Networks and Groups Within the Genus *Neisseria*: Analysis of *argF*, *recA*, *rho*, and 16S rRNA Sequences from Human *Neisseria* Species

Noel H. Smith,Edward C. Holmes, Georgina M. Donovan, Gillian A. Carpenter, and Brian G. Spratt

School of Biological Sciences, University of Sussex, Brighton, U.K.; Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, Oxford, U.K.; and School of Biological Sciences, University of Birmingham, Birmingham, U.K.

To understand the pattern of nucleotide sequence variation among bacteria that frequently exchange chromosomal genes, we analyzed sequences of the *recA*, *argF*, and *rho* genes, as well as part of the small-subunit (16S) rRNA gene, from about 50 isolates of human commensal *Neisseria* species and the pathogenic *N. meningitidis* and *N. gonorrhoeae*. Almost all isolates of these species could be assigned to five phylogenetic groups that are found for all genes examined and generally are supported by high bootstrap values. In contrast, the phylogenetic relationships among groups varied according to the gene analyzed with notable incongruences involving *N. cinerea* and *N. lactamica*. Further analysis using split decomposition showed that for each gene, including 16S rRNA, the patterns of sequence divergence within *N. meningitidis* and closely related species were inconsistent with a bifurcating treelike phylogeny and better represented by an interconnected network. These data indicate that the human commensal *Neisseria* species can be separated into discrete groups of related species but that the relationships both within and among these groups, including those reconstructed using 16S rRNA, have been distorted by interspecies recombination events.

Introduction

With the application of molecular techniques to bacterial populations, it has become increasingly apparent that biochemical and serological tests, the traditional methods for identifying bacterial species, have failed to recognize monophyletic groups of bacteria (Ochman et al. 1983; Selander et al. 1990). Which species concept best applies to bacteria, indeed, the very existence of bacterial species, has recently become a subject for debate (Dykhuizen and Green 1991; Cohan 1994; Mallet 1995; Maynard Smith 1995). One problem is that the extent of recombination within different bacteria appears to vary widely, and bacteria can have population structures ranging from the fully sexual (nonclonal) to the highly clonal (Maynard Smith et al. 1993), suggesting that no single species concept will be universally applicable. *Neisseria meningitidis* (the meningococcus) provides one of the best examples of a bacterium in which recombinational exchanges are frequently, permitting the transfer of genes between these naturally transformable species. *Neisseria* are unusual among transformable bacteria in being fully competent for DNA uptake throughout their life cycle; several studies have demonstrated that transformation can result in recombinational exchanges between *Neisseria* species that differ by as much as 25% in nucleotide identity (Spratt et al. 1989; Bowler et al. 1994; Feil, Carpenter, and Spratt 1995; Zhou, Bowler, and Spratt 1997).

Historically, the systematics of *Neisseria* have been problematic and have required frequent readjustment as new techniques became available. Initially, DNA-DNA hybridization divided *Neisseria* into two subgroups: the closely related pathogens *N. gonorrhoeae* and *N. meningitidis*, and a related group of commensal *Neisseria* species (*N. perflava*, *N. subflava*, *N. sicca*, *N. flavescens*, and *N. flava*; Kingsbury 1967). The species *N. flava*, *N. subflava*, and *N. perflava* have since been consolidated into the species *N. subflava* (Reyn 1974), and *N. lactamica* and *N. polysacchareae* have been added to the group of commensal *Neisseria* species (Guibourdenche, Popoff, and Rou 1986). Analysis of ribosomal RNA genes by hybridization divided *Neisseria* into three main subgroups (Rossau et al. 1989). One group comprised *N. flavescens*, *N. subflava*, and *N. sicca*, while the *N. gonorrhoeae* group included *N. gonorrhoeae*, *N. meningitidis*, *N. lactamica*, *N. polysacchareae*, and *N. cinerea*. Strains of *N. mucosa* occupied a third subgroup.
More recently, Barrett and Sneath (1994) presented a numerical taxonomic study of the genus *Neisseria* based on an examination of 315 strains for 155 phenotypic characters. This extensive analysis showed that some species within the genus *Neisseria* formed reasonably distinct and isolated entities (designated “phenons”). However, certain phenons were not as distinct as those that would normally be accepted in a numerical taxonomic analysis, and some well-established species were misplaced. Of particular interest is the placement of the species *N. cinerea* within the *N. meningitidis* group, and that of some strains of *N. elongata* with *N. meningitidis* and *N. gonorrhoeae*. Although the pathogenic *Neisseria* species were reasonably distinct in this analysis, the authors concluded that few of the human commensal species can be well defined phenotypically.

