Spread of Serogroup A Meningococci: a Paradigm for Bacterial Microevolution

Abstract

*Neisseria meningitidis* serogroup A bacteria have been subdivided into clonal groupings called subgroups. The global epidemiology of certain subgroups has been investigated and compared with sequence data for variable cell surface or secreted proteins. The results document occasional instances of clonal replacement where recombinant DNA has become established within a bacterial population in a very short period of time.

Bacterial strain collection

A total of about 800 strains of serogroup A *Neisseria meningitidis*, isolated from diverse epidemics since 1915, have been analyzed by multilocus enzyme electrophoresis (MLEE). Initial analyses using 7 cytoplasmic enzymes and 2 outer membrane proteins (OMPs) distinguished 50 electrophoretic types (ETs) differing by at least one allele among 423 isolates from different sources (Olyhoek et al., 1987). Closely related ETs which differed by only 1–2 alleles were grouped into 21 clones belonging to four ‘subgroups’. Later work has failed to confirm the existence of the ‘clones’ (Wang et al., 1992), whereas the subgroup assignments have been supported by various independent analyses. In contrast to this diversity, 75 bacteria isolated from diseased patients and healthy carriers, during (1982–83) and after an epidemic in The Gambia, West Africa, belonged to one single ET within subgroup IV-1 (Crowe et al., 1989). Approximately 70 additional subgroup IV-1 bacteria, isolated between 1963 and 1990 from various countries within the sub-Saharan Sahel region of Africa, allow a comparison of genetic variation within these 3 decades. Over 200 strains of subgroup III serogroup A meningococci, isolated from 2 pandemic waves between 1966 and 1994, have been compared with subgroup IV-1 bacteria for variable cell surface antigens (Achtman et al., 1992). An improved MLEE analysis was performed using 15 cytoplasmic enzymes and 4 OMPs for 290 strains representing the above collection plus an additional 165 strains from additional sources, primarily China (Wang et al., 1992). 84 ETs were recognized which were assigned to the 9 clonal groupings called subgroup I to III, IV-1, IV-1, and V through VIII (Wang et al., 1992).
Epidemiological sources

Subgroup II contains six strains isolated in the 1930's and 1940's in the USA and in the 1960's in Djibouti and may have disappeared. Subgroup IV-2 was isolated from epidemics in World Wars I and II and thereafter from Russia in the 1970's. Subgroups VII and VIII contain a few bacteria isolated from China between 1956 and 1979 while subgroup VI contains endemic isolates from former East Germany, Scandinavian countries and Russia. The limited numbers of bacteria available from these subgroups suggests that, with time, additional subgroups may be identified when endemic isolates from a greater variety of countries are tested. In contrast, the epidemiology of the other subgroups described below is well defined and numerous isolates are available.

Subgroup I: The following is a reinterpretation of the data presented by Olyhoek et al. (1987) forced by the recognition that the subgroup is the basic unit of epidemic spread in serogroup A organisms (Wang et al., 1992). The oldest subgroup I strain available was isolated in the UK in 1941. In the early 1960's, subgroup I bacteria were isolated from an epidemic in Niger and from endemic disease in Algeria, Chad and among US army personnel stationed in West Germany. In 1967, epidemic disease caused by subgroup I flared in North Africa and the Mediterranean and spread by 1968–1972 throughout West Africa. In the early 1970's, subgroup I caused outbreaks among native Americans in Canada followed by outbreaks among 'Skid Road' inhabitants in the Pacific Northwest of the United States (Counts et al., 1984; Olyhoek et al., 1987). Epidemics in Nigeria and Rwanda in the late 1970's were caused by subgroup I (Olyhoek et al., 1987) as was an outbreak among Maoris and Pacific Islanders in New Zealand in 1985 (Schwartz et al., 1989) and an outbreak among aboriginals in central Australia in the late 1980's (Patel et al., 1993) (D. A. Caugant, pers. comm.). During the 1970's, subgroup I meningococci were isolated globally from endemic disease (Olyhoek et al., 1987).

Subgroup IV-1: Subgroup IV-1 is associated with different epidemiological patterns than subgroups I or III. Almost all endemic isolates from West Africa since the early 1960's belong to subgroup IV-1 (Olyhoek et al., 1987) despite two waves of epidemic disease caused by subgroup I in the 1960's and 1970's. In addition, some isolates from an epidemic in Niger in the early 1960's and all isolates from epidemics in West Africa in the early 1980's belonged to subgroup IV-1 (Olyhoek et al., 1987). All subgroup IV-1 bacteria have been isolated in West Africa (Olyhoek et al., 1987; Wang et al., 1992), except for 4 strains isolated in India (Bjorvatn et al., 1992).

