# CHAPTER 1

# Introduction

*Nor do I doubt that the most formidable armies ever*
*here upon earth is a sort of soldier who for their smallness are not visible.*
*[on microbes, 1640]. [Sir] William Petty 1623-1687 (London: Constable)*

## 1.1 Ovine Footrot - The disease

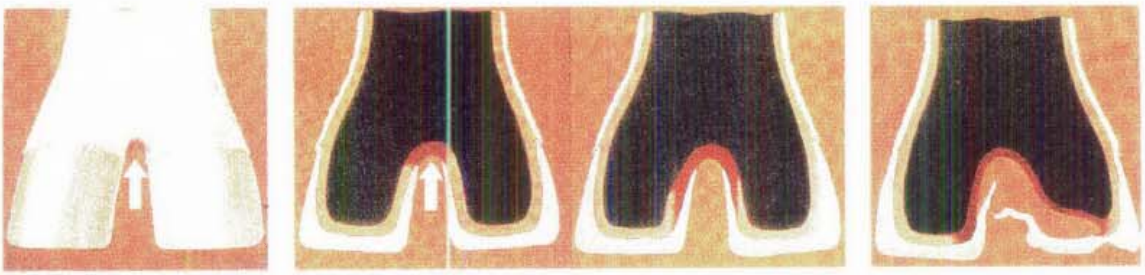
Footrot is a contagious and economically-significant disease of sheep, present in most countries in which commercial flocks are maintained, and thought to have been introduced into Australia with the first importations of British sheep in the 19th century (Stewart, 1989a). Footrot is characterised by a mixed bacterial infection of the hoof and is thought to involve synergy between several species of bacteria (Egerton & Roberts, 1969). The principal causative agent of footrot is the Gram-negative anaerobe *Dichelobacter nodosus* (formerly *Bacteroides nodosus*) (Dewhirst *et al.*, 1990) since it is only in the presence of *D. nodosus* that the disease is able to be transmitted from one animal to another (Beveridge, 1941).

The financial costs of footrot can be substantial and are attributable to decreased wool quality and production, decreased bodyweight, decreased fertility, lost market opportunities, stock losses from blowfly strike, decreased value of sheep when sold, and labour intensive and costly treatment, control and eradication programs (Plant & Walker, 1994; Stewart, Clark & Jarrett, 1984; Walker, 1988). These effects cost the sheep industry millions of dollars per annum; despite the implementation of the NSW Footrot Strategic Plan in 1988 (Plant & Walker, 1994), in the state of New South Wales alone, losses are estimated to be between A$43 million (Egerton & Raadsma, 1990; Marshall, Walker & Coveny, 1991) and A$50 million (Plant & Walker, 1994), making footrot one of the most economically-important endemic diseases of sheep. In Australia, about 50-60 million sheep are at risk of footrot (Stewart, 1989a).

### 1.1.1 Clinical Features

The disease profile for footrot consists of a continuum of clinical features, ranging from inapparent disease to severe disease. As a result, *D. nodosus* strains are classified as benign, intermediate, or virulent dependent upon the severity of the disease they cause in sheep.

Benign disease is the least contagious form of footrot and so tends to affect only a small percentage of the flock. Like intermediate and virulent disease, for the first 7-14 days, benign disease presents as interdigital dermatitis of macerated interdigital tissue which then extends into the avascular epidermal matrix of the hoof. Benign lesions rarely progress to 'underrunning', a process in which the underlying epidermal tissue of the hoof is progressively eroded (Figures 1.1 & 1.2), and, where underrunning does occur, it is mild with the occasional separation from the heel and posterior sole and very little accumulation of the necrotic exudate. As a result, such lesions result in only mild lameness which, upon the resumption of dry environmental conditions, will tend to self-heal. Production costs are therefore less severe than they are for virulent infections (Stewart, 1989a).



Footrot starts in the skin between the claws (interdigital cleft).

It then progressively spreads down the inside wall of the claw.

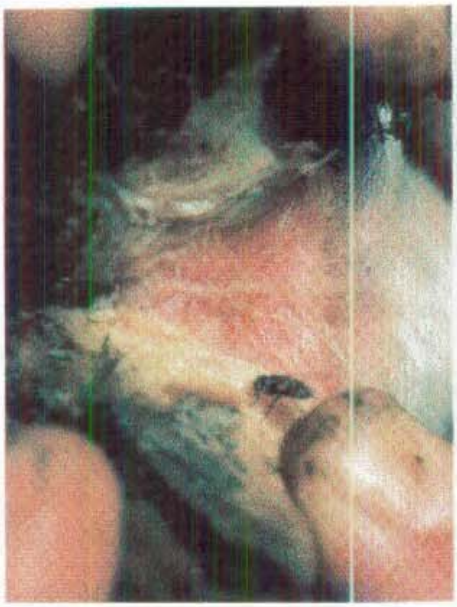
It then spreads across the sole, causing severe underrunning and separation; in severe cases it extends up the outside wall as well.

**Figure 1.1**: A schematic representation of the progression of a footrot lesion from interdigital dermatitis to severe underrunning of sole and horn of the hoof (modified from Walker, 1988).

Intermediate strains of *D. nodosus* exhibit virulence less severe than virulent strains but more severe than benign strains. Lameness, natural regression of disease, underrunning of the abaxial wall of the hoof, and production costs are dependent on the severity of the lesion which may vary greatly within this group.

Virulent strains cause the most contagious form of footrot, resulting in a rapid spread and development of lesions, and in the most severe cases, may affect up to 90% of the flock. Virulent disease is characterised by underrunning, eventually leading to physical separation of the hoof from the underlying epidermis, beginning at the heel, extending along the sole to the toe and, in severe cases, down the abaxial wall. In addition, a distinctive and foul smelling necrotic exudate is present. The severity of virulent disease results in inflammation, pain, severe lameness (Emery, 1988), and an infection that will persist unless intensive treatment is undertaken (Stewart, 1989a). In addition, virulent disease usually affects both digits and more than one foot, and results in a considerable loss in body weight of about 5 to 10 kg, a decrease in tensile strength and greasy weight of the wool, and increases susceptibility of sheep to fly strike (Stewart, 1982).

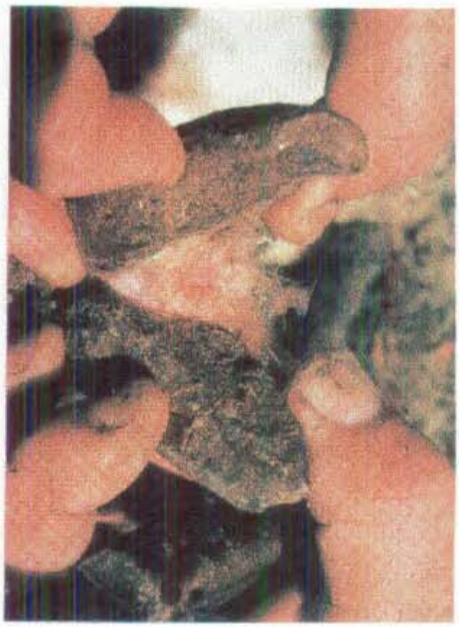
**Figure 1.2**: (pages 4 and 5 following). Footscoring of lesions in the clinical diagnosis of footrot. Benign strains of *D. nodosus* (scores of 1 to 2) are distinguishable from intermediate and virulent strains (scores of 3-5) (Plant & Walker, 1994).



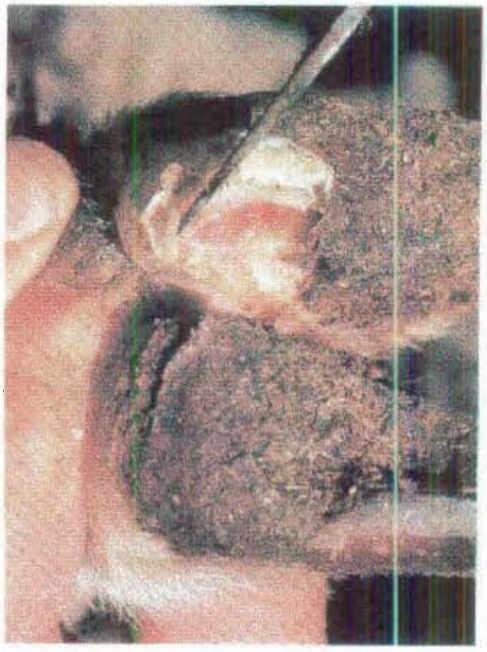
**Normal foot.** There is normal skin between the claws, with no reddening or inflammation and no loss of hair. There is no exudate present.

**Score 1:** Slight to moderate inflammation with some erosion between the claws. There is no underrunning or erosion of the skin or horn.

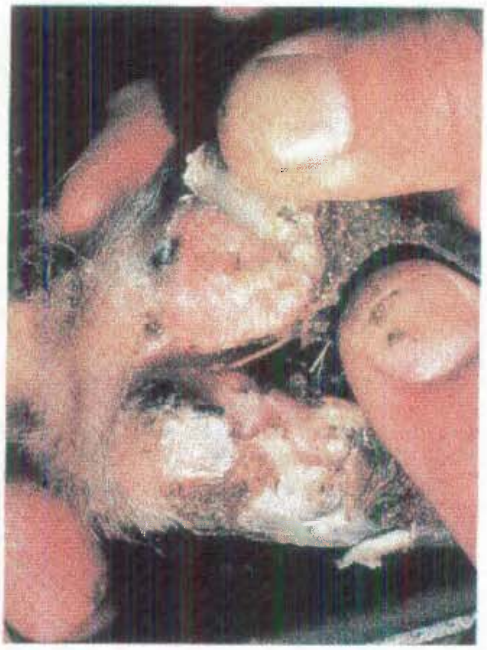
**Score 2:** The skin between the claws is inflamed and raw. This condition may involve part, or all, of the soft horn inside of the claws. There is no underrunning of the horn.



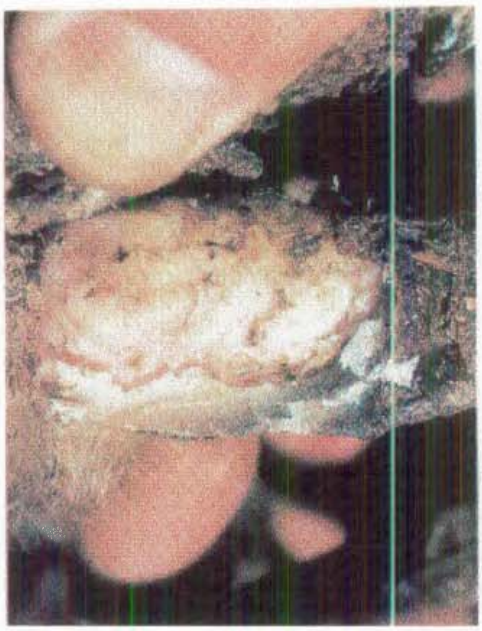
**Score 3a:** Separation of the skin horn junction, with underrunning extending no more than 5 mm.



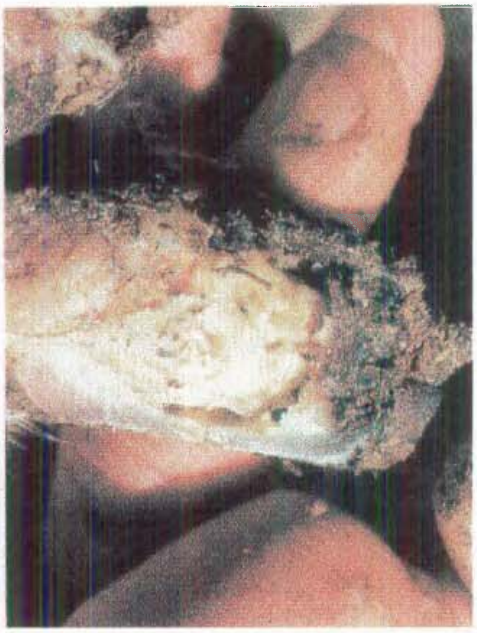
**Score 3b:** Underrunning no more than halfway across the heel or sole.



**Score 3c:** more extensive underrunning of the heel or sole but not extending to the outside edge of the sole of the claw.



**Score 4:** The underrunning extends to the outside edge of the sole of the claw and involves hard horn.



**Score 5:** This is a severe form of disease involving the sole, with extensive inflammation and underrunning of the hard horn of the hoof.

### 1.1.2 Epidemiology

The severity of footrot is complicated by many factors which include the temperature and humidity of the season, the type of pasture, the breed of sheep and the virulence of the strain of *D. nodosus*. Footrot is endemic in regions of Australia with a temperate climate (warm and humid) and high rainfall (above ~500 mm per annum). Outbreaks tend to occur seasonally, favouring the months of April/May and October/November in Australia, when conditions are warm and moist and there is lush pasture growth (Stewart, 1989a; Stewart *et al.*, 1984). Furthermore, Stewart *et al* (1984) observed that greater rates of transmission were apparent when lush clover was the dominant pasture.

Several studies have shown that virulent strains of *D. nodosus* are less pathogenic in British breeds of sheep than in Merinos. Suffolk and Dorset Merino crossbreeds (Beveridge, 1941), Merino and Border Leicester crossbreeds (Beveridge, 1941; Egerton & Roberts, 1969) and Romney, Border Leicester and Dorset Horn sheep (Skerman, Erasmuson & Morrison, 1982) were observed to be less susceptible to footrot than purebred Merinos. However, a later study (Parker, Cross & Hamilton, 1985) observed that where the interdigital skin was damaged prior to challenge, the incidence and severity of footrot were similar in all breeds. These results suggested that the integrity of the interdigital epithelium is the important factor in a resistant phenotype and showed that scarification of the interdigital skin upon challenge can substantially increase the pathogenicity of a low virulence strain (Stewart, 1989a).

Infective material is transferred directly from exposed lesions from the feet of chronically affected animals or symptomless carriers onto soil, mud, and pasture, thereby contaminating the feet of previously-unaffected livestock. If some form of damage to the interdigital skin has occurred, penetration and infection by *D. nodosus* may follow. In addition, it has been proposed (Stewart, 1989a) that the social behaviour of sheep, that is, walking in single file, may also contribute to spreading footrot.

Wheels, boots, blow flies, irrigation water and feet of other animals have been suggested as a means of transmission of footrot within and between properties, though this has not been demonstrated and is considered unlikely since *D. nodosus* is a fastidious organism that requires very specific conditions for infection to be successful (Plant & Walker, 1994; Stewart *et al.*, 1984; Walker, 1988).

Numerous investigations have established that footrot is not confined to sheep (Lambell, Hides & Blunden, 1991; Stewart, 1979; Walker, 1988), and *D. nodosus* can cause footrot in other ruminants including cattle (Egerton & Parsonson, 1966; Richards *et al.*, 1980; Toussaint & Cornelise, 1971; Wilkinson, Egerton & Dickson, 1970), goats (Claxton & O'Grady, 1985) and deer (Skerman, 1983). More recently, *D. nodosus* has been implicated in ulcerative pododermatitis in free-ranging African elephants (Keet *et al.*, 1997).

Virulence studies of *D. nodosus* isolates from other ruminants have been undertaken, since it had been recognised that other ruminants had the potential to act as passive physical carriers or reservoirs for ovine footrot, and hence could be important epidemiologically for control and eradication of footrot in sheep. Morphological examinations have revealed that *D. nodosus* isolated from the lesions in cattle, goats and deer differs from ovine isolates (Skerman, 1983), and biochemical analysis of *D. nodosus* proteases (Section 1.2.2) showed that their proteolytic activity was mild or negligible in cattle and deer, and severe in goats (Egerton, 1989).

Although many workers have isolated *D. nodosus* from cattle, these isolates, in general, have been found to be avirulent in sheep (Richards *et al.*, 1980; Toussaint & Cornelise, 1971; Walker *et al.*, 1973; Wilkinson *et al.*, 1970), an important finding particularly for Australia and New Zealand since sheep and cattle are often run together (Egerton, 1989). The failure to isolate from cattle strains of *D. nodosus* which are virulent

in sheep suggests that virulent transmission from cattle to sheep occurs only under exceptional circumstances (Richards *et al.*, 1980). According to the NSW Department of Agriculture, there have been only three cases where virulent strains of *D. nodosus* have been isolated from cattle (Plant & Walker, 1994), and one published report (Egerton & Parsonson, 1966). Only benign strains of *D. nodosus* have been isolated from deer (Skerman, 1983), and there has been no evidence of natural transmission from deer to sheep. In contrast to ovine footrot, the infection in cattle and deer tends to be restricted to the epidermis of interdigital tissues irrespective of the virulence of the isolate (Egerton, 1989), and presents as inflammation, superficial maceration, necrosis, fissuring and hyperkeratosis; there is no deep penetration or infection of interdigital tissue as there is in virulent ovine footrot lesions.

Goats are naturally susceptible to footrot, and both virulent and benign strains of *D. nodosus* have been isolated from them. However, the clinical features in goats are distinct from those observed in sheep in that, whether the infection is caused by a virulent or benign strain of *D. nodosus*, severe lameness is exhibited, though underrunning of the soles is uncommon even with virulent infections, and this is thought to be attributable to the higher interdigital cleft present in goats (Walker, 1988). Natural transmission of footrot from goats to sheep has been observed in New South Wales and Victoria (Walker, 1988).

## 1.2 Bacteriology of the causative agent

### 1.2.1 Classification

*D. nodosus* (formerly *Bacteroides nodosus*) is a strictly anaerobic, non-sporeforming, non-flagellated, slow-growing Gram-negative rod of 1 to 1.7 μm (Egerton, 1989) (Figure 1.3A). The cells are further characterised by knob-like swellings (Figure 1.3B) at their termini and polar fimbriae (or pili). *D. nodosus* dies rapidly away

from the host, due to oxygen sensitivity, however it can survive in the environment (soil, faeces, pasture) for 4 to 5 days (Laing & Egerton, 1981).

In the natural environment, *D. nodosus* is only capable of proliferation within the avascular epidermal tissues of the hoof in the host, and infects the host in an area somewhat remote from host immune/defence systems (Mattick, 1989). Thus, without treatment, more virulent infections tend to persist. Although synergistic relationships exist between fusiform and spirochaete-like bacteria associated with *D. nodosus* in footrot lesions, their exact nature and how they contribute to the pathogenic process is not well understood (Beveridge, 1941; Egerton & Roberts, 1969).

