MiniReview

Virulence regions and virulence factors of the ovine footrot pathogen, *Dichelobacter nodosus*

Stephen J. Billington 1, Joanne L. Johnston, Julian I. Rood *

Department of Microbiology, Monash University, Clayton, Victoria 3168, Australia

Received 21 June 1996; revised 9 September 1996; accepted 11 September 1996

Abstract

Ovine footrot is a debilitating and highly infectious disease that is primarily caused by the Gram-negative, anaerobic bacterium *Dichelobacter nodosus*. The major antigens implicated in virulence are the type IV fimbriae and extracellular proteases. The fimbriae show sequence and structural similarity to other type IV fimbriae, this similarity extends to genes that are involved in fimbrial biogenesis. Several acidic and basic extracellular serine proteases are produced by both virulent and benign isolates of *D. nodosus*. Subtle functional differences in these proteases appear to be important in virulence. In addition, there are two chromosomal regions that have a genotypic association with virulence. The partially duplicated and rearranged *vap* regions appear to have arisen from the insertion of a plasmid into a tRNA gene via an integrase-mediated site-specific insertion event. The 27 kb *vrl* region has several genes often found on bacteriophages and has inserted into an *ssrA* gene that may have a regulatory role in the cell. The determination of the precise role that each of these genes and gene regions has in virulence awaits the development of methods for the genetic analysis and manipulation of *D. nodosus*.

Keywords: Footrot; Virulence; Fimbria; Protease; *vap* region; *vrl* region; *Dichelobacter nodosus*

1. Introduction

*Dichelobacter nodosus* is the essential causative agent of ovine footrot, a disease that is of major economic significance in temperate climates. The disease is characterised by a mixed bacterial infection that results in the creation of an anaerobic environment and the growth of the aerotolerant anaerobe *D. nodosus* in the lesion. The resultant infection leads to the separation of the horn of the hoof from the underlying soft tissue, which results in lameness and loss of body condition [1]. Virulent, intermediate and benign footrot are three generally recognised forms of the disease which vary in severity depending on the nature of the causative *D. nodosus* isolate. *D. nodosus* is a relatively slow growing Gram-negative rod. The organism was previously known as 'Bacteroides' nodosus but comparative sequence analysis of its 16S rRNA revealed that it belonged in the gamma subgroup of the Proteobacteria and was more closely related to genera such as *Escherichia* and *Pseudomonas* than it was to members of the genus *Bacteroides* [2,3]. It was therefore placed into the new genus *Dichelobacter*, which means 'rod of the cloven hoof', in the family Cardiobacteriaceae.
Class specific genes

Class I
aroA  fimA  fimB  clpB

Class II
aroA  fimA  fimC  fimD  fimZ  clpB

Fig. 1. Genetic arrangement of the fimbrial gene regions from class I and class II isolates of D. nodosus. The extent of the open reading frames is shown by black arrows. Class specific genes are indicated.

Although there has been a great deal of molecular work done on D. nodosus in recent years, most of these studies have been carried out in Escherichia coli and Pseudomonas aeruginosa since there are no defined or recombinant methods of genetic analysis available for D. nodosus. The objective of this review is to provide an overview of these molecular studies, with an emphasis on virulence-associated gene regions and virulence factors.

2. Polar type IV fimbriae

The fimbriae of D. nodosus are long proteinaceous appendages composed of polymers of a single pilin subunit that is encoded by the fimA gene. They are highly immunogenic, with agglutination reactions involving the pilus antigen providing the basis for the classification of D. nodosus isolates into nine major serogroups, designated A–I [4]. Vaccination with whole cells or purified fimbriae protects against disease, but this protection is serogroup specific [5]. Sequencing of the genes of fimbrial subunits from the nine major serogroups has allowed the division of D. nodosus isolates into two major classes based on structural variation within the FimA protein and the genetic organisation of the fimbrial gene region [6,7]. Class I isolates (serogroups A–C, E–G and I) contain an additional gene, fimB, downstream of the fimbrial subunit gene. By contrast, strains in class II (serogroups D and H) possess three additional genes, fimC, fimD and fimZ, adjacent to fimA (Fig. 1).