Subgroup V: Similar to subgroup IV-1, subgroup V bacteria have only been isolated in China, where they caused epidemic disease in the 1970's and were occasionally isolated in the 1980's (Wang et al., 1992). The epidemiology of these two subgroups shows that even epidemic bacteria are not necessarily able to spread extensively between continents, for unknown reasons.

Subgroup III: Subgroup III bacteria were first isolated from China during a huge epidemic in the mid-1960's and spread to cause epidemics in Moscow.
(1969), an outbreak in Norway (1969), and epidemics in Finland and Brazil (mid-1970's) (Achtman et al., 1992; Wang et al., 1992). These bacteria were also responsible for most of the endemic meningococcal disease in Sweden during the 1970's and beginning of the 1980's, but they did not cause any epidemics or outbreaks there (Salih et al., 1990). During the 1970's, subgroup III disappeared from China, to be replaced by subgroup V, but they reappeared in China (and in Nepal) in the early 1980's. Thereafter in 1987, they caused an epidemic in Mecca, Saudi Arabia during the annual Haj pilgrimage (Moore et al., 1988, 1989). Healthy pilgrims carried these bacteria back to their countries of origin, including the USA, UK and France but subgroup III disappeared from most countries within a few years. In contrast, epidemics caused by subgroup III did break out in eastern Africa in 1988 and have continued to affect one African country after another since then (Achtman, 1995a).

The mechanisms responsible for epidemic disease are unknown but several theories based on coinfection with viruses or other bacteria (summarized in (Achtman, 1995a, 1995b; Cartwright, 1995) have been proposed. The results presented above show that it will be difficult to apply population genetic models based on steady state competition to epidemic bacteria.

**Genetic variation**

The serogroup A subgroups consists of related ETs which differ by a few alleles in MLEE. Such variation is often assumed to reflect evolutionary changes during clonal descent of bacteria derived from a common ancestor. However, the work needed to identify electrophoretic variation is sufficiently laborious that only limited numbers of strains can be investigated. We have used an alternative, less labor-intensive, approach to screen for variation, namely analysis of cell-surface antigens detected by reactivity with monoclonal antibodies (MAbs). Numerous bacteria have been screened for epitopes within pilin, the PorA and PorB porins (also called Class 1 and 3 proteins, respectively), Opa and Opc proteins (collectively called Class 5 proteins), and IgA1 protease (IgA1P). Variation was then analyzed by DNA sequencing, some of which has not yet been published. The results reveal that genetic variation differs with the antigen and with the subgroup and indicate that epidemic spread results in loss of antigenic variation and reduction of clonal diversity. The results with PorA (Wang et al., 1992; Suker et al., 1994) are reviewed in M. Maiden's chapter elsewhere in this volume.

**Opc protein**

A formerly undescribed Class 5 outer membrane protein, formerly called 5C but now called Opc, was expressed by some of the subgroup IV-1 meningococci from The Gambia (Crowe et al., 1989). The purified Opc protein consists of a trimer or tetramer of 28 KDal subunits (Achtman et al., 1988). The opc gene is
expressed after cloning in *Escherichia coli* and the recombinant Opc protein is localized on the cell surface (Olyhoek *et al.*, 1991). Expression of Opc enables meningococci to adhere to and invade endothelial or epithelial cells (Virji *et al.*, 1992, 1993). Binding to the apical surface of endothelial cells is dependent on serum factors, such as vitronectin, which bind through their RGD sequences to eukaryotic β-integrins (Virji *et al.*, 1994). The Opc protein is expressed at varying levels by different colonies of the same strain, varying from non-detectable to being one of the major outer membrane proteins. When a single colony is spread on a plate to single colonies, several will differ from the majority phenotype in regard to Opc expression and variation is correlated with changes in the length of a poly-cytidine stretch within the Opc promoter (Sarkari *et al.*, 1994). Poly-C stretches of 12–13 nucleotides are correlated with very efficient transcription and protein expression, stretches of 11 or 14 with intermediate expression and either shorter or longer poly-C stretches with lack of expression.