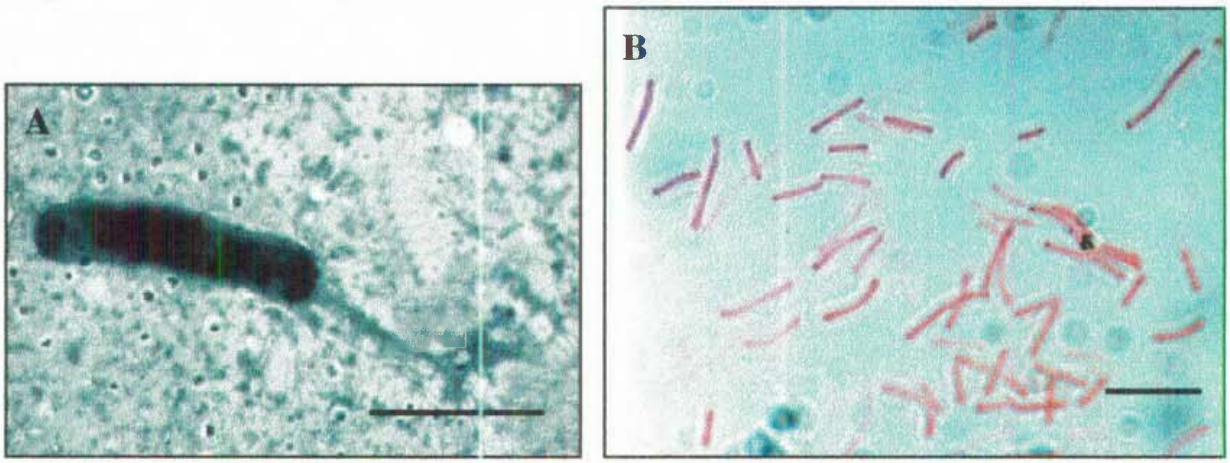



**Figure 1.3:** Electronmicrograph of a single cell of *D. nodosus* strain H1215 stained with phosphotungstic acid (**A**) (provided by Garry Bloomfield). Scale bar indicates 1 μm; Gram stain of *D. nodosus* strain C305 (**B**). Scale bar indicates 2 μm.

### 1.2.2 Virulence Determinants in *D. nodosus*

Microbial pathogenesis is typically multifactorial and therefore complex (Finlay & Falkow, 1989). Pathogens tend to have several biochemical mechanisms which may act together or individually to result in infection and disease. Removal of any such mechanisms may or may not render the organism avirulent. For this reason, observations by many workers have found that most diseases are caused by a very small number of strains within a species, namely those which have all of the necessary virulence determinants. These

observations are consistent with the continuum of virulence characteristics exhibited by strains of *D. nodosus*.

The main direction of recent research on footrot has been to identify, isolate and characterise antigens and virulence factors of *D. nodosus* with the aim of identifying immunogens which could lead to the development of more potent vaccines, or lead to the development of more efficacious diagnostic techniques to differentiate the more virulent strains of *D. nodosus* from the less virulent strains (McKern, Elleman & Hoyne, 1989). Antigens of *D. nodosus* include fimbriae (pili) (Anderson *et al.*, 1986; Elleman, 1988; Emery, Stewart & Clark, 1984b; Every, 1977), extracellular proteases (Gordon, Yong & Woodward, 1985; Green, 1985; Kortt, Burns & Stewart, 1983; Lilley, Stewart & Kortt, 1992; Moses *et al.*, 1989; Yong & Gordon, 1986), lipopolysaccharide endotoxins (Stewart, 1977) and outer membrane proteins (Emery, Clark & Stewart, 1984a). These antigens are briefly discussed below.

### (i) Fimbriae

Fimbriae are long proteinaceous appendages which are commonly found in Gram-negative bacteria, and although the role of these fimbriae in *D. nodosus* has not yet been established, it is suggested that, like many other fimbriae found in a polar location on the cell, *D. nodosus* fimbriae are not involved in conjugation, but function in surface translocation *via* twitching motility which, in turn, may have a role in attachment and colonisation of the epidermal hoof matrix by the bacterium (Every, 1979; Anderson *et al*, 1984; Depiazzi *et al*, 1985; Darymple *et al*, 1987; Moses *et al*, 1989b).

*D. nodosus* fimbriae are classified as type 4 *N*-methylphenylalanine fimbriae, and are common to other Gram-negative pathogenic bacteria including *Pseudomonas*, *Neisseria*, *Moraxella*, *Bacteroides*, and *Vibrio* (Ottow, 1975). Type 4 fimbriae typically have a polar location on the cell, a high conservation of structural proteins which constitute the fimbrial strand (Hobbs *et al.*, 1991; Mattick, 1989) and exhibit twitching motility that is greater in

virulent strains than in benign strains. These type 4 fimbriae typically have a methylphenylalanine *N*-terminus followed by a highly conserved region of 25 to 30 amino acids (Mattick, 1989). The significance of this conserved portion is thought to be related to a common mechanism of fimbrial biosynthesis in such organisms (Finlay & Falkow, 1989).

Both the basal and structural components of *D. nodosus* fimbriae are the principal antigenic structures mediating the K-agglutination reaction (Anderson *et al.*, 1984; Every, 1977; Lee, Alexander & McGowan, 1983; Stewart, 1978; Thorley & Egerton, 1981; Walker *et al.*, 1973) which has been used to classify *D. nodosus* cells into nine major serogroups designated A to I, which include eighteen subsidiary serotypes (Claxton, 1989; Claxton, Riberio & Egerton, 1983; Egerton, 1973; Lee *et al.*, 1983; Schmitz & Gradin, 1980). Geographical studies (Claxton *et al.*, 1983) indicate that serogroups A to H are present in all countries studied, with serogroup B being the predominant serogroup.

*D. nodosus* fimbrial subunit genes were isolated by immunological screening of a genomic DNA library from the virulent strain A198. Fimbrial genes were subsequently cloned into *E. coli* (Mattick *et al*, 1984; Anderson *et al*, 1984; Elleman & Hoyne, 1984), and although fimbrial subunit genes were expressed in *E. coli*, they were not assembled into mature fimbriae, and instead remained lodged in the inner membrane. This complication was eventually overcome by using *Pseudomonas aeruginosa* as an alternative host for the *D. nodosus* fimbrial subunit gene. The primary rationale for the selection of *P. aeruginosa* as an alternative host was that, like *D. nodosus*, *P. aeruginosa* possesses type 4 fimbriae, and so may have the cellular machinery necessary for maturation of the fimbrial subunit of *D. nodosus* (Elleman *et al.*, 1986b; Mattick *et al.*, 1987a).

Indeed, in *P. aeruginosa* the fimbrial subunit of *D. nodosus* was assembled properly and vaccination trials showed that fimbriae produced by recombinant *P. aeruginosa* cells were as efficacious as whole cell *D. nodosus* vaccines (Mattick, 1989). The nucleotide coding sequences of fimbrial subunit genes of all serogroups have now been determined

(Billington & Rood, 1991; Elleman & Hoyne, 1984; Elleman *et al.*, 1986a; Mattick *et al.*, 1991; Mattick *et al.*, 1987b; McKern *et al.*, 1989) and has led to the division of *D. nodosus* serogroups into two classes, Class I and Class II, based on the identity, structure and organisation of the fimbrial genes (Figure 1.4).

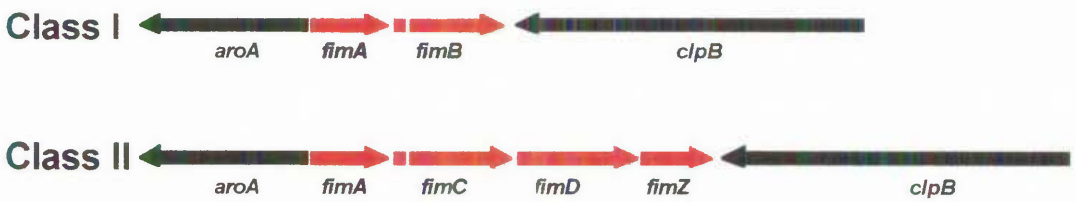



**Figure 1.4**: The genetic organisation of fimbrial subunit genes of Class I and Class II isolates of *D. nodosus*. Open reading frames and their direction is indicated by arrows: class specific genes are shown in red, whilst genes flanking the 5'- and 3'- ends of the class specific genes are shown in black.

Class I includes serogroups A, B, C, D, E, F and I, whilst serogroups G and H belong to Class II. Class I isolates contain an additional gene, *fimB*, downstream of the fimbrial subunit gene, *fimA*. In contrast, Class II isolates carry three additional genes, *fimC*, *fimD* and *fimZ*, downstream of *fimA* (Billington, Johnston & Rood, 1996a). *D. nodosus* fimbriae are composed of a single helical polypeptide subunit, and range from 6.5 kDa to 19 kDa for the different serogroups (Anderson *et al.*, 1986; McKern *et al.*, 1989), encoded by the gene *fimA*. The fimbriae of serogroup H are the only exception, the fimbriae being comprised of two smaller polypeptides of 6 and 10 kDa respectively (Moses & Yong, 1989). Upon isolation of the fimbriae, a trace amount (about 2%) of another polypeptide of about 80 kDa is also found which cannot be separated from the fimbrial strands. After electron microscope examination, it was suggested that the 80 kDa protein constitutes a basal protein thought to play a role in anchoring the pili to the outer membrane (Moses & Yong, 1989). *fimA* encodes a single pilin subunit, whilst *fimB* encodes a 29.5 kDa membrane protein thought to play a role in interactions with FimA during pilin export. The *fimD* gene is thought to produce a 45 kDa protein functionally analogous to the

FimB protein of Class I strains. The *fimC* gene product is thought to have a role in the acetylation of FimA subunits in class II strains of *D. nodosus* (Strom & Lory, 1993), whilst *fimZ* gene is thought to encode a redundant fimbrial subunit (Billington *et al.*, 1996a).

Within each class there is approximately 60% nucleotide identity between the fimbrial genes, whereas between the two classes there is only 40% nucleotide identity (McKern *et al.*, 1989). Irrespective of class, the fimbrial subunit genes (*fimA*) encode proteins with the same positively-charged short leader sequence and a highly conserved amino terminus. However, Class I mature subunits are of 152-156 amino acids whilst Class II mature subunits consisted of 149 amino acids; both classes contain a central variable region with small deletions present between the conserved domains (Mattick *et al.*, 1991). The class specific gene regions are flanked by *arcA* (Alm, Dalrymple & Mattick, 1994), located upstream of *fimA* and *clpB* downstream of either *fimB* (Class I) or *fimZ* (Class II) (Billington *et al.*, 1996a) (Figure 1.4).

In addition to the fimbrial subunit genes described above, a number of fimbrial biogenesis genes have been identified from *D. nodosus* strain A198 (Johnston *et al.*, 1995). Four of these genes, *orf197*, *fimN*, *fimO*, and *fimP*, encode proteins that have significant similarity to *Pseudomonas aeruginosa* fimbrial assembly proteins OrfX, PilB, PilC and PilD respectively. The genetic organisation of this region resembles that of the *P. aeruginosa pilBCD* and *orfX*, except that the *D. nodosus fimA* and adjacent genes are located in a different chromosomal region than the fimbrial biogenesis genes (Johnston *et al.*, 1995; La Fontaine & Rood, 1997).

### (ii) Proteases

Protease activity in *D. nodosus* was first identified by Beveridge (1941) and prompted further investigations into the molecular analysis of protease function. *D. nodosus* proteases are chymotrypsin-like serine proteases which require divalent calcium ions for activity (Depiazzi & Rood, 1984; Kortt *et al.*, 1982), and in virulent strains readily

degrade elastin (Green, 1985; Kortt *et al.*, 1982; Stewart, 1979; Thomas, 1964), keratin, fibrinogen, hemoglobin and collagen (Green, 1985). Although the precise role of these *D. nodosus* proteases in the disease process has not been determined, it is likely that they play an important role in the pathogenesis of footrot and may be involved in obtaining nutrients from tissues (Stewart, 1989a). This hypothesis is supported by the observation that physicochemical differences exist between proteases secreted by benign and virulent, intermediate strains of *D. nodosus* (Emery, 1988; Lilley *et al.*, 1992; Moses & Yong, 1989).

Biochemical studies have shown that virulent strains of *D. nodosus* secrete four extracellular acidic serine protease isoenzymes (V1-V3 and V5 with p*I* values of 5.2-5.6) and a basic protease (BprV, p*I* of 9.5). Benign strains of *D. nodosus* produce five acidic serine proteases (B1-B5) and a basic protease (BprB, p*I* of 8.6) that has 97% amino acid identity to BprV but a lower p*I* (Kortt *et al.*, 1994; Lilley *et al.*, 1995). The proteases secreted by benign strains exhibit different electrophoretic mobilities, are more thermolabile, and have decreased elastase activity (Billington *et al.*, 1996a; Gordon *et al.*, 1985; Kortt *et al.*, 1993; Moses *et al.*, 1989). These differential physicochemical characteristics form the basis of many reliable diagnostic tests to differentiate between benign, intermediate and virulent strains of *D. nodosus* (Section 1.3). The molecular basis for these differences is as yet unknown.

The first *D. nodosus* protease gene to be cloned and sequenced was *bprV* from the reference virulent strain A198, which encodes a 603 bp residue protease precursor (Lilley *et al.*, 1992). BprV contains an amino-terminal pre- and pro-peptide domain and a carboxy-terminal extension, all three of which are cleaved to form a mature protease of 344 aa. The benign homologue, *bprB*, from strain C305 has also been cloned and sequenced (Lilley *et al.*, 1995). In addition, the V5 and B5 acidic protease (p*I* of 5.2) genes, *aprV* and benign homologue *aprB* (99% amino acid identity to *aprV*) have been cloned and sequenced from *D. nodosus* strains A198 and C305 respectively (Kortt *et al.*, 1993; Riffkin *et al.*, 1993).

V5 and B5 proteases have a similar structure to BprV, having comparable prepropeptides and carboxyterminal extensions, 25-30% amino acid similarity to the subtilisin family of proteases with particularly strong conservation of the amino acid sequences around the catalytic site. The V5 and B5 acidic proteases have 64% amino acid identity to the basic serine protease (BprV) of *D. nodosus* (Kortt *et al.*, 1993; Lilley *et al.*, 1992), and like bprV show identity to the serine proteases of the plant pathogen *Xanthomonas campestris*; there is 50% amino acid identity for the basic protease (Lilley *et al.*, 1992) and 53% amino acid identity for V5 and B5 respectively (Kortt *et al.*, 1993). Biochemical analyses of the other proteases suggest that V1-V3 and B1-B4 proteases are closely related to each other but distinct from the other two groups of proteases which include V5/B5, and BprV/BprB (Billington *et al.*, 1996a; Riffkin *et al.*, 1995).

It should be noted that there is considerable controversy in the literature regarding the molecular sizes of the extracellular proteases (Gordon *et al.*, 1985), and these variations can probably be attributed to the conditions under which cultures were grown, and biochemical and physical conditions employed to determine the molecular sizes.

### (iii) Lipopolysaccharides

*D. nodosus* isolates examined by Stewart (1977) were found to contain a hydrophilic lipopolysaccharide (LPS) which exhibited physicochemical and biological characteristics common to the LPS found in the cell walls of other Gram-negative bacteria. These include the presence of lipid A, an R-specific polysaccharide core, and O-specific units, which together with the underlying protein constitute an endotoxin complex. Although Stewart (1979) showed that *D. nodosus* LPS-endotoxin induced pyrogenicity, leucopaenia, leucocytosis and inflammation, typical of other LPS endotoxins, the toxicity of the *D. nodosus* LPS endotoxins was much lower relative to the LPS endotoxins characteristic of Enterobacteria. Studies have attributed this lower toxicity to the absence of 3-keto-2-deoxyoctanoic acid and heptose present in the LPS endotoxins of other Gram-

negative bacteria (Davis *et al.*, 1990). It was concluded that, since *D. nodosus* LPS was relatively non-toxic, the LPS endotoxin alone could not account for the difference in virulence exhibited by virulent and benign strains.

In addition, the possibility of an LPS-based vaccine was investigated (Stewart, 1978). These experiments found that, although sheep did develop anti-LPS precipitating antibodies, the sheep were not protected against footrot.

### (iv) Outer Membrane Proteins

Examination of the outer membrane proteins (OMPs) from depiliated *D. nodosus* cells led to the identification of five proteins of molecular weights 75 to 80, 50, 38, 34.5 and 26.5 kDa (Emery *et al.*, 1984a) and designated 1-5 in order of decreasing molecular weight. All five OMPs are exposed on the external surface of the cell (Moses & Yong, 1989). OMPs 1, 3 and 4 were shown to be glycoproteins, present in greater quantities, and have molecular weights conserved across all serogroups. OMP 1 (75 to 80 kDa) is the putative basal protein which is coprecipitated with *D. nodosus* fimbriae, and is thought to anchor the fimbriae to the cell wall of *D. nodosus*. Four structurally-linked genes, designated *omp1A*, *omp1B*, *omp1C*, and *omp1D*, that collectively encode the Omp1 protein, have been cloned and sequenced (Moses *et al.*, 1995). The PCR-mediated cloning of inverted *omp* genes, and the identification of *nodosus* inversion cross-over sequences (*nix*) upstream from these *omp* genes, led to a multiple site-specific inversion model for the expression of *omp1* genes (Moses *et al.*, 1995). The conservation of these *omp1* genes in *D. nodosus* and the potential for inversion-mediated antigenic variation, suggests that the Omp1 protein may be important in survival, and perhaps the disease-causing potential of *D. nodosus*, as antigenic variation is in other pathogens (Borst & Greaves, 1987; Seifert & So, 1988; Borst, 1991; Brunham *et al*, 1993). However, since a *D. nodosus* infection is remote from the blood supply, and consequently the immune system, it was proposed (Mattick *et al.*, 1991) that such antigenic variation may allow *D. nodosus* to avoid its own pathogens. Nevertheless, a

possible role in immune avoidance has not been disregarded (Moses *et al.*, 1995), since this protein is a dominant antigen in natural infection.

The OMP 2 protein varied in molecular weight depending upon the serogroup and is thought to be involved in strain-specific functions (Egerton *et al*, 1984). The OMPs tend to exhibit similar antigenicity across all strains and they are known to be targets of the ovine immune system (O'Donnell, Stewart & Clark, 1983). However the serological reactions observed *in vitro* are not observed *in vivo*, where the OMPs are weak immunogens (Moses & Yong, 1989). It is thought that OMPs are not related to virulence (Depiazzi *et al.*, 1991) although the function of these OMPs has yet to be elucidated.

