The population structure of *Neisseria meningitidis* serogroup A fits the predictions for clonality

Aldert Bart\(^a\),\(^*\), Christian Barnabé\(^b\), Mark Achtman\(^c\), Jacob Dankert\(^{a,d}\),
Arie van der Ende\(^{a,d}\), Michel Tibayrenc\(^b\)

\(^a\) Department of Medical Microbiology, Academic Medical Center, University of Amsterdam, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands
\(^b\) Centre d’Études sur le Polymorphisme des Microorganismes (CEPM), Centre IRD, CEPM/UMR CNRS-IRD 9926, 911 avenue Agropolis, B.P. 5045, 34032 Montpellier Cedex 1, France
\(^c\) Max-Planck Institut für Infektionsbiologie, Schumannstrasse 21/22, D10117 Berlin, Germany
\(^d\) Reference Laboratory for Bacterial Meningitis, University of Amsterdam/RIVM, P.O. Box 22700, 1100 DE Amsterdam, The Netherlands

Received 5 November 2000; received in revised form 17 April 2001; accepted 29 May 2001

**Abstract**

The population structure of *Neisseria meningitidis* is supposedly epidemic according to Maynard Smith et al. (1993). The model predicts that linkage disequilibrium in *N. meningitidis* populations is only temporary and arises due to the outgrowth of highly successful clonal genotypes from an essentially sexual population. These clones should disappear after a few years because of frequent recombination. In contrast, multilocus enzyme electrophoresis (MLEE) data had previously been interpreted as showing that serogroup A meningococci are truly clonal and possess only limited genetic variability (Wang et al., 1992). The two interpretations are contradictory.

In order to elucidate the true population structure of serogroup A meningococci, we analyzed data for a representative group of 84 serogroup A isolates obtained by MLEE, random amplified polymorphic DNA (RAPD) and multilocus sequence typing (MLST). Analysis of linkage disequilibrium and bootstrap analyses of cluster analysis showed a strongly structured population with highly significant linkage disequilibrium. This was not due to the overrepresentation of certain genotypes, in contrast to the expectations for an epidemic population. The analyses identify two main clades, within each of which linkage disequilibrium was also highly significant, thus, excluding a cryptic speciation model. These observations support a population structure based on clonal evolution, in which clones are much more stable than expected for epidemic clonality. We propose that serogroup A meningococci may possess a different population structure from other serogroups of *Neisseria meningitidis*. © 2001 Elsevier Science B.V. All rights reserved.

**Keywords:** Molecular epidemiology; Multilocus enzyme electrophoresis; Random amplified polymorphic DNA; Multilocus sequence typing; *Neisseria meningitidis*; Population structure

1. Introduction

*Neisseria meningitidis* is a common resident of the human nasopharynx, that only occasionally causes invasive disease when it reaches the bloodstream and cerebrospinal fluid (Peltola, 1983). Different serogroups have been distinguished on the basis of the antigenic properties of the capsular polysaccharide, among which serogroup A isolates have often been associated with recurrent epidemics of meningococcal disease in developing countries. Since World War II, large epidemics caused by serogroup A have not occurred in Europe or the USA, although these strains have been isolated from endemic disease and occasional outbreaks in these areas.

\(^*\) Corresponding author. Tel.: +31-20-5664868; fax: +31-20-6979271.
E-mail address: a.bart@amc.uva.nl (A. Bart).

Similar to many other bacteria, *Neisseria* exchange genes horizontally via natural DNA transformation (Feil et al., 1995; Saker et al., 1994; Morelli et al., 1997; Feil et al., 1999; Linz et al., 2000; Feil et al., 2001; Maiden et al., 1996) which should result in increased linkage equilibrium. However, hypervirulent clonal groupings have been described (Caugant et al., 1986; Wang et al., 1992; Maiden et al., 1998), an indication for linkage disequilibrium. Maynard Smith et al. (1993) proposed an epidemic population structure for *N. meningitidis* to account for linkage disequilibrium within a recombining population. In this model, linkage disequilibrium results from the temporary overrepresentation of highly successful clonal genotypes which have multiplied so rapidly that they have not yet been disrupted by horizontal genetic exchange. Indeed, MLEE data showed strong linkage disequilibrium only when all isolates were considered and much less linkage disequlibrium...
brium when a single representative of each electrophoretic type was tested. Maynard Smith et al. (1993) also suggested that panmixis would result whenever genetic variants were introduced two to four times more frequently by horizontal genetic exchange than by mutation. However, even within the hypervirulent clonal groupings, recombination is at least five-fold as frequent as mutation (Feil et al., 1999).