The fimB gene encodes a potential 29.5 kDa membrane protein, has sequence similarity to TraX, a protein required for N\(^\circ\)-acetylation of the pilin subunit of the conjugative F pilus [8]. FimC may be involved in the acetylation of FimA subunits in class II strains, however, classical type IV fimbrial subunits from other bacteria are all N-methylated, not N-acetylated [9]. The third class II specific protein, FimZ, has 50\% identity to the FimA proteins of class II strains [6]. FimZ has a typical type IV signal sequence and is thought to represent a redundant fimbrial subunit. However, FimZ monomers are not likely to be assembled into fimbriae in D. nodosus, since they are not assembly competent when expressed in P. aeruginosa (Johnston, J.L., Billington, S.J., Mattick, J.S. and Rood, J.I., unpublished results).

In both classes, fimA is transcribed from rpoN-dependent promoters [6]. The class-specific genes appear to be cotranscribed with fimA on low level transcripts which read through the fimA transcriptional terminator to the end of the class-specific region. This region is defined by the end of the clpB gene, which encodes the regulatory subunit of an ATP-dependent protease [6] (Fig. 1). The aroA gene, located upstream of fimA, defines the other end of the class-specific gene regions (Fig. 1).

The FimA subunits of class I and class II strains are highly homologous over the amino-terminal one-third of the protein [7]. This region contains a short, positively charged leader sequence and a hydrophobic domain that is highly conserved in the pilin subunits of all type IV fimbriate bacteria. Along with the presence of an N-methylphenylalanine residue as the N-terminal amino acid, a polar distribution, and an association with twitching motility, these characteristics provide the basis for the designation of the fimbriae into the type IV or N-MePhe group [9]. The class-specific fimbrial subunits diverge over the
The fimO protein belongs to a family of proteins thought to reside in the cytoplasmic membrane [12]. EpsE, a *Vibrio cholerae* homologue of FimN, is required for protein secretion and has been shown to interact with EpsL to form a complex that is associated with the cytoplasmic membrane [13]. It is likely that FimN may interact with FimO or an EpsL homologue in a similar manner in *D. nodosus*. Although FimN and FimO are thought to be functionally analogous to their *P. aeruginosa* homologues, PilB and PilC, complementation experiments in *P. aeruginosa* suggest that they lack the structural domains which are required for interaction in the *P. aeruginosa* fimbrial biogenesis system (Johnston, J.L., Billington, S.J., Haring, V. and Rood, J.I., unpublished results).

The fimNOP gene region has a similar genetic organisation to the equivalent regions in *P. aeruginosa* and *Aeromonas hydrophila* (Fig. 2) except that in *D. nodosus* the fimbrial subunit gene is located at a different genomic location with respect to the accessory genes [14]. The fimN, fimO and fimP genes appear to comprise an operon that is transcribed from a σ70 type promoter located upstream of ORF197, a putative gene thought to be involved in tryptophan biosynthesis (Fig. 2). In *P. aeruginosa*, the pilBCD genes appear to have their own promoters, although pilIC and pilID may be arc cotranscribed. ORF197, which is located downstream of fimP, is also cotranscribed with fimNOP. In *Neisseria gonorrhoeae*, pilD and ORFX, a homologue of ORF197, are transcriptionally linked. This conservation of the organization of the fimbrial gene region provides evidence for horizontal gene transfer between these bacterial species [11].

### 3. Extracellular proteases

Biochemical studies have shown that isolates of *D. nodosus* that cause virulent footrot produce four extracellular acidic serine protease isoenzymes (V1–V3 and V5, with pI values of 5.2–5.6) and a basic protease (BprV, pI 9.5) [15,16]. Five acidic proteases are produced by isolates that cause benign footrot but these proteases (B1–B5) may have different electrophoretic mobilities, are relatively more heat labile, and have decreased elastase activity on elastin agar...
The first protease gene from *D. nodosus* to be cloned and sequenced was the *bprV* gene (1809 bp) from strain A198, which encodes the 603 residue basic protease precursor [15]. Analysis of the deduced amino acid sequence revealed that it has recognisable amino-terminal pre- and pro-peptide domains and a carboxy-terminal extension. All of these domains are cleaved to form the mature BprV protease of 344 amino acid residues. The prepeptide domain consists of a potential 21 amino acid hydrophobic signal sequence whereas the propeptide and carboxy-terminal domains have 111 and 127 residues, respectively. Comparative analysis revealed that BprV has 18–38% amino acid sequence identity to members of the subtilisin protease family although it is 70 residues longer [15]. It also contains conserved residues that have been shown to be catalytic in other subtilisins. BprV is of similar length
and has significant sequence identity (49%) to the serine protease from _Xanthomonas campestris_. This similarity extends to the respective propeptides. The homologous gene, _bprB_, from the benign strain 305 has also been cloned and sequenced [18].

