGGAAT CATCGTTGAC ACGTTGTTCG TCGGGCCGAC
     TACTCGTTCT TGCGGCCTTA GTAGCAACTG TGCAACAAGC AGCCCGGCTG

 51  GCGTCCCCAC GATGTTATGG GGCGTGACGT GGCAAGCCTT CGTCATCAAC
     CGCAGGGGTG CTACAATACC CCGCACTGCA CCGTTCGGAA GCAGTAGTTG

101  ATCATCGTGA CTATGGAAGC CTTCATCTGG ACACGCGATC TGCGCTGGCT
     TAGTAGCACT GATACCTTCG GAAGTAGACC TGTGCGCTAG ACGCGACCGA

                                                      SphI
                                                      ~~~~~~
151  TTTATTGTTT GTCCCCGATT CACGGCATCT GCTTCCCTGA TCTGCATGCG
     AAATAACAAA CAGGGGCTAA GTGCCGTAGA CGAAGGGACT AGACGTACGC

201  CGATCCACGC ACGTTTGAAT TGCTCATGCA GTGGGGGCCG ACGAAAGGCT
     GCTAGGTGCG TGCAAACTTA ACGAGTACGT CACCCCCGGC TGCTTTCCGA

                                                        StyI
                                                        ~~
251  TCGCCTTCTT CGGCAATTTC CCATACTGGC GAGCCGCCAC ATTCAGTCCC
     AGCGGAAGAA GCCGTTAAAG GGTATGACCG CTCGGCGGTG TAAGTCAGGG

     StyI
     ~~~~
301  TTGGAACTTC ACACCCCCCC CAAACGATCC TTCTTCCGTA GGAGGAAATC
     AACCTTGAAG TGTGGGGGGG GTTTGCTAGG AAGAAGGCAT CCTCCTTTAG

351  GAGGAAGGTG CCATAATGA
     CTCCTTCCAC GGTATTACT
```

d) ORF4

```
  1  ATGAAAGCGA CGGTAGATGC AAAGCGCGAG CGTGCCATTC TGCGTGAGCC
     TACTTTCGCT GCCATCTACG TTTCGCGCTC GCACGGTAAG ACGCACTCGG

 51  AGCTTTGTCG AAGAACATCC CTTACTCCGT ACACCTGACA CCGACTGCTA
     TCGAAACAGC TTCTTGTAGG GAATGAGGCA TGTGGACTGT GGCTGACGAT

                                                      NheI
                                                      ~~~
101  TCCAGACGGA GCACCATGAT TACCTCATGG TGCTGCGCTT GACTGGCGCT
     AGGTCTGCCT CGTGGTACTA ATGGAGTACC ACGACGCGAA CTGACCGCGA

     NheI                          HindII
     ~~~                           ~~~~~~~
151  AGCTTCGAGA GCGCCGACGA TGAGCAAGTC AACAACTGGC ACCATCGTCT
     TCGAAGCTCT CGCGGCTGCT ACTCGTTCAG TTGTTGACCG TGGTAGCAGA

201  GAATGGCCTT CTGCGAAGCA TCGCATCGCA CAATGTCGCG ATTTGGCAGC
     CTTACCGGAA GACGCTTCGT AGCGTAGCGT GTTACAGCGC TAAACCGTCG

251  ACATCGTTCG CCGACCGGAA AATAAATATC CAGACGGGGA ATTTCCCGAA
     TGTAGCAAGC GGCTGGCCTT TTATTTATAG GTCTGCCCCT TAAAGGGCTT

301  GGCTTTGCCG CCGATCTCAA TAAGAAGTAC GCAGCTCGCG TGAGCGGCGA
     CCGAAACGGC GGCTAGAGTT ATTCTTCATG CGTCGAGCGC ACTCGCCGCT
```

```
                HindII
                ~~~~~~
 351  ATTGCTTATG GTCAACGAAC CTGCATTTGA CTGGTCCGTG TATCGACCTC
      TAACGAATAC CAGTTGCTTG GACGTAAACT GACCAGGCAC ATAGCTGGAG

 401  AGCCCACAAT TGGTCGGCAA AGCCTTTGGT CGTTGGTTTC CTCGCCCACT
      TCGGGTGTTA ACCAGCCGTT TCGGAAACCA GCAACCAAAG GAGCGGGTGA

 451  CCCCAAGCGT TCGCGCAAGA ACGGGCAGAG TCAATCGACG CGCTCGAAAA
      GGGGTTCGCA AGCGCGTTCT TGCCCGTCTC AGTTAGCTGC GCGAGCTTTT

 501  GGTAGTGCGT GAAGTCGAAT CGTCGCTGTC TCGTTACGAC GTTGAACGGC
      CCATCACGCA CTTCAGCTTA GCAGCGACAG AGCAATGCTG CAACTTGCCG

 551  TCTGCATCTA TGAGCACAAC GGCGTGTATT TCAGTGAGCC CCTAGAGTTA
      AGACGTAGAT ACTCGTGTTG CCGCACATAA AGTCACTCGG GGATCTCAAT

 601  TTTGCATTCC TGGTCAATAG TGAGTGGCAA CGTATTCCGC TCGCCCAAGC
      AAACGTAAGG ACCAGTTATC ACTCACCGTT GCATAAGGCG AGCGGGTTCG

 651  ACCGTTGCGC ACTCTAATTC CCACTACGCG CCCCTTTTTT GGTAACGAAG
      TGGCAACGCG TGAGATTAAG GGTGATGCGC GGGGAAAAAA CCATTGCTTC

 701  CAATCGAACT GCGCTCGCCG ACGAAAACGA CGTATGGCGC GATGCTCGGC
      GTTAGCTTGA CGCGAGCGGC TGCTTTTGCT GCATACCGCG CTACGAGCCG

 751  ATCAACGCTT ACCCCCCCGA ATCAAAGTCG GTGTTCTTGA ATCACTTGCT
      TAGTTGCGAA TGGGGGGGCT TAGTTTCAGC CACAAGAACT TAGTGAACGA

 801  CACGCAGCCG TTTTCATTTG TGCTGTCGCA ATCGTTCAGC TTCCTTCAAA
      GTGCGTCGGC AAAAGTAAAC ACGACAGCGT TAGCAAGTCG AAGGAAGTTT

 851  TGGAAAGCGC ACGCTGGAAG CTGAAACTGT CGAAGAATCG GATGATTAAC
      ACCTTTCGCG TGCGACCTTC GACTTTGACA GCTTCTTAGC CTACTAATTG
                                        ClaI
                                        ~~~~~~
 901  GCAGGCGATG ATGCCCTCTC GCAAGTCGAT GAAATCGATG ATGCGGTCGA
      CGTCCGCTAC TACGGGAGAG CGTTCAGCTA CTTTAGCTAC TACGCCAGCT

 951  TGATCTAACC GCACGTCGGT GGGTGATGGG CGACCATCAC TTTAGCTTGT
      ACTAGATTGG CGTGCAGCCA CCCACTACCC GCTGGTAGTG AAATCGAACA

1001  TCGTGAAAGC CGGCAGCTTG CGCGAACTGA ACGATCACAT CGCCGAAGCG
      AGCACTTTCG GCCGTCGAAC GCGCTTGACT TGCTAGTGTA GCGGCTTCGC

1051  CGAACGGCTC TGTCAGAAGG CGGCATCACC GCTGCGCGCG AAGACCTCGC
      GCTTGCCGAG ACAGTCTTCC GCCGTAGTGG CGACGCGCGC TTCTGGAGCG

1101  GATTGCTTCG GCCTTTTGGG CGCAGCTACC GGCGCAGTTC AAGTTCCGTC
      CTAACGAAGC CGGAAAACCC GCGTCGATGG CCGCGTCAAG TTCAAGGCAG

1151  CCCGCCTGTC GCCGATCAAC AGCAAAAACA TGGCCGGCTT CGCCCCGTTG
      GGGCGGACAG CGGCTAGTTG TCGTTTTTGT ACCGGCCGAA GCGGGGCAAC

                              StyI
                              ~~~~~~
1201  CACAACTTCC CGCAAGGCCG GCGCCATGGC AACCATTGGG GCGATGCACT
      GTGTTGAAGG GCGTTCCGGC CGCGGTACCG TTGGTAACCC CGCTACGTGA

1251  CACGATGTTC ATCACGTCGG CCAACACGCC CTACTATTTC AGCTTCCACG
      GTGCTACAAG TAGTGCAGCC GGTTGTGCGG GATGATAAAG TCGAAGGTGC
```

```
1301  CGGCAGACCC CTTTGACGAA AGCGGCGGCA CGAAAAAGGA CGTTGGACAT
      GCCGTCTGGG GAAACTGCTT TCGCCGCCGT GCTTTTTCCT GCAACCTGTA

1351  ACCCTCGTGC TCGGCCCGAC CGGCAGCGGC AAGACCGCGC TCATTGCCTT
      TGGGAGCACG AGCCGGGCTG GCCGTCGCCG TTCTGGCGCG AGTAACGGAA

1401  CCTGCTGTGC ATGTTGCAGA AGTTTGGCGT GACTTCGGTT CTCTTCACGA
      GGACGACACG TACAACGTCT TCAAACCGCA CTGAAGCCAA GAGAAGTGCT

1451  AGGACCGCGA TACCGAACTG GTCATTCGTG CCCTCGGTGG CACCTACTAC
      TCCTGGCGCT ATGGCTTGAC CAGTAAGCAC GGGAGCCACC GTGGATGATG


1501  CCGATCAAAC CAGGCGAGCC GACCGGCTGG AATCCGTTCT GGCTCGACCC
      GGCTAGTTTG GTCCGCTCGG CTGGCCGACC TTAGGCAAGA CCGAGCTGGG

1551  GGCCAAACCT GGCAATGTGC AGTATCTGAA TCGCTTCGTG CGACGCCTTT
      CCGGTTTGGA CCGTTACACG TCATAGACTT AGCGAAGCAC GCTGCGGAAA

1601  GCACCCGCCC CTCTCAGACG TTGAGCGTCA CGGACGAAAT CGAAATTGAA
      CGTGGGCGGG GAGAGTCTGC AACTCGCAGT GCCTGCTTTA GCTTTAACTT

                                                        EagI
                                                         ~~~
1651  CAAGCAATCA ATGCGGTGCT TCGCATGGAC CTCGAACATC GCCGGCTCGG
      GTTCGTTAGT TACGCCACGA AGCGTACCTG GAGCTTGTAG CGGCCGAGCC

      EagI
      ~~~
1701  CCGCGTGCTC GACTTCATGA CGAAAGACCG CAGCGGCATC TATGCGCAGT
      GGCGCACGAG CTGAAGTACT GCTTTCTGGC GTCGCCGTAG ATACGCGTCA

1751  TGCAGCGGTG GTGCTATGCA CGCGAGCACG GCAAGCCCGA TGGCCCGAAC
      ACGTCGCCAC CACGATACGT GCGCTCGTGC CGTTCGGGCT ACCGGGCTTG

                                          ClaI
                                          ~~~~~~
1801  GCTTGGCTCT TCGATAATCC GCGCGACACG CTTATCGATA ACTTCGGTTC
      CGAACCGAGA AGCTATTAGG CGCGCTGTGC GAATAGCTAT TGAAGCCAAG

1851  GGCACTGACT ACCGGCTTCG ACGTAACTTC ATTCCTGAAA GATGATGAAT
      CCGTGACTGA TGGCCGAAGC TGCATTGAAG TAAGGACTTT CTACTACTTA

1901  TACGCTCGCC GATCAACATG CACCTGTTCC ACCTGACCGA AAGCCTCATC
      ATGCGAGCGG CTAGTTGTAC GTGGACAAGG TGGACTGGCT TTCGGAGTAG

                                        EcoRI
                                        ~~~~~~
1951  GACGGTCGCC GCCTCGCACT GTTTATCGCC GAATTCTGGC GTGCGTTGGG
      CTGCCAGCGG CGGAGCGTGA CAAATAGCGG CTTAAGACCG CACGCAACCC

2001  CGACCCAGAA ATGGCTGACT TCGCGAAGGA TAAGTTGAAA ACGATCCGTA
      GCTGGGTCTT TACCGACTGA AGCGCTTCCT ATTCAACTTT TGCTAGGCAT

2051  AAAAGAACGG CTTCGTCGTG CTCGATTCGC AGTCGCCTAG CGACGCTCTG
      TTTTCTTGCC GAAGCAGCAC GAGCTAAGCG TCAGCGGATC GCTGCGAGAC

2101  AACCATCGGA TCAGCCGAAC ACTGATCGAG CAAACGCCGA CGAAGATTCT
      TTGGTAGCCT AGTCGGCTTG TGACTAGCTC GTTTGCGGCT GCTTCTAAGA

2151  ATTCTCCAAC CCTGATGCCG TGTACAGCGA ATACACAAGC GGCGGTCTTA
      TAAGAGGTTG GGACTACGGC ACATGTCGCT TATGTGTTCG CCGCCAGAAT
```

```
                      EcoRI
                      ~~~~~~
2201  ACTGCTCGGA TCGTGAATTC GATCTCGTCA AGAAGCATCC CAGAGGCTCC
      TGACGAGCCT AGCACTTAAG CTAGAGCAGT TCTTCGTAGG GTCTCCGAGG

2251  CGTATGTTCC TTGTCAAGCA AGGGCATCAC TCTGTCGTCG CCAAGCTAGA
      GCATACAAGG AACAGTTCGT TCCCGTAGTG AGACAGCAGC GGTTCGATCT

2301  CCTCCAGGGC TTCGACCGTG AGTTGGCTCT GCTGTCCAGC CGCGAGGCCA
      GGAGGTCCCG AAGCTGGCAC TCAACCGAGA CGACAGGTCG GCGCTCCGGT

2351  ACATCGAAGT CGTCCAGCAG TTGATCGCGC AGTTCGGGCA AGACCCTATT
      TGTAGCTTCA GCAGGTCGTC AACTAGCGCG TCAAGCCCGT TCTGGGATAA

2401  AAATGGCTTC CCCACTTTGA CCAACACAGG AGAGCAGCAT GA
      TTTACCGAAG GGGTGAAACT GGTTGTGTCC TCTCGTCGTA CT
```

e) ORF5

```
   1  ATGAAGAACG CCAAGCATAC CGCCCTTGCC GTGGCCCTTT TCCTTTCCGT
      TACTTCTTGC GGTTCGTATG GCGGGAACGG CACCGGGAAA AGGAAAGGCA

  51  AGCAGTCACG CCGTTGTCTT CGTTCGCGAC CGGTATACCG GTGGTGGACA
      TCGTCAGTGC GGCAACAGAA GCAAGCGCTG GCCATATGGC CACCACCTGT

 101  TTGCAGCCAT CGAGCAAGCG ATCCAGCAAG TCAGCTACTT GCAGCAGCAG
      AACGTCGGTA GCTCGTTCGC TAGGTCGTTC AGTCGATGAA CGTCGTCGTC

 151  TTAAAGCAGA TGAAAAACCA GTTGGATGCG ATGACCGGTG ATCGCGGCAT
      AATTTCGTCT ACTTTTTGGT CAACCTACGC TACTGGCCAC TAGCGCCGTA

 201  GGCTGGCCTT CTCTCCGGTC AAAACCGCAA CTACTTACCC GCTGACTGGA
      CCGACCGGAA GAGAGGCCAG TTTTGGCGTT GATGAATGGG CGACTGACCT

 251  ACAGCGCCAT GAACGTACTG AACAGCGGCG GCGGTTCGTT TGGCTCGCTC
      TGTCGCGGTA CTTGCATGAC TTGTCGCCGC CGCCAAGCAA ACCGAGCGAG

          EagI
          ~~~~~~
 301  GCATCGGCCG CCCAGCAAAT CAAACAGATG CAGTCGGTTC TTTCCAGTTC
      CGTAGCCGGC GGGTCGTTTA GTTTGTCTAC GTCAGCCAAG AAAGGTCAAG

 351  TGACCTGAGC CGGCTGTCTC CGCAAATGCA GCAGTATATT GACAAGGTAC
      ACTGGACTCG GCCGACAGAG GCGTTTACGT CGTCATATAA CTGTTCCATG

 401  GAAGCGTCTC AGCCTCCCAA CAGGCATTGG GGCAACAGGC TTACTCCACG
      CTTCGCAGAG TCGGAGGGTT GTCCGTAACC CCGTTGTCCG AATGAGGTGC

 451  GCCAGCCAGC GCGTCAATCT CCTGCAAACA CTCACCAATC AGATTTCCTC
      CGGTCGGTCG CGCAGTTAGA GGACGTTTGT GAGTGGTTAG TCTAAAGGAG
                  StyI
                  ~~~~~~
 501  AGCTACCGAT CCCAAGGCCG TGTGGGACTT GCAGGCGCGT ATCCAGTCCG
      TCGATGGCTA GGGTTCCGGC ACACCCTGAA CGTCCGCGCA TAGGTCAGGC
                                                    HindII
                                                    ~
 551  AGCAATCGCA GTTGCAGAAC GATCAGTCGC GGCTTCAAAG CGTGGCGCAG
      TCGTTAGCGT CAACGTCTTG CTAGTCAGCG CCGAAGTTTC GCACCGCGTC
```

```
      HindII
      ~~~~~
601   TTGACACAAG CGCAAAGCGT GGCGACCAAG CAAATAGCTA ACGAACTGCG
      AACTGTGTTC GCGTTTCGCA CCGCTGGTTC GTTTATCGAT TGCTTGACGC

651   CTCGCAAACC AGCGGCACTG GCAATTTCCC GTCGCTCGAT ACATCAGTCG
      GAGCGTTTGG TCGCCGTGAC CGTTAAAGGG CAGCGAGCTA TGTAGTCAGC

701   GCACAAACTA A
      CGTGTTTGAT T
```

f) ORF6

```
1     ATGGCCGGCC CATACGAACA GGTCTTCACC TATGTAACGA ATGTATGCGA
      TACCGGCCGG GTATGCTTGT CCAGAAGTGG ATACATTGCT TACATACGCT

51    TAGCTACATC GGTTCGAGCG TCGCGGCGGT CGCCGCAGCC ATTGCCCCCG
      ATCGATGTAG CCAAGCTCGC AGCGCCGCCA GCGGCGTCGG TAACGGGGGC

                                                        SphI
                                                         ~~
101   CCGCTTACAC GCTGCTAGGC GTCTATATCA TGCTTTGGGG TTTAGCCAGC
      GGCGAATGTG CGACGATCCG CAGATATAGT ACGAAACCCC AAATCGGTCG

      SphI
      ~~~~
151   ATGCGCGGCT TAATTCAAGA GCCGATCATG GAAGCCGCCG TTCGTATGAT
      TACGCGCCGA ATTAAGTTCT CGGCTAGTAC CTTCGGCGGC AAGCATACTA

201   AAAGATCGCT TTCATCTTTG GCATCGGTAT CAAGCTGGCC CAGTACAACG
      TTTCTAGCGA AAGTAGAAAC CGTAGCCATA GTTCGACCGG GTCATGTTGC

