Epidemiol. Infect. (2003), **131**, 823–833. © 2003 Cambridge University Press
DOI: 10.1017/S0950268803008847 Printed in the United Kingdom

Molecular epidemiology of group B streptococci in Ireland: associations between serotype, invasive status and presence of genes encoding putative virulence factors

N. DORE[1,2], D. BENNETT[1,2], M. KALISZER[3], M. CAFFERKEY[1,2,4] AND C. J. SMYTH[5]*

[1] *Irish Meningococcal and Meningitis Reference Laboratory, The Children's University Hospital, Temple Street, Dublin 1, Ireland*
[2] *The Rotunda Hospital, Parnell Street, Dublin 1, Ireland*
[3] *Department of Community Health Trinity College, The Adelaide and Meath Hospital, Tallaght, Dublin 24, Ireland*
[4] *Department of Clinical Microbiology, Royal College of Surgeons in Ireland, York Street, Dublin 2, Ireland*
[5] *Department of Microbiology, Moyne Institute of Preventive Medicine, Trinity College, University of Dublin, Dublin 2, Ireland*

(*Accepted 2 April 2003*)

SUMMARY

Group B streptococcal isolates ($n=159$) from the three Dublin maternity hospitals, were serotyped and analysed for the *bac*, *bca*, *hylB*, *pepB*, and *rib* genes. The serotype distribution of the isolates was Ia, 19·5%; Ib, 18·9%; II, 10·7%; III, 29·5%; IV, 1·9%; V, 15·1%; non-typeable, 4·4%. There was a statistically significant association between the serotype and invasive status (carriage or infection) of isolates ($P<0·005$), but no significant association between serotype and degree of invasiveness was demonstrated. The presence or absence of each analysed gene was not associated with the invasive status of isolates. Statistically significant associations were revealed between *bca* and *hylB* (IS*1548*) ($P=0·0004$) and between *bac* and *bca* ($P=0·0l4$). The *bac*, *bca*, *hylB* (IS*1548*) and *rib* genes and the numbers of tandem repeats in the *bca* gene showed significant associations with serotype. Almost 50% of serotype III isolates possessed at least one of the *bac* and *bca* genes and 55–65% of strains of serotypes Ia, Ib and II possessed the *rib* gene. Most serotype III isolates had IS*1548* in their *hylB* genes. Serotype Ib was the only serotype in which more than half of the strains contained more tandem repeats in the *bca* gene than the overall mean for the GBS population studied of 7·4 repeats. These findings indicate that some previously reported associations between putative virulence factors and GBS disease require further study and clarification.

INTRODUCTION

Streptococcus agalactiae, the group B streptococcus (GBS), is a normal commensal of the human intestine [1, 2]. Since its emergence as an important pathogen in the 1970's [3], it has remained the primary cause of invasive disease in newborn infants and up to 3 months of age [1–4]. In the USA, the Centers for Disease Control and Prevention projected, on the basis of eight sites involved in collaborative surveillance, that there were 17 400 cases of neonatal GBS infection in the USA in 1998 with a case fatality rate of 9·5% [5].

Despite receiving considerable attention over the past few years, the pathogenic mechanisms of GBS

* Author for correspondence.
This work was presented in part at the National Medical Scientific Meeting, Royal College of Physicians in Ireland, Dublin, Ireland, March 2000, Ir J Med Sci 2000; 169 Suppl 2: 69 [abstract]; 40th Interscience Conference on Antimicrobial Agents and Chemotherapy, Toronto, Canada, September 2000 [abstract 1433]: 423.

remain poorly understood. Several GBS components are thought to be associated with its pathogenicity, including the presence and type of capsule [6], a number of surface antigens [7–13], several hydrolytic enzymes [14–21] and the β-haemolysin [22]. Studies on phenotypic properties of GBS have been complemented by molecular genetic analysis of their respective operons and genes [9–11, 13, 15, 16, 21, 23]. Based on *in vitro* and *in vivo* animal model studies of GBS disease, the capsular polysaccharides have been identified as a major virulence factor [6]. Thus far, nine capsular polysaccharide serotypes are known – Ia, Ib, II, III, IV, V, VI, VII and VIII [3, 4, 6]. Other factors thought to be associated with virulence include the surface-associated C proteins, α and β, encoded by the *bca* (group B, C protein alpha) and *bac* (beta-antigen C protein) genes, respectively, and the Rib protein encoded by the *rib* gene [7–13]. Extracellular enzymes considered to be putative virulence factors include hyaluronidase (hyaluronate lyase), oligopeptidase and C5a-peptidase encoded by the *hylB*, *pepB* and *scpA* genes, respectively [14–21]. In GBS lacking hyaluronidase activity an insertion sequence, designated IS*1548*, was reported in the *hylB* gene [16].

A variety of studies have demonstrated diversity in the distributions of recognized and putative virulence factors among GBS [e.g. 14, 16, 20, 24–26]. Associations have been suggested between the presence or absence of some of these factors and the virulence of GBS [e.g. 14, 16, 20, 22, 24, 27]. By and large, such studies have examined the distribution of each putative virulence factor or of its gene separately without accounting for possible associations between virulence factors or without considering in greater depth the genomic background of the isolates. Furthermore, some investigators in recent years have attempted to use DNA fingerprinting techniques, such as ribotyping, restriction fragment length polymorphism (RFLP) analysis, pulsed-field gel electrophoresis (PFGE), and multilocus enzyme electrophoresis, to reveal associations between evolutionary lineages or infection status and genes encoding putative virulence factors [20, 28–31].

Whether particular strains of GBS are potentially more virulent and thus pose an increased risk to newborns requires the pinpointing of useful genetic or phenotypic markers. Such identification would be of considerable value in optimizing the peripartum management of GBS carrier mothers. The importance of molecular genetic characterization has been highlighted by the observation that strains that are phenotypically negative by sensitive serological tests for the Bca and Bac proteins may harbour the corresponding *bca* and *bac* genes [32, 33]. The further finding that the dominant antigenic epitope of a Bca protein with many tandem repeats is conformational and absent from a Bca protein with few repeats illustrates potential problems in GBS phenotyping using antisera raised to reference and prototype strains [34].