The sequence of the acidic protease V5 was determined by direct amino acid sequencing and shown to have 64% amino acid sequence identity to BprV and 53% identity to the _X. campestris_ serine protease [19]. It also is a member of the subtilisin family. The genes, _aprV5_ and _aprB5_, encoding the V5 protease from strain A198 and the homologous B5 enzyme from strain 305, respectively, were subsequently cloned and sequenced [16]. These genes have a similar structure to _bprV_ in that they encode comparable prepropeptides and carboxy-terminal extensions. Comparison of _aprV5_ and _aprB5_ reveals that they have 99% identity, resulting in only two different amino acids, one in the propeptide and one in the mature protein. Biochemical analysis of the other proteases indicates that the V1–V3 and B1–B4 acidic proteases are closely related and distinct from the V5 and B5 enzymes [19–21]. It therefore appears that there are three types of related extracellular proteases in _D. nodosus_, as typified by the V2, V5 and BprV proteases. Note that recent studies have shown that _aprV5_ is located approx. 1 kb upstream of _bprV_ (Riffkin, M. and Stewart, D.J., personal communication).

4. Virulence-associated gene regions

4.1. Isolation and relationship to virulence

The previous sections of this review have dealt with virulence factors identified by their phenotypic association with virulence. The remainder of this review will deal with loci identified by their genotypic association with virulence in _D. nodosus_. As such these loci do not necessarily represent true virulence factors but gene regions that have an association with virulence and that can be used as an indicator of virulence. These loci were identified from comparative hybridization studies carried out on plasmid libraries from the virulent isolate A198, using labelled chromosomal DNA from strain A198 and the prototype benign isolate C305 [22]. Two regions of DNA, which appeared to be preferentially associated with virulent isolates of _D. nodosus_, were identified and subsequently designated as the _vap_ (virulence-associated protein) regions [23] and the _vrl_ or virulence-related locus [24].

Hybridization analysis of over 800 _D. nodosus_ isolates has indicated the presence of _vap_ sequences in 95% of virulent/high intermediate isolates and 88% of intermediate strains but in only 38% of low intermediate/benign strains [22,25]. By contrast, _vrl_ sequences are present in 77% of virulent/high intermediate strains, but in only 13 and 7% of intermediate and low intermediate/benign strains, respectively [22,25]. Therefore, while the _vap_ sequences are present in the majority of virulent strains, the presence of _vrl_ is a more likely indication that a _D. nodosus_ isolate is virulent. Other workers have repeated the initial studies but each of the virulence-associated probes identified [26] have been shown to be part of either the _vap_ or _vrl_ loci [25], suggesting that these loci represent the major differences between the genomes of virulent and benign _D. nodosus_ strains.

4.2. The _vap_ gene regions

The _vap_ regions are virulence-associated sequences present in single or multiple copies in the genomes of the majority of _D. nodosus_ isolates [22]. In the strain A198 chromosome there are three copies of the _vap_ sequence, designated as _vap_ regions 1–3 [22,27]. Most _vap_-related studies have revolved around _vap_ regions 1 and 3, which are closely linked on the A198 chromosome [27]. Using a series of probes from across _vap_ regions 1 and 3 it was shown that while there were small regions of DNA within the locus that were not virulence associated, these two regions did in fact encompass a single virulence associated locus of 11.8 kb [28].

Nucleotide sequencing of the entire _vap_ region 1/3 indicated that it had a genetic organization similar to that of prophages [28]. At the extreme left end of the locus is a putative integrase gene, _intA_, with similarity to a number of bacteriophage integrase genes. In addition, the non-virulence associated sequences to the left of _intA_ show similarity to tRNA-ser genes, suggesting that the _vap_ sequences have integrated into the 3' end of a tRNA gene, a favoured site.
for the integration of bacteriophages. To complete the similarity to prophages, vap region 1/3 is bound by 19 bp repeats with 89% identity, similar to the duplicated attachment (att) sequences of prophages [28].

