The site-specific integration of genetic elements may modulate thermostable protease production, a virulence factor in *Dichelobacter nodosus*, the causative agent of ovine footrot

Gabrielle Whittle, Garry A. Bloomfield,† Margaret E. Katz and Brian F. Cheetham

Author for correspondence: Brian Cheetham. Tel: +61 2 6773 3394. Fax: +61 2 6773 3267. e-mail: bcheetha@metz.une.edu.au

**INTRODUCTION**

The Gram-negative anaerobe *Dichelobacter nodosus* is the causative agent of footrot in sheep. The authors have previously characterized two genetic elements, the *intA* (*vap*) and *intB* elements, which integrate into the genome of *D. nodosus*. In the virulent strain A198 there are two copies of the *intA* element. One copy is integrated into the 3' end of the *tRNA-ser*-GCU gene, close to the aspartokinase (*askA*) gene, and the second copy is integrated into the 3' end of the *tRNA-ser*-GGA gene, next to the polynucleotide phosphorylase (*pnpA*) gene. In this study, a new genetic element was identified in the benign strain C305, the *intC* element, integrated into the 3' end of the *tRNA-ser*-GCU gene, next to *askA*. The *intC* element was found in most *D. nodosus* strains, both benign and virulent, which were examined, and was integrated into *tRNA-ser*-GCU in most strains. Between the *askA* and *tRNA-ser*-GCU genes, a gene (designated *glpA*), was identified whose predicted protein product has very high amino acid identity with RsmA from the plant pathogen *Erwinia carotovora*. RsmA acts as a global repressor of pathogenicity in *E. carotovora*, by repressing the production of extracellular enzymes. In virulent strains of *D. nodosus* the *intA* element was found to be integrated next to *pnpA*, and either the *intA* or *intC* element was integrated next to *glpA*. By contrast, all but one of the benign strains had *intB* at one or both of these two positions, and the one exception had neither *intA*, *intB* nor *intC* at one position. The loss of the *intC* element from the virulent strain 1311 resulted in loss of thermostable protease activity, a virulence factor in *D. nodosus*. A model for virulence is proposed whereby integration of the *intA* and *intC* genetic elements modulates virulence by altering the expression of *glpA*, *pnpA*, *tRNA-ser*-GCU and *tRNA-ser*-GGA.

**Keywords:** footrot, virulence, pathogenicity, *vap*, pathogenicity island

---

**Molecular and Cellular Biology, School of Biological Sciences, University of New England, Armidale, NSW 2351, Australia**

†Present address: Roche Diagnostics Australia Pty Ltd, 31 Victoria Avenue, Castle Hill, NSW 2154, Australia.

Abbreviation: HSL, N-(3-oxohexanoyl)-L-homoserine lactone.

The GenBank accession number for the *intC* element in strain C305 beginning at the left-most EcoRI site (Fig. 1) is Y15939.
Multiple copies of the intA element, termed vap regions 1, 2, and 3, are present in the virulent D. nodosus strain A198 (Katz et al., 1994). Regions 1 and 3 form part of a contiguous 11.9 kb virulence-associated region (Fig. 1; Cheetham et al., 1995). Region 1 is integrated into a tRNA-ser<sup>GCU</sup> gene close to the aspartokinase (askA) gene, while region 2 is integrated into a different tRNA-ser gene, tRNA-<sup>GAG</sup> next to the polynucleotide phosphorylase (ppnA) gene (Bloomfield et al., 1997).

The genes which have been identified in the vap regions of the virulent D. nodosus strain A198 are designated intA, vap (virulence-associated protein) A–H, and toxA (Fig. 1; Katz et al., 1992; Cheetham et al., 1995; Bloomfield et al., 1997). The intA gene product is an intein belonging to the lambda family of site-specific recombinases (Cheetham et al., 1995). VapA and ToxA show amino acid similarity to the HigA and HigB proteins from the killer plasmid Rts1 (Tian et al., 1996), which encode a toxin and antidote molecule, respectively, and act to prevent loss of the plasmid Rts1 in Escherichia coli. VapA and ToxA may have a similar function in D. nodosus (Bloomfield et al., 1997). The vapB and vapC genes are similar to vagC and vagD from the virulence plasmid of Salmonella dublin. The products of these genes are believed to coordinate replication of the virulence plasmid with cell division (Pullinger & Lax, 1992). VapD and VapE have amino acid similarity to the products of plasmid-borne genes of unknown function from a variety of bacterial species (Katz et al., 1992; Cheetham et al., 1995). The remaining genes, vapF, vapG and vapH, have no similarity to previously identified genes.

