Delineation of the virulence-related locus (vrl) of *Dichelobacter nodosus*

Volker Haring,† Stephen J. Billington, Cathy L. Wright,† Andrea S. Huggins, Margaret E. Katz‡ and Julian I. Rood

Author for correspondence: Julian I. Rood. Tel: +61 3 9905 4825. Fax: +61 3 9905 4811.
e-mail: JULIAN.ROOD@med.monash.edu.au

*Dichelobacter nodosus* is the primary pathogen implicated in ovine footrot. In this paper we have delineated a 27 kb locus, termed the virulence-related locus (vrl), that was essentially specific for virulent *D. nodosus* isolates. The precise ends of this locus were mapped and the sequences of the junction regions from the virulent strain A198 were compared to corresponding sequences from the benign isolate C305. The left end of the vrl locus was located in a sequence similar to that of the small stable 10Sa RNA molecule of *Escherichia coli*, next to a phage-attachment-site-like sequence, which indicated that the vrl locus might have arisen by the integration of a phage. However, no attachment-like sequence could be found at the right end of the vrl locus. In the chromosome of the benign strain the sequences bordering vrl were not contiguous but were separated by about 3 kb. It was concluded that the divergence of the benign and virulent strains at this locus was a multi-step process. Several potential ORFs were identified at the junction regions but only one ORF, encoding a 126 kDa protein, was expressed in a T7 expression system in *E. coli*.

**Keywords:** footrot, virulence, *Dichelobacter nodosus*, att site, small stable RNA

INTRODUCTION

Ovine footrot is an economically significant contagious disease, primarily affecting the feet of sheep (Stewart et al., 1984). It is a mixed bacterial infection with the main causative agent being *Dichelobacter nodosus*, a strictly anaerobic Gram-negative rod (Beveridge, 1941; Dewhirst et al., 1990). *D. nodosus* isolates are classified as virulent, intermediate, or benign depending on the severity of infection they can cause under optimal climatic and environmental conditions (Stewart, 1989; Stewart et al., 1986). However, little is known about the genetic differences between benign and virulent strains. It is likely that virulence is dependent upon several factors, with extracellular proteases and type 4 fimbriae being implicated in the disease (Depiazzi et al., 1986, 1991; Elleman, 1988; Moses & Yong, 1989; Thomas, 1962).

The study of potential virulence factors in *D. nodosus* is hampered by the lack of a genetic system which could be used to analyse the effect of defined mutations on virulence. There are currently no transformation or conjugation methods available for this organism. Therefore, the analysis of genetic elements potentially involved in virulence is primarily carried out by cloning and characterizing DNA regions unique to benign or virulent strains of *D. nodosus*.

Differential genetic analysis of such isolates led to the isolation of three DNA fragments that were present in the reference virulent strain A198, but not in the benign strain C305 (Katz et al., 1991). When used as DNA probes these fragments were found to be useful for the differential diagnosis of *D. nodosus* isolates (Katz et al., 1991). In dot-blot hybridization experiments one of the plasmids, pJIR318, hybridized with DNA from virtually all virulent and intermediate isolates as well as 33% of the benign strains. The other two fragments, carried in pJIR313 and pJIR314B, hybridized with virtually all virulent strains,
and with 36% of the intermediate isolates but with only two (6%) benign isolates (Katz et al., 1991). These results suggest that the genetic loci carried by pJIR313 and pJIR314B may be involved in the virulence of D. nodosus. Moreover, the almost identical results obtained with these plasmids suggest that they are closely linked.

In this paper we present data showing that the inserts in plasmids pJIR313 and pJIR314B were derived from a genetic locus that encompasses about 27 kb of essentially virulent-specific DNA. The exact point of divergence between virulent and benign sequences was determined. Our findings indicate that the transition in this locus between virulent and benign strains can not be explained by a simple insertion or deletion event.