The most extensive data for hyperinvasive clonal groupings have been accumulated for epidemic serogroup A meningococci (Olyhoek et al., 1987; Wang et al., 1992; Maiden et al., 1998; Achtman et al., 2001). Extensive strain collections have been established containing bacteria that were isolated from diverse global sources since 1915 (Olyhoek et al., 1987; Zhu et al., 2001). Similar to other meningococci, sequence changes are more often due to recombination after import than to mutation, both in housekeeping genes (Feil et al., 1999) and in highly polymorphic antigens (Zhu et al., 2001). Furthermore, independent analyses of genetic relationships with the same isolates keeping genes (Feil et al., 1999) and in highly polymorphic antigens (Zhu et al., 2001). Furthermore, independent analyses of genetic relationships with the same isolates have been performed using MLEE (Wang et al., 1992), RAPD (Bart et al., 1998) and MLST (Maiden et al., 1998). Analysis of these data provided the potential to determine whether the population structure of serogroup A meningococci represents epidemic clonality, cryptic speciation or clonal evolution.

2. Materials and methods

2.1. Data on N. meningitidis strains

MLEE: allele designations from 15 cytoplasmic enzymes and four outer membrane proteins for 84 serogroup A strains of N. meningitidis were from Wang et al. (1992). These isolates are single representatives of each of the 84 electrophoretic types (ET) detected in over 500 serogroup A meningococci isolated from epidemics, endemic disease and carriers. RAPD: banding pattern assignments were based on the presence or absence of individual bands generated with four different primers for the same 84 strains (Bart et al., 1998). MLST: data for six housekeeping gene fragments for the hypervirulent clonal groupings, recombination is at least five-fold as frequent as mutation (Feil et al., 1999).

The statistical tests f and g described previously (Tibayrenc et al., 1990; Tibayrenc et al., 1993; Tibayrenc, 1995) were used to test departures from panmictic expectations. These recombinations tests are based on the null hypothesis of random genetic exchange, and appraise different consequences of linkage disequilibrium between loci. Test f estimates the probability with which the observed values of linkage disequilibrium are predicted. The “extended” g test (Tibayrenc, 1995) measures the correlation between the genetic distances generated by independent sets of genetic markers with a nonparametric Mantel test (Mantel, 1967). Both f and g tests are based on Monte Carlo simulations, with 10^6 iterations. Tests d1, d2 and e were not used in the present study, because they are only applicable to raw data containing multiple isolates per genotype, unlike the current dataset where only one isolate per MLEE genotype was included.

3. Results

3.1. Phylogenetic analyses

Wagner bootstrap analyses of MLEE (Fig. 1A) and RAPD (Fig. 1B) data distinguished two main clades. Two clades were also observed when MLEE, RAPD and MLST data were combined for the 33 isolates for which all these data were available. Clade 1 contains the isolates previously (Wang et al., 1992) assigned to subgroups I, II, V, VI and VII and clade 2 contains subgroups III, IV-1, IV-2 and VIII. The same clade assignments were observed in all three trees. The pooled data from MLEE, RAPD and MLST (Fig. 1C) yielded high bootstrap values for clade 1 (81%) and clade 2 (85%) whereas only the bootstrap values for clade 2 were over 50% when the data from MLEE (Fig. 1A) or RAPD (Fig. 1B) were tested independently.