In addition to the well characterised antigens and virulence factors discussed above, it is of interest to note that immunoelectrophoretic studies by Yong *et al* (1986) have provided evidence that extracellular proteins exist which are unique to virulent or benign strains of *D. nodosus*, and do not correspond to the *D. nodosus* proteases, pili, or OMPs. These results suggest that antigens other than proteases, pili, and OMPs are important in determining whether a *D. nodosus* isolate is benign or virulent; the identity of these other virulence factors has not yet been determined and is worthy of further investigation and characterisation.

## 1.3 Diagnosis

The clinical diagnosis of footrot is of limited value since, in the early stages of the disease (first 7-14 days), it is very difficult to differentiate between benign (Stewart, 1989a), intermediate and virulent infections of *D. nodosus*, as all three present initially as interdigital dermititis. Although repeated physical examination and observation of the flock may be carried out, it is an inefficient and tedious method of diagnosis. Hence, laboratory identification of virulence is an increasingly important component in helping field

veterinarians and farmers to evaluate the potential of isolates to cause disease (Plant & Walker, 1994; Skerman, 1989). The differentiation of benign strains of *D. nodosus* from intermediate and virulent strains is important because of the considerable economic costs associated with virulent footrot. By contrast, benign footrot is associated with minimal economic costs (Stewart, 1989a). It is notable, that a number of strains (benign, intermediate, and virulent) and serogroups of *D. nodosus* may be present in a single flock (Claxton *et al.*, 1983), and this may complicate both clinical and laboratory diagnoses.

Laboratory diagnosis comprises two steps, firstly the isolation and identification of *D. nodosus* from the lesion material (Walker, 1988) followed by the determination of the virulence of the isolated strain. Tests have been developed to more efficiently differentiate between benign, intermediate and virulent strains of *D. nodosus*, and most of them exploit phenotypic characteristics of the different strains of *D. nodosus*, including colony morphology and piliation (Every & Skerman, 1983; Skerman, Erasmuson & Emery, 1981; Thorley, 1976), cell motility (Depiazzi et al, 1985) or protease characteristics (Depiazzi & Richards, 1979; Depiazzi *et al.*, 1991; Egerton & Parsonson, 1969; Every, 1982; Kortt *et al.*, 1983; Liu & Yong, 1993a; Stewart, 1989a; Stewart & Egerton, 1979).

Intermediate and benign strains of *D. nodosus* have lower motility and smaller colonies than virulent strains (Stewart, 1989a). However, the range of colony sizes between benign and virulent strains of *D. nodosus* were found to overlap, and were difficult to discern with accuracy. Therefore diagnosis based on colony morphology is considered unreliable and hence is not in use (Depiazzi, Henderson & Penhale, 1990; Depiazzi & Richards, 1979; Stewart, 1989a; Stewart & Egerton, 1979).

The degree of fimbriation does not correlate with the virulence of the strain of *D. nodosus* (Depiazzi *et al*, 1985; Moses *et al*, 1989; Rood *et al*, 1989) since some virulent strains have been observed to be poorly fimbriated whilst some benign strains are highly

fimbriated and *vice versa* (Stewart *et al.*, 1986). However, it has been observed that virulent strains are highly motile and form large spreading colonies which contrasts with benign strains which characteristically exhibit low motility and form small colonies (Depiazzi & Richards, 1979). Again, these differences are not reliable enough to be used diagnostically.

On the basis of physiochemical characteristics of proteases, benign and virulent isolates of *D. nodosus* can be differentiated reliably. Tests employed to assay these properties include:

(i) the elastase test which differentiates between benign strains which exhibit reduced or delayed elastase activity compared to intermediate and virulent strains (Liu & Yong, 1993b; Stewart, 1979). This test involves the monitoring of elastin-containing plates for the appearance of a clear zone around the bacterial colonies, indicating proteolytic digestion of insoluble elastin. However, a major drawback of the elastase test is that it takes up to 35 days from the time of sampling to produce results (Liu & Yong, 1993b), and may be influenced by bacterial metabolism and protease production. Although the results from elastase tests correlate well with the clinical classification of strains (Liu & Yong, 1993b), the use of elastase tests has declined because the test involves a very protracted process (Plant & Walker, 1994);

(ii) a degrading proteinase test (Depiazzi *et al.*, 1985) is available which distinguishes thermolabile proteases of benign strains from the thermostable proteases of virulent and intermediate strains. These tests use hide azure powder as a substrate to assay protease activity in a *D. nodosus* culture supernatant before and after exposure to high temperatures (Depiazzi *et al.*, 1990; Depiazzi & Rood, 1984; Green, 1985; Kortt *et al.*, 1982). Although the degrading protease test yields results at the time of sampling, it has disadvantages which include that intermediate and virulent strains cannot be distinguished, the test is laborious, and takes approximately 14 days before results are available (Depiazzi *et al.*, 1991);

(iii) an electrophoretic zymogram test is employed to provide qualitative information about the protease-isoenzyme composition of a particular strain. Virulent and intermediate strains exhibit similar isoenzyme patterns which can be differentiated from the benign zymogram; however, the limitations of this method are that it is technically demanding, the process is slow (taking about 14 days for results to be available), and it is difficult to test large numbers of isolates. Therefore, this test is not widely used (Liu & Yong, 1993b; Sabelnikov, 1994; Stewart, 1989a). In addition, although the benign zymogram pattern indicates that the strain lacks virulence, a virulent pattern does not guarantee virulence (Kortt *et al.*, 1983) and is unable to differentiate between intermediate and virulent strains of *D. nodosus*. Also, the zymogram test compares test sample isoenzyme patterns to only three isoenzyme patterns - V1, V2 and B as determined by Kortt *et al* (1983) and, unfortunately, these three patterns are not representative of all the patterns which exist in the field; in fact, according to Liu *et al* (1993), there are at least eight (for virulent strains S1 and S2, and for benign strains U1 to U6);

(iv) the gelatin plate test measures clearing of gelatin from plates with varying temperature. Virulent strains of *D. nodosus* have been observed to result in complete clearing, whilst intermediate strains exhibit incomplete clearing and benign strains no clearing (Liu & Yong, 1993b);

(v) the gelatin gel test is similar to the degrading proteinase test described (ii), except that a cheaper gelatin substrate is substituted for hide azure powder to assay protease activity; the heat treated culture supernatant is aliquoted into wells of a gel containing the gelatin, and clear zones of proteolysis around the inoculated wells are measured with callipers (Palmer, 1993). Benign strains show no clearing around the well, whereas virulent strains produce obvious clearing.

Although there is considerable variation in the laboratory tests employed by different

diagnostic laboratories, in Australia, the gelatin gel test is becoming more widely utilised (Plant & Walker, 1994; Rood *et al.*, 1996). This is primarily because the gelatin substrate is inexpensive, it requires less equipment and time than other tests, and it has simplified the testing of multiple samples (Palmer, 1993). Furthermore, there is a high level of agreement between the gelatin gel and elastase tests (Links & Morris, 1996). However, because of the slow growth rate of *D. nodosus*, the gelatin gel test may take at least 10 to 14 days to complete. Like the other diagnostic tests, the gelatin gel test is not absolutely reliable; one Armidale flock with 15% of the flock with a footrot lesion score of score 4 or 5, and hence recorded clinically as virulent, gave benign gelatin gel tests (Plant & Walker, 1994). The problem of false negatives has been documented by others (Links & Morris, 1996), but has been attributed to isolates growing poorly.

(vi) an ELISA (enzyme-linked immunosorbent assay) that detects antibodies to *D. nodosus* antigens present in serum of infected sheep, is currently in the developmental stages. The aim of such an assay would be to detect carriers and infected flocks during periods when the disease is subclinical (Whittington, 1996). Although the magnitude of anti-*D. nodosus* antibodies provides a relatively accurate indication of the severity of clinical footrot, interpretation of results is complicated by cross-reactions between the KSCN-extracted antigens of *D. nodosus* (comprising 52 components), with those of other bacteria (Whittington & Egerton, 1994). Furthermore, the use of vaccines in footrot control programs may decrease the value of serology as a diagnostic test in sheep, because vaccination results in marked and persistent antibody levels in sheep. By contrast, in natural infections, antibody levels decrease when lesions are treated (Whittington & Egerton, 1994).

(vii) A PCR test based on the variable regions of the 16S rRNA of *D. nodosus* was developed for the identification of *D. nodosus* isolates directly from lesion material. Although this PCR test is not able to distinguish benign and virulent strains of

*D. nodosus* it was proposed that such a test would be useful for the detection of *D. nodosus* in carrier sheep without culturing the organism (La Fontaine, Egerton & Rood, 1993).

In general, these methods are very expensive, have limitations in precision and time, the elastase and colony morphology tests are somewhat difficult to control and quantitate, and they are further complicated by the stringent requirements and slow growth characteristic of *D. nodosus*. As a result, there is a well documented need for rapid and reliable methods for the differential diagnosis of ovine footrot. Consequently, tests based on gene probes and PCR (Katz *et al.*, 1991; Liu & Yong, 1993a; Raadsma *et al.*, 1995; Rood & Yong, 1989) are being developed with the aim of providing cheaper, more rapid and precise diagnostic tools (Sabelnikov, 1994; Stewart, 1989a) than the conventional tests based on phenotypic characteristics. The progress that has been made in this direction will be discussed in section 1.5.

It should be noted that the lack of reliability in any one diagnostic test is possibly attributable to the nature of bacterial virulence. As observed in other pathogens, it is likely that virulence in *D. nodosus* is multifactorial, with various virulence factors working together collectively, to result in disease. This explains the observed continuum from benign through to virulent in *D. nodosus* strains, and consequently the inability of a single diagnostic test to give an absolute correlation with strain virulence.

## 1.4 Treatment, Control and Eradication

Since footrot is a relatively superficial infection between the avascular epidermis of the hoof and keratinized horn, there are many options available for its treatment. These options for control and eradication may be used alone, however they are most successful when used collectively. These options are three-fold and are described below.

(i) changes in management practices

Control and eradication strategies have included slaughter ("depopulation") of infected animals from the flock, leaving unaffected ones, a strategy which ultimately aims to select genetically resistant flocks; selling of the whole flock for slaughter; or treatment and segregation. Also, regional eradication programs in Western Australia (since 1949), in South Australia (since 1957), in the New England Protected Area (since 1960) and in the Footrot Control Area of Western Victoria (since 1968) have been implemented. The implementation of these programs has led to a significant reduction in the prevalence of footrot in these areas (Stewart, 1989a; Stewart, 1989b).

(ii) chemotherapy

Chemotherapy may include footbathing and topical or parenteral treatment with antibiotics. Of these, footbathing is the preferred alternative to the direct application of antibiotics or parenteral treatments, being more practical, less labour intensive and more economical than either topical or parenteral treatments. Footbaths contain antiseptic agents and may consist of 5-10% formalin, 10% zinc sulfate/water, or 10% copper sulfate. Of these, zinc sulfate is most favoured since it is relatively cheap, does not stain wool like copper sulfate, is non-toxic unlike formalin, and has a long shelf life. However, foot paring is not required prior to footbathing with copper sulfate unlike zinc sulfate (Stewart, 1989a). The cure rate after footbathing with zinc sulfate is about 80% (Emery, 1988) and is improved when combined with antibiotic treatment and/or vaccination. In contrast, topical treatments first require paring of the horn of the hoof so that bactericidal agents can be applied to the lesion directly. These agents may include 10% chloramphenicol-alcohol, oxytetracycline hydrochloride or 20% cetrimide. Parenteral treatment consists of an intramuscular injection, usually of streptomycin-penicillin. Such use of antibiotics introduces the risk of developing resistant microbial populations, and is extremely expensive, and therefore is used only when valuable stud animals are at risk.

(iii) vaccination

First-generation commercial vaccines, developed in the 1960s, and made commercially available in 1972, comprised whole cells (O'Meara, Egerton & Raadsmaa, 1993). However, first-generation commercial vaccines were of extremely low efficacy due to the poor condition of cells (Stewart, 1989b), the omission of key serogroups from the formulae and the low concentration of fimbriae present in the formulae. Due to the low efficacy, first-generation vaccines were taken off the market in 1979. Since the 1960's, efforts have concentrated on genetic and structural studies of *D. nodosus* antigens with the aim of developing a more inexpensive and efficacious vaccine against footrot (McKern *et al.*, 1989). A better understanding of serological relationships in *D. nodosus* led to the development of second-generation vaccines which have been available since 1981 (McKern *et al.*, 1989).

*D. nodosus* fimbriae, having two to four epitopes (Emery, 1988), are excellent immunogens (McKern *et al.*, 1989) being equally as immunogenic as whole cell preparations (Stewart, 1978). They form the basis of the second- and third-generation vaccines.

Second-generation, multistrain vaccines have much higher efficacy since all serogroup fimbrial antigens are represented in the formulae, and multiple serogroup infections are common (O'Meara *et al.*, 1993). Some of these vaccines include Footvax® (1981, Coopers), Pro-vac® (1982), Vaxall Norot® (1994, Websters), and Vaxicare® (1994, Smith Kine Beecham). Current licensing standards require that these commercial vaccines have cure rates of not less than 65% six weeks after administration and protection rates of 70% twelve weeks after administration (O'Meara *et al.*, 1993; Stewart, 1989b).

According to Stewart (1989), it should be possible to control footrot *via* vaccination without employing foot-paring or footbathing. However, footrot vaccines are

expensive because of the fastidious growth requirements of *D. nodosus*, the multiserogroups required, and the difficulty involved in obtaining stable fimbriae under the liquid fermentation conditions employed for commercial production (Mattick, 1989; Moore *et al.*, 1990). As a result, of the expense, acceptance within rural communities has been relatively low, and traditional, laborious and less efficacious methods such as foot-paring and footbathing have continued to be the predominant method of control and eradication.

The economics of vaccination has prompted scientists to develop third-generation vaccines using recombinant DNA technology or synthetic peptides. It is considered that cloning *D. nodosus* fimbrial genes into a host which is less fastidious and more suitable for commercial production will reduce costs. Initial trials which employed *E. coli* as the host organism were unsuccessful since, although the fimbrial genes were expressed in *E. coli*, they were not assembled into mature fimbriae (Mattick *et al.*, 1984). Subsequent work using *Pseudomonas aeruginosa* as a host proved more successful (Mattick *et al.*, 1987a; Mattick *et al.*, 1987b) as large amounts of the *D. nodosus* fimbrial subunit was expressed and assembled on the surface of the cell; vaccination trials showed that immunogenicity was at the same level as for *D. nodosus* fimbrial-based vaccines (Emery, 1988). Field trials are being carried out, but there is no indication in the literature as to when these recombinant vaccines will be made available for use.

Other work is centred on decreasing the number of serotypes required to produce comprehensive protection by identifying fimbrial epitopes using monoclonal antibodies and, from this information, synthesising peptides which contain all the epitopes of the fimbriae of different serotypes (Emery, 1988). Furthermore, such peptides may offer distinct advantages over fimbriae-based vaccines due to their purity and biological safety (McKern *et al.*, 1989).

Some work has assessed the potential of using proteases as antigens suitable for a vaccine. However, it was observed that the active infections persisted despite high antiprotease titres (Moses & Yong, 1989).

Future improvements in vaccines would involve determination of the immunological mechanisms involved in *D. nodosus* infections, optimisation of intensity and duration of response, and identification of other antigens which used in conjunction with the fimbrial antigens may be able to stimulate a broader response (Mattick, 1989).

### 1.4.1 The NSW Footrot Strategic Plan

The NSW Footrot Strategic Plan commenced in 1988 with the aim of eradicating virulent footrot, and consequently making NSW a Footrot Protected Area by December 2000. For the purposes of the Strategic Plan, NSW is divided into 'Protected', 'Control' and 'Residual' areas. Those areas classified as Protected have a prevalence of virulent footrot of less than 1% of flocks, whilst in Control and Residual areas the prevalence of virulent footrot is between 1-10% and >10% respectively (Plant & Walker, 1994).

In December 1997, an audit in one of the oldest protected areas, the New England Area, was undertaken to independently verify the prevalence of virulent footrot in the area (Plant & Walker, 1994). The audit revealed virulent footrot (including both virulent and intermediate strains) in 11.9% (12 of 101 flocks) of flocks in Armidale, and in 22.2% (8 of 36 flocks) of flocks in Glen Innes (Plant & Walker, 1994), which compares poorly with 4.9% prevalence (8339 flocks), in NSW as a whole. Results of the New England survey indicate that virulent footrot has been able to survive unnoticed, at significant levels, in a Protected area. Subsequently, the New England Area has been downgraded to the Residual area classification.

It has been suggested that the lack of detection of virulent footrot in the New England

Area has most likely been due to the period of extended drought. Dry conditions are unfavourable for the spread, and hence detection of virulent footrot (Plant & Walker, 1994). The results of this survey therefore suggest that, similarly, there may be other Protected and Control areas under the NSW Strategic Plan, that have 'Residual' levels of virulent footrot, which have been masked by drought conditions.

## 1.5 Isolation of Virulence-Associated DNA regions in the genome of *D. nodosus*

Work to develop gene probes suitable for the quick differential diagnosis of footrot began in 1991 (Katz *et al.*, 1991) with the isolation of virulence-associated DNA regions from the *D. nodosus* genome. DNA sequences present in the virulent reference strain A198, but absent from the benign strain C305, were isolated. This was done by differential hybridisation of labelled genomic DNA from virulent strain A198, and benign strain C305, to a gene library of A198 genomic DNA in the plasmid pUC18. Three recombinant plasmids, pJIR318, pJIR313 and pJIR314B which hybridised to genomic DNA from the virulent but not the benign strain, were isolated. Restriction maps for pJIR318, pJIR313 and pJIR314B differed from previously cloned regions encoding fimbrial subunits (Anderson *et al.*, 1984; Elleman & Hoyne, 1984) and proteases (Vaughan *et al.*, 1990), and hence the cloned sequences may encode genes associated with protease secretion, twitching motility, fimbrial assembly or some other function in virulence (Katz *et al.*, 1991). Alternatively, because these regions have a genotypic association with virulence rather than a phenotypic one, these loci may not encode true virulence factors but instead may encode regions that are indicative of virulence.

pJIR318 contains sequences that comprise the virulence-associated protein (*vap*) regions, whilst pJIR313 and pJIR314B contain sequences that comprise the virulence-related

locus (*vrl*) region. These plasmids were used to probe the genome of 101 *D. nodosus* isolates with known virulence characteristics (Katz *et al.*, 1991). The results, in Table 1.1, indicate that sequences similar to pJIR318, pJIR313, and pJIR314B were present in almost all virulent strains but absent from most of the benign strains. Since these sequences were not randomly distributed throughout virulent, intermediate and benign strains, it was suggested that these sequences may be necessary, but not sufficient for virulence (Katz *et al.*, 1991). Results from this work suggested that pJIR318, pJIR313, and pJIR314B could be utilised in the development of a rapid gene-probe-based method for the differential diagnosis of ovine footrot (Katz *et al.*, 1991).