**METHODS**

**Bacterial strains.** The D. nodosus strains A198 (virulent) and C305 (benign) were obtained from J. Egerton, University of Sydney, Australia. Other isolates were from the Regional Veterinary Laboratories at Hamilton, Victoria (HA207, HA212, HA233, HA240, HA304, HA646, HA659, HA733 and HA734), and Wagga Wagga, New South Wales (WW849-4A and WW1100-1A), or from J. Egerton (VCS1008, VCS1040, and VCS11172). D. nodosus isolates were grown in an atmosphere of 10% (v/v) H2 and 10% (v/v) CO2 in N2 at 37 °C on hoof agar (Thomas, 1958) or in TAS broth (Skerman, 1975, 1989). All *Escherichia coli* strains were derivatives of DH5α (Bethesda Research Laboratories) or K38 (Russel & Model, 1984) and were cultured in 2x YT medium (Miller, 1972) supplemented with ampicillin (100 µg ml⁻¹) and/or kanamycin (40 µg ml⁻¹). The D. nodosus A198 DNA library in λ GEM12 was obtained from E. Moses, Victorian Institute of Animal Science, Australia (Katz et al., 1994). The host for the recombinant λ bacteriophage, LE392 (Sambrook et al., 1989), was cultivated in LB broth supplemented with 1% (w/v) maltose (Sambrook et al., 1989).

**DNA isolation and general molecular techniques.** Genomic *D. nodosus* DNA was isolated as described by Anderson et al. (1984). Plasmid DNA from *E. coli* was prepared by previous methods (Birnboim & Dolj, 1979; Holmes & Quigley, 1981). Oligonucleotides were synthesized with an automatic DNA synthesizer (Applied Biosystems, model 381A) and dissolved in glass-distilled water. General molecular techniques were as described by Sambrook et al. (1989). Exclusive radiolabelling of plasmid-encoded proteins was carried out using the T7 RNA polymerase/promoter system (Tabor & Richardson, 1985). The host *E. coli* strain K38 harbouring pGPl and either the vector pTZ18U (Pharmacia), or recombinant plasmids, was used in this procedure.

**Hybridization analysis.** Dot-blots were prepared as previously described (Katz et al., 1991) except that *D. nodosus* cells were grown on hoof agar for 4-7 d and the lysed cells were applied to the nylon membranes (Hybond N+, Amersham) with the aid of a dot-blot apparatus (Schleicher & Schuell). The membranes were washed in 2x SSC, air-dried, and UV-crosslinked for 3 min. For Southern blots, genomic *D. nodosus* DNA (1–2 µg) was digested with the appropriate restriction endonuclease and fractionated by electrophoresis through a 0.8% (w/v) agarose gel. The DNA was denatured in 0.5 M NaOH, 1.5 M NaCl and transferred to a nylon membrane in 10 x SSC. Blots were hybridized with plasmid DNA, or DNA fragments, as before (Katz et al., 1991), or using a DIG-labelling kit (Boehringer) according to the manufacturer’s instructions. The λ library was screened as described previously (Katz et al., 1994).

**Construction of a D. nodosus C305 DNA library.** DNA libraries from *D. nodosus* C305 were constructed by digesting 1-5 µg of C305-derived DNA to completion with the appropriate restriction endonuclease, purifying the DNA by adsorption to glassmilk (Vogelstein & Gillespie, 1979), and ligating it with 1 µg of pUC18 digested with the same enzyme. The ligation mixture was either used directly to transform RbCl-treated *E. coli* cells (Hanahan, 1983) or, after an ethanol precipitation step, to transform glycerol-treated cells by electroporation.

**Nucleotide sequencing.** All sequencing reactions were done using either a T7 DNA polymerase sequencing kit (Pharmacia) or an Automated Sequencing System 373A using a PRISM cycle sequencing kit (Applied Biosystems). The reported sequences were obtained independently from both strands either by constructing and sequencing overlapping subclones or by using specifically designed oligonucleotides as primers in the sequencing reactions.

