

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Leo A. Calvo-Bado, Laura E. Green, Graham F. Medley, Atiya Ul-Hassan, Rose Grogono-Thomas, Nicky Buller, Jasmeet Kaler, Claire L. Russell, Ruth M. Kennan, Julian I. Rood and Elizabeth M.H. Wellington

Article Title: Detection and diversity of a putative novel heterogeneous polymorphic proline-glycine repeat (Pgr) protein in the footrot pathogen *Dichelobacter nodosus*

Year of publication: 2010

Link to published article: <http://dx.doi.org/10.1016/j.vetmic.2010.06.024>

Publisher statement: Copyright © 2010 Elsevier B.V. Calvo-Bado, L. A., (2010). Detection and diversity of a putative novel heterogeneous polymorphic proline-glycine repeat (Pgr) protein in the footrot pathogen *Dichelobacter nodosus*. *Veterinary Microbiology*.

1 **Detection and diversity of a putative novel heterogeneous polymorphic**
2 **proline-glycine repeat (Pgr) protein in the footrot pathogen *Dichelobacter***
3 ***nodosus***

4
5 Leo A. Calvo-Bado^{a*}, Laura E. Green^a, Graham F. Medley^a, Atiya Ul-Hassan^a, Rose
6 Grogono-Thomas^b, Nicky Buller^c, Jasmeet Kaler^a, Claire L. Russell^b, Ruth M. Kennan^d,
7 Julian I. Rood^d, Elizabeth M.H. Wellington^a.

8 ^a *Department of Biological Sciences, University of Warwick, Gibbet Hill Road, Coventry*
9 *CV4 7AL, UK.*

10 ^b *School of Clinical Veterinary Science, University of Bristol, Langford House, Langford*
11 *BS40 5DU, UK.*

12 ^c *Animal Health Laboratories, Department of Agriculture and Food, Western Australia, 3*
13 *Baron-Hay Court, South Perth, WA6151, Australia.*

14 ^d *Australian Research Council Centre of Excellence in Structural and Functional Microbial*
15 *Genomics, Department of Microbiology, Monash University, Victoria 3800, Australia.*

16

17 *Corresponding author: Tel.: +44(0)2476 522431; Fax: +44(0)2476 523701. *E-mail address:*
18 *l.a.calvo-bado@warwick.ac.uk (Leo A. Calvo-Bado).*

19

20 **Keywords:** *Dichelobacter nodosus*, footrot, sheep, strain variation, repeat regions

21

22 **Short Title:** A novel proline-glycine repeat (Pgr) protein in *Dichelobacter nodosus*

23

24

25

26 **ABSTRACT**

27 *Dichelobacter nodosus*, a Gram negative anaerobic bacterium, is the essential causative agent
28 of footrot in sheep. Currently, depending on the clinical presentation in the field, footrot is
29 described as benign or virulent; *D. nodosus* strains have also been classified as benign or
30 virulent, but this designation is not always consistent with clinical disease. The aim of this
31 study was to determine the diversity of the *pgr* gene, which encodes a putative proline-
32 glycine repeat protein (Pgr). The *pgr* gene was present in all 100 isolates of *D. nodosus* that
33 were examined and, based on sequence analysis had two variants, *pgrA* and *pgrB*. In *pgrA*,
34 there were two coding tandem repeat regions, R1 and R2: different strains had variable
35 numbers of repeats within these regions. The R1 and R2 were absent from *pgrB*. Both
36 variants were present in strains from Australia, Sweden and the UK, however, only *pgrB* was
37 detected in isolates from Western Australia. The *pgrA* gene was detected in *D. nodosus* from
38 tissue samples from two flocks in the UK with virulent footrot and only *pgrB* from a flock
39 with no virulent or benign footrot for >10 years. Bioinformatic analysis of the putative PgrA
40 protein indicated that it contained a collagen-like cell surface anchor motif. These results
41 suggest that the *pgr* gene may be a useful molecular marker for epidemiological studies.