In addition to intA, 15 ORFs have been identified within vap region 1/3, namely vapA-I, vapA', vapG', vapE' and orf118 [23,28] (Fig. 3). It appears that duplication and rearrangements of a portion of vap region 1 have resulted in the generation of variant copies of vapA, vapG and vapE, designated as vapA', vapG' and vapE', respectively [28], and a second identical copy of vapD within vap region 3 [27]. Despite the structural similarity between vap region 1/3 and prophage genomes, many of the vap-encoded genes have similarity to genes encoded on bacterial plasmids. For example, VapD has strong similarity to the products of ORF5 from the N. gonorrhoeae cryptic plasmid [22] and an ORF from an Actinobacillus actinomycetemcomitans rolling circle plasmid [27]. VapD has been shown to be expressed by D. nodosus isolates that carry vap sequences, but antibodies to VapD do not react with its N. gonorrhoeae homologue [23]. In addition, the nucleotide sequence of vapD shows similarity to a sequence on pTD1, a cryptic plasmid from Treponema denticola. VapB and VapC show similarity to putative products encoded in the trbH region of the F plasmid [23], and putative plasmid maintenance proteins, VagC and VagD [29], encoded by the Salmonella dublin virulence plasmid (Cheetham, B.F. and Whittle, G., personal communication). The VapE and VapE' proteins have similarity to the product of ORF2 from the cyanobacterial plasmid pMA1.

Our recent database searches have revealed intriguing similarity between ORF118 and the killer protein HigB [30] which is associated with the plasmid Rts1 killer gene system. Interestingly, VapA, VapA'
and VapI, which all have amino acid sequence similarity (Cheetham, B.F., personal communication), all show similarity to the antidote protein of this system, HigA [30]. Examination of the sequence upstream of vapI indicates the presence of a second ORF encoding a protein with similarity to HigB (Katz, M.E. and Cheetham, B.F., personal communication). Therefore, it appears that vap region 1/3 carries at least two complete copies of a plasmid maintenance function.

The similarity of vap-encoded products to plasmid-encoded proteins, in addition to the presence of a putative origin of replication within vap region 1/3 [28], suggests that the vap sequences may have evolved from the site-specific insertion of an integrative plasmid rather than a bacteriophage. This hypothesis has been given credence by the recent discovery that a region of DNA that is similar to vap region 1/3 can replicate as a plasmid in D. nodosus strain AC3577 [31]. This plasmid, pJIR896 (Fig. 3), appears to be a circular form of vap region 1/3 with a single att site, as would be expected prior to the insertion and duplication of the att site. However, pJIR896 lacks the duplicated vapA', vapF, vapG', vapE' genes and the second copy of vapD. In place of these regions is a 1.7 kb putative insertion sequence, IS1253, which has similarity to a family of unusual insertion sequences [31]. The presence of a putative IS element in the plasmid at the site where duplication is predicted to have occurred in vap region 1/3 implicates IS1253 in this duplication event and thus in the evolution of the chromosomal sequence. IS1253 is found in strain A198 at a site adjacent to the ompI locus, which is not associated with vap region 1/3.

It is postulated that vap region 1/3 originated from the insertion of an integrative plasmid such as pJIR896 into the 3' end of a tRNA gene. The subsequent deletion or transposition of IS1253 may have resulted in at least the initial duplication event, giving rise to the vapA'-vapE' region and the second copy of vapD (Fig. 3). The arrangement of these genes in vap regions 1 and 3 suggests that this duplication was followed by rearrangements, resulting in vap region 1/3 as present in strain A198 [28].

The evolution of vap region 2 is more difficult to deduce as less information is available. The order and arrangement of genes in vap region 2 appears similar to vap region 1 with the exception that vap region 2 has an additional 600bp sequence located between vapC and vapD [27]. It is clear that minor variations, including insertions or deletions, can occur between vap sequences of different strains [25]. It was originally proposed that vap region 2 arose from a duplication event [27]. However, it now seems likely that this region is derived from a separate integration event since vap region 2 and vap region 1/3 are integrated within the 3' end of different tRNA molecules [32].

4.3. The vrl locus

The vrl locus is a large, 27 kb region of virulence-associated DNA [24,33], which is present in a single copy, primarily in virulent D. nodosus isolates, including strain A198 [22]. The nucleotide sequences which lie to the left and right of vrl sequences in strain A198 are not adjacent to each other in the benign strain C305 [33], or in eight other strains we have tested (Billington, S.J. and Rood, J.I., unpublished data), but are separated by an unrelated sequence of approx. 3 kb (Fig. 4). These results suggest that vrl may not have arisen in virulent strains by a simple insertion event, but by an exchange mechanism involving the replacement of intervening sequences present in benign strains.