3.2. Association between geographical origin, time of isolation and genotype

The χ^2-tests were performed to determine whether there was a significant correlation between clade and continent (Asia, Europe or Africa) or date (1960s, 1970s or 1980s)
Fig. 1. Inferred phylogenies using bootstrap analysis (Wagner) of data from MLEE (A), RAPD (B) and combined data from MLEE, RAPD and MLST (C). Clade 1 is indicated by a square; clade 2 is indicated by a circle; Bootstrap values of $\geq 50\%$ are indicated.
of isolation. Significant associations ($P < 0.05$) were not found for either the MLEE or the RAPD data.

### 3.3. Linkage disequilibrium analysis

Concordance between the assignments to clades 1 and 2 by all three sets of data suggests that the serogroup A population structure is characterized by high linkage disequilibrium. Indeed, statistically significant departures from panmictic expectations were observed using both tests $f$ and $g$. According to the $f$ test, the probability of detecting the observed levels of linkage disequilibrium from a panmictic population for the individual loci tested was $2 \times 10^{-4}$ for either the MLEE or the RAPD data. The genetic distances from the MLEE data correlated significantly ($P = 10^{-4}$, g test) with the genetic distances from the RAPD data. Similar levels of significance were also obtained in g test comparisons between MLST and RAPD, MLEE, or RAPD plus MLEE for the 33 strains for which data were available.

These results allow rejection of the epidemic model (Maynard Smith et al., 1993) and indicate that the population structure of serogroup A meningococci is one of either clonal evolution or cryptic speciation. In order to distinguish between these possibilities, the combined data from MLEE, RAPD and MLST were analyzed by test $f$ for each of the two clades identified by the bootstrap analyses. Linkage disequilibrium should disappear if the population structure reflected cryptic speciation of two panmictic populations. Instead, significant ($P = 0.005$) departures from the panmictic values were obtained for both clade 1 (20 strains) and clade 2 (12 strains). Similarly, significant linkage ($P = 10^{-4}$) was observed by comparing MLEE data with RAPD data by the g test with clade 1 (46 strains) or clade 2 (37 strains).

### 4. Discussion

The population structure of bacteria has been the subject of numerous reports since the early 1980s. Initial analyses indicate that many bacterial species consist of clonal groupings (see review by Selander et al., 1986). Maynard Smith et al. (1993) summarized data for panmictic structures in several bacterial species but could not readily assign the population structure of *N. meningitidis* to either the clonal or the panmictic class. Instead, they suggested that these bacteria possess an epidemic population structure, in which high linkage disequilibrium is due to the ephemeral propagation of successful clonal genotypes in a basically sexual species. The key distinction between such a structure and that of a clonal species is that these clonal genotypes should disappear quickly due to continued recombination. Since 1993, it has become clear that population structures need not be uniform within an entire species. For example, *Neisseria gonorrhoeae* are panmictic (Maynard Smith et al., 1993) but HAU gonococci are clonal (Gutjahr et al., 1997). Similarly, *Helicobacter pylori* exhibits free recombination (Suurballe et al., 1998) but different populations exist in different continents (Achtman et al., 1999). The data presented here indicate that serogroup A isolates of *N. meningitidis* are largely clonal, even if the entire species is not.

The serogroup A isolates were assigned to the same two clades by bootstrap analysis of data from three independent typing methods, indicating a lack of panmixis (Tibayrenc et al., 1993). Significant linkage disequilibrium was determined by statistical tests in all three datasets. Furthermore, all these data were obtained using single representatives of each genotype, and cannot be caused by the repeated isolation of individual isolates from short-lived epidemic clones.

Alternative explanations to epidemic clonality for distinct clades are geographical and/or temporal separation (Wald-hund effect). However, we found no association between genotype and geographical source or date of isolation. Linkage disequilibrium was also not simply attributable to cryptic sexuality, because linkage disequilibrium persisted when these two clades were analyzed separately, as suggested by Maynard Smith et al. (1993). The existence of these two clades over many decades contrasts with the predictions of an epidemic population structure, under which epidemic genotypes should lose coherence after a few years. Thus, the data indicate that the population structure of serogroup A *N. meningitidis* is clonal even though these bacteria frequently import foreign DNA (Suurballe et al., 1994; Morelli et al., 1997; Feil et al., 1999; Linz et al., 2000; Zhu et al., 2001). The patterns found here correspond to a structured clonal model, in which a group of microorganisms is separated into two or more sharply-defined genetic subdivisions (Discrete Typing Units or DTU; Tibayrenc, 1998).