**Table 1.1**: Hybridization of pJIR318, pJIR313, pJIR314B with virulent, intermediate and benign strains of *D. nodosus*. (Katz *et al.*, 1991).

| Virulence of Strains | Dot Blot Hybridisation with | | |
|---|---|---|---|
| | **pJIR318** | **pJIR313** | **pJIR314B** |
| **Virulent** | 100% (29/29) | 94% (27/29) | 100% (29/29) |
| **Intermediate** | 94% (34/36) | 36% (13/36) | 36% (13/36) |
| **Benign** | 33% (12/36) | 6% (2/36) | 6% (2/36) |

Other workers have repeated the experiments originally done by Katz *et al.* (1991), however, each of the DNA probes identified (Liu & Yong, 1993a) contain DNA fragments from the *vap* or *vrl* regions (Rood *et al.*, 1996), consistent with and supporting the original work (Katz *et al.*, 1991), which suggested that these loci represent major differences between virulent and benign strains of *D. nodosus*. In addition, many *D. nodosus* strains carry an uncharacterised locus, designated the benign-specific region, that is not present in *D. nodosus* strains that carry the *vrl* , and may also be useful as a gene probe for avirulence

(Liu & Yong, 1993a).

In a collaborative study (Rood *et al.*, 1996), 771 *D. nodosus* isolates from four Australian Veterinary Diagnostic Laboratories, were analysed using conventional diagnositic tests (elastase test, gelatin gel test, and protease electrophoretic mobility, as discussed in section 1.3) and using gene probe tests based on the virulence-associated subclones pJIR318, pJIR314B and pJIR313B. The aim of this study was to evaluate the efficacy of the gene probes, for the differentiation of benign from intermediate and virulent strains of *D. nodosus*. This study showed that there was good correlation between conventional test results and the results obtained from the gene probe tests (Rood *et al.*, 1996). As a result, a multiplex PCR test based on the gene probe plasmids was designed and investigated, and although this test needs to be evaluated further in order to determine the feasibility of directly examining lesion material, such a test would potentially enable laboratory results to be available within 48-72 hours rather than 10-14 days (Rood *et al.*, 1996).

Since virulence in *D. nodosus* appears to be multifactorial, laboratory-based diagnosis should involve more that one potential parameter, and hence it has been recommended that a multiplex PCR test based on the virulence-associated subclones be used in conjunction with conventional tests based on the analysis and detection of extracellular proteases of benign and virulent isolates of *D. nodosus* (Rood *et al.*, 1996).

### 1.5.1 pJIR313 & pJIR314B and the Virulence Related Locus (*vrl*)

As described in Table 1.1, virulence-associated plasmids pJIR313 and pJIR314B hybridised with nearly all intermediate and virulent strains but not with benign strains, and they were found to be present in only one copy in the *D. nodosus* genome (Rood *et al.*, 1993). Southern hybridisation analysis revealed that pJIR313 and pJIR314B were divided by 9 kb of virulence-specific DNA. The 9 kb fragment was cloned and used to construct

three plasmids, pJIR590, pJIR589, pJIR592, which were used to probe nine virulent and nine benign strains of *D. nodosus*; only the virulent strains hybridised to these plasmids. The region including the 9 kb virulence-specific DNA, pJIR313, and pJIR314B has since been designated the virulence-related locus (*vrl* region, Figure 1.5) (Haring *et al.*, 1995; Rood *et al.*, 1993).

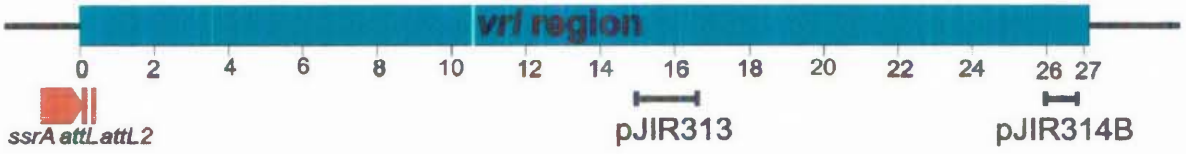



**Figure 1.5**: The *vrl* region of *D. nodosus* strain A198. The numbers indicate the distance in kb. The relative positions of virulence-associated subclones pJIR313 and pJIR314B, and of the 10Sa RNA gene (ssrA), *attL* and *attL2* sites are indicated {modified from (Haring *et al.*, 1995)}.

In order to test the hypothesis that the *vrl* is an integrated genetic element Haring *et al.*, (1995) identified and then analysed the left- and right-hand ends of the *vrl* region by comparing these junction regions from virulent strain A198 to the corresponding sequences in benign strain C305. It was predicted that the left and right limits of the *vrl* would be juxtaposed in benign strains, instead they found the ends were separated by approximately 3 kb (Haring *et al.*, 1995), which suggests that the *vrl* did not arise by a simple insertion event in virulent strains, but by an exchange mechanism in which the intervening regions present in benign strains were replaced (Billington *et al.*, 1996a).

Comparison of the sequence flanking the left-hand end of the *vrl* with sequences in the GenBank and EMBL databases showed significant sequence similarity to an *E. coli ssrA* gene, which encodes a regulatory 10Sa RNA molecule (Kirby, Trempy & Gottesman, 1994). In the 3'-end of this putative 10Sa RNA gene, a putative *attL* was identified by its similarity to the *attL* of several prophages (Kirby *et al.*, 1994). On this basis the *vrl* region has been designated as a putative prophage. 300 bp downstream from *attL* a second

putative attachment site designated *attL2* was identified. However, no concomitant *attR* site was found. This suggested that the right-end of the *vrl* region does not represent the right end of the integrated element. This was attributed to either the deletion of *attR* post-integration, or the possibility that *attR* may lie outside the sequenced region (Haring *et al.*, 1995).

The function of small stable RNAs are not well understood, although these RNAs have been implicated in gene regulation and interaction with DNA-binding proteins (Kirby *et al.*, 1994). Since a putative *attL* site is located within the 3'-end of the putative 10Sa RNA gene at the left-end of the *vrl* region, it is possible that although the *vrl* region itself may not encode virulence functions, excision or integration of the putative bacteriophage may affect virulence functions indirectly by altering small stable RNA expression (Haring *et al.*, 1995).

To date, the sequence of over 18 kb of the *vrl* region has been determined, from which 10 open reading frames have been identified. Two of these genes, *orf1243* and *orf666*, encode proteins with similarity to the *S. coelicolor* phage growth limitation system (Pgl), functions in phase-variable bacteriophage ϕC31 resistance (Bedford, Laity & Buttner, 1995). *orf1243* encodes a protein which has 15.5% aa identity to PglY, whilst *orf666* has 22% aa identity to the PglZ protein (Billington *et al.*, in press). Another open reading frame, *orf282* has 32%-34% identity to the C-terminal domain of several plant chloroplast proteins with 85% similarity to the regulatory subunit ClpA of the ClpA ATP-dependent protease of *E. coli* (Gottesman *et al.*, 1990; Billington *et al.*, in press).

The putative product of *orf929* has aa sequence similarity to several N6-adenine-specific methyltransferases (Billington *et al.* in press) including M.*Cer*M from *Caulobacter crescentus* (Zweiger, Marczynski & Shapiro, 1994) M.*Dpn*II from *Streptococcus pneumoniae* (de la Campa *et al.*, 1987) and M.*Hin*fI from *Haemophilus influenzae* (Chandrasegaran *et al.*, 1988). Both *orf1130* and *orf965* gene products have similarity to

the DEAH superfamily of ATP-dependent helicases (Haring *et al*, 1995; Billington *et al*, in press). The *orf1130* product has 34% identity in 481 aa to antiviral protein SK12, from *Saccharomyces cerevisiae* (Widner & Wickner, 1993) which functions as a helicase that blocks translation of viral transcripts; and *orf965* has similarity to putative helicases Yqh from *Bacillus subtilis* (GenBank Accession No D84432), HepA from *H. influenzae* (Fleischmann *et al.*, 1995) and HepA from *E. coli* (Bork & Koonin, 1993). *orf28*, *orf495*, *orf32* and *orf251* have no similarity to sequences in the databases and no orfs have been identified that can be directly related to virulence (Billington *et al*, in press).

It is significant that hybridisation analysis of over 800 isolates of *D. nodosus* has failed to identify a single isolate that contains the *vrl* region but not the *vap* sequences (Katz *et al.*, 1991; Rood *et al.*, 1996). This may indicate a fundamental interaction between acquisition of *vap* and *vrl* sequences.

## 1.6 Molecular analysis of the *vap* regions of *D. nodosus*

The ongoing characterisation of the *vap* regions, and the elucidation of the role of these *vap* genes in virulence, is the primary aim of work carried out in the Cheetham Laboratory at UNE. The progress that has been made towards achieving the aforementioned aims, both prior to, and in conjunction with the work for this thesis, are summarised in the following sections.

### 1.6.1 Analysis of pJIR318

The genomic regions which hybridised with pJIR318 were called 'virulence-associated protein' regions or *vap* regions. Southern blotting, using pJIR318 to probe the genomic DNA from the virulent strain A198, showed that there were three copies of the *vap* region, on 4.6 kb, 6.2 kb and 3.5 kb *Hin*dIII fragments, which were designated *vap* regions

1, 2 and 3, respectively (Figure 1.6). pJIR318 was also used to probe genomic DNA of strains from the other serogroups, and some variation as to the number of copies of sequences related to pJIR318 was observed. Members of serogroups A, C, F and I have three copies, D and E have one copy on a 6.2 kb *Hin*dIII fragment, whilst B, G and H have one copy on a 6.6 kb *Hin*dIII fragment (Katz *et al.*, 1991).

Sequence analysis of pJIR318 (Katz, Strugnell & Rood, 1992) led to the identification of four open reading frames designated as *vap* genes *A*, *B*, *C* and *D*, which were preceded by ribosome binding sites (Shine & Dalgarno, 1974). In addition, an open reading frame *orf118* (later renamed *toxA*), was identified, which was not preceded by a ribosome binding site.

### 1.6.2 Characterisation of *vap* regions 1, 2 and 3 of *D. nodosus* strain A198

The three *vap* regions in A198 were isolated by the construction of a bacteriophage lambda (EMBL3) gene library of the A198 genome and probing it with pJIR318. Analysis of *vap* regions 1, 2 and 3 using restriction mapping and Southern blotting of *D. nodosus* strain A198 showed that regions 1 and 3 were adjacent in the A198 genome (Katz *et al.*, 1994), whilst pulsed-field gel electrophoresis revealed that region 2 was separated from regions 1 and 3 by at least 150 kb and confirmed that the *vap* genes had a chromosomal location (Rood *et al.*, 1993).

On probing isolated *vap* regions 1, 2 and 3 of the A198 genome, *vap* regions 1 and 2 were observed to contain *vapA*, *vapB*, *vapC*, *vapD* and *toxA* (formerly *orf118*) in the same order and orientation and therefore these regions are closely related, whilst *vap* region 3 only contained *vapD* (Figure 1.6). Further investigation (Katz *et al.*, 1994) into the sequence of events responsible for the evolution of the *vap* regions indicated that simple insertion and deletion events were not sufficient to generate the arrangement of the *vap* regions 1, 2 and 3

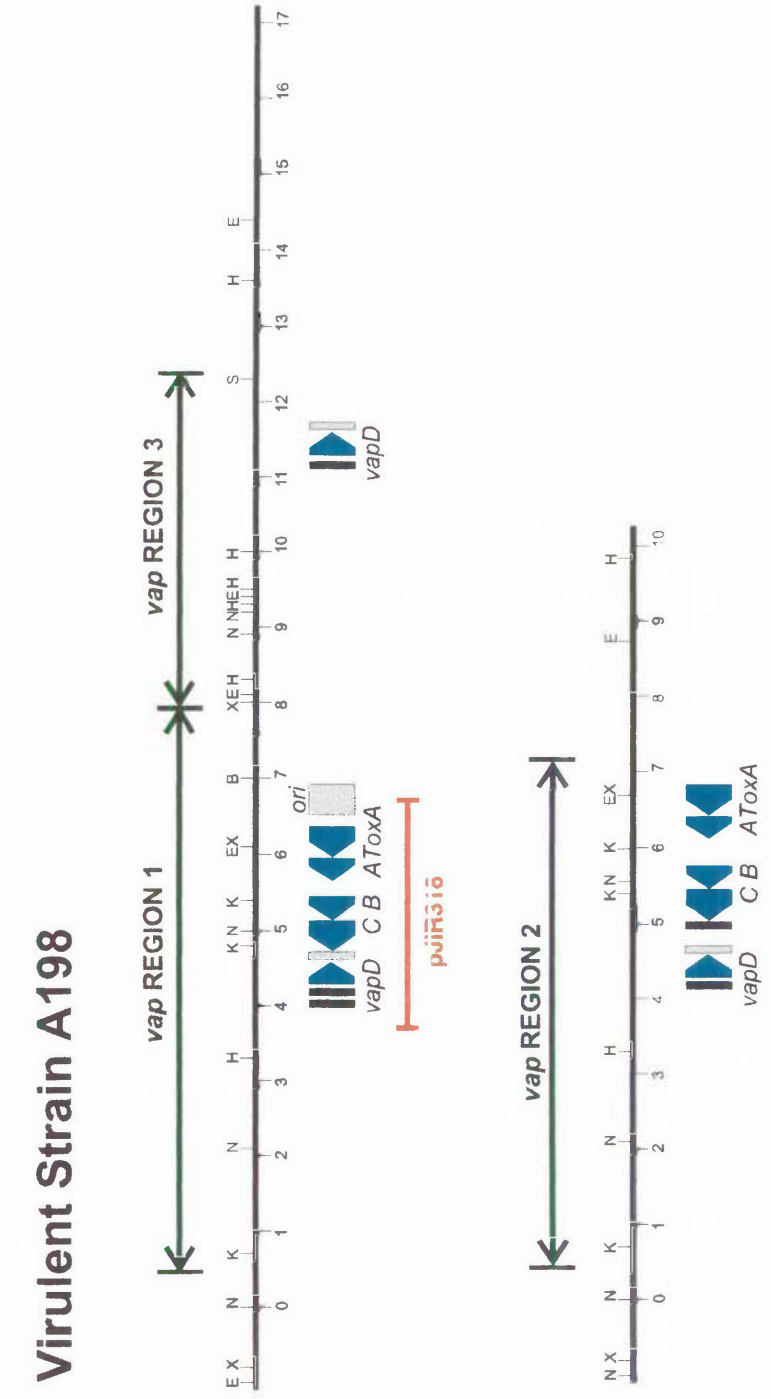



**Figure 1.6:** Hybridisation of virulence-associated probe, pJIR318 (**red**) to three *Hin*dIII fragments in *D. nodosus* strain A198, comprising part of *vap* region 1, 2 and 3 respectively as defined by Cheetham *et al.*, 1995. Sequences hybridising to pJIR318 are shown: open reading frames and their direction (**blue** arrows), 102 bp repeats or partial copies thereof (black boxes), 103 bp repeats, or partial copies of (narrow grey boxes), and a putative origin of replication (grey box). The numbers show the distance in kb from the left-most *Nru*I site. Restriction sites shown include: *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Nru*I (N), *Sac*I (S) and *Xho*I (X).

in the *D. nodosus* strain A198, and led to the construction of a preliminary model of how the *vap* regions might have evolved (Katz *et al.*, 1994).

Although the presence of all three *vap* regions is the most common arrangement in the *D. nodosus* genome, there are some virulent strains which appear to contain *vap* region 2 alone (Katz *et al.*, 1991) for which there were at least three possible explanations: (i) *vap* region 2 did not evolve from *vap* regions 1 and 3, and may therefore be the primordial *vap* region: or (ii) after the evolution of *vap* region 2, a large deletion of *vap* regions 1 and 3 occurred; or (iii) *vap* regions 1 and 2 have arisen *via* two separate independent integration events involving a conjugative plasmid or integrative bacteriophage.