42

43

44

45

46

47

48

49

50

51 1. INTRODUCTION

52 The main causative agent of ovine footrot is *Dichelobacter nodosus*, a Gram-negative
53 anaerobic bacterium with a small genome (Myers et al., 2007). The clinical manifestation of
54 footrot is dependent on both environmental conditions (Depiazzi et al., 1998) and the
55 virulence of the causative *D. nodosus* strain. Three laboratory tests have been developed to
56 classify virulence of *D. nodosus* isolates. These assays are the protease thermostability or
57 gelatin-gel test (Palmer, 1993), a test for the presence of the genetic element *intA* (Cheetham
58 et al., 2006) and a test for the presence of the genomic islands *vap* and *vrl* (Rood et al., 1996).
59 These tests do not always correlate, suggesting that either the tests do not test absolute
60 markers for virulence or that virulence is complex and linked to more than one process.

61 From bioinformatic analysis of the VCS1703A genome of *D. nodosus* and
62 comparative genome hybridization (CGH) of nine isolates (six virulent and three benign
63 strains), a gene encoding a putative large, repetitive secreted protein, DNO_0690, was
64 identified; designated here as Pgr for proline-glycine repeats. It was apparently present only
65 in virulent strains and it was suggested that this gene might encode a virulence factor that
66 could be involved in adhesion to the extracellular matrix (Myers et al., 2007).

67 The current study describes the isolation and diversity of DNO_0690 gene (*pgr*) in
68 isolates from the UK, Australia and Sweden.

69

70 2. Materials and methods

71

72 2.1. Isolates

73 One hundred isolates from the UK (32), Australia (64) and Sweden (4) were analysed
74 (Table I and Table II). The isolates included the nine Australian serogroup A-I prototype
75 strains (VCS1001, VCS1006, VCS1008, VCS1172, VCS1137, VCS1017, VCS1220,

76 VCS1687, and VCS1623, respectively) and two virulent strains (VCS1703A and VCS1001
77 (also known as A198), one benign isolate (C305). There were 34 clinical isolates from
78 Western Australia isolated from the feet of sheep at an abattoir (Buller et al., 2010). There
79 were two isolates of bovine origin (C809 and C910) from the UK.

80

81 2.2. Growth and DNA extraction

82 Growth of *D. nodosus* and DNA extraction were done as described previously (Moore
83 et al., 2005). DNA from the Australian isolates was extracted using Prepman Reagent
84 (Applied Biosystems). DNA was extracted from the interdigital skin of the feet of three sheep
85 each from two flocks with virulent footrot, and one flock with no virulent or benign footrot
86 for >10 years. Tissue (130-160 mg) was treated with collagenase (10mg/ml) to release
87 bacterial cells. Then the supernatant containing the extracted cells was pelleted at 15900 x g
88 for 15 min. DNA was extracted using MagMAX™ Express Magnetic Particle Processors
89 (AMBION, Applied Biosystems, Inc.) according to the manufacturer's recommendations.

90

91 2.3. PCR amplification, cloning and sequencing

92 The *pgr* gene from the sequenced genome of strain VCS1703A was used as a template
93 to generate PCR primers. The *pgr* gene was detected in all isolates.

94 The complete *pgr* gene was amplified in 8 isolates (Fig. 1) with primers *pgrF* (5'-
95 ATGGCAGTGATTACATTAATGTTTCGCTCCGCTAC-3') and *pgrR* (5'-
96 TTAGATGATTATGGAGCCAACGTGCCCATGTCAT-3') at an annealing temperature of
97 55°C with KOD DNA polymerase master mix containing KOD Hot Start DNA polymerase
98 (1.0 U/ µl), 10 × PCR Buffer for KOD, 25 mM MgSO₄, dNTP Mix (2 mM each), 1 × PCR
99 buffer for KOD (Novagen). Amplicons were cloned into the pGEM-T Easy vector system
100 (Promega) according to the manufacturer's recommendations. All amplicons were gel

101 purified (Qiagen) and sequenced using an ABI PRISM 3130xl Genetic Analyser (Applied
102 Biosystems). Sequences were edited, aligned with SeqMan II (Lasergene 6) and MegAlign
103 DNASTAR software. Two variants of the *pgr* gene, *pgrA* and *pgrB*, were observed. These were
104 differentiated by PCR with primers *pgrAF1* (5'- CCTGCACCATGCTTGTTAAA -3') and
105 *pgrAR1* (5'- GCTGTTGGTGGTTTGGCTAT -3') at an annealing temperature of 60°C and
106 *pgrBF3* (5'-AKCATCRGGAAAGGTGA-3') and *pgrBR2* (5'-GACGGCATCAGCAGCA-3')
107 at annealing temperature of 55°C. The annealing temperatures were optimised for all PCR
108 primers. Fifty isolates were tested against each set of primers for the *pgrA* and *pgrB* genes.
109 Positive and negative controls were always included in each PCR reaction.