Meningococcal DNA was subjected to hybridization with the cloned opc gene: all serogroup A meningococci gave a positive signal as did many serogroup B, C, 29E, Y and W strains. However, some serogroup B and C bacteria, in particular all ET-37 complex strains, did not give a signal and lack an opc gene (Sarkari *et al.*, 1994). PCR products from 110 meningococci containing an opc gene have been tested for restriction site polymorphism within the 819 bp region encoding opc and the 293 bp region upstream of the coding region. Variant RFLP patterns were identified and sequenced as were representatives of each of the serogroup A subgroups (Seiler, A., *et al.*, man. in prep.) The results show about 1% sequence variation within the coding region, including non-synonymous changes, and 4% within the upstream non-coding region. (None of these sequence polymorphisms affect the site recognized by monoclonal antibody B306 and it is anticipated that this antibody will recognize all Opc protein variants in meningococci.) Some strains contain a 250 bp insert at a particular nucleotide within the upstream region, and that insert is itself subject to about 10% sequence variation. Although at first glance, these results resemble those expected of evolutionary variation by slow accumulation of mutations, much of the sequence polymorphism occurs in mosaic blocks with different combinations of the same nucleotides and it seems as if the rearrangement of these sequence variants reflects recombinational exchange. The insertion is also present or absent in otherwise identical alleles indicating that it too is subject to recombinational exchange.

Among serogroup A meningococci, the same allele of opc was found in all subgroups except subgroup IV-2, where it differed by 1 nucleotide. No sequence differences were found within the 293 bp upstream region, except for a 1 nucleotide difference in subgroups V and VII, which were identical, and another single nucleotide difference in subgroup IV-2. These results are compatible with the notion that all serogroup A meningococci are descended from one common ancestor, which contained a primordial opc allele, and that only limited evolution has occurred or been maintained since then. Interestingly, one subgroup IV-2 variant, isolated in Britain in 1942, has been identified in which the whole
1.1 Kb region has been replaced by a different allele containing numerous nucleotide exchanges as well as the 250 bp insert in the upstream region. Because this opc allele was only found once while other subgroup IV-2 bacteria isolated both considerably earlier and later followed the pattern described above, this variant seems not to have spread extensively.

The Opc protein is immunogenic in humans (Rosenqvist et al., 1993) and bactericidal antibodies were stimulated after immunization with an outer membrane vesicle vaccine. It might have been expected that an immunogenic protein would be under selection pressure and more sequence variation should have been observed. Two possible explanations are i) that the opc gene has only recently been introduced into meningococci and has not yet had time to evolve and ii) that variable expression of Opc prevents bactericidal antibodies from being effective because only few of the infecting bacteria express sufficient Opc for the antibodies to be effective.

Opa proteins

opA genes consist of conserved sequences surrounding two highly variable regions, called HV1 and HV2, which are exposed on the cell surface. Whereas 11 or 12 opA genes are present on the chromosome of N. gonorrhoeae, meningococci only contain 3 or 4 (Aho et al., 1991; Hobbs et al., 1994). Serogroup A meningococci from The Gambia contained 3 opA genes, called opA, opB and opD (Hobbs et al., 1994), in addition to opc. opA and opB have an identical HV1 region while opB and opD have an identical HV2 region. In addition, a 9 bp repeat is duplicated in tandem within opB. Individual Opa proteins were screened by SDS PAGE for size differences and by reactivity with MAbs recognizing specific HV regions. Within The Gambia, almost all Opa proteins were indistinguishable from OpA, OpB or OpD but an exceptional protein was found (5e) which migrated differently (Achtman et al., 1991). This opA gene consisted of a variant opA in which the tandem 9 bp duplication had been resolved, with the corresponding loss of 3 amino acids (Hobbs et al., 1994). Sequencing of other variant strains revealed 1 strain in which opB had translocated to the locus of opA, replacing it and yielding a chromosome with opD plus 2 copies of opB. In still another strain, DNA had been imported from an unknown source, leading to an altered opD locus containing novel HV1 and HV2 regions. The novel HV1 region had then translocated to replace that of opB. Unpublished results (Cannon, J.C., pers. comm.) confirm that most other serogroup A meningococci isolated from The Gambia possessed indistinguishable opA genes from those present in the reference strain.

The Opa proteins expressed by subgroup IV-1 bacteria differed with the country of origin (Achtman, 1994). One strain isolated in Ghana in 1970 possessed an indistinguishable opA gene but possessed an opAF allele at the opD locus and an opAG allele at the opB locus. opAF differs from opD in the HV2 region, which is almost identical to that of an opA gene in N. gonorrhoeae strain MS-11. opAG differs from opB in the HV1 region, which is almost identical to that of
an opa gene in serogroup C meningococci of the ET-37 complex (Hobbs et al., 1994).

