Variation in the Number of Tandem Repeats and Profile of Surface Protein Genes among Invasive Group B Streptococci Correlates with Patient Age

Yueh-Ren Ho¹,⁷, Chien-Ming Li², Hsin-Pi Su³, Jane-Hong Wu⁵, Yu-Ching Tseng¹, Yuh-Jyh Lin⁶, Jiunn-Jong Wu⁴,⁷∗

Departments of Biochemistry and Molecular Biology,¹ Medical Laboratory Science and Biotechnology,⁴ Microbiology and Immunology,⁵ Pediatrics⁵ and Institute of Basic Medical Sciences,⁷ College of Medicine, National Cheng Kung University, Tainan, Taiwan

Division of Infectious Diseases, Department of Internal Medicine,² Chi Mei Medical Center, Tainan and Laboratory for Bacteriology and Mycology,³ Center for Research and Diagnostics, Center for Disease Control, Taipei, Taiwan

Keywords: Tandem repeat, rib, GBS, age

Corresponding author: Jiunn-Jong Wu, Department of Medical Laboratory Science and Biotechnology, College of Medicine, National Cheng Kung University, No. 1 University Road, Tainan, 701, Taiwan

Tel: 886-6-2353535 ext 5605

Fax: 886-6-2363956

Email: jjwu@mail.ncku.edu.tw

Running head: Tandem repeats in rib gene of GBS and patient age
ABSTRACT

The average number of tandem repeats of the *rib* gene (encoding the Rib surface protein) in invasive group B streptococci from 29 neonates was smaller than those from 20 adults (6.8 vs. 8.6; *P* < 0.05), implying a distinct contribution of immunity towards this age-related variation.
Group B streptococcus (*Streptococcus agalactiae*, [GBS]) is a major pathogen causing meningitis of neonates and sepsis of nonpregnant adults (3-7), with Ia, Ib, and II-VIII serotypes based on the capsular polysaccharides (13). Surface proteins of GBS are virulence factors and alternative vaccine candidates (11, 14). Alpha-like-protein (Alp) family of surface proteins, including Alpha C, Rib, Alp2, Alp3, Alp4, and epsilon/Alp1, exhibit a ladder-like pattern in PCR amplification of DNA (17). Alpha C protein is often expressed by non-type III GBS, and contributes to resistance to opsonophagocytosis, and elicits protective immunity (15); while Rib protein is mostly expressed by serotype III, and confers protective immunity in animal experiments (15). Both proteins exhibit size variability corresponding to length of tandem repeats in the encoding genes (12); such size variation may arise through *recA* independent slipped-strand mispairing during DNA replication (18).

The number of tandem repeats of Alpha C is important in the pathogenesis of GBS diseases (15). Several studies have indicated that internal deletion of the *bca* gene encoding Alpha C protein has been observed in GBS collected from mice pretreated with anti-serum, and renders an escape from antibodies elicited to the native protein (9, 10, 16); in other words, fewer tandem repeats can offer GBS immune evasion (9). Nevertheless, clinical isolates often contain multiple tandem repeats, suggesting that GBS with large number of repeats has enhanced virulence in humans (15). The aims of the present study were to subtype invasive GBS of bloodstreams from neonates and adults by surface protein genes of the Alp family
and capsular polysaccharides, and characterize the tandem repeat numbers within the rib gene.