The objectives of the present investigations were (i) to determine the serotype distribution of Irish GBS isolates, including strains associated with invasive disease and carriage isolates, (ii) to study the potential relationships between genes encoding putative virulence factors, either alone or together, with serotype, (iii) to determine relationships between the invasive status (carriage or infection) of isolates and serotype, (iv) to analyse potential relationships between genes encoding virulence factors and the invasive status of isolates, and (v) to examine any association between the number of tandem repeats in the *bca* gene and serotype or invasive status of isolates.

MATERIALS AND METHODS

Subjects and GBS isolates

One hundred and fifty nine GBS isolates were studied. These isolates were recovered during 1997–9 from infants and women with invasive disease and from asymptomatic maternity or gynaecology patients attending the three Dublin maternity hospitals (The Rotunda, The National Maternity Hospital and The Coombe Womens' Hospital). Carriage isolates were chosen at random from among all GBS isolates collected at intervals over the time period of the study, while all positive blood culture and other invasive GBS isolates were routinely chosen over the entire survey period. Asymptomatic infant carriage was not investigated. Isolates were recovered from clinical specimens including blood, high vaginal swabs, umbilical cord swabs, placenta (chorial plate) swabs, and breast milk. Swabs were incubated in 'Bacto™' Todd-Hewitt broth (Difco Laboratories, Detroit, Michigan) for 18–24 h before subculture onto sheep blood agar plates for a further 18–24 h. Blood cultures were processed using the BacTAlert system (Organon Technika). Breast milk samples were cultured on blood agar and incubated in 10% CO_2. GBS isolates were preliminarily identified on the basis of colony morphology, β-haemolysis and Gram staining

https://doi.org/10.1017/S0950268803008847 Published online by Cambridge University Press

and confirmed by Lancefield group antigen determination using a commercial latex agglutination kit (Streptex kit, Murex Biotech). All isolates were stored in serum broth on beads in Cryo-vials (Protect) at −70 °C. *Streptococcus agalactiae* ATCC 13813 (synonyms: NCTC 8181; A. Stableforth 019) was used as a reference GBS strain throughout these studies.

Classification of isolates according to invasive status

Each GBS isolate was classified into one of five invasive categories based on its isolation site and whether or not there was microbiological and histological evidence of infection. Isolates from vaginal swabs with no eventual morbidity to the mother or child were classified as carriage isolates with '0' grade invasion. Organisms were termed invasive if they were recovered from normally sterile body sites and were graded between (a) and (d) according to the signs of infection observed at their recovery site. Isolates recovered from the uterine cervix or placenta (chorial plate) with no histological signs of inflammation or no prominent pus formation were termed (a). Isolates recovered from a wound infection, an infected intra-uterine contraceptive device (IUCD) or from the placenta with histological signs of inflammation were designated (b). An isolate was termed (c), when it was recovered from a patient with mastitis. Isolates of groups (b) and (c) were considered to be of the same invasive status. Isolates recovered by blood culture were indicative of severe invasive disease and were termed (d). Such (d) isolates were designated as the primary cause of sepsis and in some cases were responsible for foetal death.

Serotyping

The 159 GBS isolates were serotyped using a latex agglutination typing kit (Mast Diagnostic) for identification of the capsular polysaccharide antigens Ia, Ib, II, III, IV, V, VI, VII and VIII, according to the manufacturer's instructions. Isolates were termed non-typeable when no agglutination was observed or when there was spontaneous agglutination with more than one of the typing antisera.

Preparation of GBS chromosomal DNA

Strains stored on cryobeads were thawed, subcultured onto sheep blood agar plates and incubated at 37 °C. A single well-isolated colony was inoculated into Todd-Hewitt broth and incubated at 37 °C for 18 h with shaking. Purified chromosomal DNA was prepared from harvested bacteria using the Wizard genomic DNA extraction kit (Promega) with one modification, namely, 60 μl of mutanolysin (10 mg/ml; Sigma) was added at the cell lysis stage.

Amplification of genes encoding putative virulence factors

PCR was used to amplify a specific target sequence of DNA within each of the genes encoding the five putative virulence factors studied. Primer sequences for each of the putative virulence genes were based on the gene sequences that were retrieved from the GenBank database (Table 1). The *bca* gene primers BcarptF and BcarptR contained tails with an *Eco*RI restriction site, namely, 5′-GTCGAATTC-3′, to enable cloning into an appropriate vector plasmid for future studies. Amplification was carried out in 50 μl reaction mixtures containing 20 pmol each of primer, 2·5 mM $MgCl_2$, 250 μM each of dATP, dCTP, dGTP, and dTTP (Promega), 2·5 U of *Taq* DNA polymerase (Promega), 10 mM Tris-HC1, 50 mM KCl, and 0·1% Triton X-100 to which approximately 10 ng of template DNA was added. The cycling programme for PCR amplification comprised 35 cycles of 1 min at 94 °C; 1 min at the appropriate annealing temperature depending on the target gene being amplified (see Table 1); 2 min at 72 °C, followed by a final extension of 7 min at 72 °C. Following amplification, the reaction mixtures were electrophoresed in 0·5 × TBE [10 × TBE: 10·8% (w/v) Trizma base, 5·5% (w/v) boric acid, 9·8% (w/v) EDTA] on 2% (w/v) agarose gels containing ethidium bromide (0·5 μg/ml). All PCR reaction sets included a positive DNA control and distilled water as a negative template and cross-contamination control.