110 The R1 variable region was PCR-amplified and sequenced with primers *pgrF1* (5'-
111 TTCCAAGTCTACCGTCA-3') and *pgrR1* (5'-GCATTGGCAAGCGCAA-3'). All 100
112 sequences were submitted to the GenBank database (GU944975 to GU945040 and
113 HM569229 to HM569262). All PCR amplifications were done using PCR-Promega master
114 mix (Promega). Each sample contained 25 µl Master mix (50 units/ml of *Taq* DNA
115 polymerase supplied in reaction buffer pH 8.5, 400 µM dATP, 400 µM dGTP, 400 µM dCTP,
116 400 µM dTTP, 3 mM MgCl₂), 10 mM of each primer, 2.5 µl of DMSO (dimethyl sulfoxide,
117 Fisher Scientific), 2 µl BSA (bovine serum albumin 10mg/ml, SIGMA) and 1-3 µl of
118 template DNA (50-100 ng) in 50 µl reaction mix using the following conditions: 1 cycle of
119 95°C for 2 min, 35 cycles of 95°C for 1 min, 55°C for 45 sec and 72°C for 2 min with a final
120 extension of 72°C for 5min.

121 DNA from tissue from the UK flocks were screened initially by PCR for the presence
122 of the *D. nodosus* specific 16S rRNA gene (*Cc* 5'-
123 TCGGTACCGAGTATTTCTACCCAACACCT-3' and *Ac* 5'-CGGGGTTATGTAGCTTGC-
124 3') (La Fontaine et al., 1993), then for presence of the *pgr* gene and then for the variants *pgrA*
125 and *pgrB*.

126 2.4. *Virulence tests*

127 Protease thermostability (Palmer, 1993), the presence of *intA* (Cheetham et al., 2006)
128 and the presence of genomic islands such as the *vap* and *vrl* regions (Katz et al., 1991; Rood
129 et al., 1996) were carried out in the laboratory of origin on some samples as described in the
130 relevant papers.

131

132 2.5. *Phylogenetic analysis*

133 For phylogenetic analyses, all nucleotide and protein sequences were aligned using
134 ClustalW (www.ebi.ac.uk/ClustalW) and subsequently edited in Bioedit
135 (www.mbio.ncsu.edu/Bioedit/bioedit.html). Neighbour joining trees were constructed using
136 the Phylip package of programs (<http://evolution.genetics.washington.edu/phylip.html>).
137 Distances between sequences were calculated in DNADIST and PROTDIST using the
138 Kimura-2-parameter and the Jones-Taylor-Thornton model of sequence evolution for
139 nucleotide and protein sequences respectively. Hypothetical basal PgrA and PgrB sequences
140 were created *in silico* based on their sequence identity (all tandem repeats were removed);
141 these sequences served as the root for tree constructions.

142

143 **3. Results**

144

145 3.1. *Detection of the pgr gene in D. nodosus isolates and comparisons with other assays*

146 All 100 isolates contained the *pgr* gene (Tables I and II). Sequence analysis revealed
147 considerable variation, with two groupings of variants: *pgrA* (38/100 isolates) and *pgrB*
148 (62/100 isolates). Of the isolates tested for protease activity, 28/31 (90%) *pgrA*-positive
149 isolates had thermostable protease activity and 19/26 (73%) of *pgrB*-positive isolates had
150 thermolabile protease activity (Table I). There was no correlation between *pgr* variant and

151 genetic elements such as *intA*, *vrl* and *vap* or serogroup. *PgrA* and *pgrB* were detected in
152 isolates from the UK, Eastern Australia and Sweden. The UK isolates BS2, BS22, BS23,
153 BS26 and BS27 that contained *pgrA* and isolates BS6, BS14, BS17, BS19, and BS24 that
154 contained *pgrB* were isolated from feet with benign footrot (interdigital dermatitis), virulent
155 footrot or Contagious Ovine Digital Dermatitis (CODD), suggesting that there was no
156 correlation between the presence of *pgrA* or *pgrB* and the type of clinical presentation in the
157 UK. All 34 isolates from WA had the *pgrB* variant; 59.5% (25) isolates had thermostable
158 protease activity and 40.5% (17) had thermolabile protease activity (Table II).