251   TTTACGTGGT TGATACGGTT TTTAATTCGC CGGAACAGCT TGCGCAGGCG
      AAATGCACCA ACTATGCCAA AAATTAAGCG GCCTTGTCGA ACGCGTCCGC

            EcoRI                       HindII
           ~~~~~~~                      ~~~~~~
301   CTGACGAATT CGACGAGTAA CCAAACCACG GTTAACAGCC TTGACAACAT
      GACTGCTTAA GCTGCTCATT GGTTTGGTGC CAATTGTCGG AACTGTTGTA

             StyI
            ~~~~~~~
351   CCTGACCCAA GGCTTCCAGG TCGGTAAGAG CTTTTGGGAC AAAGGCGGCA
      GGACTGGGTT CCGAAGGTCC AGCCATTCTC GAAAACCCTG TTTCCGCCGT

401   TCCTGGACGG CGACTTCGGC ATGTATCTCA TCGCTATCTC ATGCTGGGCG
      AGGACCTGCC GCTGAAGCCG TACATAGAGT AGCGATAGAG TACGACCCGC

451   ATCACCATCG TCGTGACAGC CTACGCTTGC TTCTTGATAA TTCTCGCCAA
      TAGTGGTAGC AGCACTGTCG GATGCGAACG AAGAACTATT AAGAGCGGTT

501   GATCGCCCTC GCGTTGATTG TCGCCCTCGG TCCTCTGTTC ATCATTTCCC
      CTAGCGGGAG CGCAACTAAC AGCGGGAGCC AGGAGACAAG TAGTAAAGGG

551   TCCTGTTCCA GCCTACGGCG AATTTCTTCA ATGCGTGGAT ACAACAACTG
      AGGACAAGGT CGGATGCCGC TTAAAGAAGT TACGCACCTA TGTTGTTGAC
```

```
                                                        AflIII
                                                        ~~~~~~
601  GCAAACTACG CTCTGCTCGT TATCTTGATC GTGTCCGCGA ACGTGTTCAT
     CGTTTGATGC GAGACGAGCA ATAGAACTAG CACAGGCGCT TGCACAAGTA

                                          EagI
                                          ~~~~~~
651  CTTGAATTTA TTCGTTCGCG CGGCGAGCAC GACGGCCGGC ATTTCGAGCA
     GAACTTAAAT AAGCAAGCGC GCCGCTCGTG CTGCCGGCCG TAAAGCTCGT

701  CGGCGCAAAT CGACCAAATC TTCCCCTTCA TCATCACCGG CGTGATTTCG
     GCCGCGTTTA GCTGGTTTAG AAGGGGAAGT AGTAGTGGCC GCACTAAAGC

751  CTCCTGGTGC TCGCTCAGCT CCCCTCGATA GCTGCCGGCC TGGCAGGCGG
     GAGGACCACG AGCGAGTCGA GGGGAGCTAT CGACGGCCGG ACCGTCCGCC

                                         NheI
                                         ~~~~~~
801  CATATCTCTA TCGTCGTACG GCGTCGGTCG GCTAGCTCTA CGCAAGCTCG
     GTATAGAGAT AGCAGCATGC CGCAGCCAGC CGATCGAGAT GCGTTCGAGC

851  GTGGTCTTTT CGACCGGAAC CGCAGTAACA ACAACCGTCG CGACCGGGAA
     CACCAGAAAA GCTGGCCTTG GCGTCATTGT TGTTGGCAGC GCTGGCCCTT

                                                        EcoRV
                                                        ~~~~~
901  GAACGCAAAA CGCCGCGCAC AGCGTCAATT CGGTATCACG CGGCTGATAT
     CTTGCGTTTT GCGGCGCGTG TCGCAGTTAA GCCATAGTGC GCCGACTATA

     EcoRV
     ~
951  CCGTCGCCCC CGTCTCCCCC CAACAAAACA TTCGCATTTG TTGAACAAAA
     GGCAGCGGGG GCAGAGGGGG GTTGTTTTGT AAGCGTAAAC AACTTGTTTT

     AflIII
     ~~~~~~
1001 ACACGTTTTT TGTTTTAGCA TAA
     TGTGCAAAAA ACAAAATCGT ATT
```

g) ORF7

```
1    ATGCTCGTAA GAACACTTCT CGCAGTGCTC TTGGCCGGCA CCCTCGTCGG
     TACGAGCATT CTTGTGAAGA GCGTCACGAG AACCGGCCGT GGGAGCAGCC

51   TTGTGCCTCG GTGCCGCCCT CATGCGACGG CTTGAACCGC CGCCCCGTCA
     AACACGGAGC CACGGCGGGA GTACGCTGCC GAACTTGGCG GCGGGGCAGT

101  ATCAGCCACC GCAAGCCGGC GTCGTGAATC AAAGCTGCGG CCATTCCGCC
     TAGTCGGTGG CGTTCGGCCG CAGCACTTAG TTTCGACGCC GGTAAGGCGG

151  ACGGCGTGA
     TGCCGCACT
```

h) ORF8

```
1    ATGGATGACT ACAAACTACA GCAGCAGTTG AACGAACTGC GGCAATTCAT
     TACCTACTGA TGTTTGATGT CGTCGTCAAC TTGCTTGACG CCGTTAAGTA
```

```
                                                    EagI
                                                    ~~~~~~~
 51  CGAGGCGCAT AAGGATGTGC AGCACTTGCT GGACAACTCG GCCGCGTGGG
     GCTCCGCGTA TTCCTACACG TCGTGAACGA CCTGTTGAGC CGGCGCACCC

101  CTGACACCCA ACGGGAACAG GATGAAAAAT CGAAAGCGCA GGCTTGGCGA
     GACTGTGGGT TGCCCTTGTC CTACTTTTTA GCTTTCGCGT CCGAACCGCT

151  GTCGCTGCCG GCGCCAGTGT CTTCGGCATG GTGGCCCTGT TCATTGCATG
     CAGCGACGGC CGCGGTCACA GAAGCCGTAC CACCGGGACA AGTAACGTAC

201  GACGGCGATC CGCACGGCCT ACGTCCCTCC CGCCGCGCCG CAAGTTCTTG
     CTGCCGCTAG GCGTGCCGGA TGCAGGGAGG GCGGCGCGGC GTTCAAGAAC

251  TGATTGACAA GACGACCGGC CACGTCGAAC CGCTGGTTTC ACTCAAAGAG
     ACTAACTGTT CTGCTGGCCG GTGCAGCTTG GCGACCAAAG TGAGTTTCTC

301  GTTCAGGAAT CGGTCGATGA AGCCGTGACG CGCCACTACA TCACCGAGTT
     CAAGTCCTTA GCCAGCTACT TCGGCACTGC GCGGTGATGT AGTGGCTCAA

351  TTTGCGTTGC CGGGAGAACT ACACCTTCGA TACAGCCGAA GAAAATTATT
     AAACGCAACG GCCCTCTTGA TGTGGAAGCT ATGTCGGCTT CTTTTAATAA

401  ATTGCGCAGC GGCCTATATG TCACCCCAAC TGCAAACGCA GTGGGCGGCT
     TAACGCGTCG CCGGATATAC AGTGGGGTTG ACGTTTGCGT CACCCGCCGA

451  TTTTGGGATA CCAAGAATCC GGATTCTCCG TACAACTATT ACAAGAACTC
     AAAACCCTAT GGTTCTTAGG CCTAAGAGGC ATGTTGATAA TGTTCTTGAG

501  GGCGACGGTC AAGATCGACA TTGACTCGAT CACACTCAAC ATCAATAGCG
     CCGCTGCCAG TTCTAGCTGT AACTGAGCTA GTGTGAGTTG TAGTTATCGC

551  ACGGAGCACA AGACACCGCG ACCGTGCGCT TCACACGCTA CGTGAAGAAG
     TGCCTCGTGT TCTGTGGCGC TGGCACGCGA AGTGTGCGAT GCACTTCTTC

601  AACGATCATC AGGAAGTCAC GCGGTGGGTT GCGACGCTGG CGTACAAGTA
     TTGCTAGTAG TCCTTCAGTG CGCCACCCAA CGCTGCGACC GCATGTTCAT

      HindII
      ~~~~~~
651  CGTCGACCAG GACATTGGCG CTGCTGCTGT ACCCATCTCG GTCGCACCAC
     GCAGCTGGTC CTGTAACCGC GACGACGACA TGGGTAGAGC CAGCGTGGTG

701  CACAATCTGA TTCACCCGTG CCGGCCCCTG TCCCGTCCTC GCCTTCCCCG
     GTGTTAGACT AAGTGGGCAC GGCCGGGGAC AGGGCAGGAG CGGAAGGGGC

                                          AflIII
                                          ~~~~~~~
751  CCTGCGCTGT CAGCACCTCA AGCGCCCGCG TCGCAACATG TTGGGACCGT
     GGACGCGACA GTCGTGGAGT TCGCGGGCGC AGCGTTGTAC AACCCTGGCA

801  GCAATGA
     CGTTACT
```

i) ORF9

```
  1  ATGAAGCTCC GAACCCTCGT CTGCTACCTC GTCCTCGTCG CTGCTCCGCT
     TACTTCGAGG CTTGGGAGCA GACGATGGAG CAGGAGCAGC GACGAGGCGA
```

```
 51  GCCGGCGCTC GCCGTGCAGC CTACACAGCC GTCGCCCGCA GACCCTCGTA
     CGGCCGCGAG CGGCACGTCG GATGTGTCGG CAGCGGGCGT CTGGGAGCAT

101  TCCGGTTCAT CGACTACGAC CCGTACAACA TCCCGGTCAT CTATGCGCGG
     AGGCCAAGTA GCTGATGCTG GGCATGTTGT AGGGCCAGTA GATACGCGCC

151  ATCGGCGGCG ATCTCATGCT CGTCTTCCAG AACGGCGAAG TAGTGAAGGA
     TAGCCGCCGC TAGAGTACGA GCAGAAGGTC TTGCCGCTTC ATCACTTCCT

201  CATGACGGGC GGCGACACCG ATGCCTGGGG CGTCGGTGTT TCGACCGCAG
     GTACTGCCCG CCGCTGTGGC TACGGACCCC GCAGCCACAA AGCTGGCGTC

251  GCAACAGCGT TTTTATGAAG CCGAAGGTCA CAAGTCCGAA CATGAACCTG
     CGTTGTCGCA AAAATACTTC GGCTTCCAGT GTTCAGGCTT GTACTTGGAC

301  CACGTCGTCA CAAACAAGCG CATCTACAGC ATCGACTTGA AGCTGGCAAG
     GTGCAGCAGT GTTTGTTCGC GTAGATGTCG TAGCTGAACT TCGACCGTTC

351  CAAGGGCCAA GTGGCATTTC AGACGATCTA CTACCGCTAC CCAGACGAAG
     GTTCCCGGTT CACCGTAAAG TCTGCTAGAT GATGGCGATG GGTCTGCTTC

                                                    StyI
                                                    ~~~~~~
401  ATCGCACCAA ACGCGAAGCA GCAAAAGCGC GTGACCTTCT GTCCCATGGC
     TAGCGTGGTT TGCGCTTCGT CGTTTTCGCG CACTGGAAGA CAGGGTACCG

451  AGCGCCGCGA CGACGACGAA CCGCAACTAC ACAATGCAAG GGTCCGACGC
     TCGCGGCGCT GCTGCTGCTT GGCGTTGATG TGTTACGTTC CCAGGCTGCG

                                        HindII
                                        ~~~~~~
501  ATTGGCGCCG CTGGAAGCGT GGGACGACGG GAAGTTGACC TACTTCCGAT
     TAACCGCGGC GACCTTCGCA CCCTGCTGCC CTTCAACTGG ATGAAGGCTA

551  TCCCGGCGAA TCGCGACGTG CCATCAATCT ACTACGTCAC CGACGACGGC
     AGGGCCGCTT AGCGCTGCAC GGTAGTTAGA TGATGCAGTG GCTGCTGCCG

601  AAGGAACACT TGGTCAATAA AGACATGGAC TCGAACTATG TGCTCACGGT
     TTCCTTGTGA ACCAGTTATT TCTGTACCTG AGCTTGATAC ACGAGTGCCA

651  CCAGAAGATA GCTAAGAAGT TCGTCTTCCG CACTGGCGAT GTTGTGACCT
     GGTCTTCTAT CGATTCTTCA AGCAGAAGGC GTGACCGCTA CAACACTGGA

701  GCATATTCAA CGAGTCCTAT AATTTGAACG CGGTGCGCAG CCCGACGAAC
     CGTATAAGTT GCTCAGGATA TTAAACTTGC GCCACGCGTC GGGCTGCTTG

751  ACCACTTCGC CCAACGTGGA GCGCGTCATC AGAGGCGGTC AGCAATGA
     TGGTGAAGCG GGTTGCACCT CGCGCAGTAG TCTCCGCCAG TCGTTACT
```

j) ORF10

```
  1  ATGAAATTTT TTTTCCGTAA AAGCAACCAA GCCAGCGAAC AGAATAGTCT
     TACTTTAAAA AAAAGGCATT TTCGTTGGTT CGGTCGCTTG TCTTATCAGA

                                                    HindII
                                                    ~~~~~~
 51  CGAAGAAGTC GCTCACATCG ACGGGGAGCG TGCCACTGCC TCAGTCAACA
     GCTTCTTCAG CGAGTGTAGC TGCCCCTCGC ACGGTGACGG AGTCAGTTGT
```

```
                SphI
              ~~~~~~~
101  AGGGCTTGGG CATGCAAACG AAGATCACAA ATTTCTTGAT CTTCGCATCC
     TCCCGAACCC GTACGTTTGC TTCTAGTGTT TAAAGAACTA GAAGCGTAGG

151  GTCATCGTCT TGGCCGTGGT CATGCTTTGG AAGTATTACG CCAACGTGAT
     CAGTAGCAGA ACCGGCACCA GTACGAAACC TTCATAATGC GGTTGCACTA

201  CGAGCAGCGG CACGCCGCAC AGCAAGCCGC GCAAAAGGAC ACGAAGGTGC
     GCTCGTCGCC GTGCGGCGTG TCGTTCGGCG CGTTTTCCTG TGCTTCCACG

251  AGCAACAAAC GATTCTCCCA CCTTTGGTCC CACCGAATCT GCCGGAAACC
     TCGTTGTTTG CTAAGAGGGT GGAAACCAGG GTGGCTTAGA CGGCCTTTGG

301  AATGCCACTA AGCATCGGC ACCGGCAGCT AACACGACGC CACAGACCGG
     TTACGGTGAT TTCGTAGCCG TGGCCGTCGA TTGTGCTGCG GTGTCTGGCC

351  GCAACAGCTT GGCCCTGATG GCAAGCCAAT CCTCACGCCG GCAGAACAGC
     CGTTGTCGAA CCGGGACTAC CGTTCGGTTA GGAGTGCGGC CGTCTTGTCG

                                          HindIII
                                          ~~~~~~~
401  AGTTACAGCG GCGCCTCACG TCGAGCGTGA AATTCAAGCT TGATGCGCCT
     TCAATGTCGC CGCGGAGTGC AGCTCGCACT TTAAGTTCGA ACTACGCGGA

451  GACCGGCAAT CTGGCAAAGC CGATGCCGAC ACAGCAGCAG CCGCTGACCC
     CTGGCCGTTA GACCGTTTCG GCTACGGCTG TGTCGTCGTC GGCGACTGGG

                              StyI
                             ~~~~~~
501  TGGCGGCTCG CCTGGCTCAG GTGGCCTTGG GGCTGGCTCC GACGATCCAC
     ACCGCCGAGC GGACCGAGTC CACCGGAACC CCGACCGAGG CTGCTAGGTG

551  TTGCCCGCTC GCTCCGCGCG ACGTACACCC CAGGCGCCGT GGCAACGCTT
     AACGGGCGAG CGAGGCGCGC TGCATGTGGG GTCCGCGGCA CCGTTGCGAA

601  CTGCACGACC GTGATTTCCT CATCACGAAG GGCGCAGTCA TACCCTGCTC
     GACGTGCTGG CACTAAAGGA GTAGTGCTTC CCGCGTCAGT ATGGGACGAG

651  CGTTGATCCG GCTCTTGATT CCAGCTTACC CGGCATCGTG ACTTGCACAG
     GCAACTAGGC CGAGAACTAA GGTCGAATGG GCCGTAGCAC TGAACGTGTC

701  GCTCGTCGGA CGTTTGGAGC ACCAACCACA AGGTCAAACT GATGGAAGCA
     CGAGCAGCCT GCAAACCTCG TGGTTGGTGT TCCAGTTTGA CTACCTTCGT

                                                 HindII
                                                 ~~~~~~
751  GGCACCAAGT ACGTGGGCGA GGCGAAACAG GGACTCTCCA AAAGTCAACA
     CCGTGGTTCA TGCACCCGCT CCGCTTTGTC CCTGAGAGGT TTTCAGTTGT

801  CAGGATGGCA ATCCTCTGGA CCCGTGCCGA GACACCCAAC GGTGTCATCA
     GTCCTACCGT TAGGAGACCT GGGCACGGCT CTGTGGGTTG CCACAGTAGT

                                      EagI
                                      ~~~~~~
851  TCGACCTTCA ATCGGTCGCG TCCGACGAAC TCGGCCGACC CGGTGTGAGT
     AGCTGGAAGT TAGCCAGCGC AGGCTGCTTG AGCCGGCTGG GCCACACTCA


901  GGTGAAATCG ACAACCACTT TTGGGACCGC TTCGGCGCGG CCATCATGCT
     CCACTTTAGC TGTTGGTGAA AACCCTGGCG AAGCCGCGCC GGTAGTACGA
```

```
 951  TAGTCTGCTC AACGACACAA GCGCGTTCAT GATCGCCCGC GAGCAAAACA
      ATCAGACGAG TTGCTGTGTT CGCGCAAGTA CTAGCGGGCG CTCGTTTTGT

                                                      HindII
                                                      ~~~~~~
1001  ACGGCAGCGG CAGCAACAAC ACCACTATCG CGTTTCCCAA CACTGTCAAC
      TGCCGTCGCC GTCGTTGTTG TGGTGATAGC GCAAAGGGTT GTGACAGTTG

1051  GGGACGCAGA ACATCGTAGG CGATGTGTTG AAACAGAATT TAGACATTCC
      CCCTGCGTCT TGTAGCATCC GCTACACAAC TTTGTCTTAA ATCTGTAAGG

                                                    ApaLI
                                                    ~~~~~~
1101  GCCGACGCTC ACGAAAAACC AGGGCGCAAA CATCAATATC TACGTGCACG
      CGGCTGCGAG TGCTTTTTGG TCCCGCGTTT GTAGTTATAG ATGCACGTGC