A native plasmid which contains vap region 1 and an insertion sequence (IS1253) has been isolated from strain AC377 of D. nodosus, and may be the progenitor of the chromosomal vap sequences found in other strains (Billington et al., 1996). IS1253 is not found in the vap regions of strain A198, but there is a copy near the outer-membrane protein (omp) gene cluster, which is responsible for production of the major outer-membrane protein of D. nodosus (Moses et al., 1995).

A second genetic element, the intB element, which may be part of a prophage or a conjugative transposon, is integrated next to vap region 3 in strain A198 (Bloomfield et al., 1997). The end of this element contains an integrase gene, intB, a regulatory gene, regA (Fig. 1), and three genes of unknown function, gepA, gepB and gepC.

In this paper, we describe a new genetic element, the intC element, in D. nodosus. In strain C305, the intC element is integrated into tRNA-ser<sup>GCU</sup>, next to askA. We show that loss of this element from one virulent strain resulted in a concomitant loss of thermostable protease activity, a virulence factor in D. nodosus (Depiazzi et al., 1990). Mechanisms by which the loss of the intC element might result in the loss of thermostable protease activity are discussed.

**METHODS**

**General methods.** Methods for the growth of D. nodosus, preparation of genomic DNA, cloning and analysis of DNA, Southern blotting, DNA sequencing and DNA sequence analysis have been reported elsewhere (Katz et al., 1994; Bloomfield et al., 1997).

**Dichelobacter nodosus strains.** Strains A198, AC390, B1006, C305, D1172, G1220, H1204, H1215, 1469, 1493 and 3138 were obtained from Professor J. I. Rood, Monash University, Clayton, Victoria 3168, Australia, and strains 819, 1169, 1311 and 2483 from Dr J. Sercon, NSW Agriculture, Private Mail Bag, Yanco, NSW 2703, Australia.

**Preparation and screening of a library of genomic DNA from D. nodosus strain C305.** Genomic DNA from D. nodosus strain C305 was partially digested with the restriction endonuclease Sau3A1 and fragments were ligated into the XhoI half-site arms of bacteriophage lambdaGEM-12 (Promega), packaged according to the directions of the manufacturer (Promega), and used to infect the lambda-sensitive E. coli strain LE392 (Sambrook et al., 1989). The library was screened by plaque-lift hybridization, on nitrocellulose membranes, using digoxigenin-labelled probes and a chemiluminescent detection kit (Boehringer Mannheim), according to the manufacturer’s instructions. DNA was prepared from the lambda clones by the method of Ausubel et al. (1989).

**Oligonucleotide primers.** Primer sequences were: primer A (askA) TAGAAGCCACGACATGAGC; primer B (intC) AGCCGCATTGTCGCCACATAAA; and primer C (intB) ATATACGCTGGTTCGATAG. PCR conditions have been described previously (Bloomfield et al., 1997).

**Assay for protease thermostability.** The supernatant from liquid cultures of D. nodosus was assayed for proteolytic activity by measuring the release of a soluble blue dye from hide powder azure (Green, 1985). Thermostability was determined in triplicate by measuring the activity before and after heating the supernatant at 60°C for 10 min.

**RESULTS AND DISCUSSION**

**Isolation from strain C305 of sequences flanking vap regions 1 and 3 in strain A198.**

The benign strain C305 does not contain the genes vapA–D or toxA, suggesting that the intA element is not present in this strain. However, a HindIII–NruI fragment located close to the left-hand end of vap region 3 (Fig. 1) in strain A198 has been shown previously to hybridize to genomic DNA from strain C305 (Cheetham et al., 1995), and repeated sequences found immediately to the left of vapD in strain A198 are also present in the benign strain C305 (Katz et al., 1994).