To investigate the genotypes of invasive GBS in southern Taiwan, a total of 156 strains were isolated from the bloodstream of 42 neonates (≤ 90 days) and 114 nonpregnant adults (≥ 16 years) at two medical centers during 1994-2004. Serotypes of capsular polysaccharides were determined using the Group-B Streptococci Typing Antisera (Denka Seiken, Nijgata, Japan). The type V strain 2603 V/R and type III strain NEM 316 were used as reference strains for analysis of serotypes and surface protein genes. The serotype distribution was as follows: Ia, 17.9%; Ib, 10.3%; II, 4.5%; III, 37.2%; IV, 0.6%; V, 20.5%; VI, 2.6%; and nontypeable, 6.4%. Genes encoding Alp family surface proteins, including bca, alp2/3, alp4, rib, and epsilon/alp1 were amplified by multiplex PCR as previously described (1), with resulting sizes of 398, 334, 110, 295, and 200 bp, respectively. The distribution of surface protein genes tested was as follows: rib, 31.4%; alp2/3, 26.3%; bca, 26.3%; epsilon/alp1, 10.3%; and none, 5.8%. Most alp2/3 genes (92.7%, 38/41; \( P = 0.0008 \); Fisher’s test) and bca genes (85.4%, 35/41; \( P = 0.042 \); Fisher’s test) were harbored in GBS from adult patients; while 59.2% (29/49) of rib genes were in GBS from neonates (\( P < 10^{-9} \); Fisher’s test). Only one of these genes was detected in each strain, and alp4 was not detected in our study. Combined with the profiles of surface protein genes and capsular polysaccharides, 25 serovariants were obtained. Of the GBS strains collected from adults, the serovariants Ib/bca
(85.7%; 12/14), III/rib (60%; 15/25), and V/valp2/3 (70%; 21/30) were the three most common subtypes. Only 9 subtypes were identified among strains from neonates, with the dominant serovariant III/rib accounting for 87.9% (29/33). Analysis of potentially age-related correlations by Fisher’s test showed that associations were statistically significant between III/rib and neonatal early-onset infections (53.8%; \( P = 0.049 \)), and late-onset infections (75.9%; \( P < 10^{-9} \)), and between GBS infections of adults aged 16-50 years and Ib/bca (18.8%; \( P = 0.04 \)) and V/valp2/3 (28.1%; \( P = 0.019 \)). These findings suggested that specific serovariants could contribute to the pathogenesis of GBS infections in different age groups.

To characterize the tandem repeat numbers in the rib gene of GBS, the full length gene was PCR amplified from all of the 49 rib-harboring isolates using primers 5’-CTGAAGTAATTTCAGGAAGTGC-3’ and 5’-ATCCTCTTTTTTTCTTAGAAACAGATA-3’. The number of tandem repeats was determined by subtracting the size of the non-repeat region from the size of the main PCR product and dividing by a rib repeat size of 237 bp. Our data showed that the number of tandem repeats ranged from 0 to 14, with an average of 7.6. To our knowledge, this is the first report concerning the distribution of the rib gene in invasive GBS. Comparatively, Gravekamp et al. (9) reported that the repeat numbers of alpha C protein followed normal distribution with a mean of 9-10.

Although the repeat numbers of Alpha C protein are not different between invasive and
 carriage strains (2), variation in repeat number can alter antigenicity and protective epitopes (8). This study revealed that the average number of repeats in rib genes was 6.8 in invasive GBS strains from neonates and 8.6 in those from adults ($P = 0.045$, one-tailed Student’s $t$-test), and their medians were remarkably different for a range of 0-14 (6 vs. 9, respectively) (Figure). This indicated that the variation of repeat numbers in rib was age-related; variants with few repeats were more often collected from neonates and larger number of repeats from adults. For alpha C protein of GBS collected from two human maternal and neonatal pairs, the latter contain fewer repeats than the former and are less susceptible to opsonophagocytic killing in the presence of alpha C protein-specific antiserum (16). Furthermore, immunogenicity and protective efficacy are inversely related to the number of repeats, and deletion of repeats in the alpha C protein can enhance the pathogenicity of GBS in immune mice (9, 10). Therefore, neonatal isolates with fewer repeats of the rib gene might reflect their advantage under the pressure of maternal immunity. On the other hand, GBS with larger number of repeats of Alp proteins can prevent the elimination attempts of the host (15), and possesses a selective advantage in the clinical isolates.

In conclusion, certain serovariants of invasive GBS were highly associated with patient age, and this age-related correlation was also found in number of tandem repeats of the rib gene, with larger number of repeats present in adult isolates and fewer repeats in neonatal ones. This variation in repeat numbers can be attributed to the complex interaction of
selective forces like bacterial virulence and counter forces like host immunity.

This work was partially supported by grants DOH93-DC-1110, DOH94-DC-1008, and DOH95-DC-1038 from the Bureau of Center of Disease Control, Department of Health, Taiwan. We are indebted to the staff of Medical Microbiology Laboratory of National Cheng-Kung University Hospital for their assistance and Robert Jonas for his comments in this article.
REFERENCES


7. Farley, M. M., C. Harvey, T. Stull, J. D. Smith, A. Schuchat, J. D. Wenger, and D. S.


13. Lancefield, R. C. 1933. Serological differentiation of human and other groups of


FIGURE LEGEND

Boxplot depicting the distribution of tandem repeat numbers in the rib gene in GBS strains collected from bloodstreams of 29 neonatal and 20 adult patients. Using the one-tailed t-test, $P = 0.046$ was obtained.
Figure.

Distribution of tandem repeat number of rib gene in 49 GBS strains.

Number of repeat units

Neonates (N = 29)  Adults (N = 20)