Phase variation of lic1A, lic2A and lic3A in colonization of the nasopharynx, bloodstream and cerebrospinal fluid by Haemophilus influenzae type b

Sarah L. Hosking,† Jane E. Craig and Nicola J. High

Author for correspondence: Nicola J. High. Tel: +44 161 275 5749. Fax: +44 161 275 5656.
e-mail: nicky.high!man.ac.uk

University of Manchester,
School of Biological Science,
1.800, Stopford Building,
Manchester M13 9PT, UK

INTRODUCTION

Haemophilus influenzae type b is an encapsulated organism which causes invasive disease, such as meningitis (reviewed by Moxon, 1985). It is also a frequent cause of both upper and lower respiratory tract infections. LPS is major virulence determinant of H. influenzae and was the first cell surface structure expressed by this organism that was shown to exhibit phase variation. LPS phase variation in H. influenzae is characterized by the spontaneous loss and gain of oligosaccharide structures present in the outer core (Kimura & Hansen 1986). Mechanisms contributing to LPS phase variation in H. influenzae is generated by slipped-strand mispairing (Levinson & Gutman, 1987), which causes the number of 5'-CAAT-3' repeats to vary from generation to generation. This results in translational frameshifting, placing upstream initiation codons in- and out-of-frame with the ORF. Through this process the expression of these genes can be reversibly switched on and off (Weiser et al., 1989).

lic1A was subsequently shown to encode phosphorylcholine kinase (ChoP) which promotes the phase-variable substitution of LPS with choline (Weiser et al., 1997). Intragenic 5'-CAAT-3' repeats were also identified in two further genes, lic2A and lic3A (Weiser et al., 1990). lic2A codes a glycosyl transferase which is required for the phase-variable expression of the Galα(1–4)βGal LPS structure (High et al., 1993). However, it is the product of lgtC, another phase-variable gene, that ultimately dictates whether this structure is synthesized (Hood et al., 1996). The function of the product of lic3A has not been determined and there is no evidence, to date, to suggest that it plays a role in LPS biosynthesis (Maskell et al., 1992). The phase-variable expression of these genes is generated by slipped-strand mispairing (Levinson & Gutman, 1987), which causes the number of 5'-CAAT-3' repeats to vary from generation to generation. This results in translational frameshifting, placing upstream initiation codons in- and out-of-frame with the ORF. Through this process the expression of these genes can be reversibly switched on and off (Weiser et al., 1989).

Keywords: Haemophilus influenzae type b, phase variation, invasive disease, lic genes.
Analysis of the genome sequence of *H. influenzae* strain Rd has identified a further nine loci which contain multiple repeats of a tetrameric sequence other than 5'-CAAT-3' (Hood et al., 1996). Tandem repeats of 5'-GCAA-3' have also been found in genes involved LPS biosynthesis (Jarosik & Hansen, 1994). All of these genes encode candidate virulence genes, indicating that *H. influenzae* has a massive potential for variation during the course of infection which may contribute to its success as a pathogen. The significance of the phase variation of each of these genes in the pathogenesis of invasive disease is unknown. In the case of LPS, phase variation is thought to enable *H. influenzae* to evade antigen-specific host immune defences by promoting variation in cell surface composition. The ability to vary the expression of multiple virulence determinants may optimize the virulence potential of *H. influenzae* by enabling the expression of the most appropriate phenotypic attributes for a given environment, or stage in pathogenesis. This may be of particular significance in the development of invasive disease which, in the case of meningitis, requires translocation of *H. influenzae* from the nasopharynx to the central nervous system. This sequence of events involves the colonisation of several distinct environments and the interaction of the organism with a variety of different host cells.

In this paper we investigate the role of the phase variation of *lic1A*, *lic2A* and *lic3A* in colonization of the major compartments colonized by *H. influenzae* during the development of invasive disease. This was achieved by monitoring the number of 5'-CAAT-3' repeats present in each gene from organisms isolated from the nasopharynx, bloodstream and cerebrospinal fluid (CSF) of infant rats infected with *H. influenzae* strain Eagan. Organisms colonizing each of these environments were shown to express different combinations of *lic* genes. Phase variation of these genes may therefore play an important role in facilitating the survival and persistence of *H. influenzae* in the diverse environments encountered during the development of invasive disease.

**METHODS**

**Bacterial strains and culture conditions.*** H. influenzae* strain Eagan was used throughout this study (Anderson et al., 1980). *H. influenzae* was grown in brain-heart infusion broth (BHI) supplemented with haematin (10 µg ml⁻¹) and NAD (2 µg ml⁻¹). Solid BHI medium was prepared by the addition of Bacto Agar (1%, w/v) and Levinthal's base (10%, v/v) (Alexander, 1965).