To investigate the hypothesis that the vap regions of strain A198 resulted from integrations of the intA element into the tRNA-ser<sup>GCU</sup> gene located between the askA and intB genes (Fig. 1), we prepared a library of genomic DNA from strain C305, and screened it with probes 1, 3 and 4A (Fig. 1). Probe 1 contained part of the askA gene, flanking the left-hand end of vap region 1 in virulent strain A198; probe 4A contained part of the intB gene flanking the right-hand end of vap region 3 in strain A198; and probe 3 contained the HindIII–NruI fragment from the left-hand end of vap region 3 in strain
Integrated genetic elements modulate virulence

**Virulent Strain A198**

![Diagram of virulent strain A198]

**Benign Strain C305**

![Diagram of benign strain C305]

*Fig. 1.* vap regions 1, 2 and 3 of *D. nodosus* strain A198 compared with the *intC* element in *D. nodosus* strain C305. The numbers show the distance in kb from the left-most *Nru*I site in strains A198 or C305. Restriction 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). The major potential genes are indicated by open arrows. The *tRNA*-ser genes are shown by black triangles. Repeated sequences (Cheetham et al., 1995) are indicated as follows: 19 bp *att* sites (small open boxes), 103 bp repeats, or partial copies thereof (small shaded boxes), 102 bp repeats, or partial copies (small black boxes) and the putative origin of replication (large shaded box). The large open boxes show regions which are not found in strain C305.

A198. A single lambda clone to which probes 1 and 3 hybridized strongly, and probe 4A hybridized weakly, was isolated (Fig. 1). We predicted that the *askA* and *intB* genes would be close together in this clone. However, they were found to be separated by a 6·6 kb DNA segment. The DNA sequence of an 8035 bp region from this lambda clone was determined, and several ORFs were identified (Fig. 1).

**A new genetic element is integrated into tRNA-ser in strain C305**

At the left-hand end of the lambda clone from strain C305 (Fig. 1), we identified a partial ORF, designated *askA*, the DNA sequence of which is almost identical to that of the partially characterized ORF located upstream of the *tRNA*-ser<sub>GCU</sub> gene at the left-hand end of *vap* region 1 in *D. nodosus* strain A198 (Bloomfield et al., 1997). Downstream from the *askA* gene, we identified a 92 bp *tRNA*-ser gene which had 97·8% nucleotide identity to *tRNA*-ser<sub>GCU</sub> from the left-hand end of *vap* region 1 of *D. nodosus* strain A198 (Cheetham et al., 1995).

The predicted protein product from an ORF, designated *intC*, located 200 bp downstream from *tRNA*-ser<sub>GCU</sub>, was found to have 54·2% amino acid identity with IntA from *vap* region 1 of *D. nodosus* strain A198 (Fig. 2a). The conserved residues included Arg<sup>251</sup>, His<sup>256</sup>, Arg<sup>258</sup> and Tyr<sup>274</sup>, which are conserved in integrases from this family (Argos et al., 1986; Abremski & Hoess, 1992).

The highest amino acid identity to other integrases from the databases was 40·4% with the integrase from retrophage ΦR73 (Sun et al., 1991).
Fig. 2. Comparison of the amino acid sequences of (a) IntA with IntC, and (b) Er. carotovora RsmA with D. nodosus GlpA. The amino acid residues which are identical (*), amino acids with conservative substitutions (.) and amino acids conserved among many bacteriophage integrases (X) are marked.

Two novel ORFs, designated orf242 and orf171, were located downstream from intC in strain C305 (Fig. 1). These two ORFs had no similarity to known genes from the vap regions of D. nodosus, nor to any sequences in the databases. These were followed by two small ORFs, designated vapG§ and vapH§, which showed high similarity to vapG and vapH from vap region 1 of strain A198. The deduced amino acid sequence of vapG§ had 55–2% identity with VapG and VapH§ had 56–2% amino acid identity with VapH.

IS1253 is next to vapH§

The insertion sequence IS1253 from D. nodosus strain AC3577 is a 1689 bp element consisting of two ORFs, orf375 and orf117, with similarity to transposase genes from a number of unusual IS elements (Billington et al., 1996). In strain C305, we found a copy of this element next to vapH§ (Fig. 1). The IS1253 found in strain C305 had DNA identity of 98–3% over 1649 bp with the IS1253 from strain AC3577, and consisted of two similar ORFs, orf374 and orf108. The copy of IS1253 in strain C305 appeared to be non-functional, as the coding region from orf374 had been disrupted by two frame-shift mutations, one of which was the deletion of a GTAA repeat. A related insertion sequence is found adjacent to virulence-associated genes in several pathogenic bacteria, including Salamonella typhimurium (Gulig et al., 1992), Helicobacter pylori (Censini et al., 1996), Vibrio cholerae (Bik et al., 1996), Clostridium perfringens (Katayama et al., 1995) and Yersinia pestis (Hu et al., 1998).