### 1.6.3 Evidence for an extra-chromosomal origin of the *vap* regions

Molecular analysis of *D. nodosus* strain A198 *vap* regions 1 and 3 by Cheetham *et al* (1995) defined the virulence-associated and non-virulence associated DNA junctions, and resulted in the identification of a number of other potential genes (Figure 1.7). Upstream of *vapD* in *vap* region 1, a *tRNA-ser* gene, an integrase gene designated *intA*, and *vapE*,*vapG*, *vapH*, and *vapA"* were identified. *vapA"* (formerly *vapI*) encodes a putative protein with 36.6% aa identity to VapA. *vap* region 3 was found to contain the gene *vapF* and numerous other genes with a high degree of amino acid similarity to the putative proteins produced by the aforementioned genes from *vap* region 1, including: *vapD* with 100% aa identity to VapD from *vap* region 1; *vapA'* whose product has 57.3% aa identity to VapA; *vapE'*, which encodes a protein with 62.6% aa similarity to VapE; *toxA"* whose predicted protein product has 31.5% aa identity to ToxA; *vapH'*, which encodes a protein that has 48.5% identity in 101 aa to VapH; and *vapG'* whose putative protein product has a 73.5% identity in 68 aa to VapG. In addition a number of *cis*-elements were found to be associated with these *vap* regions, including putative attachment sites (*att*), a putative origin of replication (*ori*) and additional copies of 102 bp and 103 bp repeats. These may have a role

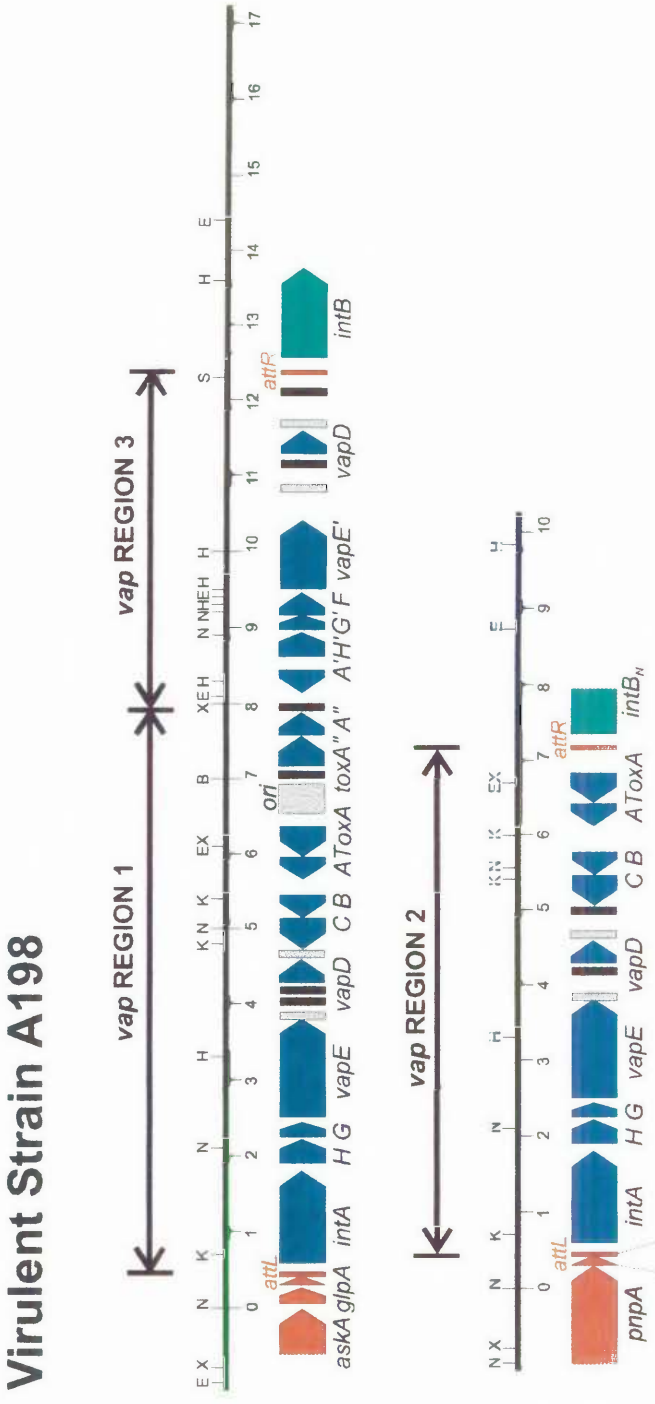



**Figure 1.7:** The *vap* regions 1, 2 and 3 of *D. nodosus* strain A198. The numbers show the distance in **kb** from the left-most *Nru*I site. Restriction sites indicated are as follows: *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Nru*I (N), *Sac*I (S) and *Xho*I (X). The location and orientation of the open reading frames are indicated by arrows, and coloured according to the following classification; belonging to the *vap* regions (**blue**), belonging to the *intB* element (**green**), genes that are not part of an integrated genetic element (**red**), and *tRNA* genes (**red triangles**). Repeated sequences include: attachment sites *attL* or *attR* (narrow red boxes), 102 bp repeats or partial copies thereof (filled boxes), 103 bp repeats or partial copies of (narrow grey boxes), and a putative origin of replication is labelled *ori* (grey box).

in the transposition of *vap* genes because of their locations, for example, all copies of *vapD* are flanked by a 103 bp repeat and by a 102 bp repeat. (Figure 1.7) (Cheetham *et al.*, 1995b). *vap* region 2 has been sequenced only in part (Cheetham *et al.*, 1995b) and appears to be almost identical to *vap* region 1, at the sequence level and in organisation. Furthermore, DNA fragments from *vap* region 2 are known to cross-hybridise with probes derived from *vap* region 1. In contrast, only DNA fragments containing *vapD* cross-hybridise with DNA fragments from *vap* region 3.

The orf designated *intA* has similarity to integrase genes from many bacteria, and is most similar to the integrase proteins (Int) of *E. coli* retronphage ϕR73 (41.5%) (Sun, Inouye & Inouye, 1991), bacteriophage P4 (35.4%) (Pierson & Kahn, 1984) and *Shigella flexneri* bacteriophage Sf6 (35.2%) (Clark, Beltrame & Manning, 1991).

In addition to temperate bacteriophage, numerous other genetic elements are known to carry integrases including plasmids, conjugative transposons, transposons, integrons and pathogenicity islands. Genetic elements which carry these integrase genes are able to integrate into the host chromosome *via* site-specific recombination between an identical, non-palindromic, short core sequence (15-20 bp) which is found both on the genetic element and on the host chromosome and determines both the polarity and the site-specificity of the recombination event. This core sequence is referred to as the attachment site, and in the genome of bacteriophage lambda this core sequence is called *attP* (phage) whilst on the *E. coli* chromosome it is called *attB* (bacterium). The site-specific recombination between *attP* and *attB* is catalysed by Int which recognises the attachment site sequence, introduces a staggered cut, and fuses *attP* and *attB* generating a co-integrate molecule. Recombination between *attP* and *attB* produces two copies of the *att* sequences often designated *attL* (left) and *attR* (right) at either end of the integrated prophage (Birge, 1994).

In *E. coli* retronphage ϕR73 (Sun, Inouye & Inouye, 1991), bacteriophage P4

(Pierson & Kahn, 1984) and *Shigella flexneri* bacteriophage Sf6 (Clark, Beltrame & Manning, 1991), the integrase genes are located adjacent to an attachment site (*attP*) (Leong *et al.*, 1986), and in integrate into the 3' ends of *tRNA*-encoding genes as do many of the other genetic elements which encode integrases closely related to *intA* (Table 1.2). Such an arrangement was found at the left of *vap* regions 1 and 2 in the *D. nodosus* genome (Cheetham *et al.*, 1995b).

Approximately 200 bp upstream from *intA*, a 19 bp sequence from the 3'-end of the predicted *tRNA-ser*$_{GCU}$ gene is repeated with 89% similarity at nt 12424. These positions correspond to the limits of the virulence-associated region at the left of *vap* region 1 and the right of *vap* region 3, and hence constitute putative attachment sites. Integration of the *vap* element re-creates the 3'-end of the tRNAser genes. Between position 6444 and 6927 a series of 21 bp tandem repeats were identified, which are preceded by a putative *DnaA* box (Fujita, Yoshikawa & Ogasawara, 1989), and flanked by A + T rich DNA regions. This region resembles the origin (*ori*) region of several plasmids and bacteriophages (Cheetham *et al.*, 1995b; Marians, 1992; Nordstrom, 1990).

Collectively, the similarity between *intA* and integrase genes that have been found on bacteriophages, plasmids (Boccard *et al.*, 1989) conjugative transposons (Salyers *et al.*, 1995) and pathogenicity islands (Hacker *et al.*, 1997); together with the location of *intA* adjacent to the 3'-end of a *tRNA* gene, and the identification of putative *att* and *ori* sites strongly suggests *D. nodosus* strain A198 acquired the *vap* regions *via* the integration of an extrachromosomal element into the 3'-end of a *tRNA* gene (Cheetham *et al.*, 1995b). Bacterial virulence factors are often located on mobile genetic elements (Finlay & Falkow, 1989; Finlay & Falkow, 1997). Indeed, recent work in which a physical and genetic map of the 1.54 Mb genome of *D. nodosus* strain A198 was constructed (La Fontaine & Rood, 1997) revealed that putative virulence genes are not clustered on the chromosome, suggesting that in *D. nodosus* virulence determinants have been acquired independently and

gradually by horizontal transfer.

In *D. nodosus* strain AC2577, Billington *et al.* (1996) have identified a 10 kb circular, double-stranded DNA molecule designated pJIR896 (Figure 1.8), which resembles a plasmid, or a bacteriophage in the replicative phase, and have proposed that this plasmid-like molecule may be the progenitor of the chromosomal *vap* region 1 sequences (Billington *et al.*, 1996b).

pJIR896 contains *(i)* genes that comprise *vap* region 1 of *D. nodosus* strain A198 including, *vapA*, *B*, *C*, *D*, *E*, *G*, *H*, *I* and *intA* (Figure 1.9); *(ii)* an insertion sequence, designated IS*1253* that is not found within the *vap* regions in strain A198, but instead is located near the outer membrane protein (*ompI*) gene cluster, responsible for the production of the major bacterial outer membrane proteins (Moses *et al.*, 1995); and *(iii)* a sequence of 600 bp from *vap* region 3 (Figure 1.8), commencing within a 103 bp repeat located immediately downstream from *vapD* (*vap* region 3) to the *attR* site.

If the *vap* plasmid were the progenitor of *vap* regions 1 and 3 of *D. nodosus* (Figure 1.8) then considerable rearrangements, duplications and deletions must have occurred to separate the 600 bp fragment from *vap* region 1. Alternatively, the IS*1253* element may have inserted into the *vap* plasmid resulting in the deletion of *vap* region 3. The IS*1253* element is not virulence-associated since it is absent from many virulent strains and, conversely, is present in many benign strains of *D. nodosus*.

The identification of a *vap* plasmid is consistent with the suggestion that an extrachromosomal element was the responsible for the introduction of the *vap* regions in *D. nodosus*, which subsequently evolved *via* additional integration events or by a series of duplications, partial duplications, multiple independent insertions, deletions or chromosomal

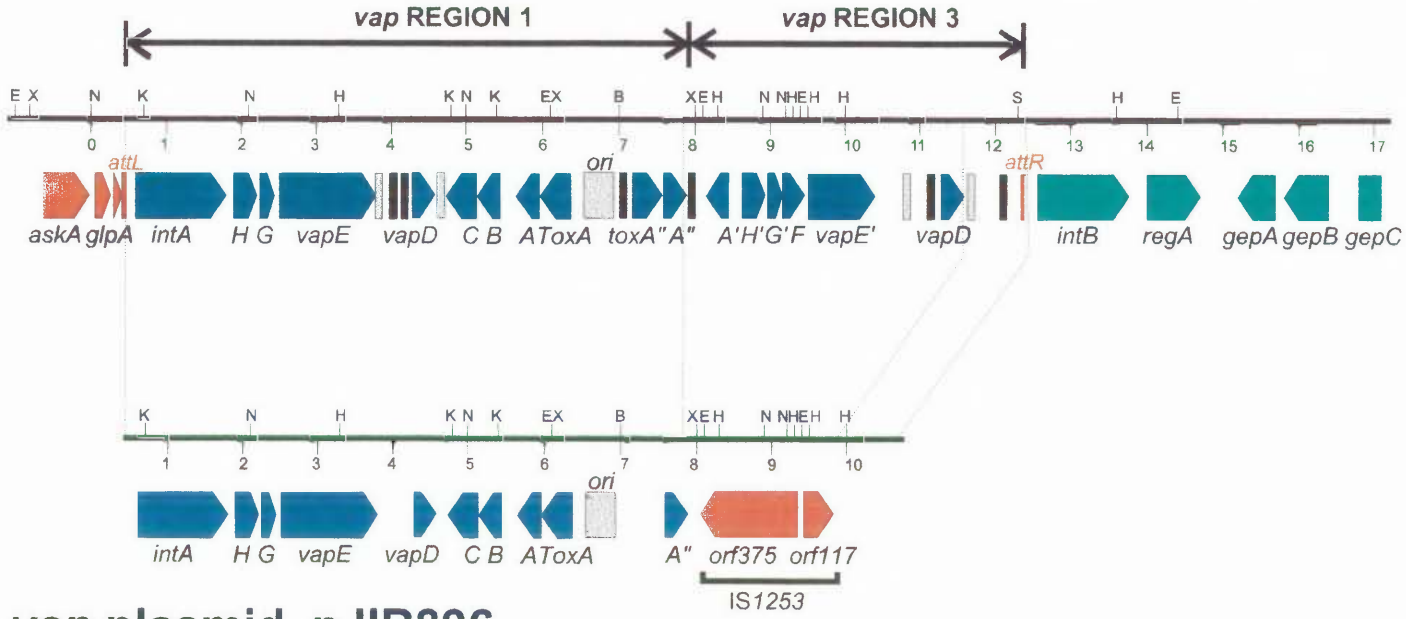



**Figure 1.8:** Alignment of a linear restriction map of *vap* regions 1 and 3 of *D. nodosus* strain A198 [Cheetham, 1995 #372], and the *vap* plasmid, pJIR896 from *D. nodosus* strain AC3577 [Billington, 1996 #343]. The numbers show the distance in kb. Restriction enzyme sites are as follows: *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Nru*I (N), *Sac*I (S) and *Xho*I (X). Open reading frames and their orientation are indicated by shaded arrows and are classified as follows: belonging to the *vap* regions (**blue**), belonging to the *intB* element (**green**), genes that are not part of an integrated genetic element (**red**), and a *tRNA* gene (**red triangle**). Repeated sequences include: attachment sites, *attL* or *attR* (narrow red boxes), 102 bp repeats or partial copies thereof (filled boxes), 103 bp repeats or partial copies of (narrow grey boxes), and a putative origin of replication is labelled *ori* (grey box). The extent of the sequences that appear similar between *vap* regions 1 and 3 and pJIR896 are indicated by dotted lines. Note that a copy of *toxA"* and the repeated sequences have been omitted from the map of pJIR896 because probes specific for these sequences were not used in Southern hybridisation analysis of pJIR896, hence although it is likely these sequences are present, it has not been confirmed experimentally. Note that 102 bp and 103 bp repeats are not shown on the map of pJIR896 because it is unknown whether these sequences are present.

rearrangements as inferred by the duplication of *vapA, D, E, G, H, toxA* and their identity to their *vapA', A", D, E', G, H', toxA"* counterparts, and the multiple copies of the *vap* regions.

In order to determine whether the *vap* regions in A198 could have arisen *via* chromosomal duplications, or multiple independent integration events, the DNA sequences flanking *vap* regions 1, 2 and 3 were analysed (Bloomfield *et al.*, 1997). Comparison of sequences to the left of *vap* regions 1 and 2 indicated that *vap* region 2 arose *via* the integration of the *vap* element into a *tRNA-*$_{serGGA}$ gene, rather than a *tRNA-*$_{serGCU}$ gene into which *vap* region 1 is integrated, and hence occurred *via* an independent integration event rather than by duplication (Figure 1.8) (Bloomfield *et al.*, 1997). In addition, sequences to the left of these *tRNA* genes were not related. Partial sequencing of the region adjacent to *vap* region 2 led to the identification of an orf designated *pnpA* (Figure 1.8), with 67% amino acid identity over 239 aa to the polynucleotide phosphorlyase genes of *E. coli* (Regnier, Grunberg-Manago & Portier, 1987), *Photorhabdus luminsecens* (Clarke & Dowds, 1994) and *Haemophilus influenzae* (Fleischmann *et al.*, 1995). In contrast, an orf designated *askA*, with 48.2% aa identity with *Mycobacterium smegmatis* aspartokinase gene (Cirillo *et al.*, 1994) was located upstream of the *tRNA* gene adjacent to *vap* region 1 (Bloomfield *et al.*, 1997).

### 1.6.4 *vap* gene similarities

At present there is no transformation system for *D. nodosus*, which precludes direct testing of the role of *vap* and other genes in the pathogenesis of ovine footrot. Attempts to develop such a transformation system, using derivatives of the *vap* plasmid (Billington *et al.*, 1996b) and the recently-identified native *D. nodosus* plasmid, pDN1 (Chapter 3), are in progress. In the interim, putative functions have been assigned to some genes on the basis of amino acid similarities to known proteins (Table 1.2).

**Table 1.2: Sequence Analysis of *vap* regions 1 and 3[a] of *D. nodosus* strain A198**

| Gene | Co-ordinates 5'-3- (nt) | Size (aa) | % identity to nt or putative protein | Homologues/Description | Accession Number | P(*n*) |
|---|---|---|---|---|---|---|
| ***intA*** | 418-1623 | 402 | 39.6/394aa | *Escherichia coli* 0157:H7 LEE pathogenicity island integrase [tRNA-sec gene] | AF071034 | 2.3e-65 |
| | | | 41.5/400aa | *E. coli* retronphage phiR73 integrase [*selCtRNA*] | M64113 | 1.5e-61 |
| | | | 34.6/405aa | *Vibrio cholerae* 0395 integrase from CTXϕ PAI [*ssrA*] | U02372 | 4.4e-51 |
| | | | 39.4/393aa | *E. coli* K12 cryptic prophage P4-57 integrase [*ssrA*] | U03737 | 1.9e-50 |
| | | | 37.6/394aa | *Mesorhizobium loti* symbiosis island integrase [*tRNAphe*] | AF049242 | 4.1e-50 |
| | | | 34.3/397aa | *Shigella flexneri* ϕSF6 integrase | X59553 | 3.9e-40 |
| | | | 34.0/396aa | *E. coli* K12 integrase [*tRNAargW*] | U11296 | 3.8e-39 |
| | | | 32.2/273aa | *E. coli* K12 integrase [*tRNAleuX*] | AE000323 | 1.3e-38 |
| | | | 35.6/135aa | *Rhodobacter capsulatus* putative prophage integrase | U57682 | 1.4e-32 |
| | | | 44.7/123aa | *Enterobacter aerogenes* putative integrase | AF039582 | 9e-19 |
| | | | 45.2/124aa | Bacteriophage P4 integrase | X05947 | 1.7e-18 |
| | | | 23.3/408aa | *Salmonella typhimurium* LT2 integrase of Gifsy-1 prophage [lepA] | AF001386 | 2.9e-16 |
| | | | 47.2/117aa | *Pseudomonas putida* integrase from clc element [*tRNAgly*] | PPAJ4950 | 1e-15 |
| | | | 22.0/239aa | Lambda Bacteriophage phi80 integrase | X04051 | 1.2e-12 |
| | | | 40.1/167aa | *Dichelobacter nodosus* integrase, IntB [$tRNAser_{GCU}$ or $tRNAser_{GCA}$] | X98547 | 1.6e-10 |
| | | | 38.9/210aa | *Salmonella enteritidis* integrase from IS-3-like insertion element | SEJ002209 | 3e-06 |
| | | | 21.9/192aa | *Streptococcus pyogenes* T12 prophage integrase [*tRNAser*] | U40453 | 5e-04 |
| | | | 22.2/239aa | *Lactococcus lactis* integrase | L27649 | 1.5e-03 |
| ***vapA*** | 5936-5628 | 103 | 50.0/60aa | Putative Orf from *Synechococcus* sp. | U04356 | 2.4e-09 |
| | | | 38.4/73aa | Orf from *Acetobacter europaeus* plasmid | AEY17109 | 2.5e-08 |
| | | | 39.7/78aa | Putative Orf from *E. coli* K12 | AE000154 | 1.1e-06 |
| | | | 22.0/70aa | HigA antidote-protein from killer plasmid Rts1 | U43847 | 2.5e-06 |
| | | | 30.0/78aa | Putative Orf from *E. coli* K12 | AE000244 | 2.6e06 |
| | | | 41.4/70aa | VapA (HI1251) from *Haemophilus influenzae* | U32805 | 6.1e-05 |
| ***toxA*** | 6227-5949 | 92 | 45.6/92aa | HigB killer protein from killer plasmid Rts1 | U43847 | 6.8e-20 |
| | | | 33.7/89aa | HI1250 from *H. influenzae* | U32805 | 1.6e-05 |