Previous studies have shown that although genetic exchange occurs frequently within *N. gonorrhoeae* and *N. meningitidis* populations (Maynard Smith et al. 1993), the rate of gene exchange between the two pathogens is not high enough to blur the boundary between them (Vázquez et al. 1993). *Neisseria gonorrhoeae*, therefore, represents a “sexual” bacterial population that appears to be genetically isolated in nature and thus conforms to the biological species concept. We have suggested that ecological isolation—of populations that primarily colonize the genital tract from those that colonize the nasopharynx—may be an important component of the barrier to gene flow between these two very closely related pathogens (Vázquez et al. 1993). The high levels of carriage of commensal *Neisseria* and meningococci in the human nasopharynx suggest that ecological barriers to the exchange of chromosomal genes should be absent, and, in contrast to the infrequency of recombination between gonococci and meningococci (Spratt et al. 1995), highly localized recombinational exchanges between human commensal *Neisseria* species and *N. meningitidis* have been observed in almost every meningococcal housekeeping gene examined, including *adk*, *argF*, *glnA*, and *aroE* (Zhou and Spratt 1992; Feil, Carpenter, and Spratt 1995; Feil et al. 1996; Zhou, Bowler, and Spratt 1997). It is therefore of interest to determine to what extent a long history of recombination between chromosomal genes has blurred the species boundaries between the human commensal *Neisseria* species and confounds attempts to provide a robust taxonomy or phylogeny of these organisms.

To address this question, we sequenced three housekeeping genes and the small-subunit (16S) rRNA gene from a range of commensal *Neisseria* species usually found in the human nasopharynx.

**Materials and Methods**

**Neisseria Strains**

The sequences of *argF*, *recA*, *rho*, and the 16S rRNA gene were obtained from 26 strains selected from the collection of Barrett and Sneath (1994). Eighteen of these were from strains identified as closely related to *N. meningitidis* (Area A of Barrett and Sneath 1994), particularly those from *N. cinerea*, *N. polysaccharea*, and *N. lactamica*. Two strains that clustered close to phenons 1–6 were selected, as well as two strains of *N. elongata* (phenon 7) that clustered within Area A. More distantly related strains, from Areas B and C of Barrett and Sneath (1994), included the type strains of *N. perflava* (phenon 10), *N. sicca* (phenon 13), and *N. subflava* (phenon 21), and a strain of *N. mucosa* from phenon 9. Strains of *N. elongata* from phenon 20 were included to provide a comparison with sequences obtained from strains of *N. elongata* of phenon 7.

Strains from Area A of Barrett and Sneath (1994) (Strains are grouped into phenons, and the strain designation of Barrett and Sneath (1994) is given in parentheses. Type strains of *Neisseria* species are marked with an asterisk.)—Phenon 4: *N. cinerea* (F1)*, *N. cinerea* (F3), and *N. cinerea* (F10). Phenon 6: *N. polysaccharea* (P1), *N. polysaccharea* (P4), *N. polysaccharea* (P7), and *N. polysaccharea* (P8). Phenon 7: *N. elongata* (B8) and *N. elongata* var. glycolytica (J2)*. Phenon 8: *N. lactamica* (L19), *N. lactamica* (L17), *N. lactamica* (L20), *N. lactamica* (L18), *N. lactamica* (L22), and *N. lactamica* (L13). Strain *N. cinerea* (F2) clustered with meningococci in our study and was reidentified by standard biochemical and serological criteria as *N. meningitidis* serogroup B (J-Y. Riou and M. Guibourdenche, personal communication); it will be referred to here as *N. meningitidis* (F2).

Strains clustering close to phenons 1–6—*N. meningitidis* (B17) and *N. mucosa* (M7). The latter strain clustered unusually in our study and was subsequently reidentified by standard biochemical tests as *N. perflava* (J-Y. Riou and M. Guibourdenche, personal communication); it will be referred to here as *N. perflava* (M7).


Strains from Area C of Barrett and Sneath (1994)—Phenon 20: *N. elongata* (I1)*, *N. elongata* (I2), *N. elongata* (I4), *N. elongata* var. intermedia (I3), and *N. flava* (U40). Phenon 21: *N. subflava* (U37)*.