Determination of the number of tandem repeats in the *bca* gene

The primers BcaF and BcaR (Table 1) were used to amplify the entire repeat region of bca^+ isolates to determine the numbers of tandem repeats within the *bca* gene. The PCR reaction mixture and cycling programme were as described above. The molecular sizes of the products were determined relative to a 1-kb marker ladder and the number of repeats calculated on the basis of a repeat size of 246 nt.

https://doi.org/10.1017/S0950268803008847 Published online by Cambridge University Press

Table 1. *Primers designed for use in amplification of virulence factor genes of group B streptococci*

Gene	Name	Primer sequence (5′–3′)	Nucleotide numbers	Reference	T_A (°C)	Product size (bp)
bac	BacF	GCAGTTCATATTGGAAGG	74–91	[9]	45	830
	BacR	AATTTCTTGATCCAGACCAGC	904–924			
bca	BcarptF	GTCGAATTCAGTACCGGATAAAGATAAAT	755–774	[9]	60	246
	BcarptR	GTCGAATTCGTTTTGGTGTCACATGAAGG	982–1000			
	BcaF (repeat region)	CCATCGATATAGTTGCTGCATCTACA				
	BcaR (repeat region)	CGGGATCCATCCTCTTTTTTCTTAGAAAC				
hylB ± IS*1548*	HylIS2F	CCGTTATCAGTTACAGGTC	771–789	[16]	45	714/2031
	Hyl3R	GTCGATGTAAGAACCGTCAGC	1464–1484			
pepB	PepBF	ATCTAGCGATAGGACGGG	816–833	[15]	45	417
	PepBR	CGTTCAGTAAAAGCACGA	1218–1235			
rib	RibF	CAGATGCCGATAAGA	755–769	[12]	39	237
	RibR	TACGCGGATCGACAA	1214–1228			

Dot-blot hybridization

Labelling of PCR products for use as probes in dot-blot hybridization was done using the AlkPhos Direct labelling system (Amersham) according to manufacturer's instructions. Chromosomal DNA was purified as described above. Of this 200 ng was blotted onto positively charged nylon membrane filters (Hybond™ N^+, Amersham). Hybridization was carried out using the AlkPhos Direct hybridization system (Amersham) according to the manufacturer's instructions. Nylon membranes were stripped of bound probe by incubating blots in 0·5% (w/v) SDS solution at 60 °C for 60 min. The blots were then rinsed in 100 mM Tris (pH 8·0) for 5 min, wrapped in Saran Wrap and stored at −20 °C between reprobings.

Automated fluorescent DNA sequencing

Fluorescence-based sequence analysis of PCR-amplified double-stranded DNA fragments [35, 36] was carried out using a BigDye terminator cycle sequencing ready reaction kit (ABI Prism, Applied Biosystems) according to the manufacturer's instructions on an automated fluorescent DNA sequencer (ABI 310, Applied Biosystems).

Statistical analyses

GBS attack rates were estimated with binomial confidence intervals and compared by Fisher's Exact test. Data for serotype and infection status in relation to the presence or absence of genes encoding putative virulence factors were analysed using frequency tables, Chi-squared (χ^2) tests of association and goodness of fit, Fisher's Exact tests, and hierarchical log-linear analysis. Analysis of the *bca* gene tandem-repeat data was done by the Kolmogorov–Smirnov Goodness of Fit test, the Mann–Whitney *U* test and Kruskal–Wallis non-parametric tests. $P<0{\cdot}05$ was considered statistically significant. The analyses were carried out with the SPSS statistical analysis package for Windows Version 10.0.

RESULTS

Serotype distribution of GBS isolates

The overall serotype distribution of the isolates studied was Ia = 19·5%, Ib = 18·9%, II = 10·7%, III = 29·5%, IV = 1·9%, and V = 15·1%. Seven isolates (4·4%) were designated as non-typeable (Table 2) and of these, four exhibited no agglutination with any of the typing antisera, and three isolates agglutinated with two antisera. Capsular polysaccharide type Ib GBS was found to be the most frequent invasive type with 22 of the 30 serotype Ib isolates (73·3%) designated as invasive. Although serotype III was represented by 29·5% of the isolates, only 36·2% of this serotype was associated with invasive disease. Excluding the non-typeable and serotype IV strains because of the small numbers of these isolates, the serotypes of the GBS isolates were significantly associated with the invasive status of isolates ($\chi^2 = 14{\cdot}9$, D.F. = 4, $P = 0{\cdot}005$). This result is mainly accounted for by serotype Ib having a higher rate of invasive isolates than expected statistically and serotypes II and III having a lower rate than expected.

https://doi.org/10.1017/S0950268803008847 Published online by Cambridge University Press

Table 2. *Serotype distribution of carriage and invasive isolates of 159 GBS isolates*

Serotype	No. of carriage isolates '0'	No. of invasive isolates*				
		'a'	'b'	'c'	'd'	Total
Ia	15	1	5	3	7	31
Ib	8	7	5	—	10	30
II	13	1	1	—	2	17
III	30	4	4	—	9	47
IV	2	1	—	—	—	3
V	14	4	3	—	3	24
NT	5	—	—	—	2	7
Totals	87	18	18	3	33	159

* For designation of codes see Materials and Methods.

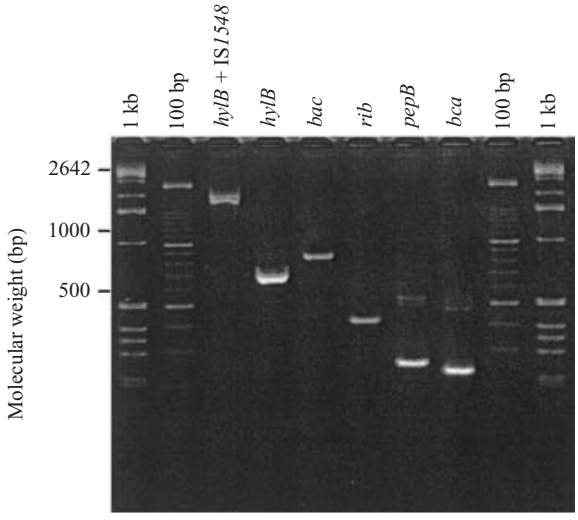

Fig. 1. PCR amplification of the *bac*, *bca*, *hylB*, *hylB* (IS*1548*), *pepB* and *rib* genes of group B streptococci.