Table 1.2 continues on next page / . . .

| | | | | | | |
|---|---|---|---|---|---|---|
| ***vapB*** | 5292-5062 | 77 | 38.7/75aa | Encoded from within orf *patA* from *Synechocystis* sp. | D90902 | 2.7e-10 |
| | | | 35.0/60aa | *Actinobacillus actinomycetecomitans* VapB protein | AF006830 | 6.2e-07 |
| | | | 36.0/75aa | VapB (HI0321) from *H. influenzae* | U32776 | 8.5e-07 |
| | | | 34.8/66aa | VagC from *Salmonella dublin* virulence plasmid | X66934 | 9.7e-07 |
| | | | 42.4/59aa | HI0953 from *H. influenzae* within orf | U32717 | 1.3e-05 |
| | | | 34.7/72aa | Encoded by an unidentified *orf* from *Agrobacterium tumefacieins* | M59852 | 1.3e-05 |
| | | | 34.7/72aa | Encoded from within orf *traD* of *E. coli* F plasmid | M29254 | 4e-04 |
| | | | 34.7/72aa | Encoded from within orf *traI* of *E. coli* F plasmid | M54796 | 7e-04 |
| | | | 33.3/72aa | Stborf1 from *S. flexineri* virulence plasmid pMYSH6000 | U82621 | 2.3e-03 |
| ***vapC*** | 5065-4658 | 136 | 40.6/133aa | Stborf2 from *S. flexineri* virulence plasmid pMYSH6000 | U82621 | 1.2e-21 |
| | | | 35.7/132aa | VapC (HI0322) from *H. influenzae* | U32776 | 1.4e-21 |
| | | | 40.9/132aa | VagD from *H. influenzae* | U32776 | 2.2e-18 |
| | | | 40.9/129aa | HI0953 from *H. influenzae* within orf | U32717 | 4e-16 |
| | | | 47.1/128aa | Encoded from within orf *patA* from *Synechocystis* sp. | D90902 | 2.2e-13 |
| | | | 38.2/85aa | Putative protein from *Synechocystis* sp. | D90902 | 1e-12 |
| | | | 42.7/82aa | VagD from *S. dublin* virulence plasmid | X66934 | 6.3e-09 |
| | | | 39.9/133aa | Encoded from within orf *traD* of *E. coli* F plasmid | X57431 | 7e-07 |
| | | | 39.9/133aa | Encoded from within orf *traI* of *E. coli* F plasmid | M54796 | 7e-07 |
| | | | 32.9/94aa | Putative protein from *Methanococcus jannaschii* | U67535 | 6.6e-06 |
| | | | 31.9/94aa | Encoded by an unidentified orf from *A. tumefaciens*[A] | M59852 | 9.3e-06 |
| ***vapD*** | 4255-4536 | 281 | 80.6/93aa | Orf2 from *A. actinomycetemcomitans* plasmid pVT736-1 | L24000 | 7.2e-40 |
| | | | 68.8/93aa | HP0309 from *Helicobacter pylori* | AE000549 | 1e-31 |
| | | | 55.8/86aa | *Riemerella antipestifer* plasmid pCFC1 VapD2 | AF048718 | 2.5e-30 |
| | | | 66.7/91aa | VapD from *H. pylori* | U94318 | 2.8e-29 |
| | | | 70.3/74aa | A protein encoded within the ORFA-ORFB of *Treponema denticola* plasmid, | M87856 | 2.5e-26 |
| | | | 46.0/91aa | pTD1 | M10316 | 2.4e-21 |
| | | | 40.7/91aa | Orf5 from cryptic plasmid pJD1 of *Nessieria gonorrhoea* | U32728 | 5.5e-20 |
| | | | 33.3/93aa | VapD from *H. influenzae* | AF048718 | 1.7e-11 |
| | | | 33.3/93aa | *R. antipestifer* plasmid pCFC1 VapD1 | AE000605 | 4.3e-07 |
| | | | 29.0/93aa | HP0969 from *H. pylori* | D88189 | 7.6e-03 |
| | | | | VapD from *A. actinomycetemcomitans* chromosome | | |
| ***vapE*** | 2304-3617 | 437 | 37.9/353aa | VapE protein from *Staphylococcus aureus tst* element pathogenicity island | U93688 | 1.4e-54 |
| | | | 30.0/307aa | Orf1 of cryptic plasmid pMA1 of cyanobacterium *Microcystis aeruginosa* | Z28337 | 9.4e-21 |
| ***vapE'*** | 9393-10234 | 280 | 37.9/219 | VapE protein from *S. aureus tst* element pathogenicity island | U93688 | 6.4e-34 |
| ***vapG*** | 2002-2208 | 69 | - | NSH[b] | - | - |
| ***vapH*** | 1718-2023 | 102 | 62.8/145 nt | *S. flexneri* IS2 *orf179* putative bacteriophage immunity region | Z23101 | 8e-02 |

a. *vap* element from *D. nodosus* strain A198 (GenBank accession number L31763).
b. NSH indicates that there is no significant homology to sequences in databases.
c. Where the site of integration is known for an integrase gene it is indicated in square brackets following the desciption of the Int homologue.

### 1.6.4.1 *vapA* and *toxA*

*vapA* and *toxA* have similarity to the *hig* genes of killer plasmid Rts1 (Tian *et al.*, 1996). VapA has 22% aa identity with the HigA antidote-protein, whilst ToxA has 45.6% aa identity to the 92 aa HigB killer-protein (Bloomfield *et al.*, 1997) (Table 1.2). The sequential arrangement of *toxA* and *vapA* genes in *D. nodosus* is the same as the arrangement of *higB* and *higA* on the plasmid Rts1. Within *vap* regions 1 and 3 there are homologous copies of *vapA* and *toxA*, designated *vapA"*, *toxA"* and *vapA'* (Figure 1.7).

*vapA"* and *toxA"* are adjacent and their protein products have 29.3% and 31.5% identity to HigA and HigB respectively. *vapA'* is not preceded by a *toxA* homologue in strain A198. *higB* and *higA* have been shown to encode a plasmid maintenance system in which a stable HigB protein selectively kills plasmid-free cells, which is neutralised only if the unstable HigA protein is present, either in *cis* or *trans* (Tian *et al.*, 1996). Experimental evidence also supports the hypothesis that *vapA* and *toxA* encode a similar plasmid maintenance system, since it was not possible to obtain a subclone containing *toxA"* without *vapA"* in *E. coli*, suggesting that the protein encoded by *toxA* is toxic in *E. coli*, and only neutralised in the presence of the VapA protein (Bloomfield *et al.*, 1997). A number of similar but unrelated plasmid maintenance systems have been identified on conjugative plasmids F (Hiraga *et al.*, 1986), R1 (Gerdes, Thisted & Martinussen, 1990), R100 (Tsuchimoto, Nishimura & Ohtsubo, 1992) and in the bacteriophage P1 genome (Lehnherr *et al.*, 1993) (Table 1.2).

VapA also has 50% identity in 60 aa to part of the *atpE* gene from cyanobacterium *Synechococcus* (Phung, Ajlani & Hazelkorn, 1994), and 39.7% identity in 78 aa to an region within a probable copper-transporting ATPase from *E. coli* (Das *et al.*, unpublished). Both of these proteins are associated with copper-transporting ATPases, hence perhaps the observed similarity is due to the presence of a common functional domain between the ATPases and *vapA*, or alternatively may be proteins that were not identified because the regions of similarity are found within the coding regions of both copper-transporting ATPase

genes. In addition, *vapA* has 41.4% identity to a protein from *Haemophilus influenzae*, which has consequently been designated *vapA* (HI1251) (Fleischmann *et al.*, 1995). 11 bp upstream from the start codon of *H. influenzae vapA* is a gene HI1250, that has 33.7% identity to *toxA* " (Fleischmann *et al.*, 1995). In addition, the 5' end of the gene encoding the VapA protein has 33.3% similarity to ParD of plasmid RK2 which inhibits ParE-associated toxicity, and similar genes in plasmids RP4 and RP1 (Gerlitz, Hrabak & Schwab, 1990; Kholodii *et al.*, 1993; Roberts & Helinski, 1992). These results indicate that *vapA* and *toxA* are genes involved in maintenance of the extrachromosomal element, *via* a postsegregational killing mechanism.

### 1.6.4.2 *vapB* and *vapC*

In *D. nodosus* strain A198, genes *vapC* and *vapB* are in an operon-like arrangement, since they overlap by two nts and are thought to be transcribed from the same promoter. Proteins encoded by *vapB* and *vapC* show similarity to VagC (34.8% identity in 66 aa) and VagD (42.7% identity in 82 aa) genes respectively from the virulence plasmid of *Salmonella dublin* (Pullinger & Lax, 1992) (Table 1.2). *vagC* and *vagD* also overlap and are thought to be transcribed from the same promoter, suggesting they are also coordinately expressed. It has been proposed that the role of *vagC* and *vagD* is to coordinate virulence plasmid replication with cell division (Pullinger & Lax, 1992); disruption of such a function could reduce copy number of the virulence plasmid, and consequently reduce the virulence characteristics of a given strain. In addition, VapB and VapC have 33.3% and 40.6% similarity respectively, to proteins encoded by *stborf1* and *stborf2* from the plasmid stability locus of *Shigella flexineri* virulence plasmid, pMYSH6000 (Radnedge *et al.*, 1997). *stborf1* and *stborf2* are thought to ensure a better-than-random distribution of plasmid copies to daughter cells, and exert incompatibility against a co-resident plasmid expressing the Stb protein (Radnedge *et al.*, 1997).

There are two pairs of genes encoding proteins with similarity to VapB and VapC in

*H. influenzae* (Fleischmann *et al.*, 1995). The first pair of these genes are designated *vagC* (HI0321, 36% aa identity to VapB) and *vapC* (HI0322, 35.7% aa identity to VapC). The other pair of genes, is located within a putative orf, HI0953, and have 42.4% and 40.9% similarity to VapB and VapC respectively. No function has been assigned to these *vapB*- and *vapC*-like genes in *H. influenzae*.

There are other examples of proteins from both plasmids and chromosomes of different organisms with similarity to *D. nodosus* proteins VapB and VapC (Table 1.2). In all cases the genes encoding these proteins are adjacent to each other as in *D. nodosus*.

As reported for genes *vapA* and *toxA*, genes with a high degree of similarity to *vapB* and *vapC* are, in general, involved in plasmid maintenance. However, the mechanism by which those VapB- and VapC-like proteins act in partitioning is not known, and is not thought to be related to the postsegregational-killing mechanisms (Radnedge *et al.*, 1997).

### 1.6.4.3 *vapD*

There are 3 identical copies of *vapD* in the genome of *D. nodosus* strain A198 in *vap* regions 1, 2 and 3 respectively (Figure 1.5) (Cheetham *et al.*, 1995b; Katz *et al.*, 1994). The *vapD* gene product has 80.6% identity (in 93aa) to *orf2* from a rolling-circle replicating plasmid, pVT736-1 from *Actinobacillus actinomycetemcomitans* (Galli & Leblanc, 1995; Galli & LeBlanc, 1997; Katz *et al.*, 1994), that has been shown to have a function in plasmid maintenance and incompatibility (Galli & LeBlanc, 1997). Although VapD shares amino acid identity with many other proteins encoded by genes on plasmids and integrated genetic elements (Table 1.2), the only VapD-like protein that has a known function, has been shown to function in plasmid maintenance and incompatibility (Galli & LeBlanc, 1997). Hence, these results indicate that *vapD*, together with *vapA-C* and *toxA* may form a functional domain involved with plasmid maintenance.

### 1.6.4.4 *vapE* and *vapE'*

The aa sequence of *vapE* has 67% identity to *vapE'* over the first 246 aa, after which the two protein sequences diverge, with the putative VapE and VapE' proteins consisting of 437 and 280 aa respectively (Cheetham *et al.*, 1995b). VapE has 30% aa similarity over 307 aa to a protein of unknown function encoded by *orf1* of the cryptic plasmid pMA1, of unicellular cyanobacterium *Microcystis aeruginosa* (Tominaga *et al.*, 1994) (Table 1.2). More recently, similarity (37.9% in 353 aa) has been noted between VapE and a potential protein from a 15.2 kb pathogenicity island from *Staphylococcus aureus*, called the *tst* element, which carries a gene encoding tox c shock toxin (Lindsay *et al.*, 1998) (Table 1.2). It is of interest that the *tst* element, like the *vap* regions of *D. nodosus*, carries an integrase gene and is flanked by putative attachment sites, and is of comparable size. In addition, at least two variants of the *tst* element have been identified in *S. aureus*, and have been designated SaPI1 and SaPI2 respectively (Lindsay *et al.*, 1998).

### 1.6.4.5 *vapF, vapG and vapH*

*vapG* from *vap* region 1 in strain A198 has 71% aa similarity to homologue *vapG'*, from *vap* region 3 but has no similarity to protein sequences available in the GenBank databases. No similarity to putative proteins VapF and VapH have been identified.

### 1.6.4.6 Conclusions from *vap* gene similarities

It is interesting that genes with similarity to the *vap* genes are found in both extrachromosomal and chromosomal locations, suggesting that genetic elements that carry *vap*-like genes can integrate into the chromosome or replicate autonomously, and thus have the potential to be horizontally transmitted. Additionally, it is notable that *vap*-like genes are found a wide range of bacteria, although they are not necessarily all present and not always clustered together as they are in the *vap* regions of *D. nodosus*. In *H. influenzae* those genes similar to *vapA/toxA* and *vapB/vapC* and *vapD* are scattered throughout the genome

(Fleischmann *et al.*, 1995). Similarly, in *Synechoccus* sp. potential proteins with similarity to VapA and VapB/VapC have been identified, but are not adjacent to one another on the chromosome. Such a different arrangement suggests a complex and long evolutionary process.

It is significant that proteins that are similar to VapA/ToxA, and VapB/VapC are arranged in a similar manner as in *D. nodosus*. In addition, it is interesting that the proteins that have been functionally analysed and are related to VapA/ToxA, VapB/VapC, and VapD all have functions involved in maintenance and stability of elements that have an autonomously replicating phase, which also indicates that *vapA-D* and *toxA* are functionally linked. Furthermore, the existence of so many highly related open reading frame pairs in such different bacterial species suggests that these genes may have a common ancient origin and there has been wide distribution of these genes by horizontal transfer and/or there is strong conservation for these genes.

### 1.6.5 Origin of the *vap* element - plasmid, bacteriophage, or transposon?

The identification of an integrase gene, putative *attL* and *attR* sites, putative origin of replication, and an extrachromosomal plasmid-like molecule, together strongly support the hypothesis that the *vap* regions do have an extrachromosomal origin. In addition, many of the genes that have similarity to the *vap* genes (Section 1.6.3) are located on mobile genetic elements, including bacteriophages, plasmids and transposons, and hence also suggest that the *vap* regions may have arisen *via* the integration of an extrachromosomal element (Katz *et al.*, 1992).

This extrachromosomal element may be a bacteriophage, a plasmid carrying an integrase gene, or a bacteriophage-like transposon. Although the P4-family of integrases are found on a variety of genetic elements (Section 1.6.3) the integrase genes most closely related to *intA* are from bacteriophages (Table 1.2). In addition, electron microscope studies

have identified putative bacteriophage or incomplete phage particles in *D. nodosus* strain A198 (Hine, 1988). Furthermore, *vapG* and *vapH* have similarity to the immunity region of bacteriophage P4 (Section 4.2)(Deho *et al.*, 1992; Ghisotti *et al.*, 1992). Bacteriophages also carry postsegregational maintenance systems (Lehnherr *et al.*, 1993), and have a circular replicative phase (Lindqvist, Deho & Calendar, 1993). If the origin of the *vap* element involved the integration of a bacteriophage, it was probably a defective bacteriophage since the *vap* regions 1 and 3 are 12 kb in size, which is of a similar to the size of defective bacteriophage P4 (11 .6 kb), but substantially smaller than bacteriophage λ (48.5 kb), P22 (43 kb) or P2 (33.8 kb) (Birge, 1994). The original *vap* region may in fact be *ca.* 8 kb, like *vap* region 2 in *D. nodosus* strain A198. It is unlikely that the *vap* element was originally larger and later was subject to deletions, since no larger variant has been found in any of the numerous *D. nodosus* strains analysed.

The progenitor of the *vap* regions could have also been a plasmid carrying an integrase gene and a bacteriophage immunity region, or an element similar to the one in which a protein with similarity to VapE was identified which is described by the authors (Lindsay *et al.*, 1998) as a bacteriophage-like transposon; both of these elements could also carry genes involved in maintenance of their circular extrachromosomal intermediates. Whether the *vap* regions were acquired by a plasmid, or a transposon, they are relatively small, and certainly not large enough to encode all the genes necessary for sex-pilus mediated conjugation, which would require a coding region of approximately 33 kb (Dorman, 1994).

At this stage it is not possible to determine whether the extrachromosomal origin of the *vap* element involved a bacteriophage, a plasmid or a transposon. However, the size of the *vap* regions of *D. nodosus* strain A198 (12 kb) indicates that a separate genetic element must have been involved in the acquisition of the *vap* element by *D. nodosus*, since the *vap* element does not seem to encode proteins that would be required for horizontal transfer/acquisition mediated by conjugation or transduction.

### 1.6.6 Analysis of the *vap* regions in four strains of *D. nodosus*

To further investigate the evolution of the *vap* regions in *D. nodosus*, and to investigate the importance of specific *vap* genes in virulence, the arrangement and presence of the *vap* genes in four strains of *D. nodosus* was determined (Figure 1.9). The four isolates studied, which included strains from serogroups B (virulent strain B1006), G (virulent strain G1220) and H (virulent strain H1215 and benign strain H1204), were compared to virulent reference strain A198 (serogroup A). The maps of the *vap* regions in these strains were constructed from the results of Southern blot experiments which were subsequently confirmed by PCR experiments and are summarised in (Figure 1.9) (Bloomfield *et al.*, 1997).