Examination of the sequence at the left junction of vrl provides some interesting clues as to the evolution of the vrl region. As for the vap sequences, there is some evidence that vrl may have arisen by a site-specific integrase-mediated event. Immediately upstream of the left end of vrl is a putative attachment site, attL, with similarity to the attachment sites of various prophages [33] (Fig. 4). This attachment site is found within the 3' end of a gene with similarity to the E. coli ssrA gene, which encodes a regulatory 10Sa RNA molecule [34-36]. Like tRNA genes the 3' ends of these genes are targets for the integration of bacteriophages. There is a sharp shift in %G+C content between the adjacent ssrA sequence (43%) and the left end of vrl (57%). Additional sequence data suggests that there are local regions within vrl which can have a G+C content as high as 75-80% (Billington, S.J., Huggins, A.S. and Rood, J.I., unpublished results), consistent with the presence of a greater number of restriction sites for enzymes with
GC-rich recognition sequences. This differential G+C content suggests that the vrl locus has originated in an organism other than *D. nodosus*, which has a G+C content of 45%.

While the *attL* site and the shift in G+C content at the left end of *vrl* favour a site-specific integration mechanism for the insertion of *vrl*, there is no corresponding attachment site at the right end of the sequence. This observation suggests that the end of the virulence-associated region may not be identical to the end of the integrated element. This conclusion is supported by the almost identical G+C content of *vrl* and non-*vrl* sequences at this end [33].

Sequence information from within the *vrl* locus has identified some ORFs with similarities to proteins encoded on bacteriophages or plasmids. Proteins that have similarity to the DEAH family of helicase-related proteins and DNA methylases as well as proteins involved in resistance to bacteriophage infection appear to be encoded within *vrl* (Billington, S.J., Huggins, A.S. and Rood, J.I., unpublished data). However, at this stage no ORFs have been identified that can be directly related to the virulence of *D. nodosus* isolates.

Hybridization analysis of over 800 *D. nodosus* isolates has failed to identify a single isolate that contains the *vrl* region but not the *vap* sequences [22,25]. Although it is possible that these results simply mirror the higher incidence of *vap*-containing strains, they may reflect a more fundamental interaction between the *vap* and *vrl* sequences. It is possible either that the *vap* integrase is required for *vrl* insertion, since sequencing of the *vrl* locus has not yet not identified a specific *vrl*-encoded integrase and there is sequence similarity between the *att* sites at the end of the *vrl* and *vap* loci, or that other *vap*-encoded products may be needed for either the integration or maintenance of *vrl*.

5. Future perspectives

Vaccines prepared against purified fimbriae will protect against infection by strains with homologous fimbrial subunits. There is also an excellent correlation between virulence and the presence of more heat-stable protease isoenzymes. These data strongly suggest that the type IV fimbriae and extracellular proteases are major factors involved in the virulence of *D. nodosus*. However, definitive proof has not been obtained because of the lack of a defined system for the genetic manipulation of this organism.

The role of the *vap* sequences in virulence has been difficult to assess for similar reasons. There is no clear evidence of similarity between the *vap* genes and virulence genes. It is possible that the association of the *vap* regions with virulent strains represents an example of coevolution, with the *vap* sequences playing no direct role in virulence but providing a genetic marker for virulent strains.

The correlation between virulence and the presence of *vrl* is much stronger than that between virulence and the *vap* regions [22-25]. However, the reason that the *vrl* region correlates with virulence remains unknown. The *vrl* locus may carry an as yet unidentified virulence factor or may have an indirect effect on genes involved in virulence. The later explanation appears more likely and a possible effector molecule is the *ssrA* gene. In *E. coli*, *ssrA* encodes a 10Sa RNA molecule whose structure has regulatory effects on several different genes. In particular, modification of the 3' end of *ssrA* leads to induction of an alternative lon protease [34], increased repression of the *lac* and *gal* operons and immunity to various λ derivatives [35,36]. In *D. nodosus*, the insertion of the *vrl* region also changes the 3’ end of the *ssrA* gene [33]. It is postulated that this alteration affects the expression of 10Sa RNA-regulated genes that are involved in virulence. This hypothesis will only be verified or disproved after the development of both transformation methods for *D. nodosus* and shuttle vectors suitable for the transfer of recombinant molecules between *D. nodosus* and *E. coli*. The identification of a plasmid in *D. nodosus* [31] hopefully represents the first important step in this process.

Acknowledgments

The work carried out in this laboratory has been supported by the Australian Government and Australian woolgrowers through research grants from the Australian Wool Research and Development Corporation and the Australian Research Council. J.J. is the recipient of an Australian Postgraduate...
Award. We thank Brian Cheetham, Margaret Katz and David Stewart for providing information prior to publication.

References