**The infant rat model of *H. influenzae* invasive disease.*** Natural litters of 5-old Sprague–Dawley rats were reduced to 12 rats at birth and randomized. Infant rats were challenged with *H. influenzae* strain Eagan by intranasal (i.n.) inoculation so that the expression status of *lic* genes during colonization of the nasopharynx could be established (Moxon et al., 1974). Although invasive disease occurs following i.n. challenge the level of bacteraemia generated is often low and does not always lead to meningitis. A separate group of animals were therefore challenged intraperitoneally (i.p.) to enable efficient isolation of organisms from the bloodstream and CSF (Smith et al., 1973). For both i.n. and i.p. challenges the inocula were prepared and quantified as described by Moxon et al. (1974) and Smith et al. (1973), and adjusted to the desired concentration. After 48 h approximately 20 µl blood was taken from a tail vein of which 5 µl was plated onto supplemented BHI agar to detect and quantify bacteraemia. CSF was also removed from each animal by cisterna magna puncture and 10% of the volume obtained plated onto supplemented BHI agar. Bacteria colonizing the nasopharynx were harvested by washing one nostril with 50 µl PBS and then withdrawing approximately 20 µl from the opposite nostril. Two micro-litres of each sample was cultured as described above, in the presence of bacitracin (1 mg ml⁻¹).

**PCR amplification of the 5'-CAAT-3' regions of *lic1A*, *lic2A* and *lic3A*.** Blood, CSF and nasal washings were boiled for 5 min. Whole-cell debris was removed by centrifugation and the supernatant decanted to a fresh tube. Samples were then treated with Hyain pre-amplification reagent, according to the manufacturer’s recommendations, to remove compounds which might inhibit the PCR. The PCR was carried out in a buffer containing 500 mM KCl, 100 mM Tris/HCl, pH 8.0, 1% (w/v) gelatin and 25 mM MgCl₂, [α-32P]dCTP (10 µCi (0.37 MBq)) was added to the dNTP mixture to increase the sensitivity of the PCR. The 5'-CAAT-3' region from lic1A was amplified using 5'-GAATGGAATGCTGATGAAG-3' and 5'-CTATAAGATTCCAGGCT-3' or, in the case of nasal lavage samples, 5'-TCTTTCTGACGACCCGAC-3'. The 5'-CAAT-3' region from lic2A was amplified using 5'-GCACTGGAACGTCGAAACAT-3' and 5'-CACCACCTTTCTCATATAAG-3'. The primers used to amplify the 5'-CAAT-3' from *lic3A* were 5'-CGGAGATGTACACTGATA-3' and 5'-AACCTTGTCCTTTATTGCTA-3'. The reaction mixture was incubated at 94 °C for 5 min and then 45 cycles of PCR were carried out, each cycle consisting of 94 °C for 1 min, 57 °C for 1 min and 72 °C for 1 min. The resultant PCR products were analysed by electrophoresis through a 6% polyacrylamide gel. The size of each product and hence the number of 5'-CAAT-3' repeats was determined by comparison with an M13 sequencing ladder. To confirm the predicted number of 5'-CAAT-3' repeats a number of samples from each PCR set was sequenced directly as previously described (High et al., 1993). Autoradiograms were scanned using a Phoretix densitometer and the captured images analysed using Phoretix ID gel analysis software. Bands were identified and densitometry measurements obtained after an automatic background subtraction had been performed.

**Statistical analysis.** Densitometry measurements obtained for each set of PCR products were compared statistically. PCR products derived from *lic1A*, which comprised two bands, were compared using a Mann–Whitney non-parametric test. PCR products derived from *lic2A* and *lic3A*, which comprised three bands, were compared using the Kruskal–Wallis non-parametric test. In both cases P values <0.05 were considered significant.

**RESULTS**

**Phase variation of *lic1A* in *H. influenzae* colonizing the nasopharynx, bloodstream and CSF of infant rats**

The distribution of 5'-CAAT-3' repeats in *lic1A*, *lic2A* and *lic3A* present in organisms isolated from the nasopharynx, bloodstream and CSF of infected animals was determined. Multiple PCR products of different intensities were obtained and their relative abundance determined by densitometry (Figs 1b, 2b and 3b). In the case of *lic1A*, which encodes ChoP, two PCR products...
of comparable intensity ($P > 0.05$) were amplified from nasal lavage samples 48 h post-infection (Fig. 1a, lanes 1–8; Fig. 1b). One PCR product contained 20 copies of 5’-CAAT-3’, which is non-permissive for expression of lic1A. The second PCR product contained 21 copies of 5’-CAAT-3’, which is consistent with expression of the ORF and phosphorylcholine substitution of LPS (Weiser et al., 1989, 1997). In the original inoculum, the PCR product containing 21 copies of 5’-CAAT-3’ was most abundant, indicating that during nasopharyngeal colonization there had been a shift in the population such that this genotype was no longer predominant (Fig. 1, lane C). In blood and CSF samples, from animals challenged via the i.p. route, the distribution of PCR products amplified was different from that obtained from nasopharyngeal lavages (Fig. 1a and b). In these samples the PCR product containing 20 copies of 5’-CAAT-3’ was consistently more abundant. An identical distribution of PCR products was also obtained from the broth-grown organisms used as the inoculum.