Southern blot experiments indicated that the arrangement of the *vap* genes was similar in benign and virulent strains of *D. nodosus* with one exception. Although all four strains contained *vapA-D* in an arrangement identical to that of strain A198, virulent strain H1215 did not contain *vapE*, but contains a copy of *vapE'*, located in the same position as *vapE* in all other strains (Figure 1.9). This is similar to the arrangement of *vap* region 3 of strain A198 except that *vap* region 3 does not contain *vapA-C* or *toxA*. Hence, it was proposed that *vap* region 3 of A198 could have evolved *via* the integration of a separate *vap* element that contained *vapE'* instead of *vapE*, followed by the loss of *vapA-C* and *toxA* rather than by the partial duplication of *vap* region 1 and subsequent divergence of *vapE* (Bloomfield *et al.*, 1997). In addition since *vapE* is absent from virulent strain H1215 these results indicate that either VapE is not essential for virulence or alternatively, suggest that VapE' is able to perform the same function as VapE. This is possible since the VapE' protein product has 67% similarity over the first 246 aa after which the aa sequence of the two proteins diverge (Cheetham *et al.*, 1995a).

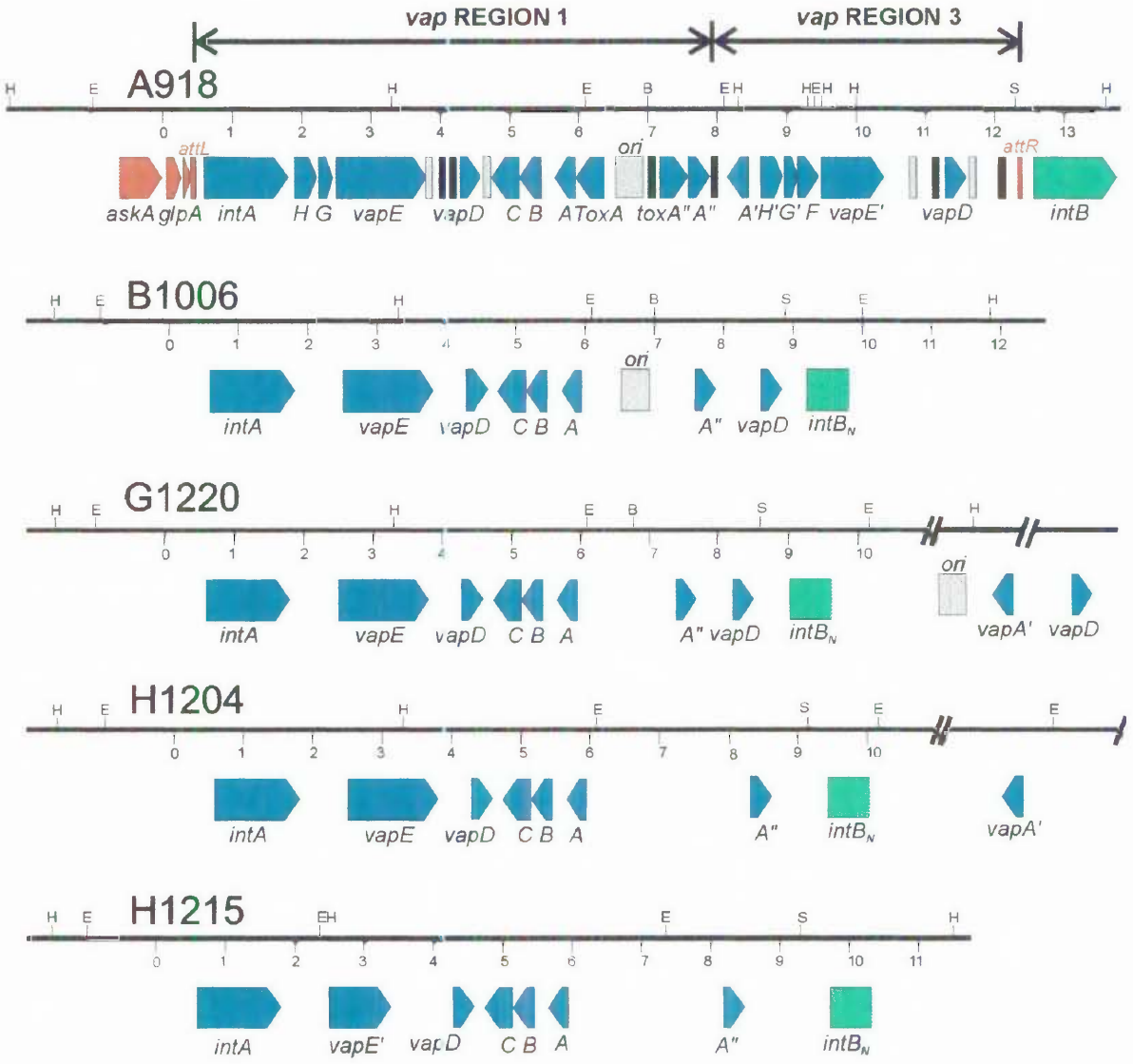



**Figure 1.9:** A map of the *vap* regions of *D. nodosus* strains A198, B1006, G1220, H1204, and H1215. Restriction sites shown include: *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), and *Sac*I (S). The positions of genes were determined by Southern blot analyses and PCR. Open reading frames are indicated by shaded arrows and are classified as follows: belonging to the *vap* regions (blue), belonging to the *intB* element (green), genes assumed not to be part of an integrated element (red), including *tRNA* genes (red triangles). Repeated sequences include: attachment sites, *attL* or *attR* (narrow red boxes), 102 bp repeats or partial copies thereof (filled boxes), 103 bp repeats or partial copies thereof (grey boxes), and a putative origin of replication is labelled *ori* (grey box). Note that in strains B1006, G1220, H1204 and H1215, only *vap* genes (with the exception of *toxA*, *toxA"*, *vapG*, *vapH*, *vapG'*) and *intB* were used as probes. As a result, other orfs from the *intB* element, and repeated sequences have not been included in this figure.

Results from these Southern blot experiments indicated that multiple copies of the *vap* genes are not required for virulence since virulent strains B1006, G1220 and H1215 all have single copies of the *vap* element. Furthermore, since benign strain H1204 contains a copy of all genes identified within the *vap* regions 1 and 3, the presence or absence of *vap* genes alone did not determine whether a given strain had a benign or virulent phenotype, and therefore lack of virulence is attributable to one or more of the *vap* gene products being non-functional or due to factors outside the *vap* regions (Bloomfield *et al.*, 1997).

### 1.6.7 Identification of the *intB* element immediately adjacent to *vap* region 3 in *D. nodosus* strain A198

Sequencing and analysis of a 4.8 kb fragment beginning at the *Sac*I site at the right-hand end of *vap* region 3 led to the identification of five potential genes including *intB*, *regA*, *gepA*, *gepB* and *gepC*, that appear to comprise part of a new genetic element, designated the '*intB* element' (Figure 1.10), that is integrated into the *attR* of *vap* region 3 in *D. nodosus* strain A198. The orf *intB*, although a pseudogene, encodes a protein with similarity to several bacterial integrases including 33.8% identity over 403 aa to *E. coli* retronphage ϕR73 (Sun *et al.*, 1991), 32.6% aa identity with the integrase from bacteriophage SF6 of *Shigella flexneri* (Clark *et al.*, 1991), and 31.3% aa identity to an integrase from *Vibrio cholerae* (Kovach & Peterson, 1994). The *regA* gene has amino acid similarity to bacterial DNA-binding proteins and bacteriophage repressor proteins, having 40.5% aa identity with the pectin lyase regulator, RdgA of *Erwinia carotovora* (Liu *et al.*, 1996), 39.2% identity to regulatory protein PrtR of *Pseudomonas aeruginosa* (Matsui *et al.*, 1993), and 36.2% to the *cI* repressor gene product of *E. coli* bacteriophage ϕ80 (Ogawa, Ogawa & Tomizawa, 1988). The gene designated *gepA* (genetic element protein A) encodes a protein with 33.3% amino acid identity to *orf4* located downstream from the *rteC* gene located near the *oriT* of a conjugative transposon from *Bacteroides thetatotaomicron* (Stevens *et al.*, 1993).

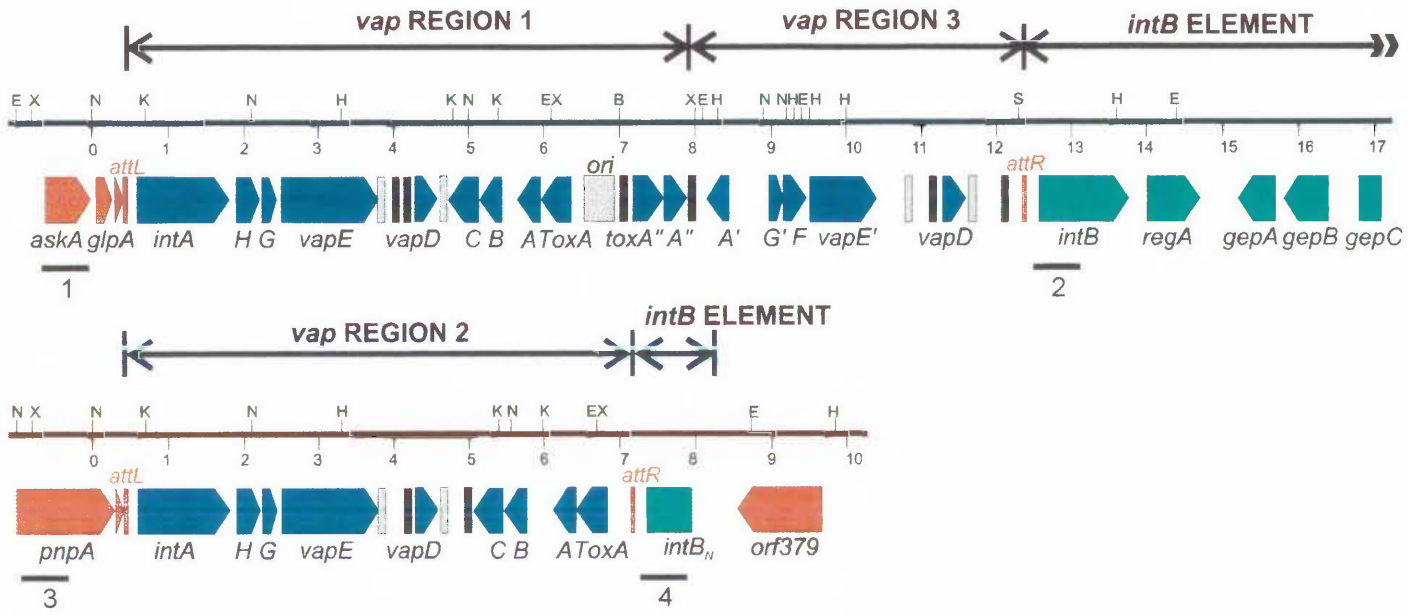



**Figure 1.10:** Restriction map of the *vap* regions of *D. nodosus* strain A198 and the *intB* element at the right of *vap* region 3 together with the partial copy of *intB*, designated *intB*$_N$, at the right of *vap* region 2. The numbers indicate the distance in kb. Restriction enzyme sites shown are *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Nru*I (N), *Sac*I (S) and *Xho*I (X). Open reading frames are indicated by shaded arrows and are classified as follows: belonging to the *intB* element (**green**), belonging to the *vap* regions (**blue**), and genes that are flank the integrated elements (**red**) including *tRNA* genes (**red triangles**). Repeated sequences are as follows: attachment sites, *attL* or *attR* (narrow red boxes), 102 bp repeats or partial copies thereof (filled boxes), 103 bp repeats or partial copies of (narrow grey boxes), and a putative origin of replication, labelled *ori* (grey box). Probe sequences are indicated as follows: *askA* (1), *intB* (2), *pnpA* (3) and *intB*$_N$ (4).

The GepB protein has 64.9% aa similarity over 188 aa to an unidentified orf in the *cob* gene cluster from *Pseudomonas denitrificans* (Crouzet *et al.*, 1991), whilst *gepC* has no similarity to genes in databases (Bloomfield *et al.*, 1997).

The arrangement of the open reading frames in the *intB* element, an *att* site followed by an integrase gene followed by a regulatory protein is characteristic of some lambdoid bacteriophages including λ and P2, suggesting that the *intB* element may be an integrated bacteriophage. However, since *gepA* has similarity to an orf from a conjugative transposon from *B. thetaiotaomicron*, and since some conjugative transposons carry integrases from the lambda integrase family (Salyers *et al.*, 1995) it is also possible that the *intB* element may be a conjugative transposon (Bloomfield *et al.*, 1997).

Comparison of the sequences to the right of the *attR* site *vap* region 3 with the sequence to the right of *attR* of *vap* region 2, indicated that a second, though partial copy of *intB*, designated $intB_N$ (Figure 1.10), is present at the right of *vap* region 2 in *D. nodosus* strain A198 (Bloomfield *et al.*, 1997).

### 1.6.8 Investigation of non-virulence-associated DNA regions adjacent to the *vap* regions of *D. nodosus* virulent strain A198, in the benign strain C305

In the virulent strain A198 *vap* region 1 is integrated into the 3'-end of a $tRNA\text{-}ser_{GGU}$ gene, upstream from which there is an aspartokinase gene, *askA*. Immediately downstream from the right-hand end of *vap* region 3 is the *intB* element (Figure 1.10, section 1.6.7).

*vap* region 2 is believed to have arisen *via* the independent integration of the *vap* element into the 3'-end of a different *tRNA* gene, $tRNA\text{-}ser_{GGA}$, which is located immediately downstream of the polynucleotide phosphorylase gene, *pnpA* (Figure 1.10). At the 3'-end of *vap* region 2 there is a second, but partial copy of the *intB* gene that consists of the first

164 aa from the amino-terminus of the 403 aa *intB* gene, and hence is designated $intB_N$ (Figure 1.10). In summary then, in *D. nodosus* strain A198 there are two copies of the *intB* gene, or part thereof, flanking the right hand ends of *vap* region 1/3 and *vap* region 2 respectively.

The benign *D. nodosus* strain C305 does not contain any copies of the *vap* regions. However, Southern blot analyses have shown that an *intB* probe hybridised to two *Hin*dIII fragments, suggesting that there are two copies, or partial copies of the *intB* gene in the C305 genome. Since there are no copies of the *vap* element in strain C305, it was hypothesised that the two copies of the *intB* gene would be located immediately adjacent to *askA* (Figure 1.11A) and *pnpA* genes respectively (Figure 1.12A).

To investigate this hypothesis, a library of genomic DNA from benign strain C305 was prepared (Bloomfield, 1997) and probed with sequences that flank the left- and right-hand ends of *vap* regions 1 (*askA*) and 3 (*intB*) of *D. nodosus* strain A198 (Figure 1.11B, Section 1.6.8.1) and the left- (*pnpA*) and right-hand ($intB_N$) ends of *vap* region 2 (Figure 1.12B, Section 1.6.8.2.).

### 1.6.8.1 An *intC* element and $intB_M$ are present adjacent to *askA* in *D. nodosus* benign strain C305

A lambda clone that hybridised to both the *askA* and *intB* probes was isolated and hybridisation analysis revealed that *askA* and *intB* were not adjacent, but were instead separated by approximately 7 kb of sequence (Figure 1.11C). The sequence of an 8035 bp region from this lambda clone (λGB321) was determined, several orfs were identified (Figure 1.11C), and the presence of a new genetic element, designated the *intC* element, was postulated (Bloomfield, 1997).

The *askA* gene in strain C305 is almost identical to the *askA* gene at the left of *vap*

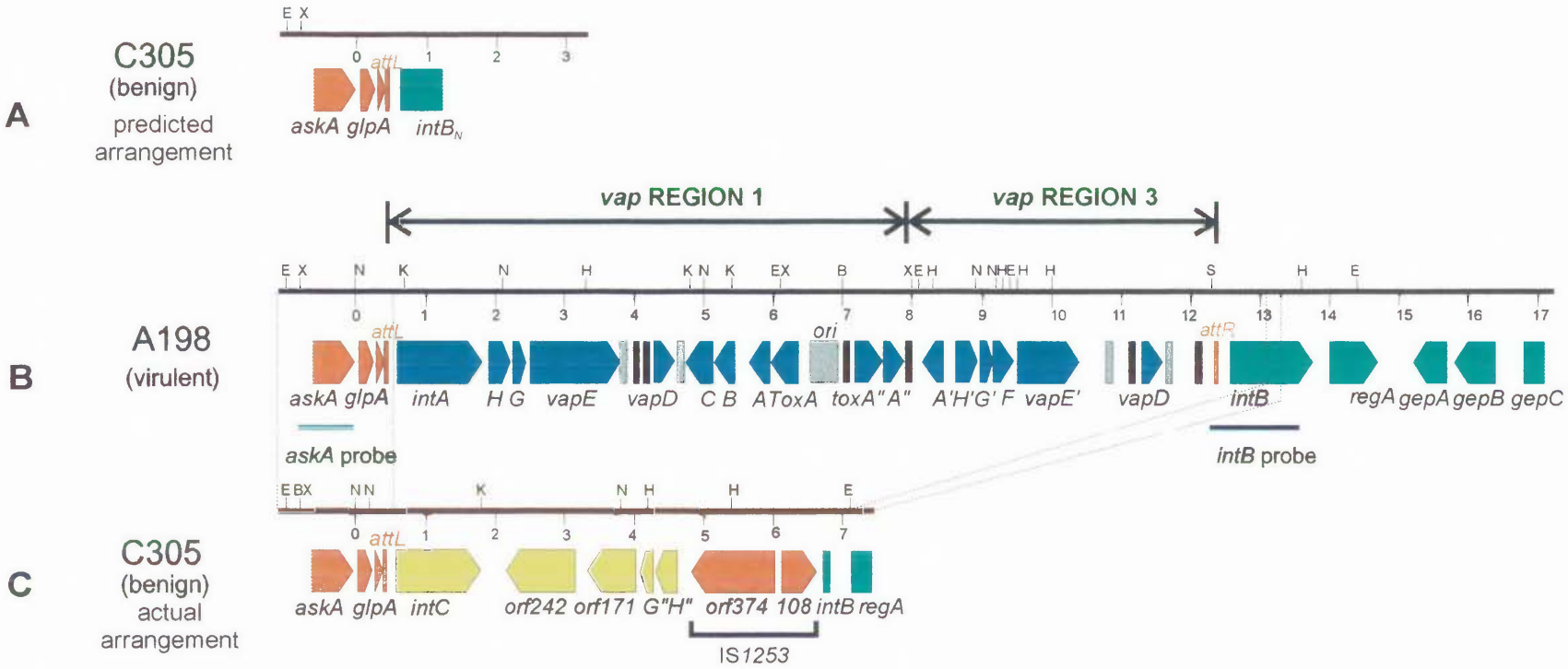



**Figure 1.11:** Alignment of restriction maps of the predicted arrangement of *askA* (**red**) and *intB* (**green**) in benign *D. nodosus* strain C305 that does not contain the *vap* element (**A**) with *askA*, *vap* region 1 and 3 (**blue**) and *intB* of *D. nodosus* virulent strain A198 (**B**), and the actual organisation of *askA* and *intB* genes in *D. nodosus* strain C305 (**C**) is shown. The *askA* and *intB* probe regions are indicated by bold lines below the corresponding sequences in *D. nodosus* strain A198. The numbers indicate fragment sizes in kb. Restriction enzyme sites shown are *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Nru*I (N), *Sac*I (S) and *Xho*I (X). Open reading frames are indicated by shaded arrows and are classified as follows: belonging to the *intB* element (**green**), belonging to the *vap* region (**blue**), belonging to the *intC* element (**yellow**) genes that are not part of *vap*, *intB* or *intC* elements (**red**), including *tRNA* genes (**red triangles**). Repeated sequences are as follows: attachment sites, *attL* or *attR* (narrow red boxes), 102 bp repeats or partial copies thereof (filled boxes), 103 bp repeats or partial copies of (narrow grey boxes), and a putative origin of replication is labelled *ori* (grey box). The extent of the sequences that are similar in strain A198 and strain C305 are indicated by dotted lines.