1151  CGACCTGGAC TTCCGCAGCG TGTACGATTT GA
      GCTGGACCTG AAGGCGTCGC ACATGCTAAA CT
```

k) ORF11

```
   1  ATGACGGGTG AGAAATCCAC GCTGGAAATT CACCTGGAAC CGCTTCGCCC
      TACTGCCCAC TCTTTAGGTG CGACCTTTAA GTGGACCTTG GCGAAGCGGG

  51  ATTCCTTGAC GATCCAGCGA ACAACGAAAT CGTCATCAAT AACCCGCTTG
      TAAGGAACTG CTAGGTCGCT TGTTGCTTTA GCAGTAGTTA TTGGGCGAAC

 101  TAGTATGGAC GGAAAGCCGG GGTCAGTGGG TCACACATGA CGTACCTTCG
      ATCATACCTG CCTTTCGGCC CCAGTCACCC AGTGTGTACT GCATGGAAGC

 151  ATCACATCGG AGTGGTGCGA CGAATTGTCA AAGCTGGTCG CTAATTTCTC
      TAGTGTAGCC TCACCACGCT GCTTAACAGT TTCGACCAGC GATTAAAGAG

 201  CGATCAGAAG ATCGACGTTG AGCATCCCAT GATCGGCTCG ACGCTACCTA
      GCTAGTCTTC TAGCTGCAAC TCGTAGGGTA CTAGCCGAGC TGCGATGGAT

 251  CGCGAGAGCG CATCCAGATC GTGATTCCGC CAGTTGTCAA AACCGTGTCC
      GCGCTCTCGC GTAGGTCTAG CACTAAGGCG GTCAACAGTT TTGGCACAGG

 301  GTCACTATTC GGCGCCCAAG CGCCGACGTG ATGACGTTCG ATGAAATCTA
      CAGTGATAAG CCGCGGGTTC GCGGCTGCAC TACTGCAAGC TACTTTAGAT

 351  TGAGCGTGGC ACTTTTGACG ACACGCGATG CGAACAGTCT TCGCGCCTAG
      ACTCGCACCG TGAAAACTGC TGTGCGCTAC GCTTGTCAGA AGCGCGGATC

 401  ATGCCGAAGA GCGCGAGGCC ATCGAATCGA CCATACCCGC CAATGACAAG
      TACGGCTTCT CGCGCTCCGG TAGCTTAGCT GGTATGGGCG GTTACTGTTC

 451  AAACTGATAT ATCTGTTTCG GAAGAAAGAT TGGAAGGAGT TCTTGCAACA
      TTTGACTATA TAGACAAAGC CTTCTTTCTA ACCTTCCTCA AGAACGTTGT

 501  GGCAGTGTTG CTGCGCAAGA ACATCATCCT ATCAGGGAGA ACCGGCTCTG
      CCGTCACAAC GACGCGTTCT TGTAGTAGGA TAGTCCCTCT TGGCCGAGAC

 551  GCAAAACGAC GCTCGCGAAC TCACTGTGCA TGATGATCCC GATCCAGGAG
      CGTTTTGCTG CGAGCGCTTG AGTGACACGT ACTACTAGGG CTAGGTCCTC

 601  CGGATCATCA CGGTAGAAGA CGCGCGCGAA GTGCGCCTGC CCCACCCGAA
      GCCTAGTAGT GCCATCTTCT GCGCGCGCTT CACGCGGACG GGGTGGGCTT
```

```
          HindII
          ~~~~~~
 651  CCAAGTCAAC CTTGTCTATT CGCGTGATAA CCGCGGCCTC GCCAAGCTGT
      GGTTCAGTTG GAACAGATAA GCGCACTATT GGCGCCGGAG CGGTTCGACA
 701  CGCCGAAGCA GATGTTCGAG TCAGGTCTAC GTATGCGGCC CGACCGCGTT
      GCGGCTTCGT CTACAAGCTC AGTCCAGATG CATACGCCGG GCTGGCGCAA

 751  CTGCCCGCCG AGCTTCGCGG CGAGGAAGCC TATTTCTTCT TCCAGAACGT
      GACGGGCGGC TCGAAGCGCC GCTCCTTCGG ATAAAGAAGA AGGTCTTGCA

 801  CGTGAACAGC GGCCACCCCG GCGCGATCAC CAGCATCCAC TCTAACACCT
      GCACTTGTCG CCGGTGGGGC CGCGCTAGTG GTCGTAGGTG AGATTGTGGA

 851  CGAAGCTGGC CTTCCGCCGC CTAACTAACA TGATCCAGTC AAGCCAAGAG
      GCTTCGACCG GAAGGCGGCG GATTGATTGT ACTAGGTCAG TTCGGTTCTC

 901  GGGCGCGGGC TTGAACAAGG CGTCATCCTC GAAATGCTGT ACGCCCTGGT
      CCCGCGCCCG AACTTGTTCC GCAGTAGGAG CTTTACGACA TGCGGGACCA

        EcoRV
       ~~~~~~
 951  CGATATCGTG ATCCAAACAG ACACGGTGCA AATCAACAGC GCCGGTGAGA
      GCTATAGCAC TAGGTTTGTC TGTGCCACGT TTAGTTGTCG CGGCCACTCT

1001  ACCGCAAGAT GGTCACAGAA ATCTATTTCG ATCCGGCTTA CGCACTGAAG
      TGGCGTTCTA CCAGTGTCTT TAGATAAAGC TAGGCCGAAT GCGTGACTTC

1051  CAAATAGGAT AA
      GTTTATCCTA TT
```

l) ORF12

```
   1  ATGTCTTTTC TACGTGCCTT GATGGTCGGC GCCGCGCTCG CGGCGCTGGC
      TACAGAAAAG ATGCACGGAA CTACCAGCCG CGGCGCGAGC GCCGCGACCG

  51  CGCCTGTACA TCCAAGCCGA CGCATATCTC GTTTCTGGAT CGCGTGCTGA
      GCGGACATGT AGGTTCGGCT GCGTATAGAG CAAAGACCTA GCGCACGACT

 101  CAGTCGAAGA GTTCACCGCG CAGTCCGACA TTCGCGACAA TGTCATCGCT
      GTCAGCTTCT CAAGTGGCGC GTCAGGCTGT AAGCGCTGTT ACAGTAGCGA

                                                      HindII
                                                      ~~~~~
 151  GAGTGCTCGT CGGACCCCGG CGATTTGAAT GATGACCCGA ACTGCGTCAA
      CTCACGAGCA GCCTGGGGCC GCTAAACTTA CTACTGGGCT TGACGCAGTT

      HindII
      ~
 201  CGCCAAACAG TCGCTTCGTG CGAGCGTGAC GGGAACCGGA AATTTTCCGT
      GCGGTTTGTC AGCGAAGCAC GCTCGCACTG CCCTTGGCCT TTAAAAGGCA

 251  CGCTTGACAT AACGCCCCCG CACAAGAAGG GCAAGCAAAT AGGGAATTGA
      GCGAACTGTA TTGCGGGGGC GTGTTCTTCC CGTTCGTTTA TCCCTTAACT
```

m) ORF13

```
   1  ATGATACTCA CCGACATTCA AAGCAAGCGA GTGAGGTCCG TGGTCATTTC
      TACTATGAGT GGCTGTAAGT TTCGTTCGCT CACTCCAGGC ACCAGTAAAG
```

```
 51  CGGTCTTCTC GTATGCGGCT TAATCCTCGG GCTGTACCTT GCCGGCTATT
     GCCAGAAGAG CATACGCCGA ATTAGGAGCC CGACATGGAA CGGCCGATAA

101  TCTTCTTGTG GAAGGTCAAG CTCAAGCCGT ACAGCGCGAC GCCGCTGACC
     AGAAGAACAC CTTCCAGTTC GAGTTCGGCA TGTCGCGCTG CGGCGACTGG

151  ATCATCGACT ATTGGGCGTA CTACAGCGAC AATCCGCAAC TTCGCAAATT
     TAGTAGCTGA TAACCCGCAT GATGTCGCTG TTAGGCGTTG AAGCGTTTAA


201  GATGAAGTAT TGCCTGGGTT TCGGCTTCAT CTTGTCCTAT GGCCTTGTAG
     CTACTTCATA ACGGACCCAA AGCCGAAGTA GAACAGGATA CCGGAACATC

251  CCATGCTGTT CATGCCGGTG CGGCGCGCTC TGCACGGCGA CGCCCGCTTC
     GGTACGACAA GTACGGCCAC GCCGCGCGAG ACGTGCCGCT GCGGGCGAAG

301  GCGAAGGACA AGGAAGTTCG TGATGCCGAC CTGCTCGGCG AACATGGGCT
     CGCTTCCTGT TCCTTCAAGC ACTACGGCTG GACGAGCCGC TTGTACCCGA

351  AATCCTCGGC AAGTGGGGCG ACCGCTTCAT CATGCTGGCC GGCCAGCTTG
     TTAGGAGCCG TTCACCCCGC TGGCGAAGTA GTACGACCGG CCGGTCGAAC

401  GCGCGATCTG CGCAGCGCCA CCTCGAACGG GCAAGGGCGC AGGCTTGGTT
     CGCGCTAGAC GCGTCGCGGT GGAGCTTGCC CGTTCCCGCG TCCGAACCAA

451  CAGCCCAACA TGCTGAACTG GCTGCAAAGC GTCGTGCTGC TGGACGTGCG
     GTCGGGTTGT ACGACTTGAC CGACGTTTCG CAGCACGACG ACCTGCACGC

                                                    AflIII
                                                      ~~~
501  GCAGGAAAGC TATCGGCTGA CCTCTGGCTT CCGCAAGATG TTCAGCGACG
     CGTCCTTTCG ATAGCCGACT GGAGACCGAA GGCGTTCTAC AAGTCGCTGC

     AflIII                          EagI
     ~~~                             ~~~~~~~
551  TGTTCCTCTT CAATCCTGTC GCCGAAGACG GCCGCACGAT GCAATGGAAT
     ACAAGGAGAA GTTAGGACAG CGGCTTCTGC CGGCGTGCTA CGTTACCTTA

601  CCACTTAGCT ATGTGAACGA CGATCCGATT CTACGGATCA ACGATCTTCA
     GGTGAATCGA TACACTTGCT GCTAGGCTAA GATGCCTAGT TGCTAGAAGT

651  GAAGATCGCA AACATGCTGT CGCCCGACCC CGCCGAAGGT GATCCGTTCT
     CTTCTAGCGT TTGTACGACA GCGGGCTGGG GCGGCTTCCA CTAGGCAAGA

701  GGCCTGCGTC GTGCCGCACG CTATTCCTCG GCCTTGCTCT GTATGTTTTC
     CCGGACGCAG CACGGCGTGC GATAAGGAGC CGGAACGAGA CATACAAAAG

751  GAGACACCTG ACACGCCACG AACATTCGGT GAGATCGTTA GGCAAATCAT
     CTCTGTGGAC TGTGCGGTGC TTGTAAGCCA CTCTAGCAAT CCGTTTAGTA

                                                  EcoRV
                                                  ~~~~~~~
801  GTATGGAGAA GGTGAGTCTG TAGGCCAGCA CTGGAAAGAT ATCATCGAAG
     CATACCTCTT CCACTCAGAC ATCCGGTCGT GACCTTTCTA TAGTAGCTTC

                 BstEII
                 ~~~~~~~~
851  AGCGCGACGC TTCAGGTAAC CCGCTCTCCC CTGCCTGCAA GGCAGCGCTC
     TCGCGCTGCG AAGTCCATTG GGCGAGAGGG GACGGACGTT CCGTCGCGAG
```

```
 901  TACGACTTCA TCTACACCAG CGGCAACACG CAATCTTCAA TTCGGAAGAC
      ATGCTGAAGT AGATGTGGTC GCCGTTGTGC GTTAGAAGTT AAGCCTTCTG

                                          StyI
                                          ~~~~~~
 951  CTTTACGGCC AAGCTCGAAC TGTGGCTGAA CCCCTTGGTA GACGCGGCTA
      GAAATGCCGG TTCGAGCTTG ACACCGACTT GGGGAACCAT CTGCGCCGAT

1001  CCAGCGGCGA TTCGTTCGAC CTGCGCGACT TTCGCCGGCG CCGCATTTCT
      GGTCGCCGCT AAGCAAGCTG GACGCGCTGA AAGCGGCCGC GGCGTAAAGA

                           EagI
                          ~~~~~~~
1051  TTGTACATCG GTATACGTCC GGCCGACCTG AGCCGGCTGC AACTCATTCT
      AACATGTAGC CATATGCAGG CCGGCTGGAC TCGGCCGACG TTGAGTAAGA

1101  GAACCTGCTT TTCCAGCAGA TTATCGACCT CAACACGGAC GAGATGCCGG
      CTTGGACGAA AAGGTCGTCT AATAGCTGGA GTTGTGCCTG CTCTACGGCC


1151  AAGACAACCC GGATCTCAAG TTCCAGTTGT TGATGATGAT GGATGAGTTC
      TTCTGTTGGG CCTAGAGTTC AAGGTCAACA ACTACTACTA CCTACTCAAG

                     SphI
                     ~~~~~~
               EagI                                      StyI
              ~~~~~~~                                    ~~~~~~
1201  ACAGCCATCG GCCGCATGCC CATCTTCGCG AAGTCGATCA GTTTCCTTGG
      TGTCGGTAGC CGGCGTACGG GTAGAAGCGC TTCAGCTAGT CAAAGGAACC

1251  CGGCTACAAC ATCCGCCCGT TCATCATCAT TCAAGGCATG TCGCAATTAC
      GCCGATGTTG TAGGCGGGCA AGTAGTAGTA AGTTCCGTAC AGCGTTAATG

1301  GATCCACATA CGGCGCCGAC GTGGCAGAAA CCATCGTCAC ATGCTGCGCT
      CTAGGTGTAT GCCGCGGCTG CACCGTCTTT GGTAGCAGTG TACGACGCGA

1351  GCGATGATCG TGTACGCCCC GAAGGAACAA CGCCACGCCA ACGAAATTTC
      CGCTACTAGC ACATGCGGGG CTTCCTTGTT GCGGTGCGGT TGCTTTAAAG

1401  CGAAATGCTC GGTTACATGA CGGTGCAAGC GAAGTCGAAG TCGCAGCAAG
      GCTTTACGAG CCAATGTACT GCCACGTTCG CTTCAGCTTC AGCGTCGTTC

                                  HindII
                                  ~~~~~~~
                                       AflIII
                                       ~~~~~~
1451  TCGGCTTCAA GCGCGTCGGT GGTTCGGTCA ACACGTCTGA CCAGCGCCGC
      AGCCGAAGTT CGCGCAGCCA CCAAGCCAGT TGTGCAGACT GGTCGCGGCG

1501  GCCCTGATGC TGCCCCAAGA GGTCAAGGAA ATCGGCAAGG AGCGGGAAAT
      CGGGACTACG ACGGGGTTCT CCAGTTCCTT TAGCCGTTCC TCGCCCTTTA

1551  CATCTTTTTG GAAAATTTTA AGCCAATTCT CGCCAGCAAG ATTTCCTATT
      GTAGAAAAAC CTTTTAAAAT TCGGTTAAGA GCGGTCGTTC TAAAGGATAA

1601  GGAAAGACAA GGCATTTAAG CGCCGGCTAC TGCCGGCTGC TGTCGTCCCG
      CCTTTCTGTT CCGTAAATTC GCGGCCGATG ACGGCCGACG ACAGCAGGGC

1651  GCCATCGACG TAAAGATGCC GGAACCCTCT CAGCGGCCCA AGAAGAAAAA
      CGGTAGCTGC ATTTCTACGG CCTTGGGAGA GTCGCCGGGT TCTTCTTTTT
```

```
1701  GAAAAAAGAG GGCGAGACAG TCAAGACAAC AGACGGCCAA CTTATCACTG
      CTTTTTTCTC CCGCTCTGTC AGTTCTGTTG TCTGCCGGTT GAATAGTGAC

1751  TTACGGAAAA AGAAATTACG GCTGACGACG TTGGCAAGCT CGATAAATTG
      AATGCCTTTT TCTTTAATGC CGACTGCTGC AACCGTTCGA GCTATTTAAC

1801  AGCCTTACTG ACTACAACGT CGATTTTGAC AGCGTGGAAG TGCCACGCGG
      TCGGAATGAC TGATGTTGCA GCTAAAACTG TCGCACCTTC ACGGTGCGCC

                                    SphI
                                   ~~~~~~~
1851  CCAGCCGATC ACCGACGACG ACATGAAGCA TGCGTTTTCG TCGTTCCTGC
      GGTCGGCTAG TGGCTGCTGC TGTACTTCGT ACGCAAAAGC AGCAAGGACG

1901  AAACCATTGA GGATGCGTAA
      TTTGGTAACT CCTACGCATT
```

n) ORF14

```
   1  ATGACCCGCC GCACACTTGC CGGTCTGTTT GTCTGCATCG CCCTACCCTG
      TACTGGGCGG CGTGTGAACG GCCAGACAAA CAGACGTAGC GGGATGGGAC

  51  CGCCGCTAAC GCCCCTTGCC GGTACTCAGG TCGAAGAACA GCTCTCGAAC
      GCGGCGATTG CGGGGAACGG CCATGAGTCC AGCTTCTTGT CGAGAGCTTG


 101  AGCGTCCGAG CCGGCCTGTC GGCTTCGTTG AGCGACAAGA CCACGCCCGC
      TCGCAGGCTC GGCCGGACAG CCGAAGCAAC TCGCTGTTCT GGTGCGGGCG

                                          HindII
                                         ~~~~~~~
 151  GCCGATATTT CCTTCCGTAG CAGCGAAAAC AGCTTGGTTG ACCGAACAAG
      CGGCTATAAA GGAAGGCATC GTCGCTTTTG TCGAACCAAC TGGCTTGTTC

                          SphI
                         ~~~~~~~
 201  GCCAAACTGT CGCCAAGCGC ATGCCTGCCG TACAGGACCG CATCGCGTTT
      CGGTTTGACA GCGGTTCGCG TACGGACGGC ATGTCCTGGC GTAGCGCAAA

 251  CTCAACACCC GTGTACTACG AAGCCCATGC GTGCTGGCCT CGATCCCGAG
      GAGTTGTGGG CACATGATGC TTCGGGTACG CACGACCGGA GCTAGGGCTC

 301  CTTGTGCTTG CTGTCATCCA GACCGAAAGC GCCTTCCGCA AGTACGCAGT
      GAACACGAAC GACAGTAGGT CTGGCTTTCG CGGAAGGCGT TCATGCGTCA

           EagI
          ~~~~~~~
 351  TTCGACGGCC GGTGCGCGCG GCTATATGCA AGTCATGCCG TTTTGGTTGA
      AAGCTGCCGG CCACGCGCGC CGATATACGT TCAGTACGGC AAAACCAACT