159

160 3.2. Structure, phylogeny, diversity of the *pgr* gene and gene product

161 The *pgrA* gene in strain VSC1703A contained two coding regions with in-frame
162 tandem repeats R1 and R2 at the 5' (18 nt) and 3' (54 nt) ends respectively. These varied in
163 copy number in other strains. R1 was absent from the benign strain C305 (Fig. 1), which has
164 the *pgrB* variant. *PgrA* had 98-100% and *pgrB* <85% sequence identity to the *pgrA* gene
165 from strain VCS1703A in the strains tested. Sequence variation in the *pgrA* derivatives was
166 due to differences in the copy number of the 18 nt R1 region
167 (GGTGAAAACC(C/A)GGTAGT), which encoded a putative glycine rich repeat,
168 GEKPGS. The number of repeats varied, for example there was one copy in isolate VCS1006
169 and six in VCS1703A; there was a maximum of 28 repeats in VCS1137 (Table I). The R2
170 region contained 54 nt

171 (GACCCGGAAAATCCGACCAACCCGGTAGATCCTGAAAATCCAGATAAACCGAC
172 T), which encoded a putative proline rich repeat, DPENPTNPVDPENPDKPT, which also
173 varied in copy number from 1 to 16 repeats, these were in isolates BS1 and VCS1703A,
174 respectively.

175 Isolates with the *pgrB* gene formed three sub-groups (Tables I and II); some isolates
176 (AC390, T9015, UNE6, UNE13, 22477, 7004b and some WA isolates) had different tandem
177 repeats that also varied in copy number; R3 (GGCGATGGAACTAAACCC; GDGTKP), R4
178 (AATCCTGGCGAAGGAACT; DDGTKP) or R5
179 (ATTTCCGGTATTACCCGTAGTACCTGAATCGCCGGTATTTCTGTGTTTCCGGTA
180 TTCCCCGTACTGCCTGAATCGCCGGT; TGDSGSTGNTGNTGNTGDSGTTGNTGN)
181 while two isolates did not have any repeats (C395 and BS4). The *pgrB* region R5 was only
182 observed in WA isolates, with 59.5% (25) isolates of these strains having this region (Table
183 II).

184 Phylogenetic analysis of the sequences of the variable regions revealed that the
185 isolates were grouped into four main clades that corresponded with the putative PgrA and
186 PgrB proteins (Fig. 2). The PgrA clade had a long branch, and strains progressively increased
187 the number of R1 repeats. The PgrB was more diverse, being subdivided into clades 1, 2 and
188 3. Clade 3 also had a long branch that reflected an increasing number of R3 repeats.

189 Bioinformatics analysis suggested that PgrA, but not PgrB, may be an outer
190 membrane protein (<http://www.expasy.org/tools/>; <http://bioinformatics.biol.uoa.gr>) with two
191 putative transmembrane helices and that the R1 and R2 regions represent collagen-like cell
192 surface motifs (GEKPGS) and collagen alpha chain motifs (DPENPTNPVDPENPDKPT)
193 respectively (<http://www.ebi.ac.uk/Tools/InterProScan>).

194

195 3.3. The *pgr* gene in samples and isolates from the UK and Western Australia.

196 All tissue samples from the three UK farms contained *D. nodosus* by 16S rRNA gene
197 specific PCR. In sheep from the flock without footrot only the *pgrB* variant was detected,
198 whilst in the other two flocks one or more *pgrA* variant with 7, 18 and 21 R1 repeats was
199 detected. Finally, the two *D. nodosus* isolates from cattle were *pgrB* variants; both strains

200 were isolated from a 5 day old Belgian Blue cross from East Sussex, UK found dead with
201 severe interdigital ulcerations on all four feet.

202

203 3.4. Design of a PCR assay to detect *pgrA* and *pgrB* variants

204 Isolates with the *pgrA* variant were discriminated from those with the *pgrB* variant by
205 annealing temperatures of 60°C and 55°C respectively (Fig. 3). Subsequently, approximately
206 50 of the total 100 isolates were examined using this assay. The results correlated with the
207 original *pgrA* and *pgrB* designations (Table I).