The intB and regA genes are disrupted in strain C305

The intB and regA genes from strain A198 are located immediately to the right of vap region 3, and form part of the intB genetic element, separate from the intA element (Bloomfield et al., 1997). In strain C305, an internal fragment of the intB coding region was located immediately following IS1253 (Fig. 3a) – a 100 bp DNA segment with 98% identity to nucleotides 410–510 of the 1203 bp coding region of the intB gene next to region 3 of strain A198 (Fig. 3b). At this position in the intB gene seven nucleotides from the 19 bp attachment (att) site found at the ends of the vap regions were absolutely conserved (Fig. 3b). Thus, this deletion may have been catalysed by intC.

Following the remnant of intB in strain C305, there was a 344 bp segment which had no similarity to other known sequences from D. nodosus, or to other sequences from the databases. After this, the DNA sequence from strain C305 was almost identical to the DNA sequence from the beginning of the regA gene of strain A198 (Fig. 3c). However, the promoter, Shine–Dalgarno sequence, and start codon of the regA gene were absent from strain C305. The next 322 bp of the sequence were almost identical, except that there was a 1 bp insertion in the sequence from strain C305, resulting...
Integrated genetic elements modulate virulence

**Fig. 3.** Comparison of the DNA sequences of strain C305 and strain A198 containing the intB gene and the regA gene. (a) Diagrammatic representation of the intB and regA genes from strains A198 and C305. (b) Comparison of sequences containing the intB gene. (c) Comparison of sequences containing the regA gene. Nucleotides which are identical in both sequences are joined by colons. Every tenth nucleotide is indicated by v. The seven nucleotides underlined in (b) are identical to seven nucleotides from the att site.

In a frameshift mutation. Hence, the regA gene, like the intB gene, did not appear to be intact in strain C305.

**Detection of the intC element in other strains of D. nodosus**

To determine whether the genes intC, orf242, orf171, vapG* and vapH* were part of an integrated genetic element, 15 strains of *D. nodosus* were analysed in Southern blot experiments using probes 6, 7, 8 and 2 (Fig. 1), which detect intC, orf242, orf171 and vapG"H", respectively. The strains included five virulent strains, one intermediate strain, seven benign strains lacking the intA element, and two benign strains which contain the intA element. The results (data not shown) showed that, in ten of these strains, these five genes were...
**Table 1. Genetic elements integrated next to askA and pnpA in 16 strains of *D. nodosus***

<table>
<thead>
<tr>
<th>Strain</th>
<th>Virulence classification</th>
<th>Located next to askA</th>
<th>Located next to pnpA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A198</td>
<td>virulent</td>
<td>intA</td>
<td>intA</td>
</tr>
<tr>
<td>B1006</td>
<td>virulent</td>
<td>intC</td>
<td>intA</td>
</tr>
<tr>
<td>D1172</td>
<td>virulent</td>
<td>intC</td>
<td>intA</td>
</tr>
<tr>
<td>G1220</td>
<td>virulent</td>
<td>intC</td>
<td>intA</td>
</tr>
<tr>
<td>1311</td>
<td>virulent</td>
<td>intC</td>
<td>intA</td>
</tr>
<tr>
<td>H1215</td>
<td>virulent</td>
<td>intB</td>
<td>intA/unknown*</td>
</tr>
<tr>
<td>AC390</td>
<td>benign</td>
<td>intB</td>
<td>intA</td>
</tr>
<tr>
<td>C305</td>
<td>benign</td>
<td>intC</td>
<td>intB</td>
</tr>
<tr>
<td>H1204</td>
<td>benign</td>
<td>intB</td>
<td>intA</td>
</tr>
<tr>
<td>819</td>
<td>benign</td>
<td>intB</td>
<td>intB</td>
</tr>
<tr>
<td>1169</td>
<td>benign</td>
<td>intB</td>
<td>intB</td>
</tr>
<tr>
<td>1311A</td>
<td>benign</td>
<td>intB</td>
<td>intA</td>
</tr>
<tr>
<td>1469</td>
<td>benign</td>
<td>intC</td>
<td>intB</td>
</tr>
<tr>
<td>1493</td>
<td>benign</td>
<td>intC</td>
<td>intB</td>
</tr>
<tr>
<td>2483</td>
<td>benign</td>
<td>intC</td>
<td>intB</td>
</tr>
<tr>
<td>3138</td>
<td>benign</td>
<td>intC</td>
<td>not intA, intB, intC</td>
</tr>
</tbody>
</table>

*Two copies of pnpA in this strain.