 401  ATGAAATTGG GAATCCTACC GATGATCTGT TCCATCTGCG GACGAATCTG
      TACTTTAACC CTTAGGATGG CTACTAGACA AGGTAGACGC CTGCTTAGAC

 451  CGATACGGCT GCACTATCCT CCGCCACTAC CTTGACACTG AGCGCGGGAA
      GCTATGCCGA CGTGATAGGA GGCGGTGATG GAACTGTGAC TCGCGCCCTT

                                           MluNI
                                          ~~~~~~~
 501  CTTGTTCCTC GCGCTTGGCC GCTACAACGG CAGTCGTGGC CAAGCCGCAT
      GAACAAGGAG CGCGAACCGG CGATGTTGCC GTCAGCACCG GTTCGGCGTA
```

```
551  ATCCGAACCT GGTACTCGCA TCTTGGTCGC AGTGGCGCTC GATGCCGAGT
     TAGGCTTGGA CCATGAGCGT AGAACCAGCG TCACCGCGAG CTACGGCTCA

601  GACTCTCAGA TAACGATGAA ATGA
     CTGAGAGTCT ATTGCTACTT TACT
```

o) ORF15

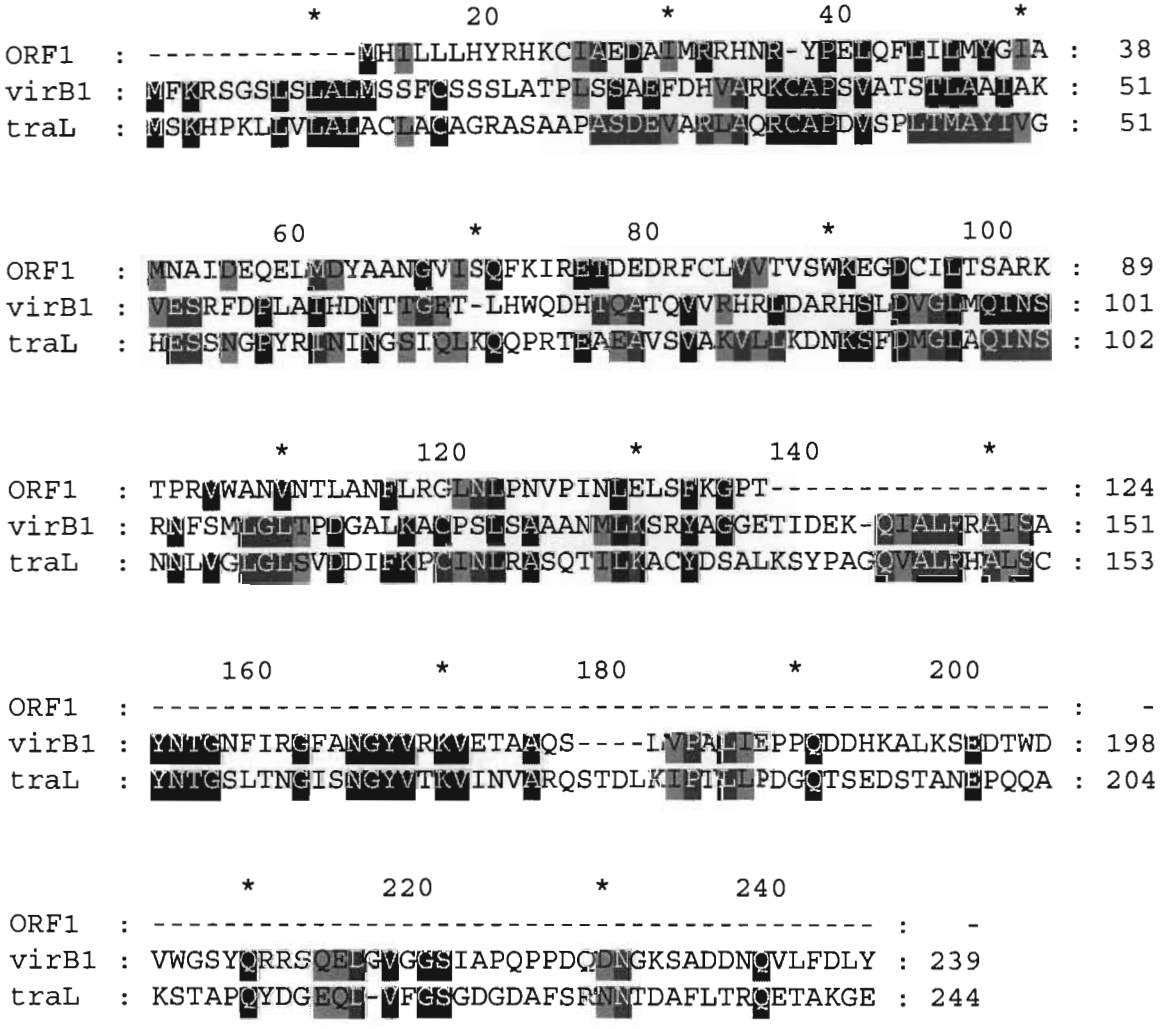
```
  1  ATGAATAAAG AAGTCACAGC TTCAGTTGTT GTCGCGGAAG ATGAACGCCG
     TACTTATTTC TTCAGTGTCG AAGTCAACAA CAGCGCCTTC TACTTGCGGC

 51  AGATTTTTTG TCCAAACATT TCGGCATCCG CTTTGCGAAA CGGGGAGAGG
     TCTAAAAAAC AGGTTTGTAA AGCCGTAGGC GAAACGCTTT GCCCCTCTCC

101  CACTCGTTTT TGCCTGGTTA CTTCGTTTAT CAAAAGTCCC GATTGAGTGG
     GTGAGCAAAA ACGGACCAAT GAAGCAAATA GTTTTCAGGG CTAACTCACC

151  ACACGATTGC AATACTACAC ACTATCAAAC AGTGGATTTT ATCTTGCTCC
     TGTGCTAACG TTATGATGTG TGATAGTTTG TCACCTAAAA TAGAACGAGG

201  TCGCGAGCTG CGCATATCAG AATGCGAACT CTCAGCAGAT GCTCTCGGCA
     AGCGCTCGAC GCGTATAGTC TTACGCTTGA GAGTCGTCTA CGAGAGCCGT

251  TTGTTGCAAC AATGCTCACA CTCCGACAAC TTGCGCATGA GTCTGCCGCG
     AACAACGTTG TTACGAGTGT GAGGCTGTTG AACGCGTACT CAGACGGCGC

301  TCTGTCGAAG CAGATTCGAC ACACGCGGCA GCGAAGCTCG CGGTAAAAGC
     AGACAGCTTC GTCTAAGCTG TGTGCGCCGT CGCTTCGAGC GCCATTTTCG

351  GTCGGTGATG TTCGCACAGA ATTATCACAA TCTCGCTGCT TACGCTGTTA
     CAGCCACTAC AAGCGTGTCT TAATAGTGTT AGAGCGACGA ATGCGACAAT

                ClaI                  PvuI
               ~~~~~~                ~~~~~~
401  AACACGCAGA ATCGATCAAC ATCTACCGGG CGATCGACTG A
     TTGTGCGTCT TAGCTAGTTG TAGATGGCCC GCTAGCTGAC T
```

p) ORF16

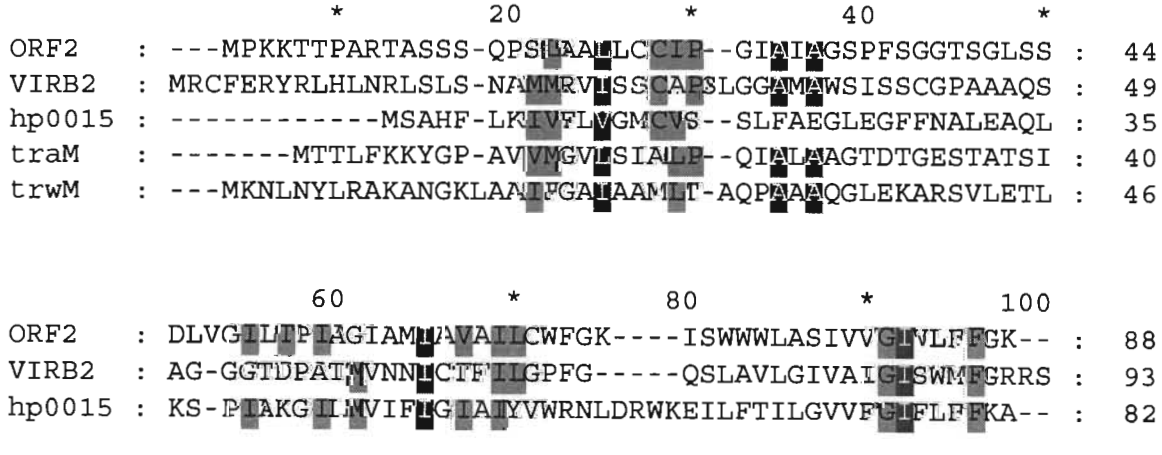
```
  1  ATGCACAACG AAATAAGACC TCAGTGCTTG CGTCCAGCCG AATGCTGGAC
     TACGTGTTGC TTTATTCTGG AGTCACGAAC GCAGGTCGGC TTACGACCTG

 51  GCAGCCGACC GCAAATGAAA TCCGCGCAGT GCTTCGCATG GCAGGGATGA
     CGTCGGCTGG CGTTTACTTT AGGCGCGTCA CGAAGCGTAC CGTCCCTACT

101  CGGGCGGCGC TGCTGCAAAA TTGCTCGGGC TTGGCGCTAA GGGCGACCGG
     GCCCGCCGCG ACGACGTTTT AACGAGCCCG AACCGCGATT CCCGCTGGCC

                                      EcoRI
                                     ~~~~~~
151  ACTATTCGGC GGTGGATCGG TGGTGACAGT CGAATTCTTT GGCAGCGCCA
     TGATAAGCCG CCACCTAGCC ACCACTGTCA GCTTAAGAAA CCGTCGCGGT

201  AATACATGAT GGTCCTGCAC ACGCCTGA
     TTATGTACTA CCAGGACGTG TGCGGACT
```

APPENDIX TWO

a) ORF1



b) ORF2



```
traM   : QT-WLSTWIPIGCALAIMVSCFMW-----MLHVIPASFIPRLVISLIG--  :  82
trwM   : QQ-ELTTIVPIAAAVILLCLGIAY-----AGRFIEKDTFVRSIGVIIA--  :  88
```