208

209 4. Discussion

210 In this study we have described the diversity of a gene encoding a putative novel and
211 highly polymorphic proline-glycine repeat protein Pgr. This gene was present in all isolates
212 of *D. nodosus* that we investigated. Sequence analysis of the *pgr* gene revealed that there
213 were two variants both variants had regions with a variable number of tandem repeats. This
214 variation, coupled with the fact that the Myers et al., (2007) results were derived from an
215 oligonucleotide-based microarray might explain why Myers et al., (2007) thought that only
216 virulent strains of *D. nodosus* contained Pgr.

217 Pgr is a putative protein with a molecular size of 100-120 kDa (881 to 1121 amino
218 acid residues). The repeat regions in PgrA contain glycine rich (R1) and proline rich (R2)
219 repeats and the hypervariability in copy number provides evidence that this protein is highly
220 polymorphic. Proline-rich regions play an important role in many protein-protein interactions
221 and cell-wall-spanning domains are preserved in many pathogenic bacteria (Williamson,
222 1994; Vanhoof et al., 1995; Kay et al., 2000; Girard et al., 2006). Collagen-like proteins are
223 common in bacteria such as *Bartonella* (Zhang et al., 2004), *Streptococcus* (Lukomski et al.,
224 2000; Paterson et al., 2008; Rasmussen et al., 2000) and *Yersinia* (Heise and Dersh, 2006)

225 and are involved in virulence, or mediating processes such as attachment, colonization and
226 internalization.

227 Both the function of the Pgr products and the significance of the repeat regions are
228 unknown. Variability in the number of in-frame tandem repeats has been associated with
229 variation in virulence in other bacteria (Gravekamp et al., 1998; Nallapareddy et al., 2000;
230 Puopolo et al., 2001; Jordan et al., 2003) and with human diseases (O'Dushlaine et al., 2005).
231 These variations create protein polymorphisms and may alter antigenicity, as illustrated by
232 the surface proteins alpha C of Group B streptococci (Gravekamp et al., 1996; Madoff et al.,
233 1996) and Lmp1 and Lmp3 of *Mycoplasma hominis* (Ladefoged, 2000).

234 We postulate that the Pgr protein is at least partially exposed at the cell surface and
235 that the observed polymorphisms might be a result of selective immune pressures in the host.
236 The mechanisms that generate and maintain tandem repeats in bacterial genomes are poorly
237 understood, but it has been suggested that they can arise from slipped strand mispairing,
238 transformation and homologous recombination, or rolling circle replication (Romero and
239 Palacios, 1997; Bzymek et al., 2001; Verstrepen et al., 2005). Since multiple strains of *D.*
240 *nodosus* co-exist in individual feet and within flocks during subclinical and clinical infections
241 (Claxton et al., 1983; Hindmarsh and Fraser, 1985; Jelinek et al., 2000; Moore et al., 2005),
242 and antigenic changes can occur in *D. nodosus* following natural transformation and
243 subsequent homologous recombination (Kennan et al., 2003), we suggest that both
244 endogenous or exogenous *pgr* genes or gene regions may be responsible for the generation of
245 these observed variants.

246 A footrot eradication programme has been in operation in WA for the last 20 years
247 and the rate of isolation of gelatin-gel positive strains has declined with an increase in the
248 prevalence of benign footrot. The 100% *pgrB* variant strains in sheep from WA suggests
249 either that the eradication programme, which focussed on culling sheep with gelatin-gel

250 positive strains of *D. nodosus* may have excluded *pgrA* strains (this might be also be the case
251 for the one UK flock with no signs of FR) alternatively, since previous data showed that *D.*
252 *nodosus* isolates from WA were often different from other Australian isolates (Rood et al,
253 1996), it is possible that the *pgrA* gene might never have be present at detectible levels in
254 WA. Cattle have been reported to harbour benign strains of *D. nodosus* (Pringel et al., 2008);
255 the two cattle isolates in this study were *pgrB*, possibly indicating a species divide for these
256 variants.