How can clonal population structures survive despite frequent recombination? One possibility is that a critical frequency of recombination is needed to disrupt clonal structure, higher than the frequency that applies to serogroup A isolates (Tibayrenc et al., 1990). In the past, this frequency has been estimated relative to the mutation frequency ($\mu/m$). Maynard Smith et al. (1993) estimated that an $\mu/m$ ratio of $2-4$ would be sufficient to produce apparent panmixis. Feil et al. (1999) identified five cases of import versus two mutations for housekeeping genes in serogroup A isolates ($\mu/m > 2$). In recent analyses, Zhu et al. (2001) identified 40 imports for four mutations ($\mu/m = 10$) among six polymorphic genes for 500 isolates of serogroup A and Linz et al. (2000) identified 17 imports for three mutations ($\mu/m = 5$) among 100 isolates of subgroup IV-1 isolated from a single epidemic. These observations suggest that $\mu/m$ ratios of $2-4$ are not sufficient to disrupt clonality, at least in serogroup A meningococci.

The results presented here show that none of the alternative models which might account for linkage disequilibrium in a frequently recombining population are satisfactory explanations for the population structure of serogroup A meningococci. The data are most compatible with clonality.
Yet, clonal descent should not be possible in the presence of frequent recombination. We suggest that gene flow into serogroup A meningococci is insufficient to disrupt clonality in part because of the geographical spread of these microorganisms. Rather than an epidemic structure which reflects temporary population expansion, we argue that an important mechanism contributing to apparent clonality in serogroup A meningococci is the frequent purification of sequence variants.

These bacteria do not persist in individual areas for more than a few years, and thus, do not have the opportunity to accumulate genetic variants indefinitely. Their survival depends on geographic spread. In turn, epidemic spread from country to country is accompanied by bottlenecks that reduce sequence variation introduced by both recombination and mutation (Morelli et al., 1997). Furthermore, many genetic variants are less fit than their parents and are removed by competition (Zhu et al., 2001). The degree to which such explanations apply to other meningococci and other pathogenic species will first become clear once the parameters of these various phenomena have been estimated quantitatively and subjected to mathematical modeling. However, it seems possible that similar considerations will apply to other "hyper-virulent clones" of _N. meningitidis_ and clonal groupings within other epidemic pathogens such as _Vibrio cholerae, Shigella_ and certain _Salmonella_. We note that it is not necessary for any of these species to manifest a single population structure. In particular, different isolates of _N. meningitidis_ might differ in their population structure in a subgroup-dependent manner due to different ecological lifestyles (Spratt et al., 1995).

In conclusion, we have tested the epidemic clonality model and the cryptic speciation model with serogroup A _N. meningitidis_ by using the very procedures proposed by the authors of those models (Maynard Smith et al., 1993). Our results do not support either of those models, but rather favor a clonal evolution model (Tibayrenc, 1995). Molecular typing by multilocus methods is a suitable approach for epidemiological surveys of predominantly clonal bacteria but is inherently unsuitable for panmictic bacteria or bacteria with transient epidemic clonality. The data presented here shows that serogroup A isolates are sufficiently clonal to warrant long-term typ- ing in order to follow these bacteria during their global spread.

**Acknowledgements**

We are grateful to Ed Feil for providing previously unpublished data at an early stage of this analysis. Aldert Bart gratefully acknowledges the Stimulationfonds voor Internationaaliserings (SRK, grant no. 15.02409) of the Netherlands Organization for Scientific Research (NWO) that enabled a visit to the Centre d’Etudes sur le Polymorphisme des Microorganismes in Montpellier (France), where this work was initiated. This publication made use of the Multi Locus Sequence Typing website (http://mlst.zoo.ox.ac.uk) developed by Dr. Man-Suen Chan and sited at the Wellcome Trust Centre for the Epidemiology of Infectious Disease, University of Oxford. The development of this site is funded by the Wellcome Trust.

**References**