**RESULTS**

**Delineation of the vrl region.** Hybridization analysis of pJIR313 (1-3 kb insert) and pJIR314B (0-6 kb insert) revealed that both fragments were present in only a single copy on the A198 genome (Katz et al., 1991). Since preliminary sequence analysis showed that both plasmids contained incomplete ORFs it was initially postulated that these plasmids were derived from the same gene. However, chromosome walking experiments subsequently carried out using a λ GEM12 library showed that the fragments carried by pJIR313 and pJIR314B were separated by 9 kb of DNA, all of which had the same dot-blot hybridization specificity, with respect to virulent and benign *D. nodosus* isolates, as the original recombinant plasmids. Further chromosome walking experiments were carried out to determine the extent of this region, which was designated the virulence-related locus or the vrl region. One end of this locus was known, since pJIR314B was derived from a larger plasmid, pJIR314, and the adjacent sequences contained in pJIR314A hybridized to both virulent and benign isolates (Katz et al., 1991). The other end was identified after four separate overlapping λ clones were isolated (Fig. 1). When appropriate subclones of these recombinant phages were constructed and tested in dot-blot hybridization experiments, the other end of the vrl region was shown to be located within pJIR744 (Fig. 1). Therefore, dot-blot experiments showed that the vrl locus consisted of about 27 kb of essentially virulence-specific DNA located between the regions carried on pJIR744 and pJIR314A (Fig. 1).

**Fine mapping of the ends of the vrl region.** If it is assumed that the difference between the virulent and benign *D. nodosus* strains in this region is due to the insertion or deletion of vrl DNA, the sequences bordering each vrl end should be juxtaposed in benign strains. Southern blots revealed that pJIR314A, which contained the right end of vrl, hybridized to a 2 kb EcoRI fragment from the benign strain C305; this fragment was subsequently cloned. The resultant plasmid, pJIR787, showed...
Virulence-related locus of *Dichelobacter nodosus*

Fig. 1. Delineation of the *vrl* region. A restriction map of the *vrl* region, as deduced from the *λ* clones, is shown at the top. The relative positions of the inserts in pJIR313, pJIR314B and the isolated *λ* clones are indicated, as well as those subclones used to demonstrate the virulence specificity of this region (see below). The extent of the virulence-related locus is indicated by the open box at the bottom. Restriction sites for BamHI (B), EcoRI (E), EagI (G), HindIII (H), SstI (S) and XhoI (X) are indicated. The orientation of the map is consistent with the large ORF found at the right end of *vrl* being transcribed from left to right. The virulence specificity of subclones of the *vrl* region is shown in the bottom section. Dot-blot analysis of virulent (V) and benign (B) *D. nodosus* isolates was probed with labeled DNA of the plasmid indicated. The benign isolates tested were, from top to bottom: HA646, HA304, C305, WW849-4A, HA733, HA734, HA233 and VCS1040. The virulent isolates tested were A198, HA659, VCS1008, VCS1172, WW1100-1A, HA240, HA212 and HA207.

Fig. 2. Restriction maps of recombinant plasmids and the PCR fragment containing the *vrl* junctions. pJIR744 and pJIR314A are the initial isolates containing the left and right ends of *vrl*; pJIR880 is a subclone of pJIR744. Plasmid pJIR787 contains the region corresponding to the right end of *vrl* from the benign *D. nodosus* strain C305. The positions of the primers used to amplify the region corresponding to the left border are indicated, as is the resultant PCR fragment.

some restriction site similarity to pJIR314A (Fig. 2). Analysis indicated that the right junction between the virulence-related locus and the chromosome was located 300–500 nt from the right end of pJIR787. However, pJIR787 had no restriction site identity with pJIR744 (Fig. 2), the plasmid containing the left end of *vrl*. Moreover, the pJIR744 insert did not hybridize with pJIR787 DNA, indicating that pJIR787 contained the
sequences bordering the right end of vrl but not the left end. It was concluded that in strain C305 these bordering sequences were not contiguous.

Southern blots of C305 DNA probed with pJIR744 revealed hybridizing 2.7 kb EcoRI and 8 kb BamHI fragments (results not shown). Despite screening over 2000 recombinant colonies from both EcoRI and BamHI libraries of C305 DNA constructed in pUC18, no clones hybridizing to pJIR744 were isolated, which suggested that this region may not be stably maintained on a multicopy plasmid in E. coli.

Comparative Southern hybridization analysis of C305 and A198 DNA showed that the sequences hybridizing to pJIR744 were isolated, which suggested that this region may not be stably maintained on a multicopy plasmid in E. coli.