```
                  *       120
ORF2   : ----DQVVSWIRGLFGDLSERLPIL---  : 109
VIRB2  : LGLVAGVVGGIVIMFGASFLGQTLTGGS  : 121
hp0015 : ----PSLANWFMGIF-------------  :  93
traM   : ----ISSASYLVSLTGVGS---------  :  97
trwM   : -----GSAVQITAMLFT-----------  : 100
```

c) ORF3

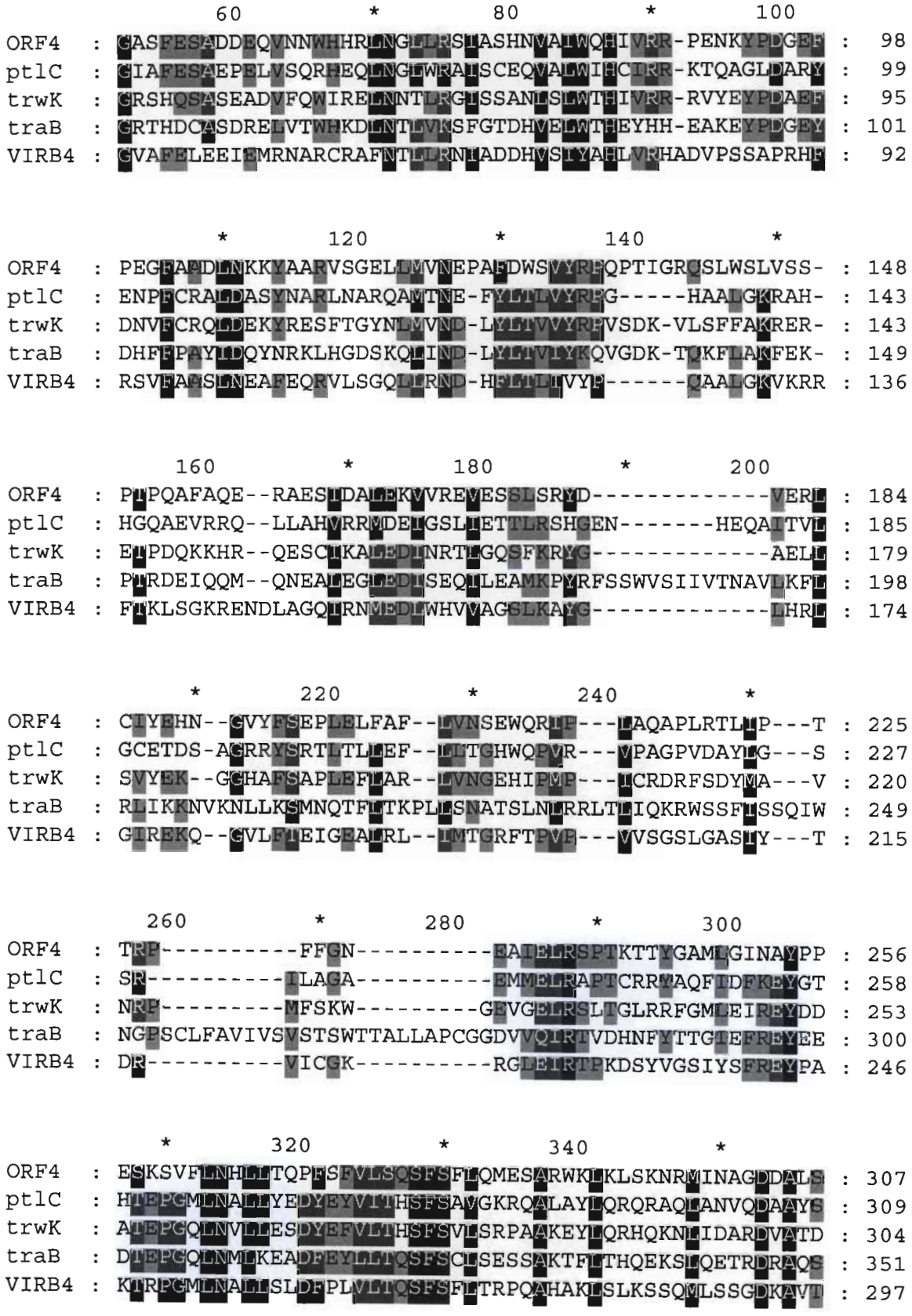
```
                 *        20         *        40         *
ORF3   : MSKNAGIIVD-TLFVGPTRPHDVMGRDVASLRHQHHRDYGSGHLDTRSALA  :  50
trwL   : MKPPQQQHEAFPLFKGATRLPTILGEPMIPLMAMVM-GVAVLALTVS-IWW  :  49
ptlB   : MRDP--------LFKGCTRPAMLMGVPATPLAVCSG-TIALLGIWFS-IAF  :  41
traA   : MFVDGKR----PLFKGATRLPRALGVPRNVAMMIFM-ISASLFMIIH-MWA  :  45
virB3  : MNDRLEEA---TLYLAATRPALFLGVPLTLAGLFMM-FAGFVIVIVQNPLY  :  47
```

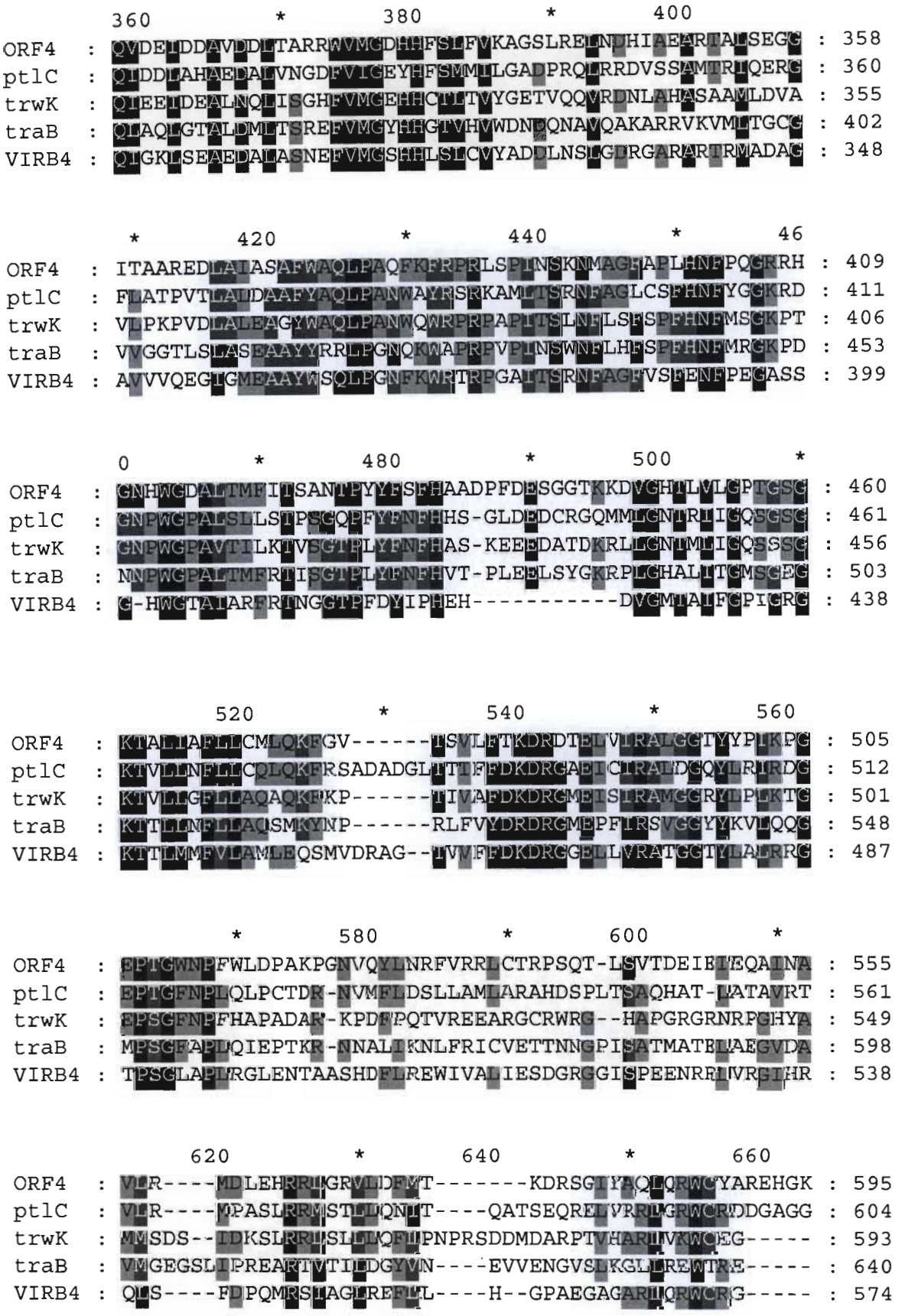
```
              60         *        80         *        100
ORF3   : FIVCPRFTASASLLCMRDPRTFELLMQWGPTKGFAFFGNFPYWRAATFSPL  : 101
trwL   : WALVPPLWFIMAQLTKNDDKALRLWWLWIDTKFR--NRNKGFWGASSYSP-  :  97
ptlB   : LALFPVALLAMRIMIRRDDQQFRLIWLYLRMRWLSRDRTHAFWQSTVYAP-  :  91
traA   : ILVFVFLRIPSAALTKYDDRMFRIMGLWLKLNSVIGLILRLSSGEDRLIP-  :  95
virB3  : EVVLAPLWFGARLIVERDYNAASVVLLFLRTAGR--SIDSAVWGGATVSPN  :  96
```

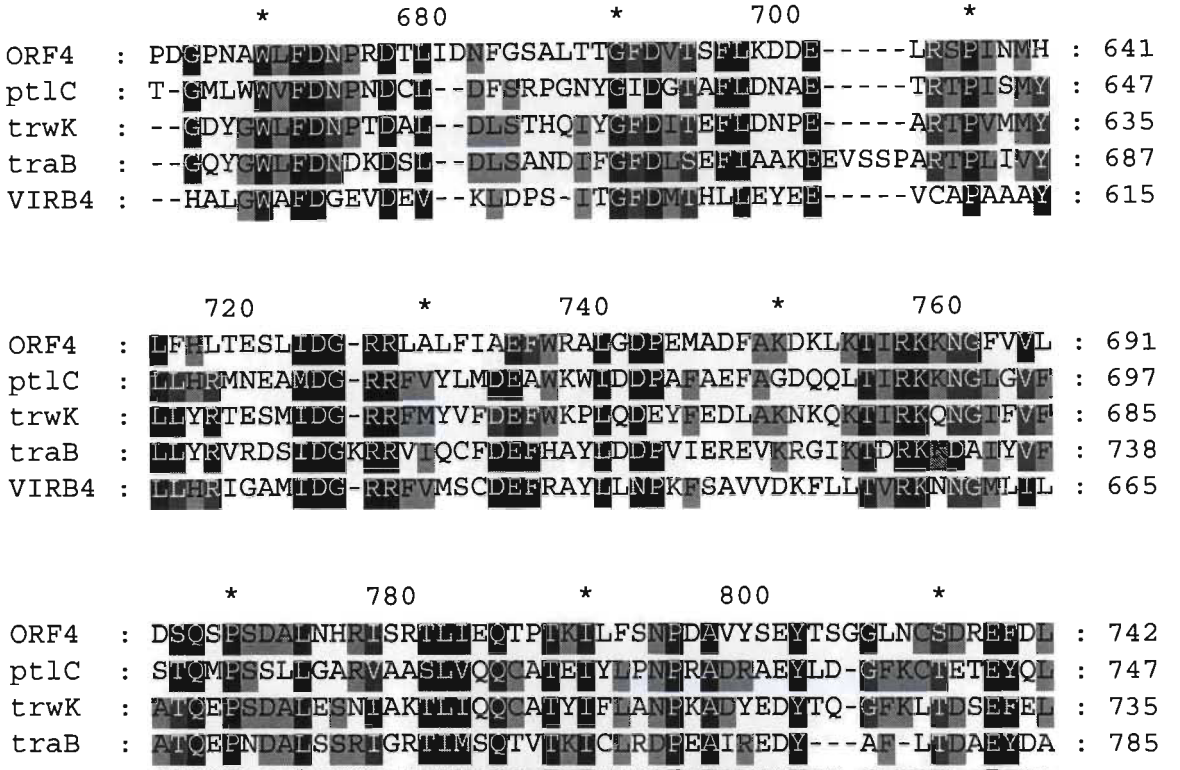
```
               *       120
ORF3   : ELHTPPKRSFFRRRKSRKVP  : 121
trwL   : -ANYRK--------------  : 102
ptlB   : -LRYAE-----RRRRLRKP-  : 104
traA   : -LLTTN-----VRV------  : 103
virB3  : PIRVPP-----RGRGMV---  : 108
```

d) ORF4

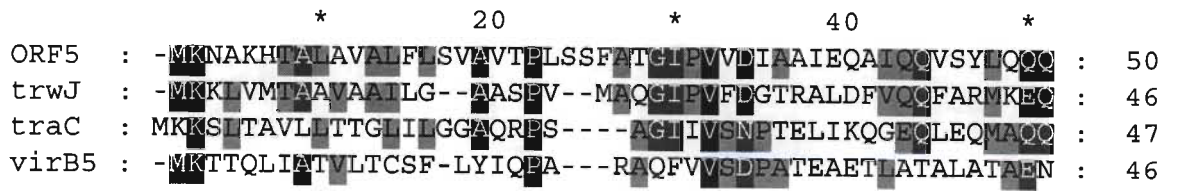
```
                  *        20         *        40         *
ORF4   : MKATVDAKRERAILR--EPALS-KNIPYSVHLIPTAIQTEHHDYLMVLRLT  :  48
ptlC   : MNRRGGQTAFAAIARN-ERAIA-AFIPYSSHLTDTTLITHGADLVRTWRVQ  :  49
trwK   : ---MGAIESRKLLAS--ETPVG-QFIPYSHHVIDTIISTKNAEYLSVWKID  :  45
traB   : MRAATATKPKKIDAYRKEPSVNKKYLPYSYHLNDYVLSMENGDLMAFFKLD  :  51
VIRB4  : MLGASG-----TT----ERSGE-IYLPYIGHLSDHIVLLEDGSIMSIARID  :  41
```


60 * 80 * 100
ORF4 : GASFESADDEQVNNWHHRLNGLLRSIASHNVALWQHIVRR-PENKYPDGEF : 98
ptlC : GIAFESAEPELVSQRHEQLNGLWRAISCEQVALWIHCIRR-KTQAGLDARY : 99
trwK : GRSHQSASEADVFQWIRELNNTLRGISSANLSLWTHIVRR-RVYEYPDAEF : 95
traB : GRTHDCASDRELVTWHKDLNTLVKSFGTDHVELWTHEYHH-EAKEYPDGEY : 101
VIRB4 : GVAFELEEIEMRNARCRAFNTLLRNIADDHVSIYAHLVRHADVPSSAPRHF : 92
* 120 * 140 *
ORF4 : PEGFAADLNKKYAARVSGELLMVNEPAEDWSVYRFQPTIGRQSLWSLVSS- : 148
ptlC : ENPFCRALDASYNARLNARQAMTNE-FYLTLVYRFG-----HAALGKRAH- : 143
trwK : DNVFCRQLDEKYRESFTGYNLMVND-LYLTVVYRFVSDK-VLSFFAKRER- : 143
traB : DHFFPAYLDQYNRKLHGDSKQLIND-LYLTVIYKQVGDK-TQKFLAKFEK- : 149
VIRB4 : RSVFAASLNEAFEQRVLSGQLLRND-HFLTLIVYF------QAALGKVKRR : 136
160 * 180 * 200
ORF4 : PLPQAFAQE--RAESIDALEKVVREVESSLSRYD----------VERL : 184
ptlC : HGQAEVRRQ--LLAHVRRMDELGSLIETTLRSHGEN-------HEQAITVL : 185
trwK : ELPDQKKHR--QESCIKALEDINRTLGQSFKRYG----------AELL : 179
traB : PLRDEIQQM--QNEALEGLEDISEQLLEAMKPYRFSSWVSIIVTNAVLKFL : 198
VIRB4 : FTKLSGKRENDLAGQLRNMEDLWHVVAGSLKAYG----------LHRL : 174
* 220 * 240 *
ORF4 : CIYEHN--GVYFSEPLELFAF--LVNSEWQRLP---LAQAPLRTLLP---T : 225
ptlC : GCETDS-AGRRYSRTLTLLEF--LLTGHWQPVR---VPAGPVDAYLG---S : 227
trwK : SVYEK--GGHAFSAPLEFLAR--LVNGEHIPMP---ICRDRFSDYMA---V : 220
traB : RLIKKNVKNLLKSMNQTFLTKPLLSNATSLNLRRLTLIQKRWSSFLSSQIW : 249
VIRB4 : GIREKQ--GVLFLEIGEALRL--LMTGRFTPVP---VVSGSLGASLY---T : 215
260 * 280 * 300
ORF4 : TRP----------FFGN----------EATELRSPTKTTYGAMLGINAYPP : 256
ptlC : SR----------ILAGA----------EMMELRAPTCRRVAQFLDFKEYGT : 258
trwK : NRP---------MFSKW---------GEVGELRSLLTGLRRFGMLEIREYDD : 253
traB : NGPSCLFAVIVSVSTSWTTALLAPCGGDVVQLRTVDHNFYTTGLEFREYEE : 300
VIRB4 : DR----------VICGK----------RGLEIRTPKDSYVGSIYSFREYPA : 246
* 320 * 340 *
ORF4 : ESKSVFLNHLLTQPFSFVLSQSFSFLQMESARWKLKLSKNRMINAGDDALS : 307
ptlC : HLEPGMLNALLYEDYEYVLTHSFSAVGKRQALAYLQRQRAQLANVQDAAYS : 309
trwK : ALEPGQLNVLLESDYEFVLTHSFSVLSRPAAKEYLQRHQKNLIDARDVATD : 304
traB : DLEPGQLNMLKEADFEYLLTQSFSCLSESSAKTFLTHQEKSLQETRDRAQS : 351
VIRB4 : KLRPGMLNALLSLDFPLVLTQSFSFLTRPQAHAKLSLKSSQMLSSGDKAVT : 297


360 * 380 * 400
ORF4 : QVDELDDAVDDLTARRWVMGDHHFSLFVKAGSLRELNDHIAEARIALSEGG : 358
ptlC : QIDDLAHAEDALVNGDFVIGEYHFSMMLLGADPRQLRRDVSSAMTRIQERG : 360
trwK : QIEELDEALNQLISGHFVMGEHHCTLTVYGETVQQVRDNLAHASAAMLDVA : 355
traB : QLAQLGTALDMLTSREFVMGYHHGTVHVWDNPQNAVQAKARRVKVMLTGCG : 402
VIRB4 : QIGKLSEAEDALASNEFVMGSHHLSLCVYADDLNSLGDRGARARTRMADAG : 348
* 420 * 440 * 46
ORF4 : ITAAREDLALASAFWAQLPAQFKFRPRLSPINSKNMAGHAPLHNFPQGRRH : 409
ptlC : FLATPVTLALDAAFYAQLPANWAYRSRKAMLTSRNFAGLCSFHNFYGGKRD : 411
trwK : VLPKPVDLALEAGYWAQLPANWQWRPRPAPITSLNFLSFSPFHNFMSGKPT : 406
traB : VVGGTLSLASEAAYYRRLPGNQKWAPRPVPINSWNFLHFSPFHNFMRGKPD : 453
VIRB4 : AVVVQEGLGMEAAYWSQLPGNFKWRTRPGAITSRNFAGFVSEENFPEGASS : 399
0 * 480 * 500 *
ORF4 : GNHWGDALTMFITSANTPYYFSFHAADPFDESGGTKKDVGHILVLGPTGSG : 460
ptlC : GNPWGPALSLLSTPSGQPFYFNFHHS-GLDEDCRGQMMLGNTRLIGQSGSG : 461
trwK : GNPWGPAVTILKTVSGTPLYFNFHAS-KEEEDATDKRLLGNTMLIGQSSSG : 456
traB : NNPWGPALTMFRTISGTPLYFNFHVT-PLEELSYGKRPLGHALITGMSGEG : 503
VIRB4 : G-HWGTAIARFRTNGGTPFDYIPHEH----------DVGMIAIFGPIGRG : 438
520 * 540 * 560
ORF4 : KTALIAFLLCMLQKFGV------TSVLFTKDRDTELVIRALGGTYYPLKPG : 505
ptlC : KTVLLNFLLCQLQKFRSADADGLITTFFDKDRGAEICIRALDGQYLRIRDG : 512
trwK : KTVLLGFLLAQAQKFKP------TIVAFDKDRGMEISIRAMGGRYLPLKTG : 501
traB : KTTLLNFLLAQSMKYNP------RLFVYDRDRGMEPFIRSVGGYYKVLQQG : 548
VIRB4 : KTTLMMFVLAMLEQSMVDRAG--IVVFFDKDRGGELLVRATGGTYLALRRG : 487
* 580 * 600 *
ORF4 : EPTGWNPFWLDPAKPGNVQYLNRFVRRLCTRPSQT-LSVTDEIEITEQAINA : 555
ptlC : EPTGFNPLQLPCTDR-NVMFLDSLLAMLARAHDSPLTSAQHAT-LATAVRT : 561
trwK : EPSGFNPFFHAPADAR-KPDFPQTVREEARGCRWRG--HAPGRGRNRPGHYA : 549
traB : MPSGFAPLQIEPTKR-NNALIKNLFRICVETTNNGPISATMATELAEGVDA : 598
VIRB4 : TPSGLAPLRGLENTAASHDFLREWIVALIESDGRGGISPEENRRLVRGIHR : 538
620 * 640 * 660
ORF4 : VLR----MDLEHRRLGRVLDFMT-------KDRSGIYAQLQRWGYAREHGK : 595
ptlC : VLR----MPASLRRMSTLLQNLT----QATSEQRELVRRLGRWCRDDGAGG : 604
trwK : MMSDS--IDKSLRRLSLLLQFLPNPRSDDMDARPTVHARLVKWCEG----- : 593
traB : VMGEGSLIPREARTVTILDGYVN----EVVENGVSLKGLLREWTRE----- : 640
VIRB4 : QLS----FDPQMRSLAGLREFLL----H--GPAEGAGARLQRWCRG----- : 574


680
700
ORF4 : PDGPNAWLFDNPRDTLIDNFGSALTTGFDVTSFLKDDE-----LRSPINMH : 641
ptlC : T-GMLWWVFDNPNDCLL--DFSRPGNYGIDGTAFLDNAE-----TRTPISMY : 647
trwK : --GDYGWLFDNPTDALL--DLSTHQIYGFDITEFLDNPE-----ARIPVMMY : 635
traB : --GQYGWLFDNDKDSLL--DLSANDIFGFDLSEFIAAKEEVSSPARTPLIVY : 687
VIRB4 : --HALGWAFDGEVDEVL--KLDPS-ITGFDMTHLIEYEE-----VCAPAAAY : 615
720
740
760
ORF4 : LFHLTESLIDG-RRLALFIAEFWRALGDPEMADFAKDKLKTIRKKNGFVVL : 691
ptlC : LLHRMNEAMDG-RRFVYLMDEAWKWIDDPAFAEFAGDQQLTIRKKNGLGVF : 697
trwK : LLYRTESMIDG-RRFMYVFDEFWKPLQDEYFEDLAKNKQKTIRKQNGIFVF : 685
traB : LLYRVRDSIDGKRRVIQCFDEFHAYLDDPVIEREVKRGIKTDRKKDAIYVF : 738
VIRB4 : LLHRIGAMIDG-RRFVMSCDEFRAYLLNPKFSAVVDKFLLTVRKNNGMLLL : 665
780
800
ORF4 : DSQSPSDALNHRISRTLIEQTPLKILFSNPDAVYSEYTSGGLNCSDREFDL : 742
ptlC : STQMPSSLLGARVAASLVQQCAIETYLPNPRADRAEYLD-GFKCTETEYQL : 747
trwK : ATQEPSDALESNIAKTLIQQCAIYLFLANPKADYEDYTQ-GFKLTDSEFEL : 735
traB : ATQEPNDALSSRIGRTIMSQTVIKICLRDPEAIREDY---AF-LIDAEYDA : 785
VIRB4 : ATQQPEHVLESPLGASLVAQCMIKIFYPSPTADRSAYID-GLKCTEKEFQA : 715
820
840
860
ORF4 : VKK-HPRGSRMFLVKQGHHSVVAKLDLQG-------------FDRELALLS : 779
ptlC : LRS-MAEDSHLFLVKQGRQAVVAQLDLSG-------------MDDELALLS : 784
trwK : VRG-LGEFSRRFLIKQGDQSALAEMNLGKFRTIVDGETVERDFDDELLVLS : 785
traB : LMS-ITEHSRQFLVKQGQQSALASFNLYP-RNSDDIDADIKTMDNVLSVLS : 834
VIRB4 : LREDMTVGSRKFLLKRESGSVLCEFDLRD-------------MREYVAVLS : 753
880
900
ORF4 : SREANIEVVQQLIAQFGQ-DPIKWLPHFDQHRRMA------ : 813
ptlC : GNARNLRCFEQALALTRERDPNDWIAVFH-RLRREASAGLR : 824
trwK : GTPDNAEIAESIIAEVGD-SPAVWLPIFLDRVKAERSDV-- : 823
traB : GEPQNAEIAHELVERLGN-DPEVWLKEYW-RLTA------- : 866
VIRB4 : GRANTVRFAARIREAQEG-NSSGWLSEFMARHHEAED---- : 789

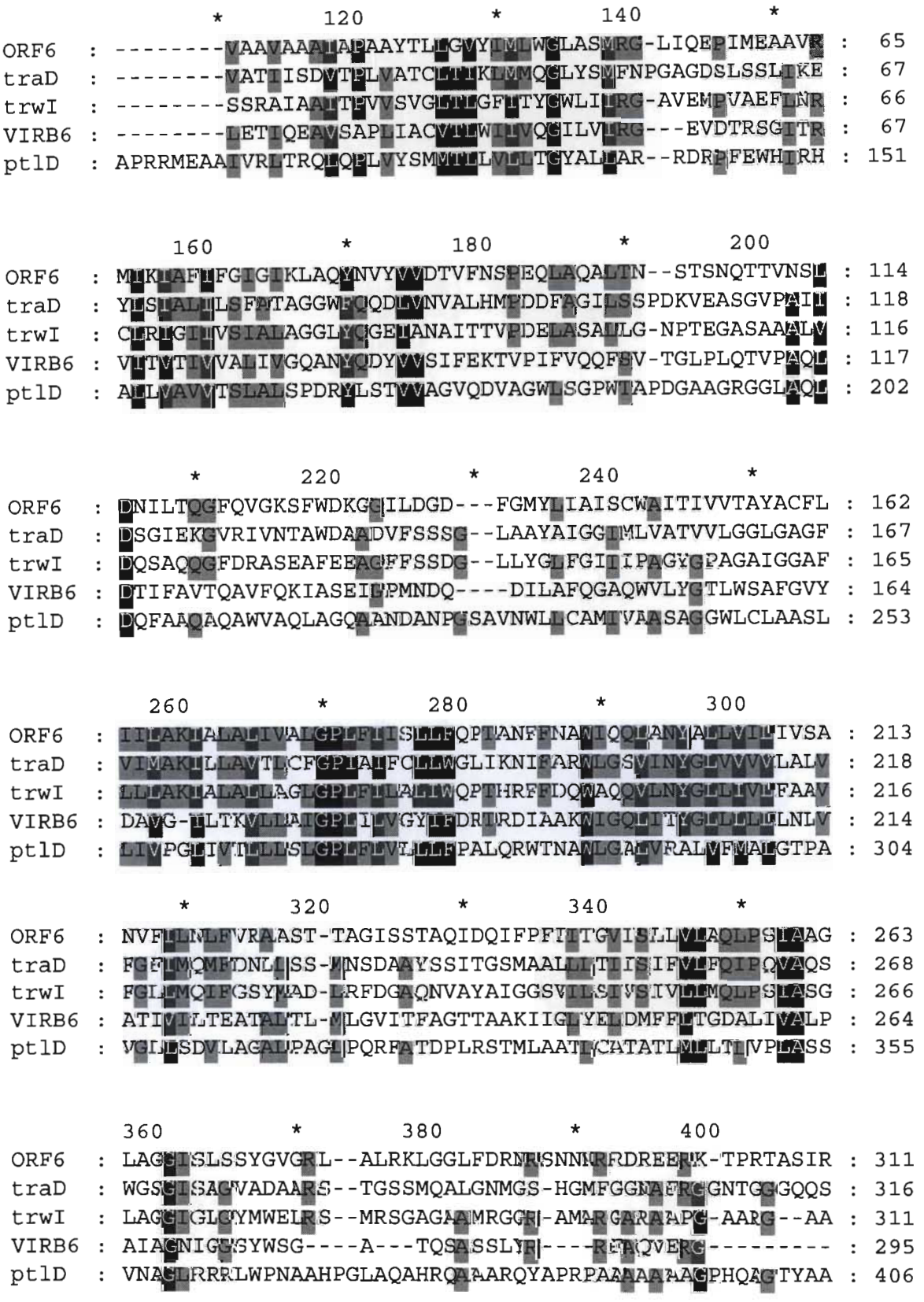
d) ORF5


20
40
ORF5 : -MKNAKHTALAVALFLSVAVTPLSSFATGIPVVDIAAIEQAIQQVSYLQQQ : 50
trwJ : -MKKLVMTAAVAAILG--AASPV--MAQGIPVFDGTRALDFVQQFARMKEQ : 46
traC : MKKSLTAVLLTTGLILGGAQRPS----AGIIVSNPTELIKQGEQLEQMAQQ : 47
virB5 : -MKTTQLIATVLTCSF-LYIQPA---RAQFVVSDPATEAETLATALATAEN : 46