---

**Site of integration of the intC element**

The intC element in strain C305 was integrated into the tRNA-ser<sub>GCU</sub> gene downstream from askA (Fig. 1). In the virulent strain A198, vap region 1 is integrated into this tRNA-ser gene, while vap region 2 is integrated into tRNA-ser<sub>GGA</sub>, downstream from the pnpA gene (Bloomfield *et al.*, 1997). Probes specific for askA and pnpA were used in Southern blot experiments to determine the site of integration of the intC element in 13 strains of *D. nodosus* known to contain intC (Table 1: all strains except A198, AC390 and 1311A). In 10 of these strains, the intC element was located next to the askA gene. None of the strains tested had the intC element located next to pnpA. In strains H1215, H1204 and 1169, the intC element was not located next to askA, and its site of integration is unknown. These three strains did not have all of the genes so far recognized as being part of the intC element, perhaps suggesting that the intC element was unstable in the alternative position.

**Loss of the intC element from virulent *D. nodosus* strain 1311**

The virulent strain 1311 carries a small native plasmid, pDN1 (G. Whittle, M. E. Katz & B. F. Cheetham, unpublished), which has not been found in any other strain of *D. nodosus*. During routine laboratory growth,
pDN1 was lost spontaneously from strain 1311, generating a strain designated 1311A. Southern blot analysis (data not shown) showed that the intC element was not present in strain 1311A, since genomic DNA from this strain did not hybridize to DNA fragments from intC, orf242 or orf171. vapG" and vapH" were also lost, together with one copy of IS1253. Thus, loss of pDN1 was accompanied by loss of the intC element. The concomitant loss of intC, orf242, orf171, vapG", vapH" and a copy of IS1253 provided further evidence that these genes were part of an integrated genetic element.

intB is next to askA in strain 1311A

If in strain 1311 the intB element were located next to the intC element, as in strain C305, then excision of the intC element in strain 1311 would place intB next to askA. Southern blot analysis of strains 1311 and 1311A using an intB probe showed that this was indeed the case (Fig. 4). Both strains contained two copies of intB, one of which is located in a different position in the two strains. In strain 1311, the intB probe hybridized to a 2.5 kb EcoRI fragment, while in strain 1311A a 3.3 kb EcoRI fragment was detected (Fig. 4, lanes C and D). In strain 1311A, the askA probe also hybridized to a 3.3 kb EcoRI fragment. These results indicated that in strain 1311, intC was located next to askA and intB was located at the right-hand end of the intC element. The loss of the intC element from strain 1311 has resulted in the relocation of the intB gene such that it was adjacent to askA in strain 1311A.

In both strains 1311 and 1311A the intA element was integrated into tRNA-serGCU, next to pnpA, and a truncated copy of intB was integrated into the attR at the right-hand end of the intA element. The intB probe hybridized to a 2.2 kb EcoRI fragment in both strains, indicating that the position of this copy of intB was unchanged by the excision of the intC element.

To confirm that intC was adjacent to askA in D. nodosus strain 1311, and that intB was adjacent to askA in strain 1311A, PCR products were amplified using primer sets specific for askA-intC (primers A and B), and askA-intB (primers A and C). As expected, a 1-kb askA-intB product was detected in strain 1311A whilst no askA-intC product was observed. In strain 1311, a 2.6 kb askA-intC product was observed, as expected.

Continuing loss of the intC element from strain 1311

Contrary to expectations, a 1-kb askA-intB product was also obtained in strain 1311, suggesting that intB was also located next to askA in this strain. However, when serial dilutions of 1311 template and 1311A template were used, the amount of the askA-intB product was much lower in 1311 than in 1311A (data not shown). This suggested that a subpopulation of cells from D. nodosus strain 1311 was undergoing loss of the intC element frequently enough to be detected. This was supported by Southern blot analysis (Fig. 4), where a faint band was detected in genomic DNA of strain 1311 by the intB probe at 3-3 kb. The loss of the intC element may be favoured under laboratory conditions.