Phase variation of lic2A in \textit{H. influenzae} colonizing the nasopharynx, bloodstream and CSF of infant rats

In all three sites sampled the predominant number of 5’-CAAT-3’ repeats in lic2A was consistent with the expression of a full-length ORF (High et al., 1993). The pattern of PCR products amplified from nasal lavage samples was identical to that obtained from the original inoculum. The predominant PCR product amplified contained 17 copies of 5’-CAAT-3’ (Fig. 2a, lanes 1–8; Fig. 2b). This number of repeats in lic2A is associated with high-level reactivity to mAb 4C4 which recognizes the phase variable Galα(1–4)βGal LPS structure (High et al., 1993). A second PCR product containing 18 copies of 5’-CAAT-3’, which is not consistent with expression of lic2A, was also amplified, but was much less abundant. In some samples a third PCR product, corresponding to 16 copies of 5’-CAAT-3’, was also detected. This format is also consistent with expression of lic2A but is associated with a lower level of reactivity.
The distribution of lic2A genotypes in *H. influenzae* strain Eagan isolated from the nasopharynx, bloodstream and CSF of infant rats. The 5′-CAAT-3′ repeat region in lic2A was amplified by PCR and the distribution of PCR products visualized by PAGE. In all experiments each lane represents PCR products obtained from an individual animal. Lane C represents PCR products amplified from broth-grown organisms used as the inoculum. (a) PCR products amplified from nasal lavage samples obtained from animals challenged by i.n. inoculation. (b, c) PCR products amplified from blood and CSF samples, respectively, obtained from animals challenged by i.p. inoculation. Lanes with corresponding numbers represent paired blood/CSF samples obtained from the same animal. (d) The intensities of each of the PCR products shown in (a)–(c) were determined by densitometry using the Phoretix software package. The mean intensity for bands containing the same number of 5′-CAAT-3′ repeats was calculated for each sample. These values are presented as a histogram. A Kruskal–Wallis non-parametric test was used to determine whether the relative quantities of PCR products in each sample were significantly different (*P* < 0.05). * indicates samples where *P* < 0.05.

to mAb 4C4. PCR amplification of blood and CSF samples, from i.p.-challenged animals, revealed that in these environments the majority of organisms had either 17 or 16 copies of 5′-CAAT-3′ and therefore had the potential to express lic2A and incorporate Gal(1–4)-βGal into their LPS (Fig. 2) (High et al., 1993). PCR products containing a non-permissive number of 5′-CAAT-3′ repeats were also detected, but were approximately three times less abundant than products amplified from organisms containing 17 copies of 5′-CAAT-3′. The apparent ratio of the PCR products obtained from blood samples was, in some cases, different from the original inoculum, in which the predominant PCR product amplified contained 17 copies of 5′-CAAT-3′. This was particularly apparent in samples 3, 5, 8 and 11 (Fig. 2) in which PCR products containing 16 copies of 5′-CAAT-3′ were amplified, a format which was not readily detectable in the original inoculum. This suggested that phase variation had occurred in vivo, following i.p. challenge, but that the ability to express lic2A had been retained. Comparison of the PCR products obtained from paired blood and CSF samples revealed that translocation of organisms from the bloodstream to the central nervous system had occurred concurrently with a change in the predominant number of 5′-CAAT-3′ repeats in some animals. In animals 2, 5 and 8 the predominant number of repeats changed from 16 to 17. In animals 7, 9 and 11 the switch was in the opposite direction (Fig. 2). In each case the observed change did not alter the ability to express lic2A. PCR products generated from organisms unable to express lic2A were consistently less abundant in each CSF sample. The fact that phase variation had occurred during translocation of organisms from the bloodstream...
to the CSF suggests that organisms with this genotype were selected against.