Within each serotype, except for serotype IV and the non-typeable strains, isolates displayed a spectrum of invasive capacity (Table 2). However when the categories of invasiveness were variously combined, no statistically significant association of degree of invasiveness with serotype was demonstrated.

Detection of the putative virulence factor genes by PCR

Figure 1 shows that PCR products of the expected sizes were obtained using the primers described in Table 1. Under the PCR conditions used the primer pair BcarptF and BcarptR gave only a PCR product corresponding to one *bca* repeat sequence with all *bca*$^+$ GBS.

PCR with primers Hy1IS2F and Hy13R flanking the *hylB* gene amplified either a 714 bp or a 2031 bp PCR product (Fig. 1). The larger PCR product was detected almost exclusively from serotype III isolates with two exceptions, one isolate of serotype II and one of serotype V. The difference in the *hylB* product sizes was identified as due to the presence of the IS*1548* element through sequencing of the two PCR products [16].

Association of genotypic traits with serotype

The GBS isolates were examined by PCR for the presence of five genes encoding putative virulence factors, namely, *bac*, *bca*, *hylB*, *pepB* and *rib*. The occurrence of all genes in strains was further confirmed by dot-blot hybridization tests using specific, sequence-verified gene probes comprising alkaline phosphatase-labelled PCR products. Expression studies were not performed. Table 3 shows the numbers and percentages of positives for each of the five genotypic traits [*bac*, *bca*, *hylB* (IS*1548*), *pepB* and *rib*] examined within each serotype. Because the *hylB* gene was detected in all the GBS isolates but one, it was excluded from Table 3. Serotype IV and the non-typeable strains were also excluded because of the small numbers (total of 10 isolates accounting for 6·3% of strains). The *bac*, *bca*, *hylB* (IS*1548*) and *rib* genes had statistically significant associations with serotype.

The percentage of isolates positive for the *pepB* gene within each serotype was relatively constant (range 75·0%–83·9%). Hence, there was no association between this gene and serotype (all adjusted residuals are small). Serotype Ib had a higher proportion than expected of *bac*$^+$ and *bca*$^+$ isolates and serotypes III and V had a very similar pattern of occurrence of *bac*$^+$, *bca*$^+$ and *rib*$^+$ isolates. However, when the association between serotype and the four patterns made of the *bac* and *bca* genes alone was examined, the highest residuals were found with joint occurrence of like signs for these genes. A clear excess occurrence over expectation was found for the pattern *bac*$^-$ *bca*$^-$ in serotypes III and V and for the pattern *bac*$^+$ *bca*$^+$ in serotype Ib.

Putative virulence genes versus invasive status

All except one (an isolate of serotype III) of the 159 GBS isolates possessed the *hylB* gene. Forty-two (26·4%) had a *hylB* gene with the IS*1548* insertion

https://doi.org/10.1017/S0950268803008847 Published online by Cambridge University Press

Table 3. *Frequencies of occurrence of genotypic traits within serotypes of 149 GBS isolates*

	No. of isolates of given genotype (%)					
Serotype	*bac*$^+$	*bca*$^+$	*hylB* (IS*1548*)$^+$	*pepB*$^+$	*rib*$^+$	Total
Ia	9 (29·0)	17 (54·8)	0	26 (83·9)	17 (54·8)	31
Ib	21 (70·0)	26 (86·7)	0	23 (76·7)	17 (56·7)	30
II	8 (47·1)	9 (52·9)	1 (5·9)	13 (76·5)	11 (64·7)	17
III	14 (29·8)	14 (29·8)	40 (85·1)	39 (83·0)	39 (83·0)	47
V	6 (25·0)	7 (29·2)	1 (4·2)	18 (75·0)	20 (83·3)	24
Total	58 (38·9)	73 (49·0)	42 (28·2)	119 (79·9)	104 (69·8)	149
P	0·002	<0·001	<0·001	0·868	0·018	

Table 4. *Interaction between the presence and absence of the* rib, bca *and* bac *genes*

Pattern	*bac*	*bca*	*rib*	Observed frequency	Expected frequency*	Residual	Adjusted residual†	No. of isolates by serotype							No. of isolates by status		
								Ia	Ib	II	III	IV	V	NT	Invasive	Carriage	Total
A	−	−	−	17	15·26	−1·74	0·61	6	0	3	5	2	1	0	7	10	17
B	−	−	+	39	32·32	6·68	1·96	5	0	3	19	0	13	0	18	22	40
C	−	+	−	16	15·85	0·15	0·05	6	3	2	2	0	2	1	6	10	16
D	−	+	+	25	33·56	−8·56	−2·50	5	6	1	7	0	3	3	11	14	25
E	+	−	−	2	9·76	−7·76	−3·08	0	1	1	0	0	0	0	2	0	2
F	+	−	+	20	20·66	−0·66	−1·21	3	3	1	9	0	3	0	7	12	19
G	+	+	−	16	10·13	5·87	2·30	2	9	0	1	1	1	2	11	5	16
H	+	+	+	24	21·45	2·55	0·81	4	8	6	4	0	1	1	10	14	24

* Expected frequencies are based on a model of complete mutual independence.
† An adjusted residual approximately >2 is statistically significant at $P=0·05$.

sequence. The 2031 bp product was amplified from 40 of the 47 (85·1%) serotype III isolates. All but one of the invasive serotype III isolates possessed this insertion sequence (16 out of 17 invasive isolates *vs.* 24 out of 30 carriage isolates).

Analysis of the results for the other four genes showed that the population of GBS was genetically diverse. Fisher's Exact tests of association between the presence of each of these genes and the invasive status of isolates demonstrated that none of these associations was statistically significant (the smallest *P* value was 0·39 for the *rib* gene).