Fig. 3. Alignment of sequences at the left end of the vrl region (A198) with corresponding sequences from the benign isolate (C305). The point of divergence is indicated by an open arrowhead. Included in this alignment is the E. coli 10Sa RNA gene (EC ssrA). Nucleotides conserved in A198 and C305 are in capitals, as are the E. coli residues identical to the D. nodosus sequences. Sequences similar to att sites are shaded and regions of similarity to the pseudouridine loop of tRNAs are overlined by a bar. The end of the mature 10Sa RNA is marked by an arrow and the end of the precursor form by an asterisk. The 14 nt boxes following the att sites are boxed and the inverted repeats are overlined with arrows in the A198 sequence.

region in C305. The sequence for primer #1129, ACCGTAAGCGACATTAAACAG, was 129 nt from the left end of the pJIR880 insert and the sequence of primer #1169, CGCCCGACCCATTACCCCGTGA, was 322 nt to the left of the EcoRI site of pJIR314A (Fig. 2). PCR analysis using these primers on C305 DNA as template yielded a product of about 3.5 kb, as expected. This product hybridized to both pJIR880 and pJIR314A probes in Southern blots.

Sequence analysis of the junction regions

The exact endpoints of vrl were determined by sequencing the A198- and C305-derived DNA fragments corresponding to the left and right junctions and comparing the resulting sequences. For the left vrl junction, the pJIR880 insert was completely sequenced and found to be 1487 nt long. The 3.5 kb PCR fragment was used to obtain relevant sequence from the corresponding C305 region; in all 329 nt, starting from primer #1129, were obtained and compared to the sequence from pJIR880 (Fig. 3). The sequences were identical for the first 156 nt (pJIR880, nt 167–322; PCR fragment, nt 1–156) then the sequences diverged. Therefore, the left end of the vrl region was
At the right junction both the virulent and the benign sequences contained two incomplete ORFs which were convergent in pJIR314 (ORF277' and ORF1130) but had the same orientation and overlapped by 4 nt in pJIR787 (ORF172' and ORF240') (Fig. 5b). The ORFs from the non-vrl region, ORF172' and ORF277', were almost identical up to the junction site and continued past this site for 51 and 4 codons in pJIR314 and pJIR787, respectively (Fig. 4). Of the eight sequence differences in the common portions of pJIR314 and pJIR787, only three resulted in a change in the amino acid sequence. To find the start of the rightmost ORF within the vrl region, sequence analysis was continued. This putative gene, ORF1130, extended through pJIR589 into the adjacent clone, pJIR589 (Fig. 5). ORF1130 was preceded by a potential ribosome-binding sequence, AGGAGG, and another incomplete open reading frame, ORF42', which overlapped ORF1130 by 11 nt. In total, 4462 bp of the right end of the vrl region were sequenced. No sequences similar to E. coli promoters were found. None of the ORFs identified showed significant similarities with entries in the databases. However, examination of the amino acid sequence of the predicted ORF1130 product revealed similarity to the seven motifs associated with the DEAH superfamily of ATP-dependent helicase-related proteins (Gorbalevya et al., 1989).

The vrl region appears to have inserted into a 10Sa RNA gene

Comparison of the DNA sequences with entries in the GenBank and EMBL databases revealed significant similarities in the non-vrl region, ORF172' and ORF277', for 51 and 4 codons in pJIR314 and pJIR787, respectively (Fig. 4). Of the eight sequence differences in the common portions of pJIR314 and pJIR787, only three resulted in a change in the amino acid sequence. To find the start of the rightmost ORF within the vrl region, sequence analysis was continued. This putative gene, ORF1130, extended through pJIR589 into the adjacent clone, pJIR589 (Fig. 5). ORF1130 was preceded by a potential ribosome-binding sequence, AGGAGG, and another incomplete open reading frame, ORF42', which overlapped ORF1130 by 11 nt. In total, 4462 bp of the right end of the vrl region were sequenced. No sequences similar to E. coli promoters were found. None of the ORFs identified showed significant similarities with entries in the databases. However, examination of the amino acid sequence of the predicted ORF1130 product revealed similarity to the seven motifs associated with the DEAH superfamily of ATP-dependent helicase-related proteins (Gorbalevya et al., 1989).