Excision of the intC element does not disrupt tRNA-serGCU

The askA-intB PCR product from strain 1311A was subcloned, and the DNA sequence determined and aligned to the left end and right end of vap region 1/3 in D. nodosus virulent strain A198. Analysis of this junction region showed that the left-hand side of the att site was identical to attL, as would be expected if the intC element were excised by homologous recombination or an integrase-catalysed reaction. Analysis of the DNA sequence also showed that the tRNA gene was not disrupted by the loss of the intC element. Disruption of the leuX tRNA gene by the deletion of a pathogenicity island has been shown to have a major effect on virulence in uropathogenic E. coli (Ritter et al., 1995).

The intC element in strain C305 may be truncated

Southern blot analysis (summarized in Fig. 5) showed that the distance between askA and intB in strain 1311 was greater than 14 kb, and that all of the sequences between tRNA-serGCU and intB were lost together from strain 1311 to form strain 1311A. By contrast, in strain C305, the distance between askA and intB was 6.6 kb (Fig. 1). These results suggested that in strain C305 part of the intC element had been deleted, perhaps due to the insertion of IS1253, which was located at the right-hand end of the intC element in strain C305.

Loss of the intC element from strain 1311 is associated with loss of a virulence characteristic

It has been established that the extracellular proteases of virulent strains of D. nodosus are more thermostable than proteases of benign strains (Depiazzi et al., 1990). The thermostability of the secreted proteases of virulent strain 1311 and its daughter strain 1311A was compared with that of the virulent strain A198 and the benign strain C305. The percentage loss of activity after heating (± standard deviation) was 29.5 ± 5.6 for virulent strain A198, 74.1 ± 11.9 for benign strain C305, 20.1 ± 11.4 for strain 1311, and 84.9 ± 9.0 for strain 1311A. These results clearly showed that the proteases from strain 1311A were much less thermostable than the proteases of strain 1311, and therefore strain 1311A behaved as a benign strain in this assay. These results suggest that the loss of the intC element was associated with loss of protease thermostability, a trait associated with virulence in D. nodosus. In addition, colonies from strain 1311A were smaller and exhibited reduced exoenzyme activity on solid medium when compared to colonies from strain 1311 (data not shown). Virulent strains of D. nodosus have, in general, larger colonies and exhibit greater exoenzyme activity than benign strains (Depiazzi & Richards, 1985).
Loss of pDN1 may occur without loss of the intC element

In an effort to determine whether the loss of intC occurred concomitantly with the loss of pDN1, the size difference between colonies of strains 1311 and 1311A was utilized to select colonies of strain 1311 that had potentially lost pDN1 and/or the intC element. Three small colonies were selected, and analysed. All three had lost pDN1 but still contained the intC element, and retained protease thermostability (data not shown). Of the 55 colonies screened, none which had lost the intC element were isolated. These results indicated that the loss of pDN1 may occur without loss of the intC element and that the loss of protease thermostability was not related to the loss of the native plasmid, pDN1, from strain 1311. However, since three colonies of reduced size were found to have lost pDN1, but not the intC element, the reduced size of 1311A colonies may have been due to loss of pDN1.

Sites of integration of the intA and intC elements in virulent and benign strains

There was no direct correlation between the presence of the intC element and virulence, since the virulent strain A198 does not contain the intC element, while most benign strains do contain it. Southern blot analysis of the integration sites of the intC and intA elements in 16 strains of D. nodosus showed that most virulent strains had the intC element next to askA and the intA element next to pnpA (Table 1). The virulent strain A198 had two copies of intA, one next to askA and one next to pnpA. By contrast, most benign strains had intC next to askA, but neither intC nor intA next to pnpA. We propose that virulent strains have the intA element next to pnpA and either the intA element or intC element next to askA. Strains which lack the intA element next to pnpA, or have neither the intA nor the intC element next to askA, would be benign if our model were correct.

Of the 16 strains studied, the only exception to the above pattern was virulent strain H1215, which has intB next to askA. Strain H1215 was atypical in several ways. Unlike all other strains studied, it contained two copies of pnpA, one of which had intA next to it. It is not known what is next to the other copy of pnpA. This strain also had two copies of intC, and their integration sites are unknown. However, it lacked orf171, vapG and vapH, suggesting that deletions had occurred. Strain H1215 also contains a complete integrated prophage (G. A. Bloomfield, unpublished), but the integration site is also unknown.