*Neisseria meningitidis* and *N. gonorrhoeae* strains—Sequences of *recA* and *argF* have previously been obtained for eight strains of *N. meningitidis* that were chosen to represent the diversity within this species uncovered by multilocus enzyme electrophoresis (Zhou and Spratt 1992). To complement these data, the sequences of *rho* and a fragment of the 16S rRNA gene were obtained from the same eight strains. Sequences of *rho* and *recA* have previously been obtained from eight strains of *N. gonorrhoeae*. All of these strains had identical *recA* sequences (Vázquez et al. 1993), and their *rho* sequences were also identical, except at one position which was outside the region that was analyzed in this work (E. J. Feil, personal communication). The sequence of *argF* was available from *N. gonorrhoeae* strain FA19 (Zhou and Spratt 1992) and the sequence of the 16S rRNA gene fragment was obtained from the same gonococcal strain.

Additional commensal strains—The *recA*, *rho*, and *argF* genes have previously been sequenced from strains of *N. cinerea* LNP1646, *N. polysaccharea* NCTC-
11858*, N. mucosa LNP405, N. flavescens LNP444*, N. lactamica NCTC10617*, N. pharyngis var. flava NCTC4590, and N. animalis NCTC10212* (Zhou and Spratt 1992; E. J. Feil, personal communication). The sequences of the 16S rRNA gene of these strains were obtained. The 16S rRNA gene was also sequenced from two further N. sicca strains (Q13 and Q28) from Area B of Barrett and Sneath (1994).

The al sequences used in this study are available in GenBank with accession numbers AJ223862–AJ223974 and AJ239278–AJ239315.

Nucleotide Sequencing

Direct sequencing of both strands was carried out on PCR fragments generated from purified chromosomal DNA; PCR primers for recA and argF have previously been described (Zhou and Spratt 1992). PCR products of the rho gene were generated from strains of groups A and C of Barrett and Sneath (1994) using primers rho1-up (5′-ATGCACGTCCTCGAATTCA-3′) and rho-down (5′-CCGGCATCAGTCAAAAGAACCTC-3′) (E. Feil, personal communication). For all other strains, rho1-up was replaced by rho2-up (5′-CAAATCCG C/A CGCTTCAA C/T CTGCA-3′). The PCR primers for the small-subunit rRNA genes, rRNA1 (5′-GCTATGA-AGCATCCGGTGGT-3′) and rRNA2 (5′-CATAAGAGTTTGTACCTGCTGCT-3′), were designed by Dr. L. Bowler from a compilation of available Neisseria 16S rRNA sequences.

The majority of the recA sequences were generated by manual sequencing of single-strand DNA generated by the lambda exonuclease method (Higuchi and Ochman 1989). The sequences of the argF, rho, and 16S rRNA genes were generated by automated sequencing (ABI Prism 377) of Qiaquick purified PCR fragments using standard techniques. Internal sequencing primers were designed as needed for each gene and group of strains. All sequences were determined on both strands.

Sequence Analysis

Phylogenetic trees of the nucleotide sequences from each gene were reconstructed using maximum likelihood. In every case, the HKY85 model of DNA substitution was used with values for the transition: transversion (Ts/Tv) ratio and α, the shape parameter of a gamma distribution (with eight categories) of rate variation among sites, estimated from the data during tree reconstruction. The values of these parameters were as follows: argF—Ts/Tv = 2.673, α = 0.392; recA—Ts/Tv = 2.945, α = 0.240; rho—Ts/Tv = 3.651, α = 0.186; 16S rRNA—Ts/Tv = 5.360, α = 0.003. As expected, the 16S rRNA gene exhibited much greater among-site rate variation than the other genes. Although some of the percentages of divergence between the sequences estimated under this substitution model were large, such that unrecognized multiple substitution may have occurred between the most distantly related taxa and especially at silent sites (Maynard Smith and Smith 1996), this will most likely affect the deep structure of phylogenies constructed, which is not the main point of interest here. Furthermore, and to assess the robustness of the groupings obtained, a bootstrap analysis was undertaken for each gene (1,000 resamplings) using trees reconstructed with the neighbor-joining clustering method, but incorporating the maximum-likelihood substitution parameters. All of these analyses were undertaken with the 4.0d64 test version of PAUP*, kindly provided by David L. Swofford.

In addition to screening for phylogenetic incongruence, we also asked how recombination might affect the evolutionary relationships among the Neisseria species using split decomposition, a method which depicts all the shortest pathways linking sequences, rather than just those that conform to a standard bifurcating tree (Bandelt and Dress 1992; Huson 1998). Split decomposition is therefore a powerful means to visualize the extent of conflicting phylogenetic signal in the data. Under this analysis, recombination will give rise to sequences linked to multiple ancestors, which will be depicted as an interconnected network of phylogenetic relationships, resembling a parallelogram in shape (Page and Holmes 1998). This analysis was carried out using the SplitsTree program, version 2.4 (Huson 1998), using pairwise distances estimated under the Kimura 3-ST model (Kimura 1981).