The vrl region appears to have inserted into a 10Sa RNA gene

Comparison of the DNA sequences with entries in the GenBank and EMBL databases revealed significant simil-
Fig. 5. Comparison of genetic maps at the left (a) and right (b) ends of vrl and the corresponding regions in C305. The black boxes indicate the sequenced regions. The ORFs identified, and their orientation, are shown as open arrows. Incomplete ORFs are indicated by the broken arrows.

Similarity between the left junction and small stable 10Sa RNA sequences from *E. coli* (Chauhan & Apirion, 1989), *Alcaligenes eutrophus* (Brown et al., 1990), and *Mycobacterium tuberculosis* (Tyagi & Kinger, 1992). The similarity in all five sequences was optimal around the site where the precursor RNA is cleaved to form the mature 10Sa RNA (Chauhan & Apirion, 1989). Interestingly, this region showed strong similarity to the pseudouridine arm of tRNA genes from different organisms (Tyagi & Kinger, 1992) and to the integration site of several prophages, *attL* (Kirby et al., 1994). The similarity with the other 10Sa RNA genes extended 6 nt beyond the 3' end of the mature 10Sa RNA and the *D. nodosus* sequences were identical for another 4 nt (Fig. 3). In pJIR880, another sequence similar to *attL, attL2*, was found in the vrl region, some 300 nt downstream of the first site. Both *att* sites were followed by sequences that were identical in 13 out of 14 positions. The distance between these 14 nt boxes and the *att* sites was 16 nt for *attL* and 62 nt for *attL2*. In addition, the left end of *vrl* was marked by a 26 bp inverted repeat. None of these sequence motifs could be found at the right end of *vrl*. The only significant structures in the *vrl* portion of pJIR314 were two 19 bp inverted repeat sequences and two 21 bp direct repeats (Fig. 4). The position of the inverted repeat, about 50 bp downstream of the last *vrl* ORF, could indicate that it functions as transcription terminator.

Expression of vrl-encoded genes

T7-expression studies were carried out to determine if any of the ORFs were expressed in *E. coli*. For ORF1130, a 7·6 kb HindIII subclone of AR29 was cloned downstream of the T7 promoter in pTZ18U in both orientations. The resultant plasmids were pJIR903, where the T7 promoter was upstream of ORF1130, and pJIR902, which was in the opposite orientation. T7-directed expression of pJIR903 resulted in the specific labelling of a protein larger than 100 kDa, as would be expected from ORF1130 (Fig. 6). No other induced protein was observed using this plasmid or pJIR902. To examine the left end of *vrl* a 9·4 kb EcoRI fragment of AR82 was cloned and tested in a similar experiment. Several induced proteins of 12–40 kDa were observed. While some of these proteins correlated with the sizes of ORFs found in the pJIR880 sequence, none of these proteins were observed in T7 expression experiments using pGEM7zf(−) recombinants containing only the insert present in pJIR880 (data not shown).
hybridization analysis showed that the gene regions they
pJIR313 and pJIR314B were derived. Subsequent
were overexpressed from the plasmids indicated using the T7
were separated by electrophoresis on a denaturing 12
Fig. 6.

were part of the 27 kb
represented were separated by about 9 kb and that they
a repeated region, designated vap, in the A198 genome
pJIR313, pJIR314B and pJIR318, carrying gene regions
DISCUSSION
Differential hybridization experiments using the virulent
and benign D. nodosus strains, A198 and C305 respectively,
resulted in the isolation of three recombinant plasmids,
pJIR313, pJIR314B and pJIR318, carrying gene regions
which were present only in the virulent strain (Katz et al.,
1991). Other studies have shown that pJIR318 represents
a repeated region, designated nap, in the A198 genome
(Katz et al., 1992, 1994). In this study we have charac-
terized the chromosomal region from which the inserts of
pJIR313 and pJIR314B were derived. Subsequent
hybridization analysis showed that the gene regions they
represented were separated by about 9 kb and that they
were part of the 27 kb vrl locus. It seems likely that this
locus has been derived from the integration of a phage,
plasmid or transposon, as such elements are known to be
potential carriers of virulence determinants (Levin &
Eden, 1990). Unfortunately, the lack of a genetic exchange
system prevents the direct genetic analysis of vrl and the
determination of its precise involvement in virulence in
D. nodosus.