Associations between the occurrences of the *bac*, *bca*, *pepB*, *hylB* (IS*1548*) and *rib* genes

Unadjusted pairwise associations (Fisher's Exact test) between the *bac*, *bca*, *hylB* (IS*1548*), *pepB* and *rib* genes revealed that most of these genes were associated to some extent. Of the 10 associations, 4 were statistically significant and 2 were borderline significant. These findings point to more complex associations. Accordingly, a hierarchical log-linear model examined partial associations. This analysis gave two statistically significant associations, *viz. bca vs. hylB* (IS*1548*) ($P=0·0004$) and *bac vs. bca* ($P=0·014$) and also indicated a potential three-way interaction between *bac vs. bca vs. rib* ($P=0·043$).

The observed and expected frequencies of the eight combinations of the presence and absence of these three genes are shown in Table 4. The combinations primarily responsible for the three-way interaction are (i) where the *bac* and *bca* genes are both absent or both present and the *rib* gene is, respectively, present or absent (patterns B and G), and (ii) where the *bca* and *rib* genes are both present or both absent and the *bac* gene is, respectively, present or absent (patterns D and E). Patterns B and G are more frequent than expected, whereas patterns D and E occur less frequently than expected. The associations among the

https://doi.org/10.1017/S0950268803008847 Published online by Cambridge University Press

Table 5. *Determination of the numbers of bca tandem repeats in 59* bca$^+$ *GBS isolates*

Number of *bca* repeats	Size of PCR products (bp)	No. of isolates by serotype					No. of isolates by status		
		Ia	Ib	II	III	V	Invasive	Carriage	Total
1	918	1	—	1	—	—	1	1	2
2	1164	3	2	—	2	2	3	6	9
3	1410	—	4	1	—	—	4	1	5
4	1656	—	1	2	2	1	2	4	6
5	1902	2	—	—	1	—	1	2	3
6	2148	3	1	—	1	—	3	2	5
7	2394	1	1	—	—	—	—	2	2
9	2886	1	2	1	3	—	4	3	7
10	3132	—	—	—	1	—	—	1	1
11	3378	2	3	—	—	—	3	2	5
12	3624	1	2	—	—	1	2	2	4
14	4116	1	5	1	—	—	5	2	7
15	4362	—	1	—	—	—	—	1	1
16	4608	—	1	—	1	—	1	1	2
Mean repeats		6·6	9·1	5·8	6·9	5·0	7·9	7·0	7·4
Median repeats		6·0	11·0	4·0	6·0	3·0	9·0	6·0	6·0

The molecular sizes of the products were determined relative to a 1-kb marker ladder and the number of repeats calculated on the basis of a repeat size of 246 bp.

genes may be modified after adjusting for their interactions with the serotypes. However, a larger data set would be required to perform this analysis.

Association of *bca* tandem repeats with serotype and invasive status

A stratified random sample of 59 of the 73 *bca*$^+$ GBS excluding serotype IV and non-typeable isolates (Table 5; Ia = 15, Ib = 23, II = 6, III = 11, V = 4) was characterized to determine the numbers of tandem repeats within their *bca* genes using primers that flanked the repeat regions. The PCR primers BcaF and BcaR gave rise to different sized amplimers depending on the numbers of tandem repeats within the *bca* gene (Fig. 2). The tandem-repeat number ranged from 1–16 and an overall mean of 7·4, and a median of 6·0, but repeat numbers 8 and 13 were absent. Serotype Ib was the only serotype in which more than half of the strains contained more tandem repeats than the overall mean (7·4 repeats). The data obtained (not shown) suggest that the occurrence of any number of repeats from 1–16 is not equally likely. When the distribution of the numbers of tandem repeats for the whole sample was compared with a uniform distribution, the difference was statistically significant (Kolmogorov–Smirnov Goodness of Fit test, $P = 0{\cdot}01$).

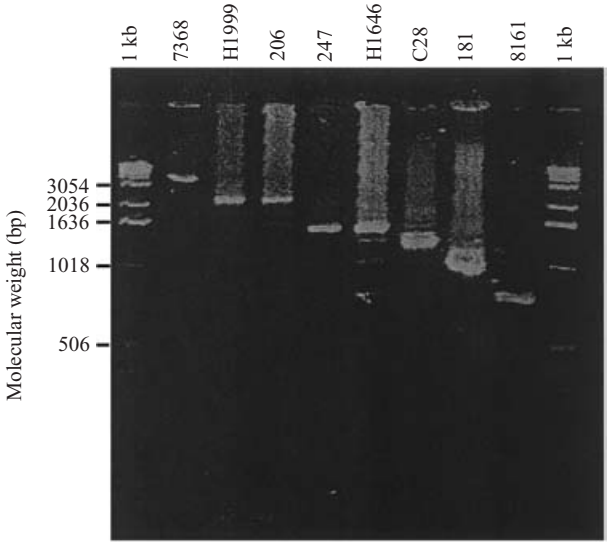

Fig. 2. PCR amplification of the tandem repeats in *bca*$^+$ GBS.

Potential associations between the number of tandem repeats and the serotype or invasive status of GBS were analysed. No difference was found in the distributions of tandem repeats between invasive and carriage isolates (Mann–Whitney *U* test, $P = 0{\cdot}7$), or among the distributions of tandem repeats in serotypes Ia, Ib, II and III (Kruskal–Wallis test, $P = 0{\cdot}26$). Serotype Ib GBS possessed higher numbers

https://doi.org/10.1017/S0950268803008847 Published online by Cambridge University Press

of *bca* tandem repeats than other serotypes, as indicated by its cluster of isolates with repeats from 11–16 and a substantially higher median, and the fact that more than half of the strains contained more tandem repeats than the overall mean of 7·4 repeats. There was a statistically significant difference between its tandem-repeat distribution and that of the other four serotypes combined (Mann–Whitney *U* test, $P = 0{\cdot}034$).