A putative global repressor (glpA) gene between askA and tRNA-ser

Between the askA and tRNA-ser genes we identified an ORF, designated glpA (Fig. 1), encoding a putative product of 64 aa which had 74% aa identity over the first 50 aa with RsmA (Fig. 2b) from the plant pathogen Erwinia carotovora (Cui et al., 1995). In Er. carotovora, RsmA acts as a global repressor of virulence. The rsmA gene product reduces the transcript level of hslI, a luxI homologue required for N-(3-oxohexanoyl)-l-homoserine lactone (HSL) synthesis (Cui et al., 1995), and...
thereby represses extracellular enzyme production, reduces motility and colony size, and so attenuates virulence (Mukherjee et al., 1996). It is interesting that *D. nodosus* strain 1311A exhibits a similar phenotype to cells expressing *rsmA* in *E. carotovora*, given the similarity between GlpA and RsmA. CsrA from *E. coli*, which is also very similar to RsmA, modulates glycogen biosynthesis by effects on mRNA stability (Liu et al., 1995). CsrA belongs to a family of proteins containing KH domains, which are involved in RNA binding. The *E. coli* *pnp* gene belongs to the same family (Liu et al., 1995), affects mRNA processing and decay (Alifano et al., 1994), and is one of the two principal enzymes involved in the degradation of bacterial mRNA to nucleotides (Li & Deutscher, 1994). Thus, in *D. nodosus*, the *inta* and the *intC* elements are integrated next to two genes whose products may be involved in gene regulation by the alteration of mRNA stability. The integration of these genetic elements may affect levels of *glpA* and *pnpA* transcripts, resulting in increased production of thermostable extracellular proteases.

Comparison of DNA sequences at the beginning of the *inta*, *intB* and *intC* elements

Analysis of the DNA sequences at the beginning of the *inta*, *intB* and *intC* elements identified several features which may be important in differential gene expression in *D. nodosus* strains containing different integrated genetic elements. These are discussed below.

Termination of transcription. The DNA sequences starting from the TGA stop codon for *glpA* for strains C305, A198 and 1311A were compared (data not shown). In all three strains, the sequence was almost identical up to the 3' end of the tRNA gene (the *attL* site). Within this sequence, there was a region of dyad symmetry (loop 1) very close to the stop codon for *glpA*, which could act as a transcriptional terminator for *glpA*. Strains C305, A198 and 1311A had the *intC*, *inta* and *intB* elements, respectively, integrated after the *attL* site. The first 42 bp of the *inta* and *intC* elements were identical, and differed substantially from the first 42 bp of the *intB* element. Within this 42 bp sequence was a second region of dyad symmetry (loop 2), which could act as an alternative transcriptional terminator for *glpA*. If this transcriptional terminator were used, the transcript would include the tRNA molecule. This dyad symmetry (loop 2) was not present in the *intB* element. Thus, termination of transcription of *glpA* may differ according to whether *intC*, *inta* or *intB* is integrated at this position.

Virulent strains had the *inta* element integrated next to *pnpA*, while most benign strains had *intB* at this position. In the virulent strain A198, the stop codon in *pnpA* was located only one base before the *tRNA-ser*<sub>GCU</sub> gene, and no putative transcriptional terminators were evident before the end of the tRNA gene. Thus, it is likely that *pnpA* transcripts include the tRNA molecule, and would differ according to the genetic element integrated at this position. The integration of the *inta* element, but not the *intB* element, next to *pnpA* would provide a transcriptional terminator.

Antisense RNA. Within the divergent region downstream of loop 2, sequences weakly resembling the consensus sequence for the *E. coli* σ<sup>54</sup> — 35 and — 10 promoter sequences (Hawley & McClure, 1983) were identified in the regions upstream of *intC* and *inta*, but were not present upstream of *intB*. If these promoter sequences are functional, the antisense transcript produced could potentially sequester those messages transcribed on the opposite DNA strand, and left of the *intC* element and *inta* element integration sites, including those transcripts for *glpA*, *pnpA*, *tRNA-ser*<sub>GCU</sub> and *tRNA-ser*<sub>GGA</sub> genes.

tRNA processing. Since the *inta* and *intC* elements integrate into the 3' ends of *tRNA-ser*<sub>GCU</sub> and *tRNA-ser*<sub>GGA</sub>, they could alter the expression of these genes, thereby affecting the levels of these two serine tRNAs. This could, in turn, affect the translation of genes which use the corresponding serine codons. The tRNA-ser may have a role in virulence, similar to that proposed in uropathogenic *E. coli*, where the LeuX tRNA molecule is thought to act as a global regulator of several virulence factors (Ritter et al., 1995).