To test the hypothesis that vrl is an integrated genetic
element, we analysed the ends of the vrl region and the
‘integration site’ in a benign strain. This analysis was
somewhat hampered by the lack of isogenic strains, since
all D. nodosus strains are field isolates with unknown
evolutionary relationships. We chose the benign strain
C305 to analyse the integration site, as this was the strain
used in the isolation of the plasmids pJIR313 and
pJIR314B (Katz et al., 1991). It was anticipated that the
sequences bordering the vrl locus would be adjacent in
strain C305. However, our results showed that these
bordering sequences were separated by about 3 kb of
DNA found in both virulent and benign strains. This
finding can either be explained by a displacement mech-
anism by which one region displaced the other in the
insertion process, or by postulating that both regions
inserted independently at the same site. The latter notion
is supported by the identification, at the left end of vrl, of
an attL site with high similarity to a site found in the gene
for the small stable 10Sa RNA, ssrA, in E. coli (Kirby et al.,
1994). Moreover, the upstream region shares significant
similarity with the E. coli ssrA gene (Fig. 3). This suggests
that both the vrl region and the insert in C305 may be
derived from the integration of a phage in a small stable
RNA gene similar to the situation found in E. coli (Kirby
et al., 1994). The lack of the concomitant attachment site,
attR, at the right end of vrl implied that the end of the vrl
region was not identical with the end of the putative
prophage. This conclusion is supported by the almost
identical mol% G+C content on both sides of the
junction in A198 and C305. The second attachment site
could be outside the sequenced region or could have been
deleted since the insertion event took place.

The function of the small stable RNA is not well
understood. It is implicated in gene regulation (Kirby
et al., 1994) and in interaction with DNA-binding proteins
(Retallack et al., 1994). The ssrA transcript is processed
within the attL site (Fig. 3) to give the mature RNA
(Chauhan & Apirion, 1989). Excision of the prophage
results in an altered 3' end of the transcript which is
sufficient in E. coli to affect its function (Kirby et al.,
1994; Retallack et al., 1994). Against this background it seems
possible that vrl does not itself encode a virulence function
but rather its insertion at ssrA activates such a function
located somewhere else on the chromosome. In this
context we noted that in A198 the attL site was followed
by an inverted repeat (Fig. 3) that could function as a
factor-independent transcription terminator. Such a
structure is also found near the end of the E. coli ssrA gene
(Kirby et al., 1994) but is absent after excision of the
prophage (Retallack et al., 1994) and is absent from C305
(Fig. 3).

The significance of the second attL-like sequence near the
left end of vrl is not clear. Although the sequence could be
fortuitous it could also be an indication that more complex
genetic processes lead to insertion of the vrl region in
A198, e.g. the integration of two phages either inde-
pendently or by one being carried by another at the time
of integration. The notion of two different genetic
elements at this site is strengthened by the significant
difference in the mol% G+C content of the sequence between attL and attL2 (51%) and the sequence to the right of attL2 (61%). The importance of the direct repeats following the attL sites in A198 (Fig. 3) is also not known. Similarly, within the att region of the P4 family of Esx. coli, conserved sites were identified (Kirby et al., 1994), which have varying distances from the att sequences. These sites were assumed to be protein-binding sites implicated in integration and/or excision of the phage or regulation of transcription (Kirby et al., 1994). Although there is no similarity between the A198 repeats and the conserved sites of the P4 family they could serve similar functions.

ACKNOWLEDGEMENTS

This work was generously supported by grants from the Australian Research Council and the Australian Wool Research and Development Corporation. A.S.H. is supported by an Australian Wool Industry Postgraduate Scholarship. We thank Pauline Howarth, Vivien Vasic and Khim Hoe for their expert technical assistance and Eric Moses for providing the λ library.

REFERENCES


Received 6 February 1995; revised 18 April 1995; accepted 4 May 1995.