257 In summary, the *pgr* gene of *D. nodosus* is diverse but has two variants, *pgrA* and
258 *pgrB*, which exhibit considerable heterogeneity. The *pgrA* variant has two repeat regions,
259 generating a protein that may have collagen-like regions and cell surface anchor structures
260 that are associated with invasion of the host in other pathogenic bacteria species.
261 Discrimination of *pgr* variants with a PCR assay offers significant potential as a molecular
262 tool in epidemiological studies.

263

264 **Conflict of interest statement**

265 None of the authors (L.A. Calvo-Bado, L.E Green, G.F. Medley, A. Ul-Hassan, R.
266 Grogono-Thomas, N. Buller, J. Kaler, C.L. Russell, R.M. Kennan, J.I. Rood, and E.M.H.
267 Wellington) has a financial or personal relationship with other people or organisation that
268 could inappropriately influence or bias the paper entitled “Detection and diversity of a novel
269 heterogeneous polymorphic proline-glycine repeat (Pgr) protein from the footrot pathogen
270 *Dichelobacter nodosus*”.

271

272 **Acknowledgments**

273 This research was supported by the Biotechnology and Biological Sciences Research
274 Council UK (BBRSC) under the Combating Endemic Diseases of Farmed Animals for

275 Sustainability (CEDFAS) initiative Grant No. BBE01870X1. RMK and JIR acknowledge the
276 research support of the Australian Research Council through the ARC Centre of Excellence
277 in Structural and Functional Microbial Genomics. We thank Brian Cheetham at the
278 University of New England (Armidale, Australia), Viveca Baveraud at the National
279 Veterinary Institute, Uppsala, Sweden, and Andrew Steventon at the Veterinary Laboratories
280 Agency Winchester for some of the strains used in this study.

281

282 **References**

283 Buller N.B, Ashley P., Palmer M., Pitman D., Richards R.B., Hampson D.J., (2010).
284 Understanding the molecular epidemiology of the footrot pathogen *Dichelobacter nodosus*
285 to support control and eradication programs. J. Clin. Microbiol. 48, 877-882.

286

287 Bzymek M., Lovett S.T., 2001. Instability of repetitive DNA sequences: The role of
288 replication in multiple mechanisms. Proc. Natl. Acad. Sci. USA 98, 8319-8325.

289

290 Cheetham B.F., Tanjung L.R., Sutherland M., Druitt J., Green G., McFarlane J., Bailey G.D.,
291 Seaman J.T., Katz M.E., 2006. Improved diagnosis of virulent ovine footrot using the *inta*
292 gene. Vet. Microbiol. 116, 166-174.

293

294 Claxton P.D., Ribeiro L.D., Egerton, J.R., 1983. Classification of *Bacteroides nodosus* by
295 agglutination tests. Aust. Vet. J. 60, 331-334.

296

297 Depiazzi L.J., Roberts W.D., Hawkins C.D., Palmer M.A., Pitman D.R., Mcquade N.C.,
298 Jelinek P.D., Devereaux D.J., Rippon R.J., 1998. Severity and persistence of footrot in

299 Merino sheep experimentally infected with a protease thermostable strain of
300 *Dichelobacter nodosus* at five sites. Aust. Vet J. 76, 32-38.
301

302 Girard V., Mourez M., 2006. Adhesion mediated by autotransporters of Gram-negative
303 bacteria: Structural and functional features. Res. Microbiol. 157, 407-416.
304

305 Gravekamp C., Horensky D.S., Michel J.L., Madoff L.C., 1996. Variation in repeat number
306 within the alpha C protein of group B streptococci alters antigenicity and protective
307 epitopes. Infect. Immun. 64, 3576-3583.
308

309 Gravekamp C., Rosner B., Madoff L.C., 1998. Deletion of Repeats in the Alpha C Protein
310 Enhances the Pathogenicity of Group B Streptococci in Immune Mice. Infect. Immun. 66,
311 4347-4354.
312

313 Heise T., Dersch P., 2006. Identification of a domain in *Yersinia* virulence factor YadA that
314 is crucial for extracellular matrix-specific cell adhesion and uptake. PNAS 28, 3375–3380.
315

316 Hindmarsh F., Fraser J., 1985. Serogroups of *Bacteroides nodosus* isolated from ovine footrot
317 in Britain. Vet. Rec. 116, 187-188.
318