DISCUSSION

GBS isolates of all serotypes have been associated with neonatal infections [2–4, 37]. However, there seems to be considerable demographic, geographic and temporal variation with respect to the predominant GBS serotypes present in the human population [37, 38]. For instance, although serotype III isolates have been the most prevalent serotype in invasive disease in the United States in the past, strains of serotype V are now emerging as a major clinical problem [39, 40], whilst in Zimbabwe serotype V colonization is as prevalent in pregnancy as serotype III [41]. Newly emerging capsular polysaccharides are also a problem, as seen in Japan where GBS of serotypes VI and VIII are now common carriage isolates [42]. There are limited data available on the serotypes of the carriage and invasive GBS population in Ireland; however, the serotype distribution of carriage GBS among pregnant women at the National Maternity Hospital, Dublin reported by Kieran et al. [43] was similar to that found here.

The present study identified serotype III as the most common colonization isolate while serotype Ib was the predominant invasive type. Attempts to correlate the presence of putative virulence factors with GBS disease or its severity have been made previously [14, 16, 20, 22, 24, 27–31]. In most of the latter studies the roles of particular putative virulence factors have been analyzed separately without taking possible linkages between factors into account. This approach may bias conclusions on the roles of putative virulence factors in disease by not accounting for the confounding effects of these factors that have not been tested for and by ignoring potential synergistic effects between these factors.

Based on serological tests, GBS strains of different polysaccharide type have been shown to express particular cell-surface proteins. For example, the Bca protein is found largely on strains of serotypes Ia, Ib and II, the Rib protein on strains of serotype III and the Alp3 protein on most isolates of serotypes V and VIII [e.g. 25, 42, 44, 45]. The genotypic data obtained in the current study are somewhat at variance with these general phenotypic trends. This study found that almost 50% of serotype III strains possessed the *bac* or *bca* genes or both, while 55–65% of the strains of serotypes Ia, Ib and II possessed the *rib* gene. In contrast, the prevalence of the *bac* and *bca* genes in the same population of serotypes Ia, Ib and II is in concordance with the PCR, hybridization and serological data in several studies over the past decade [e.g. 25, 26, 33, 44, 45]. The same is true of the prevalence of the *rib* gene and the Rib protein in serotype III strains.

The genotyping observations discussed above are surprising and merit further study and clarification by other workers. There are three possible explanations for these findings. The genotypic data may represent novel observations that have been missed heretofore because of emphasis on phenotypic testing. Alternatively, these observations may indicate the presence of silent *bac*, *bca* and *rib* genes in these serotypes. Finally, the findings may indicate the presence of genes encoding as yet unrevealed proteins with sequence similarities to the *bca*, *bac* and *rib* genes.

In the present study, isolates were screened for five genes using PCR amplification to assess if there was any association between the presence or absence of genes encoding putative virulence factors and serotype or invasiveness. Correlation between the distributions of the *bac*, *bca*, *hylB* (IS*1548*) and *rib* genes and serotype revealed significant associations. Moreover, the joint occurrence of the *bac* and *bca* genes was associated with serotype Ib and their joint absence with serotypes III and V, confirming published epidemiological data [25, 33, 44, 45]. No phenotypic characterization of isolates for gene products was performed in the present study. Thus, it might be argued that to analyse associations with invasiveness based solely on genotypic characteristics could be misleading. However as pointed out by Maeland et al. [33], lack of expression of a cell-surface protein *in vitro* does not mean absence of expression *in vivo*, since the environmental stimuli and the nature of the regulatory cascades controlling GBS gene expression *in vivo* are unknown. The present study also investigated the mutual interactions of these putative virulence genes in a hierarchical log-linear analysis. This analysis pointed to a potential three-way interaction between the presence or absence of the *bac*, *bca* and *rib* genes. These three genes were found to be present or absent

in specific patterns in some serotypes. No statistically significant associations were demonstrated between invasiveness and serotype or between the presence of each of the five genes sought and invasive status.

Interestingly, the majority of the Irish serotype III isolates, whether associated with invasion or carriage, were found to possess a *hylB* gene bearing the insertion sequence IS*1548*. The inactivation of the *hylB* gene by this insertion element [16] and the frequency of occurrence of the *hylB* (IS*1548*) gene in both carriage and invasive isolates suggest that hyaluronidase does not play an important role in the invasiveness of the serotype III isolates.

The numbers of repeats in the C-terminus of the *bca* gene were examined herein. No such study has been published previously. Given the experimental evidence for a small number of repeats being associated with avoidance or thwarting of the immune response [34, 46–48], a correlation between lower numbers of repeats and invasiveness might be hypothesized. This, however, was not the case. The only statistically significant difference found was the higher tandem-repeat distribution in serotype Ib compared to the other serotypes combined. Thus, it remains to be established whether the published *in vitro* and experimental infection observations on repeat number are relevant in human infections.

Stemming from the GBS population studies of Musser et al. [20], which identified two lineages of serotype III isolates based on MLEE analysis of 11 genetic loci, other workers have attempted to identify virulence-associated properties or alleles of genes with these two clusters of strains [e.g. 16, 29–31, 49]. While hyaluronidase production was associated with the so-called high virulence ET1 clone of Musser et al. [20], more recent studies have associated an IS-inactivated *hylB* gene and a hyaluronidase-negative phenotype with serotype III strains from invasive neonatal infections and endocarditis [16, 29, 31, 50]. While it is very difficult to compare frequencies of occurrence of the *hylB* (IS*1548*) gene between studies because of the differing populations of strains studied, a surprising feature of the Dublin serotype III isolates was the high incidence of this gene among both carriage and invasive isolates. In addition, 22 of our 47 serotype III isolates very closely resemble the division I Danish strains of Hauge et al. [29], being additionally *bac*$^-$ *bca*$^-$. Three of the so-called 'low virulence strains' of Musser et al. [20] possessed the *hylB* (IS*1548*) and clustered with the Danish isolates [29]. Thus, the *hylB* (IS*1548*) gene may be a useful marker for the pathogenic potential of serotype III clones for which the risk of invasive infection appears to be high, but definition of clear virulence determinants remains elusive.