More recent analyses (Malorny, B. et al., man. in prep.) have concentrated on opa genes from subgroup III and other subgroup IV-1 strains from different countries. Subgroup III serogroup A meningococci isolated prior to the Mecca epidemic of 1987 variably expressed Opa proteins formerly called 5a, 5f and/or 5h while those isolated during or after that epidemic expressed 5a, 5f and/or 5i (Achtman et al., 1992). Proteins indistinguishable from 5f and 5h were also expressed by some subgroup IV-1 strains isolated in the 1960's (Achtman et al., 1988). The sequences of these opa genes, opaF and opaH, have now been determined. opaF is indistinguishable to that described above while opaH is a novel, formerly undescribed gene. Subgroup III meningococci contain 4 opa alleles (plus opc) and the sequences of the opa genes from the bacteria isolated prior to Mecca correspond to opaA, opaF, opaH plus a new gene, opaJ. The former 3 sequences are identical to those in subgroup IV-1 bacteria from the 1960’s, indicating that such sequences can be inherited over very long periods of time, and that subgroup III and IV-1 are descended from a common ancestor. The opa genes have also been sequenced from one subgroup III strain isolated after the Mecca epidemic. opaA and opaJ are unchanged. opaF differs by 1 nucleotide and opaH has been replaced by a novel gene, opaI. The mechanisms leading to nucleotide exchange are unclear but the genetic replacement can only have happened by import of foreign DNA in the mid-1980’s. OpaH (5h) was never expressed by any subgroup III meningococci isolated in Africa after 1988 and it is amazing that such a genetic recombination event can have been fixed within this short period of time.

*IgA1 Protease*

IgA1 protease is a highly immunogenic (Brooks et al., 1992) protein secreted by all meningococci and gonococci. Within serogroup A meningococci, MAbs to 3 variable epitopes distinguished 6 antigenic variants, which were uniform within individual subgroups, except subgroup III (Morelli et al., 1994). Subgroup III bacteria isolated prior to the Mecca epidemic expressed one serological variant of IgA1P while bacteria isolated later expressed a different variant. During the few months after the Mecca outbreak, 4 strains were isolated which still expressed the ancestral serological variant. Two of these four strains also expressed OpaH while the two others expressed neither OpaH nor OpaI and have not yet been tested by DNA techniques.

The 3 Kb segment of the iga gene encoding the mature IgA1 protease has been sequenced from 4 meningococci (Morelli et al., man. in prep.) and shows minor sequence variation throughout the gene. The iga sequences from meningococci isolated prior to Mecca and post Mecca differ in two regions by numerous nucleotides, indicating that at least one and possibly two recombinational events have occurred. T-track comparisons of numerous iga genes from post-Mecca subgroup III meningococci showed no differences: all contained the
recombinant sequence (Morelli et al., man. in prep.). However, diversity was found within 6 strains isolated in China in 1966. 2 of the sequences contain a T at a position where the other 6 and subsequent isolates possessed a C. Thus, at least two sequential events of clonal replacement have occurred in these bacteria, with bacteria isolated after 1966 all possessing the C nucleotide and bacteria isolated post-Mecca containing the recombinant sequence.

**Clonal expansion and replacement**

The most surprising aspect of the work described above is the speed with which serogroup A populations can be replaced by variants which differ in one or more alleles. These results contrast with the relative uniformity of these bacteria within any one epidemic or most epidemic waves. Several examples have also been described where sequences have been uniform since the evolution of the individual subgroups. Yet, over just a few years, other genes can be totally replaced leading to a uniform recombinant population.

The phenomenon of epidemic spread has been reviewed elsewhere (Achtman, 1994, 1995b) and the differences to periodic selection emphasized. Although the role of selection in clonal replacement remains uncertain, I believe that spread of only few bacteria from country to country coupled with a high rate of recombination during epidemics is at least as likely to explain clonal replacement as are selection pressures (Achtman, 1994, 1995b).

The analysis of genetic phenomena in meningococci complements analyses of bacterial populations that do not spread rapidly and that do not seem to possess efficient methods of genetic exchange. Our data with meningococci illustrate several of the phenomena predicted by population biologists (see chapters by Milkman and Selander in this volume) and show that these occur continuously within relatively short periods of time, decades rather than millennia. Correlation with the epidemiology of infectious disease may allow a better understanding of microevolution in bacteria and meningococci may well be a paradigm for studying such phenomena.

**References**


parison of the variable antigens expressed by clone IV-1 and subgroup III of *Neisseria meningitidis* serogroup A. *J. Infect. Dis.* **165**, 53–68.


Achtman, M., 1995a - Global epidemiology of meningococcal disease. 159–175.


Achtman, M., 1995 - Global epidemiology of meningococcal disease. 159–175.