```
                60          *          80          *          100
ORF5   : LKQMKNQL-------DAMTQDRGMAGLLS---GQNRNYLPADWNSAMNVLN :  91
trwJ   : LDTAKDQLAEAQRMYEAVTGGRGLGDLMRN--AQLREYLPDDLRTVYDSAN :  95
traC   : LEQLKSQLETQKNMYESMAKTTNLGDLLGTSTNTLANNLPDNWKEVYSDAM :  98
virB5  : LTQTIAMV-------TMLTSAYGVTGLLTS--LNQKNQYPST-KDLDNEMF :  87

                *          120         *          140         *
ORF5   : SGGGSFGSLASAAQQIKQMQSVLSSSDLS--RLSPQMQQYIDKVRSVSASQ : 140
trwJ   : -GGGYSGISGSINDILRDERLNGSVADMR--RSIEERSR---TARTTDKAV : 140
traC   : --NSSSSVTPSVNSMMG--QFNAEVDDMSPRQAIAYMKQ---KLDGKRCLR : 142
virB5  : --SPRMPMSTTARAITSDTDRAVVGSDAE---ADLLRSQ----ITGSANSA : 129

              160         *          180         *          200
ORF5   : QALGQQAYSTASQRVNLLQTLTNQISSAT--DPKAVWDLQARIQSEQSQLQ : 189
trwJ   : GCGPTKGHSNASRR------LKDEISRTQ--DQKAIEELQARIAGEQAAIQ : 183
traC   : PCDGRKAYNNQMQELSDMQALTEQIKSTP--DLKSLADLQARIQTSQGAIQ : 191
virB5  : GIA-ADNLETMDKRLTANADTSAQLSRSRNIMQATVTNGLLLKQIHDAMIQ : 179

               *          220         *          240         *
ORF5   : NDQSRLQSVAQLTQA---QSVATKQIA-NELRSQTSGTGN----FPSLDTS : 232
trwJ   : NETTKLQMIAQLRQA---EQALISEQR-RERNMPILSSGNQG--MPTIQ-- : 226
traC   : GEQATWNLMNMLQQS---QDKLLRAQKDRATRNFVFGTGGDVTASPSIN-- : 237
virB5  : NVQATNLLTMATAQAGLHEAEEAAAQR-KEHQKTAVIFGA----LP----- : 220

ORF5   : VGTN : 236
trwJ   : ---- :   -
traC   : ---- :   -
virB5  : ---- :   -
```

f) ORF6

```
                    *          20          *          40          *
ORF6   : ------------------------------------------------------- :   -
traD   : ------------------------------------------------------- :   -
trwI   : ------------------------------------------------------- :   -
VIRB6  : ------------------------------------------------------- :   -
ptlD   : MAGLSRILLSCTLACLLAGQAAQASVDDPTRAGGDNRVRALRADQARRDVL :  51

                60          *          80          *          100
ORF6   : --------------------MAGPYEQ----VFTYVTNVCDSYLGSS--- :  23
traD   : --------------------MAFTLVQD---IFAKVLGAITSMVSAN--- :  24
trwI   : --------------------MAFELFTP---LFNKILQTTATYVTDI--- :  24
VIRB6  : --------------------MNFTIPAPFTAIHTIFDVAFTTGLDSM--- :  27
ptlD   : LTACRDDPGHRRGEPDCVNAERAQALQQWQAAAMTSVDAAFSDLAGALRNA : 102
```


ORF6 : --------VAAVAAAIAPAAYTLLGVYIMLWGLASMRG-LIQEPIMEAAVR : 65
traD : --------VATTISDVTPLVATCLTIKLMMQGLYSMFNPGAGDSLSSLIKE : 67
trwI : --------SSRAIAAITPVVSVGLTLGFITYGWLILRG-AVEMPVAEFLNR : 66
VIRB6 : --------LETIQEAVSAPLIACVTLWILVQGILVLRG---EVDTRSGITR : 67
ptlD : APRRMEAAIVRLTRQLQPLVYSMMTLLVLLTGYALLAR--RDRPFEWHIRH : 151
ORF6 : MIKIAFIFGIGIKLAQYNVYVVDTVFNSPEQLAQALTN--STSNQTTVNSL : 114
traD : YLSIALLLSFATAGGWFQQDLVNVALHMPDDFAGILSSPDKVEASGVPAII : 118
trwI : CLRIGIIVSIALAGGLYQGEIANAITTVPDELASAILG-NPTEGASAAALV : 116
VIRB6 : VITVTIVVALIVGQANYQDYVVSIFEKTVPIFVQQFSV-TGLPLQTVPAQL : 117
ptlD : ALLVAVVTSLALSPDRYLSTVVAGVQDVAGWLSGPWTAPDGAAGRGGLAQL : 202
ORF6 : DNILTQGFQVGKSFWDKGQILDGD---FGMYLIAISCWAITIVVTAYACFL : 162
traD : DSGIEKGVRIVNTAWDAADVFSSSG--LAAYAIGGIMLVATVVLGGLGAGF : 167
trwI : DQSAQQGFDRASEAFEEAGFFSSDG--LLYGLFGIIIPAGYGPAGAIGGAF : 165
VIRB6 : DTIFAVTQAVFQKIASEIGPMNDQ----DILAFQGAQWVLYGTLWSAFGVY : 164
ptlD : DQFAAQAQAWVAQLAGQAANDANPGSAVNWLLCAMIVAASAGGWLCLAASL : 253
ORF6 : IILAKIALALIVALGPLFIISLLFQPTANFFNAWIQQLANYALLVIIIVSA : 213
traD : VIMAKILLAVTLCFGPIAIFCLLWGLIKNIFARWLGSVINYGLVVVVLALV : 218
trwI : LLFAKIALALLAGLGPLFILALIWQPTHRFFDQWAQQVLNYGLLIVLFAAV : 216
VIRB6 : DAVG-ILTKVLLAIGPLILVGYIFDRIRDIAAKWIGQLITYGLLLLIFLNLV : 214
ptlD : LIVPGLIVILLLSLGPLFLVILLFPALQRWTNAWLGALVRALVFMALGTPA : 304
ORF6 : NVFLLMLFVRAAST-TAGISSTAQIDQIFPFIITGVISLLVLAQLPSIAAG : 263
traD : FGFIMQMFDNLISS-VNSDAAYSSITGSMAALLLTLISIFVLFQIPQVAQS : 268
trwI : FGLLMQIFGSYMAD-LRFDGAQNVAYAIGGSVILSIVSIVLLMQLPSLASG : 266
VIRB6 : ATIVLLTEATALTL-MLGVITFAGTTAAKIIGLYELDMFFLTGDALIVALP : 264
ptlD : VGLLSDVLAGALPAGLPQRFATDPLRSTMLAATLYCATATLMLLTLVPLASS : 355
ORF6 : LAGSISLSSYGVGRL--ALRKLGGLFDRNRSNNQRRDREEQK-TPRTASIR : 311
traD : WGSGISAGVADAARS--TGSSMQALGNMGS-HGMFGGNAERGGNTGGGQQS : 316
trwI : LAGGIGLGYMWELRS--MRSGAGAAMRGGR-AMARGARAAPG-AARG--AA : 311
VIRB6 : AIAGNIGGSYWSG----A---TQSASSLYR----RTAQVERG--------- : 295
ptlD : VNAGLRRRLWPNAAHPGLAQAHRQAAARQYAPRPAAAAAAAGPHQAGTYAA : 406

```
          *       420          *       440          *       46
ORF6   : YHAADIRRPRLPPTKHSHLLNKNTFFVLA---------------------- : 340
traD   : AGGGSGSNSGGSSGSNLSGKARG-SRGKKAA-------------------- : 346
trwI   : VGAANMAKTVATGGAGVARAAAGYFRGRKAG-------------------- : 342
VIRB6  : ------------------------------------------------- :   -
ptlD   : SATPAPAPARPAPSFPAHAYRQYALGGARRPPPRVRRDDRPAPAPDRRVLP : 457

          0
ORF6   : ------ :   -
traD   : ------ :   -
trwI   : ------ :   -
VIRB6  : ------ :   -
ptlD   : RKPNLP : 463
```

g) ORF7

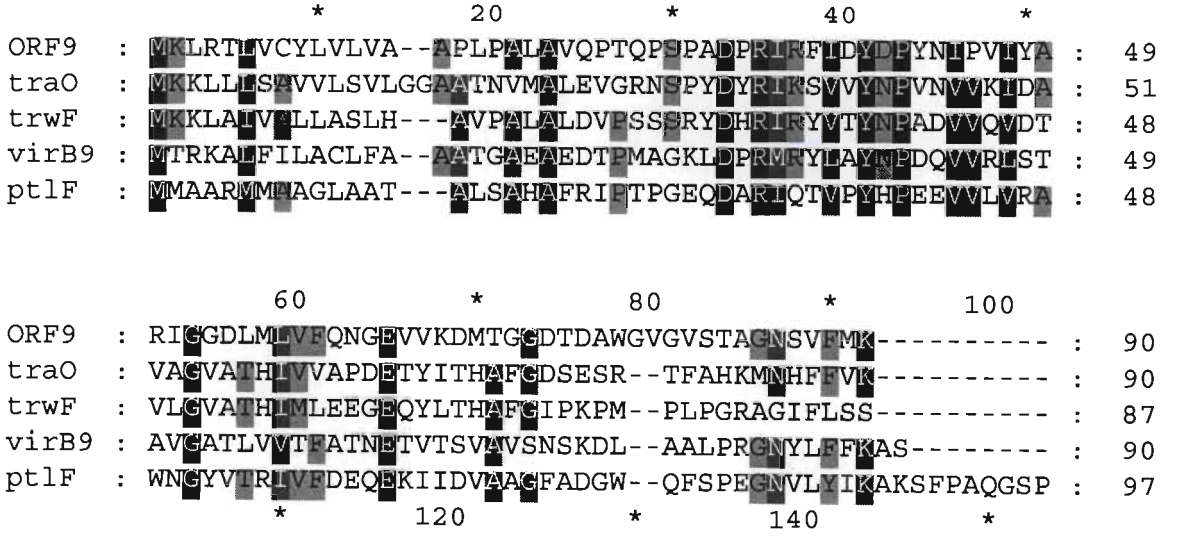
```
                  *         20          *         40          *
ORF7   : MLVRTLLAVLLAGTLVGCASVP-PS--CGGLNRRPVNQPPQAGVVNQS---CG : 47
traN   : -MRSLLLMGVL--LISACSSGHKPPPEPDWSNTVPVN--KTIPVDTQ----GG : 44
trwH   : -MKTILFAILMTGLLSACASAPKPK-QPSDFNREPVN--KIVPVEIQ----RG : 45
virB7  : -MKYCLLCLVV--ALSGCQQTNDTIASCKGPIFPLNVGRWQPTPSDLQLRNSGG : 50

ORF7   : HSATA : 52
traN   : RNES- : 48
trwH   : AL--- : 47
virB7  : KYDGA : 55
```

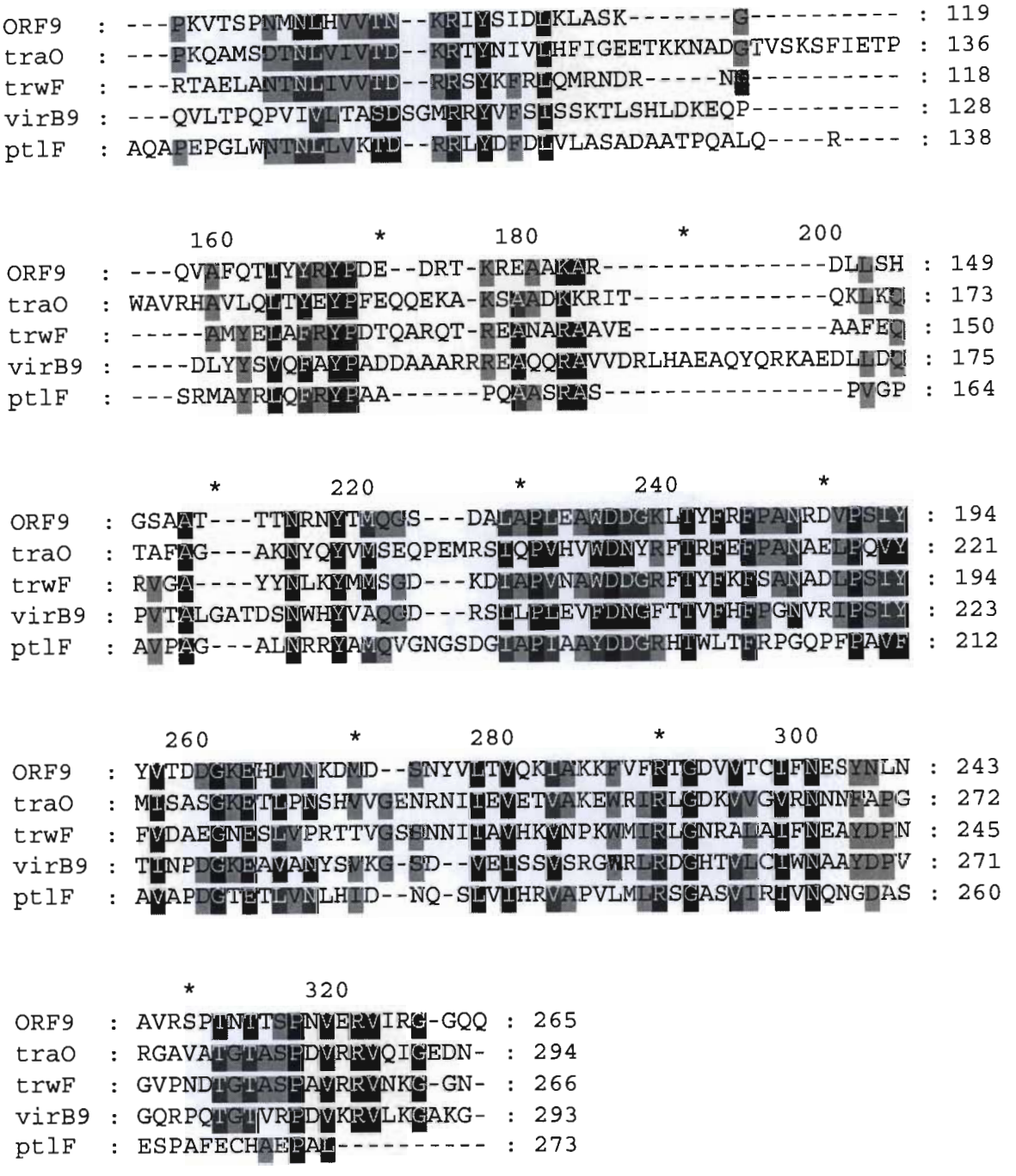
h) ORF8