It is clear from the present and other studies [20, 29–31] that the pathogenicity of GBS strains is either multifactorial or that some new or as yet undiscovered virulence factor may be important [51, 52]. In conclusion, our genotypic findings suggest that some previously established associations between putative virulence factors and GBS disease require further clarification and that individual gene markers are not necessarily good predictors of invasive capacity, as previously shown for PCR detection of the *bac* gene [27]. PCR detection of several putative virulence factor genes, along with serotyping, may eventually prove to be a more appropriate method for identifying potentially invasive GBS isolates in the clinical laboratory [32, 33, 45, 51].

ACKNOWLEDGEMENTS

We thank the laboratory staff of The Rotunda, The National Maternity Hospital and The Coombe Women's Hospital for their cooperation and assistance with this study.

REFERENCES

1. Schuchat A, Schrag S. Group B streptococci: from emerging infection to prevention success story. In: Scheld WM, Craig WA, Hughes JM, eds. Emerging pathogens 4. Washington, DC: ASM Press, 2000: 107–120.
2. Edwards MS, Baker CJ. Group B streptococcal infections. In: Remington JS, Klein JO, eds. Infectious diseases of the fetus and newborn infant, 5th edn. Philadelphia: WB Saunders Co., 2001: 1091–1156.
3. McCracken GH. Group B streptococci: a new challenge in neonatal infections. J Pediatr 1973; **82**: 703–706.
4. Boyer KM. Neonatal group B streptococcal infections. Curr Opin Pediatr 1995; **7**: 13–18.
5. Schuchat A, Hilger T, Zell E, et al. Active bacterial core surveillance of the emerging infections program network. Emerg Infect Dis 2001; **7**: 92–99.
6. Jennings HJ. Capsular polysaccharides as vaccine candidates. Curr Top Microbiol Immunol 1990; **150**: 97–127.
7. Jerlström PG, Talay SR, Valentin-Weigand P, Timmis KN, Chhatwal GS. Identification of an immunoglobulin A binding motif located in the β-antigen of the C protein complex of group B streptococci. Infect Immun 1996; **64**: 2787–2793.
8. Lancefield RC, McCarty M, Everly WN. Multiple mouse-protective antibodies directed against group B

streptococci. Special reference to antibodies effective against protein antigens. J Exp Med 1975; **142**: 165–179.
9. Michel JL, Madoff LC, Kling DE, Kasper DL, Ausubel FM. Cloned alpha and beta C-protein antigens of group B streptococci elicit protective immunity. Infect Immun 1991; **59**: 2023–2028.
10. Hedén LO, Frithz E, Lindahl G. Molecular characterisation of an IgA receptor from group B streptococci: sequence of the gene, identification of a proline-rich region with unique structure and isolation of N-terminal fragments with IgA-binding capacity. Eur J Immunol 1991; **21**: 1481–1490.
11. Maeland JA, Brakstad OG, Bevanger L, Kvam AT. *Streptococcus agalactiae* β gene and gene product variations. J Med Microbiol 1997; **46**: 999–1005.
12. Stålhammar-Carlemalm M, Stenberg L, Lindahl G. Protein Rib: a novel group B streptococcal cell surface protein that confers protective immunity and is expressed by most strains causing invasive infections. J Exp Med 1993; **177**: 1593–1603.
13. Wästfelt M, Stålhammar-Carlemalm M, Delisse A-M, Cabezon T, Lindahl G. Identification of a family of streptococcal surface proteins with extremely repetitive structure. J Biol Chem 1996; **271**: 18892–18897.
14. Jackson RJ, Dao ML, Lim DV. Cell-associated collagenolytic activity by group B streptococci. Infect Immun 1994; **62**: 5647–5651.
15. Lin B, Averett WF, Novak J, et al. Characterization of PepB, a group B streptococcal oligopeptidase. Infect Immun 1996; **64**: 3401–3406.
16. Granlund M, Öberg L, Sellin M, Norgren M. Identification of a novel insertion element, IS*1548*, in group B streptococci, predominantly in strains causing endocarditis. J Infect Dis 1998; **177**: 967–976.
17. Milligan TW, Baker CJ, Strauss DC, Mattingly SJ. Association of elevated levels of extracellular neuraminidase with clinical isolates of type III group B streptococci. Infect Immun 1978; **21**: 738–746.
18. Pritchard DG, Lin B. Group B streptococcal neuraminidase is actually a hyaluronidase. Infect Immun 1993; **61**: 3234–3239.
19. Pritchard DG, Lin B, Willingharn TR, Baker JR. Characterization of the group B streptococcal hyaluronate lyase. Arch Biochem Biophys 1994; **315**: 431–437.
20. Musser JM, Mattingly SJ, Quentin R, Goudeau A, Selander RK. Identification of a high-virulence clone of type III *Streptococcus agalactiae* (group B *Streptococcus*) causing invasive neonatal disease. Proc Natl Acad Sci USA 1989; **86**: 4731–4735.
21. Chmouryguina I, Suvorov A, Ferrieri P, Cleary PP. Conservation of the C5a peptidase genes in group A and B streptococci. Infect Immun 1996; **64**: 2387–2390.
22. Gibson RL, Nizet V, Rubens CE. Group B streptococcal beta-hemolysin promotes injury of lung microvascular endothelial cells. Pediatr Res 1999; **45**: 626–634.
23. Yamamoto S, Miyake K, Koike Y, et al. Molecular characterization of type-specific capsular polysaccharide synthesis genes of *Streptococcus agalactiae* type Ia. J Bacteriol 1999; **181**: 5176–5184.
24. Chun CSY, Brady U, Boyle MDP, Dillon HC, Ayoub EM. Group B streptococcal C protein-associated antigens: association with neonatal sepsis. J Infect Dis 1991; **163**: 786–791.
25. Johnson DR, Ferrieri P. Group B streptococcal Ibc protein antigen: distribution of two determinants of wild-type strains of common serotypes. J Clin Microbiol 1984; **19**: 506–510.
26. Suvorov A, Dmitriev A, Ustinovitch I, Schalen C, Totolian AA. Molecular analysis of clinical group B streptococcal isolates by alpha and beta gene probes. FEMS Immunol Med Microbiol 1997; **17**: 149–154.
27. Berner R, Bender A, Rensing C, Forster J, Brandis M. Low prevalence of the immunoglobulin-A-binding beta antigen of the C protein among *Streptococcus agalactiae* isolates causing neonatal sepsis. Eur J Clin Microbiol Infect Dis 1999; **18**: 545–550.
28. Takahashi S, Nagano Y, Nagano N, Fujita K, Taguchi F, Okuwaki Y. Opsonisation of group B streptococci and restriction endonuclease digestion patterns of their chromosomal DNA. J Med Microbiol 1993; **38**: 191–196.
29. Hauge M, Jespersgaard C, Poulsen K, Kilian M. Population structure of *Streptococcus agalactiae* reveals an association between specific evolutionary lineages and putative virulence factors but not disease. Infect Immun 1996; **64**: 919–925.
30. Quentin R, Huet H, Wang F-S, Geslin P, Goudeau A, Selander RK. Characterization of *Streptococcus agalactiae* strains by multilocus enzyme genotype and serotype: identification of multiple virulent clone families that cause invasive neonatal disease. J Clin Microbiol 1995; **33**:2576–2581.
31. Rolland K, Marois C, Siquier V, Cattier B, Quentin R. Genetic features of *Streptococcus agalactiae* strains causing severe neonatal infections, as revealed by pulsed-field gel electrophoresis and *hylB* gene analysis. J Clin Microbiol 1999; **37**: 1892–1898.
32. Maeland JA, Bevanger L, Iversen G, Valsoe Lyng R. *bca*, beta gene, and gene product divergency in reference and prototype strains of *Streptococcus agalactiae*. Clin Diag Lab Immunol 1999; **6**: 986–988.
33. Maeland JA, Brakstad OG, Bevanger L, Krokstad S. Distribution and expression of *bca*, the gene encoding the C alpha protein, by *Streptococcus agalactiae*. J Med Microbiol 2000; **49**: 193–198.
34. Gravekamp C, Horensky DS, Michel YE, Madoff LC. Variation in repeat number within the alpha C protein of group B streptococci alters antigenicity and protective epitopes. Infect Immun 1996; **64**: 3576–3583.
35. Rao VB. Strategies for direct sequencing of PCR-amplified DNA. PCR Methods Appl 1994; **4**: S15–S23.
36. Rao VB. Direct sequencing of polymerase chain reaction-amplified DNA. Anal Biochem 1994; **216**: 1–14.
37. Christensen KK, Christensen P, Ferrieri P. Neonatal group B streptococcal infections. Antibiot Chemother 1985; **35**: 1–17.
38. Paoletti U, Bradford J, Paoletti LC. A serotype VIII strain among colonizing group B streptococcal isolates