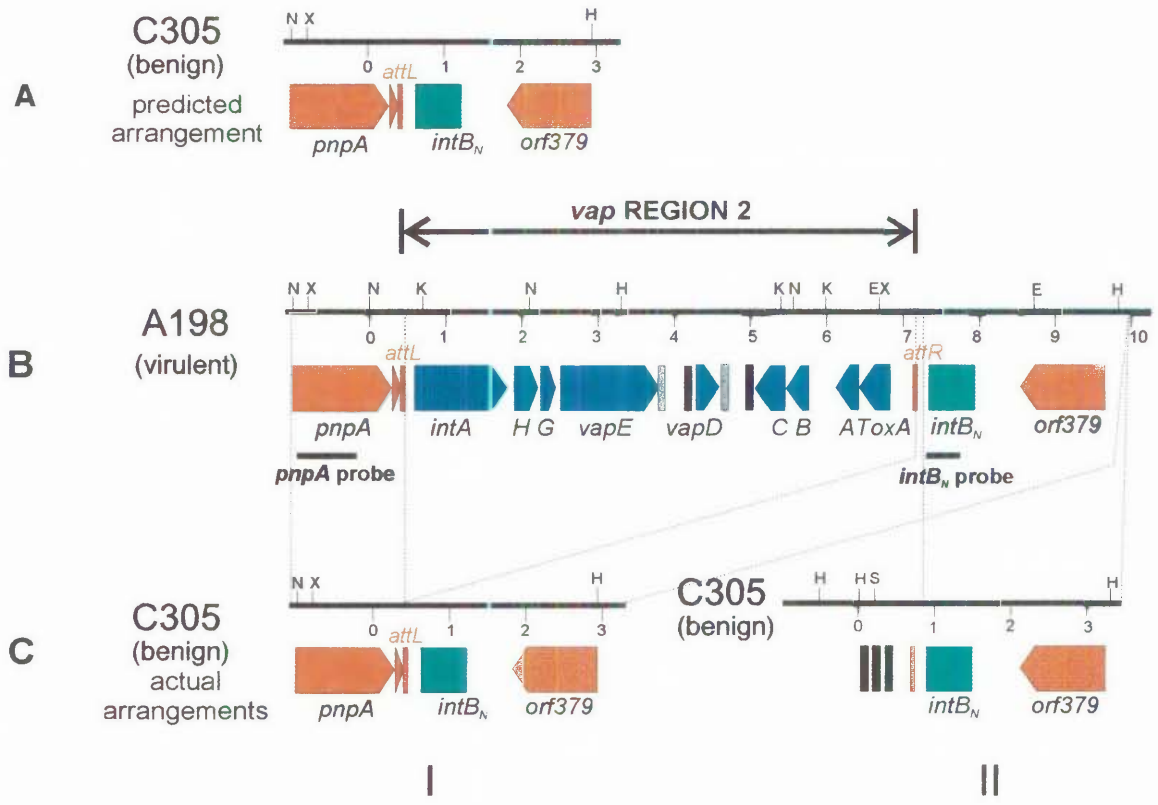



**Figure 1.12:** Alignment of restriction maps of the predicted arrangement of *pnpA* (**red**) and *intB*$_N$(**green**) in benign *D. nodosus* strain C305 that does not contain the *vap* element (**A**) with *pnpA*, *vap* region 2 (**blue**) and *intB*$_N$ of *D. nodosus* virulent strain A198 (**B**), and the actual arrangements (**I** and **II**) of *pnpA* and *intB*$_N$ genes in *D. nodosus* strain C305 (**C**) is shown. The *pnpA* and *intB* probe regions are indicated by bold lines below the corresponding sequences in *D. nodosus* strain A198. The numbers indicate fragment sizes in kb. Restriction enzyme sites shown are *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Nru*I (N), *Sac*I (S) and *Xho*I (X). Open reading frames are indicated by shaded arrows and are classified as follows: belonging to the *intB* element (**green**), belonging to the *vap* region (**blue**), genes that are not part of an integrated element (**red**), including *tRNA* genes (**red triangles**). Repeated sequences are as follows: attachment sites, *attL* or *attR* (narrow red boxes), 102 bp repeats or partial copies thereof (filled boxes), 103 bp repeats or partial copies of (narrow grey boxes), and a putative origin of replication is labelled *ori* (grey box). The extent of the sequences that are similar in strain A198 and strain C305 are indicated by dotted lines.

region 1 in *D. nodosus* strain A198 (Bloomfield *et al.*, 1997). Downstream from *askA*, a *tRNA-ser*$_{GCU}$ gene with 97.8% nt identity to the *tRNA-ser*$_{GCU}$ located at the left-hand end of *vap* region 1 of *D. nodosus* strain A198 was identified. In addition, an orf of 64aa, designated *glpA* was identified between *askA* and the *tRNA-ser*$_{GCU}$ (Whittle *et al.*, unpublished) (Figure 1.11C). GlpA has 74% aa identity to RsmA from the plant pathogen *Erwinia carotovora* (Cui *et al.*, 1995). In *E. carotovora* RsmA acts as a global repressor of virulence, and GlpA may have a similar role in *D. nodosus* (Chapter 7).

200 bp to the right of the *tRNA-ser*$_{GCU}$ an orf designated *intC* was identified and shown to have similarity to several bacterial integrases. IntC has 54.2% aa identity to IntA from *vap* region 1 of *D. nodosus* strain A198 and 40.4% identity with the integrase from retronphage ϕR73 (Sun *et al.*, 1991). Downstream from *intC* two orfs designated *orf242* and *orf171* were identified, and have no similarity to any sequences in the GenBank databases. Adjacent to *orf171* are two small orfs designated *vapG"* and *vapH"*, whose predicted protein products have 55.2% aa identity with VapG and 56.2% aa identity with VapH respectively, from *vap* region 1 of strain A198.

Downstream from *vapH"* is an insertion sequence with 98.3% identity over 1649 nt with the previously-identified IS*1253* on the *vap* plasmid in *D. nodosus* strain AC3577 (Billington *et al.*, 1996b), which consists of 2 orfs, *orf374* and *orf108*. In C305 the IS*1253* element appears to be non-functional due to two frameshift mutations (Bloomfield, 1997). Immediately following the IS*1253* element there is 100 bp with 98% identity to part of the coding region from the *intB* gene (positions 410-500) from the right end of region 3 in strain A198. Since this partial copy of the *intB* gene is derived from the middle of the *intB* gene sequence, it has been designated *intB*$_M$. *intB*$_M$ is followed by 344 bp of sequence with no similarity to other sequences from *D. nodosus* or sequences from the databases. After this point, similarity to the *intB* element from strain A198 resumes from the beginning of the *regA* gene, though the Shine-Dalgarno sequence and start codon for *regA* are absent

(Bloomfield, 1997). Collectively, these results suggested that a second integrated element, called the *intC* element, is located between *askA* and *intB* in *D. nodosus* strain C305. This hypothesis was the subject of further investigations that are discussed in Chapter 7.

### 1.6.8.2 $intB_N$ is integrated adjacent to *pnpA* in *D. nodosus* benign strain C305

Two additional lambda clones that contained copies of the *intB* gene sequence on 3.3 kb (λGB300.1) and 4.2 kb (λBA1) *Hin*dIII fragments respectively, were also isolated from the C305 genomic DNA library (Bloomfield *et al*, 1997; B. Aspbury, unpublished; Shaw, 1997].

The 3.3 kb *Hin*dIII fragment hybridising to the *intB* probe (Figure 1.12C, II) was found to contain three copies of the 102 bp repeats found within the *vap* regions of strain A198; 10 bp from the *attR* identified at the right of *vap* region 2 in strain A198; and a copy of $intB_N$, encoding the first 164 aa from the N-terminal of the 403 aa *intB* gene, and is almost identical to $intB_N$ from strain A198. Sequences downstream from $intB_N$ in both strains are almost identical for as far as the sequences have been determined (Bloomfield, 1997; Bloomfield *et al.*, 1997) (Figure 1.12C, II). In strain C305, *orf379*, which has no similarity any sequences in databases is found downstream from $intB_N$.(Bloomfield *et al.*, 1997). This copy of $intB_N$ that was isolated on a 3.3 kb *Hin*dIII fragment is no longer detectable in Southern blot analysis using *intB* to probe DNA from current laboratory stocks of *D. nodosus* strain C305 and instead *intB* probes hybridise to a 4.2 kb *Hin*dIII fragment.

The 4.2 kb *Hin*dIII fragment that hybridised to sequences that flank the left- (*pnpA*) and right-hand ends of *vap* region 2 ($intE_N$) of strain A198 (Figure 1.12, I) was isolated (B. Aspbury, unpublished), and a 1.0 kb *Cla*I/*Hpa*I fragment from it sequenced (Shaw, 1997). This experiment showed that the 4.2 kb *Hin*dIII fragment detected in current laboratory stocks of C305 contains *pnpA*, $tRNA\text{-}ser_{GGA}$ and $intB_N$ in consecutive order, as

hypothesised (Figure 1.12C, I).

These results suggest that a sequence previously separating the *pnpA* gene and $intB_N$ sequences was lost from, or moved in the original laboratory strain of C305 (herein called C3051) to generate the current laboratory strain of C305 (herein called C3052). Thus, at the time the library was made, genomic DNA had been isolated from of a mixed population of C3051 and C3052 cells. This hypothesis is supported by Southern blot analyses done in our laboratory which show that prior to 1994 *intB* gene probes hybridised to 6.1 kb and 3.3 kb *Hin*dIII fragments, whilst post 1994 and currently strain C305 hybridises to 6.1 kb and 4.2 kb *Hin*dIII fragments. Regions to the left of the 3.3 kb *Hin*dIII fragment containing this copy of $intB_N$ have been the subject of further investigations that will be discussed in Chapter 6.

### 1.6.9 Identification of native bacteriophage, DinoH1

The possibility that bacteriophages may have been the mobile genetic elements that carried the *vap* (Cheetham *et al.*, 1995a), *vrl* (Haring *et al.*, 1995) and *intB* elements (Bloomfield *et al.*, 1997) into the *D. nodosus* genome prompted attempts to induce bacteriophages from *D. nodosus* (Bloomfield, 1997). *D. nodosus* strains A198, AC3577, C305, B1006, G1220, H1204, H1215 were treated with mitomycin C and ultraviolet light in an effort to induce prophage excision from these putative hosts.

Treatment of virulent strain H1215 with UV light was successful in inducing a bacteriophage, designated DinoH1. Electron microscopy revealed that DinoH1 has a long non-contractile tail with a claw-like base plate and an icosahedral head, and analysis of the phage genome indicated that it consists of a 43 kb linear dsDNA molecule, and hence has been classified as belonging to the family Siphoviridae (Figure 1.13) (Bloomfield, Katz & Cheetham, unpublished; Murphy *et al.*, 1995).

**Figure 1.13**:

Electron micrograph of native *D. nodosus* bacteriophage, DinoH1 stained with phosphotungstic acid (electron micrograph provided by Garry Bloomfield). Scale bar = 50 nm.

Southern blot analysis showed that the DinoH1 genome does not contain *intA*, *intB*, or the *vrl* region, thus the prophage is a different integrated genetic element. Only part of the phage genome is present in strains A198, AC3577 and H1204. In addition, it was found that in strain A198 DinoH1 is integrated next to the *vrl* region, and hence suggests that DinoH1 may integrate into a similar site as the *vrl* and/or may have been involved in the acquisition of the *vrl* region by *D. nodosus* (Bloomfield *et al.*, unpublished). It is at this stage unknown whether DinoH1 has any role in virulence, though it seems unlikely given that virulent strain A198 contains the same fragments of DinoH1 as benign strains C305 and H1204 (Bloomfield *et al.*, unpublished).

Partial sequence analysis of DinoH1 failed to identify any similarity to genes encoded by characterised bacteriophages, however since most phages are specific for the bacterial host species or genus (Ackermann & DuBow, 1987), and as DinoH1 is the first bacteriophage identified from the genus *Dichelobacter*, a lack of similarity is not entirely unexpected (Bloomfield *et al.*, unpublished).

The host-range of *D. nodosus* was investigated by infecting several strains of *D. nodosus* with DinoH1. These experiments provided evidence that DinoH1 was at least

able to attach to all strains of *D. nodosus* tested, however it was unclear whether the phage genome was able to enter the cell. After propagation of infected cells on plate media, phage DNA fragments acquired post-infection, were no longer detectable by Southern hybridisation indicating that Dino1 did not stably lysogenise the host bacterium. An inability to stably lysogenise host cells may be attributable to inappropriate infection or growth conditions, or the expression of bacteriophage immunity regions previously present in recipient genomes (Bloomfield *et al.*, unpublished).

A preliminary restriction map of the phage genome was constructed (Bloomfield *et al.*, unpublished), which will be useful for further work which will aim to identify the site of DinoH1 integration in the H1215 genome, and to further characterise DinoH1 in an effort to determine whether DinoH1 has a role in virulence, or is involved with the acquisition of other extrachromosomal elements.

## 1.7 Project Aims

### 1.7.1 Characterisation and modification of native *D. nodosus* plasmid, pDN1 for the development of a transformation system

Although a number of genes encoding potential virulence determinants have been identified in *D. nodosus*, there is no transformation system for *D. nodosus*, which precludes direct testing of the role of these and other genes in virulence. The development of a transformation system has been impeded by the lack of native bacteriophages and plasmids, and attempts to develop a transformation system using plasmids derived from other bacteria have so far been unsuccessful. A native *D. nodosus* plasmid pDN1 was identified in our laboratory. Hence it was the aim of this work to isolate, sequence, characterise, and modify pDN1 such that it would be useful as a cloning vector for *D. nodosus*. Subsequently, the intention was to use these pDN1 derivatives in

transformation experiments with *D. nodosus*, and develop a transformation system so that the role of *vap* and other virulence associated genes in virulence could be determined directly.

### 1.7.2 Northern blot analysis of the *vap* genes in *D. nodosus*

In the absence of a transformation system for *D. nodosus*, one is confined to using more indirect methods in order to determine whether the *vap* genes of *D. nodosus* have a role in virulence, thus previous investigations have concentrated on analyses at the DNA level. Southern blot experiments indicated that *(i)* multiple copies of the *vap* genes are not required for virulence; *(ii)* that the arrangement of the *vap* genes was similar in benign and virulent strains of *D. nodosus*; and *(iii)* that the presence or absence of *vap* genes alone, did not determine whether a given strain exhibited a benign or virulent phenotype. Previous work showed that *vapD* is expressed in all strains of *D. nodosus* that contain *vapD*. In this study, Northern blot experiments were undertaken to determine whether the other *vap* genes are expressed differentially at the RNA level in the virulent and/or benign strains in which they are present.

### 1.7.3 Further characterisation of the *intB* element in *D. nodosus* strain A198

It was previously proposed that a genetic element, called the *intB* element was integrated into the *attR* site of *vap* region 3 of *D. nodosus* virulent strain A198 (Bloomfield *et al.*, 1997). It was unknown whether sequences adjacent to *intB*, *regA*, and *gepA* corresponded to part of an integrated element or otherwise. In an effort to further characterise the *intB* element, and to determine whether sequences adjacent to *intB*, *regA* and *gepA* corresponded to part of an integrated element, chromosome walking and sequencing to the right of *gepC* was undertaken. In addition, the prevalence and integrity of *intB* element sequences in seventeen different strains of *D. nodosus* was determined.

### 1.7.4 Identification and characterisation of part of an *intD* element from *D. nodosus* strain C305

A sequence previously separating the *pnpA* gene and $intB_N$ was lost from, or moved position in, the original laboratory strain of C305 (C3051) to generate the current laboratory strain of C305 (C3052). A lambda clone that contained 3.7 kb of sequences upstream of $intB_N$ in strain C3051 had been isolated previously. In order to determine whether these sequences were lost or had moved in strain C3052, and in an effort to investigate the nature and origin of the sequences that separated *pnpA* and $intB_N$ in strain C3051, a sequence of 3.7 kb immediately upstream of $intB_N$ from the lambda clone λGB300 (Bloomfield, 1997) was to be determined and analysed. The prevalence of these sequences in the seventeen different strains of *D. nodosus,* including benign, intermediate and virulent isolates was also to be determined.

### 1.7.5 Characterisation of an *intC* element in *D. nodosus*

The presence of a new genetic element in *D. nodosus* strain C305, designated the *intC* element was postulated previously (Bloomfield, 1997). In this work, Southern blot analyses and PCR experiments were undertaken in order to determine whether the genes *intC*, *orf242*, *orf171*, *vapG"*, *vapH"* are present in other strains of *D. nodosus* and if so, whether genes *intC*, *orf242*, *orf171*, *vapG"*, *vapH"* are clustered together as they are in *D. nodosus* strain C305, and therefore constitute part of a single genetic element or otherwise. In addition, in order to determine whether there is a correlation between the presence of certain genetic elements in *D. nodosus* and virulence, the integration sites for the *vap*, *intB* and *intC* elements were investigated in seventeen strains of *D. nodosus*.