```
                     *         20          *         40          *
ORF8   : ---------------------------------MDDYKLQQQLNELRQFIEA :  19
traE   : ----------------------------------------MKANKKTGLT :  10
trwG   : ----------------------------------------MSKKQPKPVK :  10
VIRB8  : ------------------------------------------------MLVA :   4
ptlE   : MGHPGHRPRHARRLRPERAARLAVSRWQAPDSHQHRRPGARLVMPDPRPLT :  51
                60          *         80          *        100
ORF8   : HKDVQHLL--DNSAAWADTQREQDEKSKAQAWRVAAGASVFGMVAUFIAWT :  68
traE   : REAIKEFN--ESRKGLEVDLMDEVLKSRRTAWMVATGSAVVTVFALSLVGY :  59
trwG   : AEQLKSYY--EESRGLERDLIGEFVKSRKTAWRVATASGLFGLLGM-VCGI :  58
VIRB8  : RESLAEHY--KEVEAFQTARAKSARRLSKIIAAVAAIAILGNVAQAFAIAT :  53
ptlE   : PDQTHGRGHAEAAVDWEASRLYRLAQSERRAWTVAWAALAVTALSLIAIAT : 102
```

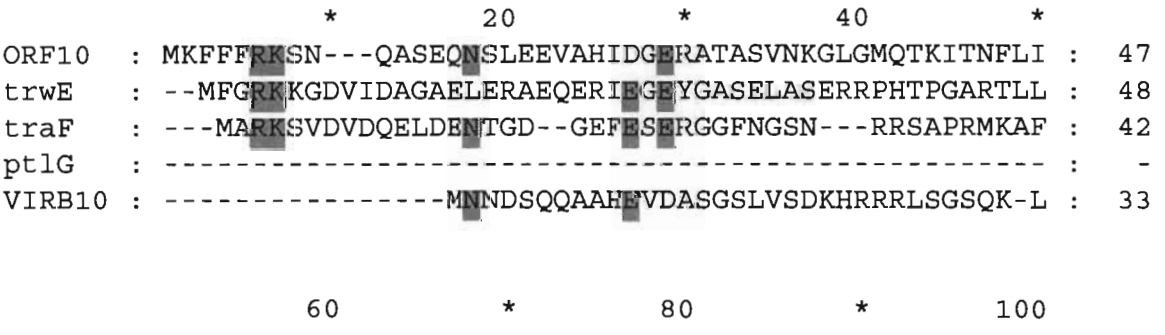

ORF8 : AIRTAYVPPAAPQVLVIDKTTGHVEPLVSLKEVQESVDEAVTR-HYITEFL : 118
traE : VVHKYSQPIPAHLIT-LNEATHEVQQVKLTRD-QTSYGDEIDK-FWLTQYV : 107
trwG : VG--FSQPAPPPLVLRVDNATGAVDVVTTLREHESSYGEVVDT-YWLNQYV : 106
VIRB8 : MV---PLSRLVPVYLWIRADGTVDSEVSISRLPATQEEAVVNASLWE--YV : 99
ptlE : ML---PLKTTIPYLIEVEKSSGAASVVTQFEPRDFTPDTLMNQ-YWLTRYV : 149
ORF8 : RCRENYTFDTAEENYYCAAAYMSPQLQTQWAAFWDTKNPDSPYNYYKNSAT : 169
traE : IHRESYDFYSVQVDYTAVALMSTPNVAESYQSKFKGRN--GLDKVLGDSET : 156
trwG : LNREAYDYNTIQMNYDTTALLSAPAVQQDYYKLEDGSN--ARDRVLGNKAR : 155
VIRB8 : RLRESYDADTAQYAYDLVSNFSAPTVRQDYQQFENYPNPSSPQVILGKRGR : 150
ptlE : AARERYDWHTIQHDYDYVRLLSAPAVRHDYETSYEAPD--APDRKYGAGTT : 198
ORF8 : VKLDIDSITLNINSDGAQDTATVRFTRY--VKKNDHQ-EVTRWVATLAYKY : 217
traE : ARVKINSVILDKP----HGVRTIRFTTVRRVRTNPVDDQPQRWIAIMGYEY : 203
trwG : ITVRVRSIQPNG-----RQQATVRFTTQQHNSNGTVE-RPQHQIATLGYTY : 200
VIRB8 : VEVEHIASNDVT-----PSTQQIRYKRT-LVVDGKMP-VVSTWTATVRYEK : 194
ptlE : LAVKTLLSAIDHG-----KGVGTVRFVRTRRDADGQGAAESSIWVATVAFAY : 244
ORF8 : VDQDIGAAAVPLSVAPPQSDSPVPAPVPSSPSPPALSAPQAPASQHVGTVQ : 268
traE : KS-------LAMNAEQ-RYVNPLGFRVTSYRVNPEVN-------------- : 232
trwG : IG-------APMRSSD-RLLNPLGFQVTSYRADPEILNN------------ : 231
VIRB8 : VT--------SLPGRL-RLTNPAGLVVTSYQTSEDTVSNVGQGAP------ : 230
ptlE : DQP------RALTQAQ-RWLNPLGFAVTSYRVDAEAGQP------------ : 276

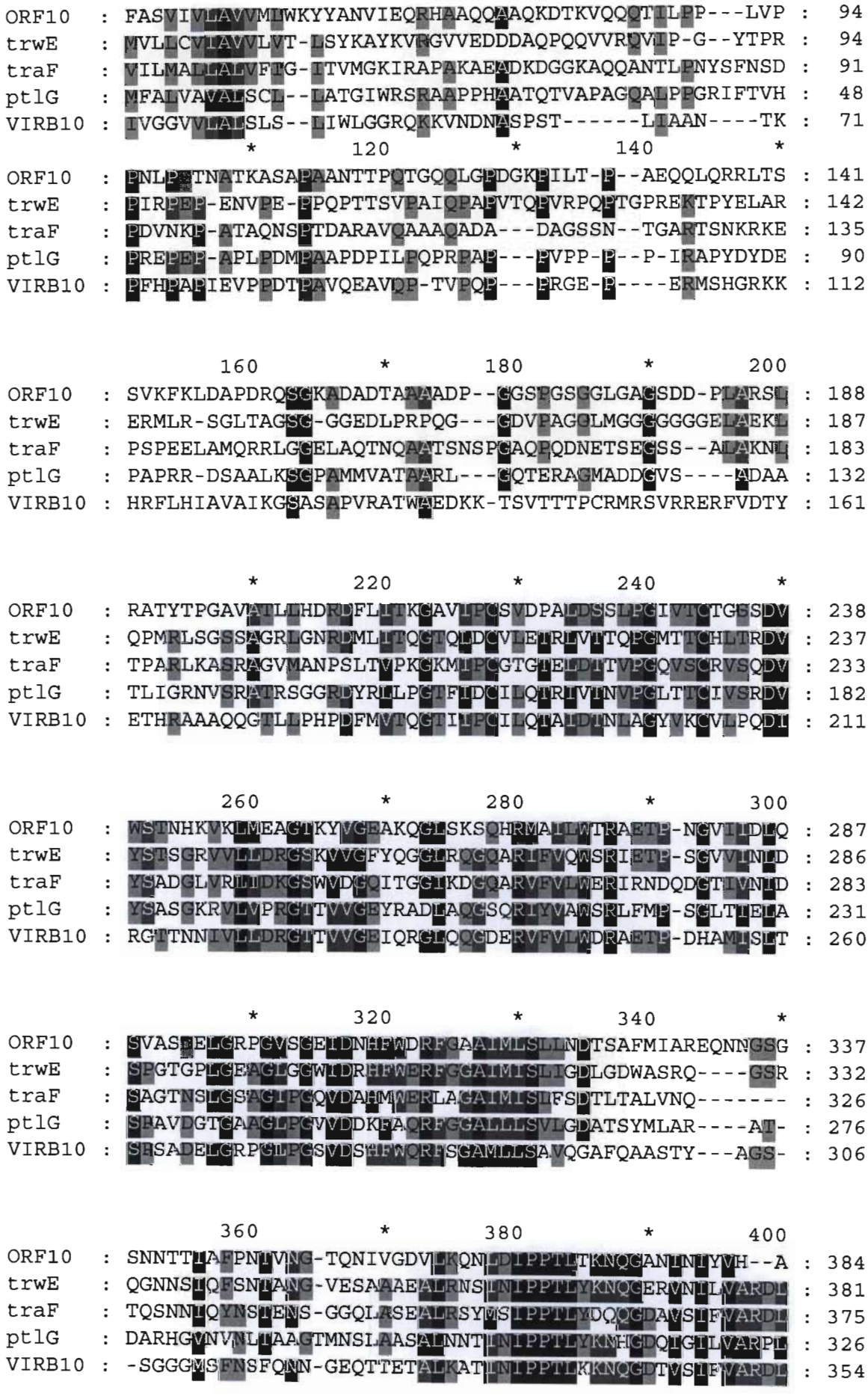
i) ORF9


ORF9 : MKLRTLVCYLVLVA--APLPALAVQPTQPSPADPRIRFIDYDPYNIPVIYA : 49
traO : MKKLLLSPVVLSVLGGAATNVMALEVGRNSPYDYRIKSVVYNPVNVVKIDA : 51
trwF : MKKLALVLLLASLH---AVPALALDVPSSSRYDHRIRYVTYNPADVVQVDT : 48
virB9 : MTRKALFILACLFA--AATGAEAEDTPMAGKLDPRMRYLAYNPDQVVRLST : 49
ptlF : MMAARMMMAGLAAT---ALSAHAFRIPTPGEQDARIQTVPYHPEEVVLVRA : 48
ORF9 : RIGGDLMLVFQNGEVVKDMTGGDTDAWGVGVSTAGNSVFMK---------- : 90
traO : VAGVATHLVVAPDETYITHAFGDSESR--TFAHKMNHFFVK---------- : 90
trwF : VLGVATHLMLEEGEQYLTHAFGIPKPM--PLPGRAGIFLSS---------- : 87
virB9 : AVGATLVVTFATNETVTSVAVSNSKDL--AALPRGNYLFFKAS-------- : 90
ptlF : WNGYVTRLVFDEQEKIIDVAAGFADGW--QFSPEGNVLYIKAKSFPAQGSP : 97



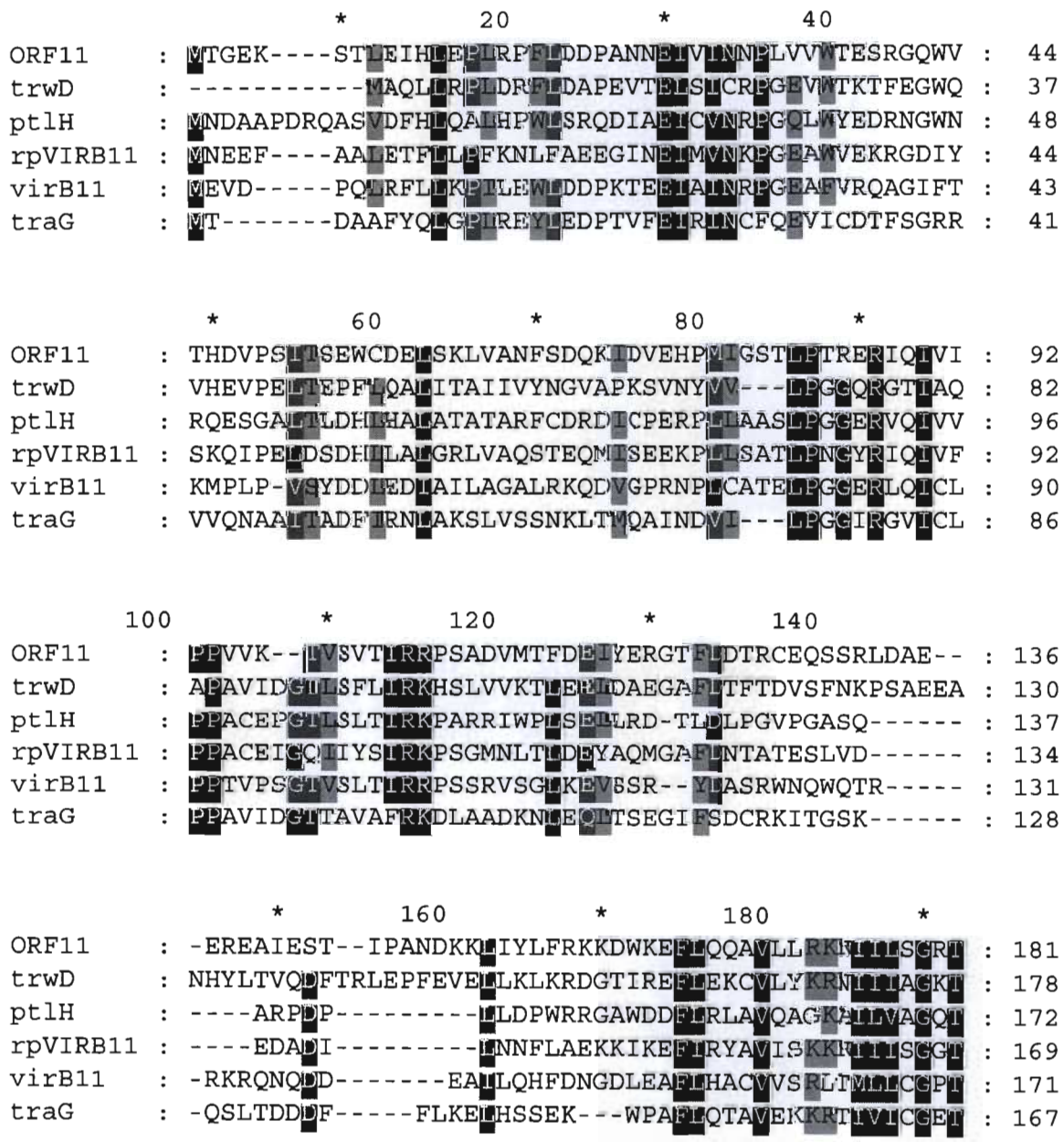
j) ORF10




ORF10 : FASVIVLAVVMLWKYYANVIEQRHAAQQAAQKDTKVQQQTILPP---LVP : 94
trwE : MVLLCVIAVVLVT-LSYKAYKVHGVVEDDDAQPQQVVRQVIP-G--YTPR : 94
traF : VILMALLALVFTG-ITVMGKIRAPAKAEADKDGGKAQQANTLPNYSFNSD : 91
ptlG : MFALVAVALSCL--LATGIWRSRAAPPHAATQTVAPAGQALPPGRIFTVH : 48
VIRB10 : IVGGVVLALSLS--LIWLGGRQKKVNDNASPST------LIAAN----TK : 71
* 120 * 140 *
ORF10 : PNLPATNATKASAPAANTTPQTGQQLGPDGKPILT-P--AEQQLQRRLTS : 141
trwE : PIRPEP-ENVPE-PPQPTTSVPAIQPAPVTQPVRPQPTGPREKTPYELAR : 142
traF : PDVNKP-ATAQNSPTDARAVQAAAQADA---DAGSSN--TGARTSNKRKE : 135
ptlG : PREPEP-APLPDMPAAPDPILPQPRPAP---PVPP-P--P-IRAPYDYDE : 90
VIRB10 : PFHPAPIEVPPDTPAVQEAVQP-TVPQP---PRGE-P----ERMSHGRKK : 112
160 * 180 * 200
ORF10 : SVKFKLDAPDRQSGKADADTAPAADP--GGSPGSGGLGAGSDD-PLARSL : 188
trwE : ERMLR-SGLTAGSG-GGEDLPRPQG---GDVPAGQLMGGGGGGGELAEKL : 187
traF : PSPEELAMQRRLGGELAQTNQAATSNSPGAQPQDNETSEGSS--ALAKNL : 183
ptlG : PAPRR-DSAALKSGPAMMVATAARL---GQTERAGMADDGVS----ADAA : 132
VIRB10 : HRFLHIAVAIKGSASAPVRATWAEDKK-TSVTTTPCRMRSVRRERFVDTY : 161
* 220 * 240 *
ORF10 : RATYTPGAVATLLHDRDFLITKGAVLPCSVDPALDSSLPGIVTCTGSSDV : 238
trwE : QPMRLSGSSAGRLGNRDMLITQGTQLDCVLETRLVTTQPGMTTCHLTRDV : 237
traF : TPARLKASRAGVMANPSLTVPKGKMLPCGTGTELDTTVPGQVSCRVSQDV : 233
ptlG : TLIGRNVSRATRSGGRDYRLLPGTFLDCILQTRLVINVPGLTTCIVSRDV : 182
VIRB10 : ETHRAAAQQGTLLPHPDFMVTQGTILPCILQTALDTNLAGYVKCVLPQDI : 211
260 * 280 * 300
ORF10 : WSTNHKVKLMEAGTKYVGEAKQGLSKSQHRMALLWTRAETP-NGVILDLQ : 287
trwE : YSTSGRVVLLDRGSKVVGFYQGGLRQGQARIFVQWSRIETP-SGVVINLD : 286
traF : YSADGLVRLLDKGSWVDGQITGGIKDGQARVFVLWERIRNDQDGTIVNLD : 283
ptlG : YSASGKRVLVPRGTTVVGEYRADLAQGSQRIYVAWSRLFMP-SGLTIELA : 231
VIRB10 : RGTTNNIVLLDRGTTVVGEIQRGLQQGDERVFVLWDRAETP-DHAMISLT : 260
* 320 * 340 *
ORF10 : SVASPELGRPGVSGEIDNHFWDRFGAAIMLSLLNDTSAFMIAREQNNGSG : 337
trwE : SPGTGPLGEAGLGGWIDRHFWERFGGAIMISLIGDLGDWASRQ----GSR : 332
traF : SAGTNSLGSAGIPGQVDAHMWERLAGAIMISLFSDTLTALVNQ------- : 326
ptlG : SRAVDGTGAAGLPGVVDDKFAQRFGGALLLSVLGDATSYMLAR----AT- : 276
VIRB10 : SRSADELGRPGLPGSVDSHFWQRFSGAMLLSAVQGAFQAASTY---AGS- : 306
360 * 380 * 400
ORF10 : SNNTTIAFPNIVNG-TQNIVGDVLKQNLDIPPTLTKNQGANINIYVH--A : 384
trwE : QGNNSIQFSNIANG-VESAAAEALRNSLNIPPTLYKNQGERVNILVARDL : 381
traF : TQSNNIQYNSIENS-GGQLASEALRSYMSIPPTLYDQQGDAVSIFVARDL : 375
ptlG : DARHGVNVNLIAAGTMNSLAASALNNTLNIPPTLYKMHGDQIGILVARPL : 326
VIRB10 : -SGGGMSFNSFQNN-GEQTTETALKATLNIPPTLKKNQGDTVSIFVARDL : 354


420
ORF10 : TWISAACTI---------- : 393
trwE : DFSDVYSLESIPTK-------- : 395
traF : DFSGVYTLADN---------- : 386
ptlG : DFSILRGTNE---------- : 336
VIRB10 : DFFGVYQLRLTGGAARGRNRRS : 376

k) ORF11

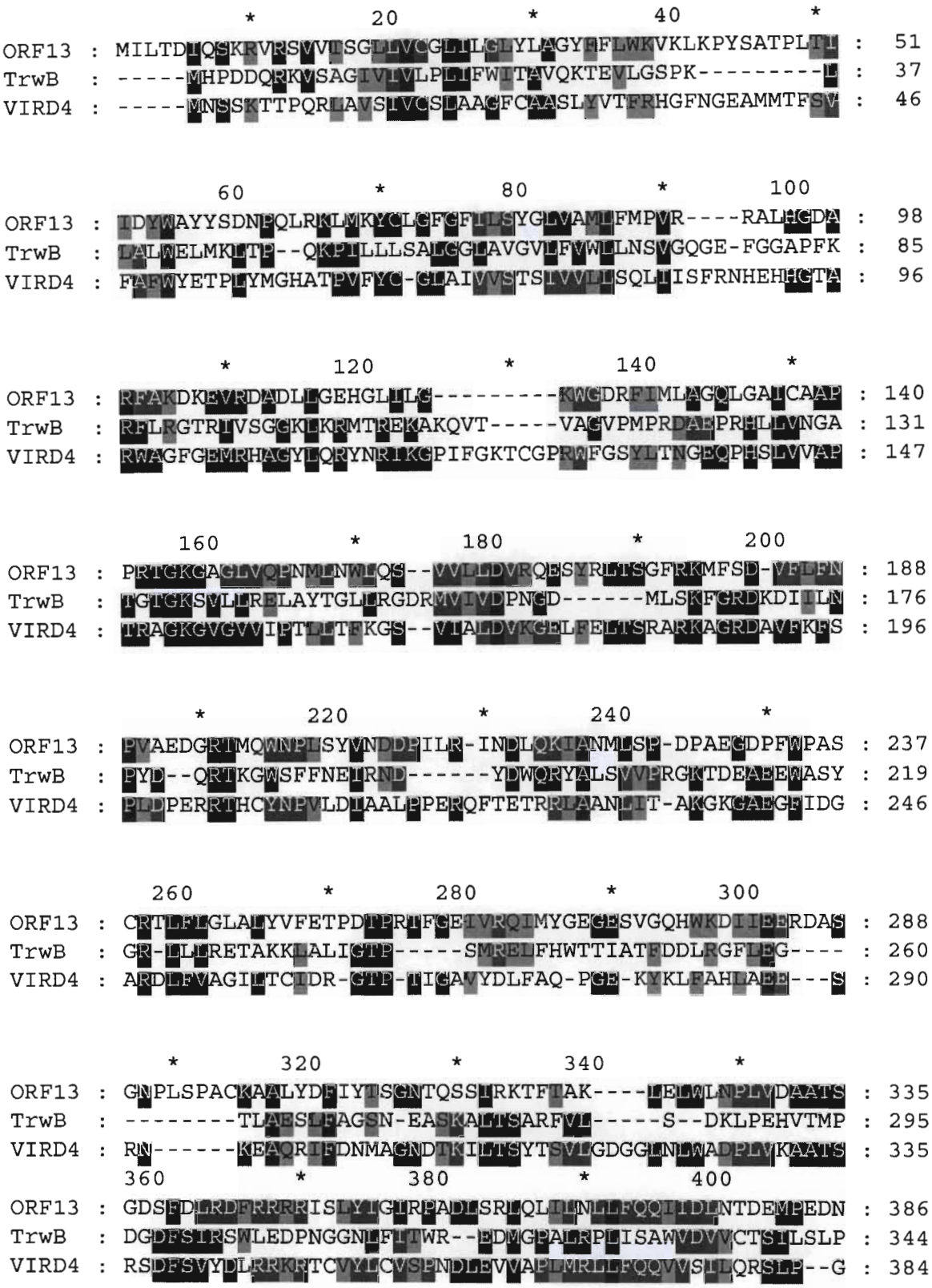

20 40
ORF11 : MTGEK----STLEIHLEPLRPFLDDPANNEIVINNPLVVWTESRGQWV : 44
trwD : ----------MAQLLRPLDRFLDAPEVTELSICRPGEVWTKTFEGWQ : 37
ptlH : MNDAAPDRQASVDFHLQALHPWLSRQDIAEICVNRPGQLWYEDRNGWN : 48
rpVIRB11 : MNEEF----AALETFLLPFKNLFAEEGINEIMVNKPGEAWVEKRGDIY : 44
virB11 : MEVD-----PQLRFLLKPILEWLDDPKTEEIAINRPGEAFVRQAGIFT : 43
traG : MT-------DAAFYQLGPIREYLEDPTVFEIRINCFQEVICDTFSGRR : 41
60 80
ORF11 : THDVPSITSEWCDELSKLVANFSDQKIDVEHPMIGSTLPTRERIQIVI : 92
trwD : VHEVPELTEPFLQALITAIIVYNGVAPKSVNYVV---LPGGQRGTIAQ : 82
ptlH : RQESGAITLDHIHAIATATARFCDRDICPERPLLAASLPGGERVQIVV : 96
rpVIRB11 : SKQIPELDSDHILAIGRLVAQSTEQMISEEKPLLSATLPNGYRIQIVF : 92
virB11 : KMPLP-VSYDDLEDLAILAGALRKQDVGPRNPLCATELPGGERLQICL : 90
traG : VVQNAAITADFTRNLAKSLVSSNKLTMQAINDVI---LPGGIRGVICL : 86
100 120 140
ORF11 : PPVVK--LVSVTIRRPSADVMTFDEIYERGTFLDTRCEQSSRLDAE-- : 136
trwD : APAVIDGLLSFLIRKHSLVVKTLEELDAEGAFLTFTDVSFNKPSAEEA : 130
ptlH : PPACEPGTLSLTIRKPARRIWPLSELLRD-TLDLPGVPGASQ------ : 137
rpVIRB11 : PPACEIGQIIYSIRKPSGMNLTLDEYAQMGAFLNTATESLVD------ : 134
virB11 : PPTVPSGTVSLTIRRPSSRVSGLKEVESR--YLASRWNQWQTR----- : 131
traG : PPAVIDGTTAVAFRKDLAADKNLEQLTSEGIFSDCRKITGSK------ : 128
160 180
ORF11 : -EREAIEST--IPANDKKLIYLFRKKDWKEFLQQAVLLRKRTLLSGRT : 181
trwD : NHYLTVQDFTRLEPFEVELLKLKRDGTIREFLEKCVLYKRNIIIAGKT : 178
ptlH : ----ARPDP---------LLDPWRRGAWDDFLRLAVQAGKAILVAGQT : 172
rpVIRB11 : ----EDADI---------LNNFLAEKKIKEFIRYAVISKKRTLLSGGT : 169
virB11 : -RKRQNQDD-------EALLQHFDNGDLEAFLHACVVSRLTMLLCGPT : 171
traG : -QSLTDDDF-----FLKELHSSEK---WPAFLQTAVEKKRTIVICGET : 167

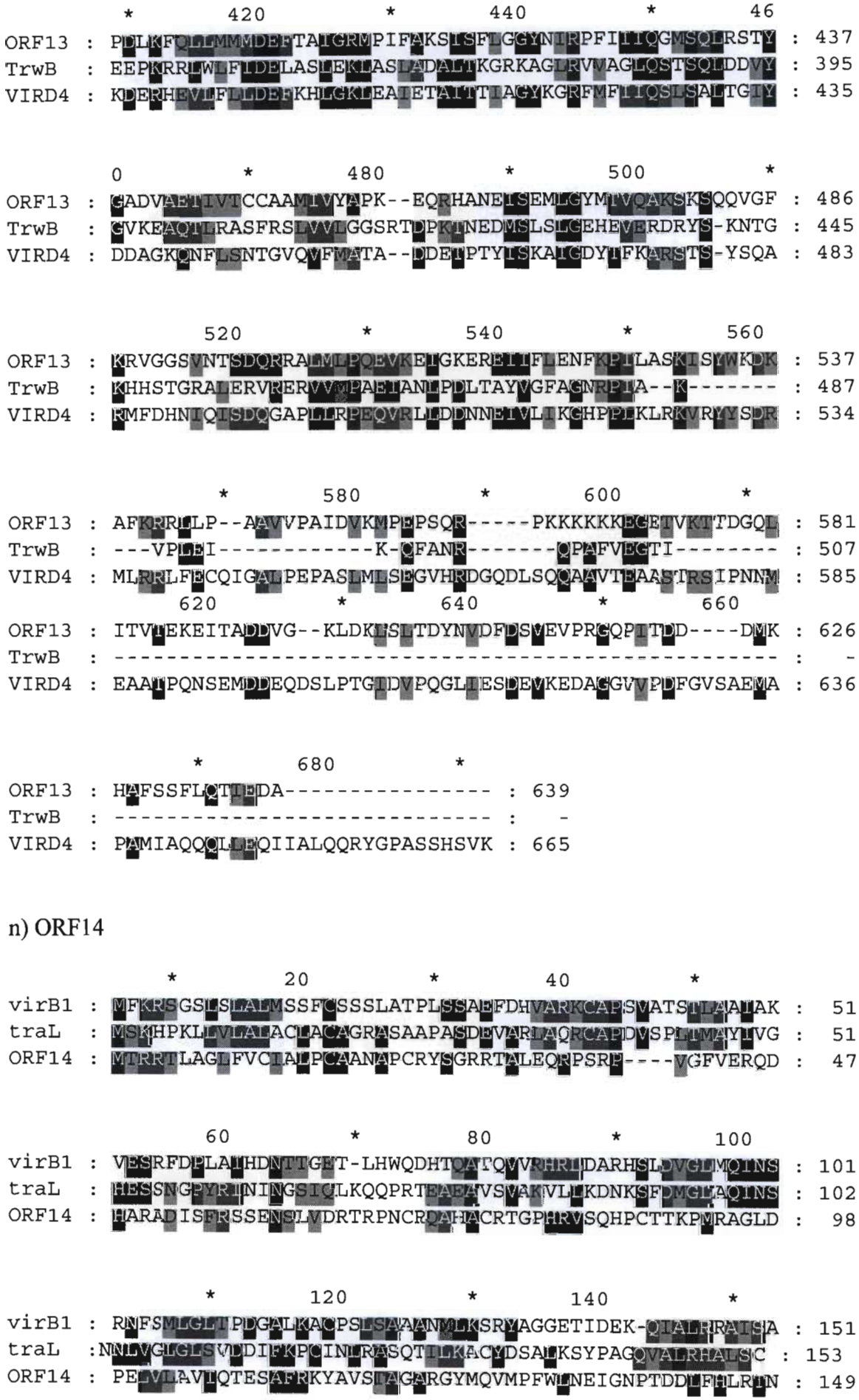
```
                200         *         220         *         240
ORF11     : GSGKTTLANSLCMMIPIQERIITTVEDAREVRLP-HPNQVNLVYSRDNR : 228
trwD      : GSGKTTFARSLIEKVPPEERIITTIEDVHELFLPNHPNRVHMLYGYG-- : 224
ptlH      : GSGKTTLMNALSGEIPPRERIVTIEDVRELRLDPATNHVHLLYGTPTE : 220
rpVIRB11  : STGKTTFTNAALTEIPAIERLITVEDAREVVLSSHPNRVHLLASKGGQ : 217
virB11    : GSGKTTMSKTLISAIPPQERLITIEDTLELVIP-HDNHVRLLYSKNGA : 218
traG      : GSGKTVLTRALLKSLHKDERVIILEDVHEVTVDHVVEAVYMMYGDAGK : 215