319 Jelinek P.D., Depiazzi L.J., Galvin D.A., Spicer I.T., Palmer M.A., Pitman D.R., 2000.
320 Occurrence of different strains of *Dichelobacter nodosus* in new clinical lesions in sheep
321 exposed to footrot associated with multi-strain infections. Aust. Vet. J. 78, 273-276.
322

323 Jordan P., Snyder L.A., Saunders N.J., 2003. Diversity in coding tandem repeats in related
324 *Neisseria* spp. BMC Microbiol. 3, 23.
325

326 Katz M.E., Howarth P.M., Yong W.Y., Riffkin G.G., Depiazzi L.J., Rood J.I., 1991.
327 Identification of three gene regions associated with virulence in *Dichelobacter nodosus*,
328 the causative agent of ovine footrot, J. Gen. Microbiol. 137, 2117–2124.
329

330 Kay B., Williamson M.P., Sudol M., 2000. The importance of being proline: the interaction
331 of proline-rich motifs in signaling proteins with their cognate domains. FASEB 14, 231-
332 241.
333

334 Kennan R.M., Dhungyel O.P., Whittington R.J., Egerton J.R., Rood J.I., 2003.
335 Transformation-mediated serogroup conversion of *Dichelobacter nodosus*. Vet. Microbiol.
336 92, 169-78.
337

338 La Fontaine S., Egerton J.R., Rood J.I., 1993. Detection of *Dichelobacter nodosus* using
339 species-specific oligonucleotides as PCR primers. Vet. Microbiol. 35, 101-117.
340

341 Ladefoged, S.A., 2000. Molecular dissection of *Mycoplasma hominis*. APMIS Suppl. 97, 1-
342 45.
343

344 Lukomski S., Nakashima K., Abdi I., Cipriano V.J., Ireland R.M., Reid S.D., Adams G.G.,
345 Musser J.M., 2000. Identification and Characterization of the *scl* Gene Encoding a Group
346 A *Streptococcus* Extracellular Protein Virulence Factor with Similarity to human
347 Collagen. Infect. Immun. 68, 6542–6553.

348 Madoff L.C., Michel J.L., Gong E.W., Kling D.E., Kasper D.L., 1996. Group B streptococci
349 escape host immunity by deletion of tandem repeat elements of the alpha C protein. PNAS
350 93; 4131-4136.
351

352 Moore L.J., Wassink G.J., Green L.E., Grogono-Thomas R., 2005. The detection and
353 characterisation of *Dichelobacter nodosus* from cases of ovine footrot in England and
354 Wales. Vet. Microbiol. 108, 57-67.
355

356 Myers G.S., Parker D., Al-Hasani K., Kennan R.M., Seemann T., Ren Q., Badger J.H.,
357 Selengut J.D., Deboy R.T., Tettelin H., Boyce J.D., McCarl V.P., Han X., Nelson W.C.,
358 Madupu R., Mohamoud Y., Holley T., Fedorova N., Khouri H., Bottomley S.P.,
359 Whittington R.J., Adler B., Songer J.G., Rood J.I., Paulsen I.T., 2007. Genome sequence
360 and identification of candidate vaccine s from the animal pathogen *Dichelobacter*
361 *nodosus*. Nat. Biotechnol. 25, 569-575.
362

363 Nallapareddy S.R., Singh K.V., Duh R.W., Weinstock G.M., Murray B.E., 2000. Diversity of
364 ace, a gene encoding a microbial surface component recognizing adhesive matrix
365 molecules, from different strains of *Enterococcus faecalis* and evidence for production of
366 ace during human infections. Infect. Immun. 68, 5210-5217.
367

368 O'Dushlaine C.T., Edwards R.J., Park S.D., Shields D.C., 2005. Tandem repeat copy-number
369 variation in protein-coding regions of human genes. Genome Biol. 6, R69.
370

371 Palmer M.A., 1993. A gelatin test to detect activity and stability of proteases produced by
372 *Dichelobacter (bacteroides) nodosus*. Vet. Microbiol. 36, 113-122.