DNA topology. It has been reported previously that the integration of genetic elements can alter gene expression by affecting DNA topology (Dorman, 1993; Ott, 1993; Zagaglia et al., 1991). It is therefore possible that an element of critical size, such as the *intC* or *inta* elements, must be integrated adjacent to both *glpA* and *pnpA*, in order to provide the appropriate topology for thermostable protease expression or secretion.

A model for the role of the *inta* and *intC* elements in the virulence of *D. nodosus*

Our analysis of 16 strains of *D. nodosus* has shown that the *inta*, *intB* or *intC* elements may integrate next to *glpA*, and either the *inta* or the *intB* elements may integrate next to *pnpA*. Our model proposes that virulent strains of *D. nodosus* have the *inta* element integrated next to the *pnpA* gene, and the *intC* or *inta* element integrated next to the *glpA* gene. The presence of these integrated genetic elements may alter the expression of *glpA*, *pnpA*, *tRNA-ser*<sub>GCU</sub>, *tRNA-ser*<sub>GGA</sub> or *askA*, and reduce the expression of thermostable proteases through effects on mRNA stability, tRNA availability, or HSL levels.

It has been observed previously that virulent strains may have more than one copy of the *inta* element, but that strains with increased numbers of copies do not show increased virulence. This is consistent with our model, since the integration of a single copy of the *inta* element next to the *pnpA* gene and the *intC* element next to the *glpA* gene is sufficient for virulence. The presence of the *inta* element in some benign strains is also consistent with this model, since these strains would be benign if they did not have either the *inta* or the *intC* element integrated next to *glpA*, and the *inta* element integrated
next to pnpA. For example, the benign strain AC390 has two copies of intA, one of which is integrated next to pnpA, but does not have the intC element, and the second copy of the intA element is not integrated next to glpA (Table 1). Based on these results, a diagnostic test to determine the virulence of strains of _D. nodosus_ could be designed, using PCR to determine which genetic elements are integrated next to glpA and pnpA.

There is one report of a small number of virulent strains which lack sequences from the _vap_ region (Rood _et al._, 1996). These strains may have a copy of the intC element integrated next to both glpA and pnpA. It will be of great interest to analyse such strains to determine whether genetic elements are integrated at these two locations.

**Similarities between the intA and intC elements**

The majority of the genes from the _intA_ and _intC_ elements are quite unrelated. However, the two integrases, IntA and IntC, have 54.2% amino acid identity, and _vapG_ and _vapH_, from the _intA_ element, and _vapG’_ and _vapH’_, from the _intC_ element, have approximately 55% amino acid identity, and 81% nucleotide identity. In addition, two sequences of 42 bp and 61 bp are almost totally conserved in the DNA segment between the _att_ site and the beginning of _intA_ and _intC_. These results suggest that there has been genetic exchange between the _intA_ and _intC_ elements, that the related sequences are required for similar functions and have been maintained by selection, or that the two elements have evolved from a common ancestor.

**Phase variation in _D. nodosus_**

In culture we have shown loss of the _intC_ element from a virulent strain of _D. nodosus_, resulting in loss of a virulence factor. Loss of the _intC_ element from _D. nodosus_ strains in natural infections may allow switching from the virulent to the benign phenotype, and may have a selective advantage. However, this has not yet been observed _in vivo_. Loss of the _intA_ element is very unlikely, due to the _vapA/toxA_ maintenance system in this element (Bloomfield _et al._, 1997). At present, we do not know whether the _intA_, _intB_ or _intC_ elements can be transferred between _D. nodosus_ strains, although a plasmid containing the genes from _vap_ region 1 has been identified in one strain of _D. nodosus_ (Billington _et al._, 1996). It is possible that strains of _D. nodosus_ may be able to switch between the benign and virulent states by the loss or gain of these genetic elements. This work has significant implications for footrot eradication programmes which are currently in place in Australia.

**ACKNOWLEDGEMENTS**

We thank J. Druitt for her excellent technical assistance, and J. Rood and J. Searson for the provision of _D. nodosus_ strains. This work was supported by a grant from the Australian Research Council. G.W. and G.B. were recipients of Australian Postgraduate Research Awards.

**REFERENCES**


Integrated genetic elements modulate virulence


Received 2 March 1999; revised 10 June 1999; accepted 1 July 1999.