                    *       260         *       280
ORF11     : G-LAKLSPKQMFESGLRMRPDRVLPAELRGEEAYFFFQNVVNSGHPGA : 275
trwD      : --AGRVSADECLAACMRQSPDRIFLAELRGNEAWEYLN-SLNTGHPGS : 269
ptlH      : GRTAAVSATELLRAALRMAPTRILLAELRGGEAFDFLQ-ACASGHSGG : 267
rpVIRB11  : G-RANVTTQDLIEACLRLRPDRIIVGELRGKEAFSFLR-AINTGHPGS : 263
virB11    : G-LGAVSAEHLLQASLRMRPDRILLGEMRDDAAWAYLS-EVVSGHPGS : 264
traG      : --IGPVSATDALRACMRLTPGRIIMTELRDDAAWDYLK-ALNTGHPGG : 260

              *       300         *       320         *
ORF11     : ITSIHSNTSKLAFRRLTNMIQSSQEGRGLEQGVILEMLYALVDIVIQT : 323
trwD      : ITTTHANNALQTFERCATLIKKSDVGRQLEMEMIKLVLYITIDVVLFF : 317
ptlH      : ISTCHAASADMALQRLTLMCMQHPNCQMLPYSTLRALVESVIDIVVVV : 315
rpVIRB11  : ISTLHADSPAMAIEQLKLMVMQADFG--MPPEEVKKYILSVVDIVVQL : 309
virB11    : ISTIHGANPIQGFKRLFSLVKSSVQGASLEDRTLIDMLSTAIDVIIPF : 312
traG      : VMSTHANSARDAFNRIGLLIKATPIGRMLDMSDIMRMLYSTIDVVVHM : 308

              340         *       360         *
ORF11     : DTVQINSAGENRKMVTEIYFDPAYAL--KQIG----- : 353
trwD      : K---------DRKLV-EVFYDPIFSK--SKMA----- : 337
ptlH      : ERRAGQG---ARRRVVDIWYRDGLPA--P-------- : 339
rpVIRB11  : KRASG-----GKRYVSEVYYKKNKNS--ESML----- : 334
virB11    : RAYE------DVYEVGEIWLAADARRRGETIGDLLNQ : 343
traG      : E---------KRKIK-EIYFDPEYKM--QCVNGSL-- : 331
```

l) ORF12

```
                    *         20          *         40          *
ORF12 : MSFLRALMVG-AALAALAACTSKPTHISFLDRVLTVEEFTAQSDIRDNVIAE :  51
eex   : ---MKKLLL--LIPFFLVACDAS----------HDVDWYKKHDKERKATIEE :  37
eexw  : ---MKKILVPLLMVSALSACKDEPT--------HDVQYYLDNPKERTEKIAE :  41

                  60          *         80          *         100
ORF12 : CSSDPGDLNDDPNCVNAKQSLRASVTGTGNFPSLDITPPHKKGKQIGN :  99
eex   : CKKN-ADELQKPDCKNAREADRQLFVLGK-EDGEINSPKI-------- :  75
eexw  : CKNNPGEKALAANCVNATEAQRKAMFQG---DG---MPKIR------- :  76
```

m) ORF13


ORF13 : MILTDIQSKRVRSVVTSGLLVCGLILGLYLAGYHFLWKVKLKPYSATPLTL : 51
TrwB : -----MHPDDQRKVSAGIVLVLPLIFWITAVQKTEVLGSPK---------L : 37
VIRD4 : -----MNSSKTTPQRLAVSIVCSLAAGFCAASLYVTFRHGFNGEAMMTFSV : 46
ORF13 : IDYWAYYSDNPQLRKLMKYCLGFGFILSYGLVAMLFMPVR----RALHGDA : 98
TrwB : LALWELMKLTP--QKPILLLSALGGLAVGVLFVWLLNSVGQGE-FGGAPFK : 85
VIRD4 : FAFWYETPLYMGHATPVFYC-GLAIVVSTSLVVLLSQLIISFRNHEHHGTA : 96
ORF13 : RFAKDKEVRDADLLGEHGLLLG---------KWGDRFIMLAGQLGALCAAP : 140
TrwB : RFLRGTRLVSGGKLKRMTREKAKQVT-----VAGVPMPRDAEPRHLLVNGA : 131
VIRD4 : RWAGFGEMRHAGYLQRYNRIKGPIFGKTCGPRWFGSYLTNGEQPHSLVVAP : 147
ORF13 : PRTGKGAGLVQPNMLNWLQS--VVLLDVRQESYRLTSGFRKMFSD-VFLFN : 188
TrwB : TGTGKSVLLRELAYTGLLRGDRMVIVDPNGD------MLSKFGRDKDIILN : 176
VIRD4 : TRAGKGVGVVIPTLLTFKGS--VIALDVKGELFELTSRARKAGRDAVFKFS : 196
ORF13 : PVAEDGRTMQWNPLSYVNDDPILR-INDLQKIANMLSP-DPAEGDPFWPAS : 237
TrwB : PYD--QRTKGWSFFNELRND------YDWQRYALSVVPRGKTDEAEEWASY : 219
VIRD4 : PLDPERRTHCYNPVLDLIAALPPERQFTETRRLAANLIT-AKGKGAEGFIDG : 246
ORF13 : CRTLFLGLALYVFETPDTPRTFGETVRQIMYGEGESVGQHWKDIIEERDAS : 288
TrwB : GR-LLLRETAKKLALIGTP-----SMRELFHWTTIATFDDLRGFLEG---- : 260
VIRD4 : ARDLFVAGILTCIDR-GTP-TIGAVYDLFAQ-PGE-KYKLFAHLAEE---S : 290
ORF13 : GNPLSPACKAALYDFIYTSGNTQSSLRKTFTAK----LELWLNPLVDAATS : 335
TrwB : --------TLAESLFAGSN-EASKALTSARFVL-----S--DKLPEHVTMP : 295
VIRD4 : RN------KEAQRIFDNMAGNDTKILTSYTSVLGDGGLNLWADPLVKAATS : 335
ORF13 : GDSFDLRDFRRRRISLYIGIRPADLSRLQLLLNLLFQQIIDLNTDEMPEDN : 386
TrwB : DGDFSIRSWLEDPNGGNLFLTWR--EDMGPALRPLISAWVDVVCTSILSLP : 344
VIRD4 : RSDFSVYDLRRKRTCVYLCVSPNDLEVVAPLMRLLFQQVVSILQRSLP--G : 384



n) ORF14

```
            160         *         180         *         200
virB1  :  YNTGNFIRGFANGYVRKVETAAQS----LVPALIEPPQDD-HKALKSEDTW  : 197
traL   :  YNTGSLTNGISNGYVTKVINVARQSTDLKIPTLLPDGQTS-EDSTANEPQQ  : 203
ORF14  :  LRYG-------CTILRHYLDTERG------NLFLALGRYNGSRGQAAYPNL  : 187

             *        220         *        240
virB1  :  DVWGSYQRRSQEDGVGGSIAPQPPDQDNGKSADDNQVLFDLY  : 239
traL   :  AKSTAPQYDGEQD-VFGSGDGDAFSRNNTDAFLTRQETAKGE  : 244
ORF14  :  VLASWSQWRSMPS--------------------DSQITMK--  : 207
```

o) ORFklca

```
                    *         20          *         40
ORFklca  :  ----MNK--EVTASVVVAEDEFRDFLSKHFGIRF--AKRGEALVFAWLL  :  41
KLCA     :  ----MNTQDQPVTASLVAEAQRLDFLPTYFGPRL--MMRGEALVYAWLR  :  43
kilC     :  MTDVMFQIPSPIVATRVAEADRLRFPPTYFGPSMPRMLRGEALVFGWMG  :  49

           *         60          *         80          *        1
ORFklca  :  RLSKVPIEWTRLQYYTLSNSGFYLAP-----RELRIS----ECELSADA  :  81
KLCA     :  RLCER-YNGAYWHYYTLSDGGFYLAPDLAERLEIEVDGNGFRGELSADA  :  91
kilC     :  RLCAA-YHGGFWHFYTLSNGGFYMAPEHDGPLRIEVDGNGHAGELSADA  :  97

             00          *        120          *        140
ORFklca  :  LGIVATMLTLRQLAHESAASVEADSTHAAAKLAVKASVMFAQNYHNLAA  : 130
KLCA     :  AGIVATLFALGQLAAEIAG---TDAAD--ALIDR---------YHFLRG  : 126
kilC     :  AGIVATLFALNQLCAELAG-----TADADALIDR---------YHHLAA  : 132

              *        160
ORFklca  :  YAVKHAESINIYRAID  : 146
KLCA     :  FAAGHPEAAAAIYRAID  : 142
kilC     :  FASEHAEAAAAIYRAID  : 148
```

p) ORFkorc

```
                    *         20          *         40          *
ORFkorc  :  --MHNEIRPQCLRPAECWTQPTANEIRAVLRMAGMTGGAAAKLLGLGAKGD  : 49
KORC     :  MTNDANIRLECLKPADGWAQPSGEEVREVLRLAGLTGGKAAKVLASAAKGD  : 51
KORC2    :  -MSDVNIRLECLRPAERWVQPTGAEIREVLHLAGLTGGQAARILGLGAKGD  : 50

                  60          *         80
ORFkorc  :  RTIRRWIGGDSRILWQRQIHDGPAHA---------  : 75
KORC     :  RTIRRWVGEDTPIPYAAWALLCDYAGLGLIWKEV-  : 85
KORC2    :  RTVRRWVGEDSPIPYAAWAILCDLAGIGAIWKGQG  : 85
```