373 Paterson G.K., Nieminen L., Jefferies J.M.C., Mitchell T.J., 2008. PclA, a pneumococcal
374 collagen-like protein with selected strain distribution, contributes to adherence and
375 invasion of host cells. *FEMS Microbiol. Lett.* 285, 170–176.
376

377 Pringle M., Bergsten C., Fernström L.L., Höök H., Johansson K.E., 2008. Isolation and
378 characterization of *Treponema phagedenis*-like spirochetes from digital dermatitis lesions
379 in Swedish dairy cattle. *Acta Vet. Scand.* 50, 40-48.
380

381 Puopolo K.M., Hollingshead S.K., Carey V.J., Madoff L.C., 2001. Tandem repeat deletion in
382 the alpha C protein of group B *Streptococcus* is recA independent. *Infect. Immun.* 69,
383 5037-5045.
384

385 Rasmussen M., Edén A., Björck L., 2000. SclA, a novel collagen-like surface protein of
386 *Streptococcus pyogenes*. *Infect. Immun.* 68, 6370-6377.
387

388 Romero D., Palacios R., 1997. Gene amplification and genomic plasticity in prokaryotes.
389 *Annu. Rev. Genet.* 31, 91–111.
390

391 Rood J.I., Howarth P.A., Haring V., Billington S.J., Yong W.K., Liu D., Palmer M.A.,
392 Pitman D.R., Links I., Stewart D.J., Vaughan J.A., 1996. Comparison of gene probe and
393 conventional methods for the differentiation methods for the differentiation of ovine
394 footrot isolates of *Dichelobacter nodosus*. *Vet. Microbiol.* 52, 127-141.
395

396 Vanhoof G., Goossens F., Meester I.D., Hendriks D., Scharpe S., 1995. Proline motifs in
397 peptides and their biological processing. *FASEB* 9, 736-744.

398 Verstrepen K.J., Jansen A., Lewitter F., Fink G.R., 2005. Intragenic tandem repeats generate
399 functional variability. Nat. Genet. 37, 986-990.

400

401 Williamson M.P., 1994. The structure and function of proline-rich regions in proteins.
402 Biochem. J. 297, 249-260.

403

404 Zhang P., Chomel B.B., Schau M.K., Goo J.S., Droz S., Kelminson K.L., George S.S.,
405 Lerche N.W., Koehler J.E., 2004. A family of variably expressed outer-membrane proteins
406 (Vomp) mediates adhesion and autoaggregation in *Bartonella quintana*. PNAS 101,
407 13630-13635.

408

409

410 **Tables**

411

412 **Table I.** Isolates, virulence parameters and source of strains

413 **Table II.** Isolates from Western Australia and virulence assays compared to *prg* test

414

415 **Figure legends**

416

417 **Fig 1.** Schematic representation of Pgr variants in *D. nodosus*. In the PgrA variants
418 (VCS1006, VCS1623, VCS1220, VCS1703A and BS1) (A), the tandem repeat regions
419 consist of 6 (R1, glycine repeats) and 18 (R2, proline repeats) amino acids. PgrB variants
420 (B) have different tandem repeats of 6 amino acids for both regions (R3 and R4) (7004b).
421 In some isolates the R3 and R4 regions are lacking (C305 and BS4).

422

423 **Fig 2.** Phylogenetic analysis of the deduced amino acid sequences of the putative Pgr protein
424 from *D. nodosus*. The C-terminal region of the *pgr* genes of 100 isolates were analysed.
425 The numbers on the branches indicate the percentage bootstrap value of 100 replicates and
426 the scale bar indicates 10% nucleotide dissimilarity. Pgr sequences were grouped into in
427 four main clades. PgrA isolates were in a single clade; PgrB were in three clades. A
428 hypothetical PgrB sequence was included to serve as a root for tree constructions (see
429 methods).

430

431 **Fig 3.** PCR assay specific for discrimination of *pgrA* and *pgrB* isolates. (a) *pgrA* primers, (b)
432 *pgrB* primers. Lanes: MWM (molecular weight marker), BS1 (*pgrA*), BS2 (*pgrA*), BS3
433 (*pgrA*), BS4 (*pgrB*), BS5 (*pgrA*), BS6 (*pgrB*), BS11 (*pgrA*), BS12 (*pgrA*), BS14 (*pgrB*),
434 water (Negative control), BS15 (*pgrA*), BS18 (*pgrA*), BS20 (*pgrA*), BS21 (*pgrB*), BS23
435 (*pgrA*), C305 (*pgrB*), C309 (*pgrB*), A198 (*pgrA*), C310 (*pgrB*).

436