Phase variation of lic3A in H. influenzae colonizing the nasopharynx, bloodstream and CSF of infant rats

Amplification of the 5'-CAAT-3' repeats in lic3A from organisms present in nasal lavage samples generated a predominant PCR product consistent with the presence of 28 repeat units in this gene (Fig. 3a, lanes 1–8; Fig. 3b). This number of 5'-CAAT-3' repeats is permissive for the expression of lic3A (Maskell et al., 1992). PCR products corresponding to 29 and 27 repeat units were also detectable in one sample (Fig. 3, lane 3), but were present at low levels. In the inoculum used to challenge these animals the majority of organisms contained 27 copies of 5'-CAAT-3', which is non-permissive for expression of lic3A, although a subpopulation containing 28 copies was also present. Colonization of the nasopharynx by H. influenzae therefore appears to involve selection of organisms which have the ability to express lic3A. In contrast, in the bloodstream and CSF the predominant population of organisms were unable to express lic3A. Three PCR products were amplified from blood and CSF samples (Fig. 3a and b). The relative distribution of these products was identical to that of organisms in the original inoculum. The most intense PCR product was consistent with the presence of 27 copies of 5'-CAAT-3' in lic3A, which does not permit translation of the ORF. Subpopulations of organisms, containing 28 and 26 copies of 5'-CAAT-3', both of which permit expression of lic3A, were also detected at low levels.

**DISCUSSION**

In this study we have investigated the role of the phase-variable expression of lic1A, lic2A and lic3A in colo-
zation of the nasopharynx, bloodstream and CSF of infant rats infected with *H. influenzae* type b strain Eagan. Each site was found to be colonized by a heterogeneous population of organisms, expressing different combinations of each gene product. This is consistent with the idea that phase variation is a stochastic process. Homogeneous populations would only be anticipated to occur in the presence of a strong selection pressure which favoured the expression of a particular combination of lic genes. The relative proportions of each lic genotype differed most between organisms colonizing the nasopharynx and organisms isolated from either the bloodstream or CSF. In the latter two environments the pattern of PCR products generated was indistinguishable. This may reflect the similar distribution of complement, antibodies and other components of the immune system in these two locations once an inflammatory response has occurred (Zwahlen *et al.* 1982; Whittle & Greenwood, 1977). The observed differences between the lic genotypes harboured in these environments suggests that *H. influenzae* may require different phenotypic attributes for colonization of the nasopharynx, bloodstream and CSF.

Evidence of a selection pressure influencing the distribution of lic genotypes during nasopharyngeal colonization was most apparent in the case of lic3A. During the 48 h period following i.n. challenge the predominant lic3A genotype changed, relative to the inoculum, resulting in a marked reduction in the proportion of organisms unable to express lic3A. This observation suggests that expression of this gene is important for nasopharyngeal colonization, although its precise role has yet to be determined. The distribution of lic1A genotypes also altered following i.n. challenge. However, this did not result in colonization by a predominant population since equal proportions of organisms which either expressed or did not express lic1A were detected. This even distribution of genotypes would suggest that during the first 48 h the expression status of lic1A is not important for nasopharyngeal colonization. In a study by Weiser *et al.* (1998), a gradual increase in the proportion of organisms able to express lic1A was shown to occur in the nasopharynx over a period of 10 d. If the distribution of lic1A genotypes had been monitored over a longer period, a gradual increase in the proportion of organisms expressing this gene might therefore have been expected to occur. In contrast to lic1A and lic3A, the distribution of lic2A genotypes in the nasopharynx was indistinguishable from that of the original inoculum. No conclusive evidence suggesting that organisms expressing this gene were selected in the nasopharynx was therefore obtained.

Following i.p. challenge, differences were observed in the distribution of 5’-CAAT-3’ repeats in blood and CSF samples, relative to that in the original inoculum. The numbers of 5’-CAAT-3’ repeats also varied between paired blood and CSF samples, indicating that phase variation of lic2A had occurred in vivo. Despite phase variation organisms expressing lic2A always formed the predominant population suggesting that this phenotype was important in the persistence of *H. influenzae* in the bloodstream and CSF. This notion is consistent with previous observations that show that *H. influenzae* mutants which lack a functional copy of lic2A have reduced virulence in the infant rat model (Cope *et al.*, 1991). No evidence was obtained, however, that phase variation followed by selection of specific lic1A and lic3A genotypes had occurred. The distribution of genotypes in organisms present in the bloodstream and CSF was indistinguishable from that of broth grown organisms used as the inoculum. If a specific genotype was essential for survival at these sites, the relative proportion of organisms with this attribute would be expected to increase. This was not the case for either lic1A or lic3A. Expression of these genes may therefore be neither an advantage nor a disadvantage in these environments.

In summary, the phase variation of lic genes has been monitored in vivo using an infant rat model of *H. influenzae* invasive disease. We have demonstrated that organisms colonizing the nasopharynx express a different combination of lic genes to those colonizing the bloodstream and CSF. Phase variation of these genes may therefore play an important role in the development of invasive disease by *H. influenzae*.

**ACKNOWLEDGEMENTS**

This work was funded by a project grant from the Wellcome Trust.

**REFERENCES**


Received 14 December 1998; revised 30 June 1999; accepted 8 July 1999.