in Boston, Massachusetts. J Clin Microbiol 1999; **37**: 3759–3760.
39. Blumberg KM, Stephens S, Modansky M, et al. Invasive group B streptococcal disease: the emergence of serotype V. J Infect Dis 1996; **173**: 365–373.
40. Hickman ME, Rench MA, Ferrieri P, Baker CJ. Changing epidemiology of group B streptococcal colonization. Pediatrics 1999; **104**: 203–209.
41. Moyo SR, Mudzori J, Tswana SA, Maeland JA. Prevalence, capsular type distribution, anthropometric and obstetric factors of group B streptococcus (*Streptococcus agalactiae*) colonization in pregnancy. Cent Afr J Med 2000; **46**: 115–120.
42. Lachenauer CS, Kasper DL, Shimada J, et al. Serotypes VI and VIII predominate among group B streptococci isolated from pregnant Japanese women. J Infect Dis 2000; **179**: 1030–1033.
43. Kieran E, Matheson M, Mann AG, Efstratiou AA, Butler K, Gorman W. Group B streptococcus (GBS) colonisation among expectant Irish mothers. Ir Med J 1998; **91**: 21–22.
44. Lachenauer CS, Madoff LC. A protective surface protein from type V group B streptococci shares N-terminal sequence homology with the alpha C protein. Infect Immun 1996; **64**: 4255–4260.
45. Kvam AI, Efstratiou A, Bevanger L, et al. Distribution of serovariants of group B streptococci in isolates from England and Norway. J Med Microbiol 1995; **42**: 246–250.
46. Madoff LC, Michel JL, Gong EW, Kling DE, Kasper DL. Group B streptococci escape host immunity by deletion of tandem repeat elements of the alpha C protein. Proc Natl Acad Sci USA 1996; **93**: 4131–4136.
47. Gravekamp C, Kasper DL, Michel JL, Kling DE, Carey V, Madoff LC. Immunogenicity and protective efficacy of the alpha C protein of group B streptococci are inversely related to the number of repeats. Infect Immun 1997; **65**: 5216–5221.
48. Gravekamp C, Rosner B, Madoff LC. Deletion of repeats in the alpha C protein enhances the pathogenicity of group B streptococci in immune mice. Infect Immun 1998; **66**: 4347–4354.
49. Sellin M, Håkansson S, Norgren M. Phase-shift of polysaccharide capsule expression in group B streptococci, type III. Microb Pathogen 1995; **18**: 401–415.
50. Sellin M, Linderholm M, Norgren M, Håkansson S. Endocarditis caused by a group B *Streptococcus* strain, type III, in a nonencapsulated phase. J Clin Microbiol 1992; **30**: 2471–2473.
51. Brodeur BR, Boyer M, Charlebois I, et al. Identification of group B streptococcal Sip protein, which elicits cross-protective immunity. Infect Immun 2000; **68**: 5610–5618.
52. Granelund M, Michel F, Norgren M. Mutually exclusive distribution of IS*1548* and GBSi1, an active group II intron identified in human isolates of group B streptococci. J Bacteriol 2001; **183**: 2560–2569.