Finally, to characterize some recombination events in more detail, the maximum chi-square test (Maynard Smith 1992) was applied to pairs of sequences, using all nucleotide sites, and the significance of any clustering of polymorphic sites was determined using 1,000 randomly generated data sets.

Results

Phylogenetic Relationships Inferred from argF Sequences

The sequence of a 696-bp internal fragment of argF (corresponding to codons 68–299 of the sequence from N. gonorrhoeae) was available from 43 isolates. Eight strains of N. meningitidis had identical nucleotide sequences, and identical sequences were also recovered from strains N. cinerea (F10) and N. cinerea (F3) and from strains N. polysaccharea (P4) and N. polysaccharea (NCTC11858). A single representative of each of these identical sequences was selected, and the final data set therefore contained 34 unique nucleotide sequences.

Figure 1 shows a maximum-likelihood tree of these sequences. Five main groups of strains, supported by high bootstrap values, have been marked on the phylogeny:

Group 1. The Meningitidis group, containing all sequences from strains of N. meningitidis, N. gonorrhoeae, N. polysaccharea, N. lactamica, and N. perflava (M7).

Group 2. The Flavescens group, containing sequences from strains of N. flavescens (LNP444) and N. mucosa (M5) and the type strains of N. perflava (U15) and N. subflava (U37).

Group 3. The Cinerea group, containing sequences from all strains of N. cinerea.

Group 4. The Pharyngis group, containing sequences from N. pharyngis var. flava (NCTC4590), N.
Fig. 1.—Maximum-likelihood phylogenetic tree of argF sequences. Bootstrap values for selected nodes (estimated on neighbor-joining trees) are shown. Groups 1–5 are indicated. For clarity, group 3 (the Cinerea group) is circled, and a group of five N. lactamica strains, distinct from the other members of group 1, are boxed. All branch lengths are drawn to scale.

argF

Fig. 2 shows a maximum-likelihood phylogenetic tree for the 40 unique nucleotide sequences of recA. A further 80 trees of equal likelihood were recovered, but they differ only in minor ways from that shown here. The five groups identified in argF are easily discernible, although the relationships among them were different. The most clear difference was that group 3 (the Cinerea group) was related, although distantly, to group 1 strains with recA but was placed very close to (in fact, derived from) group 4 (the Pharyngis group) in the analysis of argF. Again, there was very little structure within the groups, although N. lactamica isolates L19 and L22 are now clearly separated from the other group 1 strains in this analysis, in contrast to the close association they show in argF. Three of the five N. polysaccharea isolates (P1, P4, and NCTC11858) were resolved from the other strains in group 1 (93% bootstrap support).

Phylogenetic Relationships Inferred from rho Sequences

The sequence of 1,026 bp of coding sequence of the rho gene (codons 40–381 of the sequence of N. gonorrhoeae [Miloso et al. 1993]) was available from 50 strains. All eight gonococcal sequences were identical. Five other pairs of identical nucleotide sequence were identified (N. elongata var. intermedia [I3] and N. elongata [I8], N. cinerea [F10] and N. cinerea [F3], N. polysaccharea [P4] and NCTC11858, N. lactamica [L19] and [L22], and N. meningitidis [B17] and [HF46]) and are represented by single isolates in the final data set. Figure 3 shows a maximum-likelihood phylogenetic tree for the 38 unique nucleotide sequences of rho.
**RecA**

Fig. 2.—Maximum-likelihood phylogenetic tree of recA sequences. The groups identified in figure 1 are indicated, including the group 3 species and the *N. lactamica* strains that are divergent in argF. All branch lengths are drawn to scale.

Again, the groups described for argF are identifiable. A major difference between the phylogeny obtained with rho and those obtained with argF and recA was the very close (and strongly supported) association of the Cinerea group with group 1, especially with a subset of strains from *N. lactamica* and *N. polysaccharea* P1 and P8. Very little structure was found within the groups although *N. polysaccharea* P4, P7, and NCTC11858 were closely related (as they were at argF, although not at recA) and were distinct from the other *N. polysaccharea* isolates.

Phylogenetic Relationships Inferred from 16S rRNA Gene Sequences

The sequence of a 1,355-bp internal fragment of the 16S rRNA gene, commencing at position 71 of the sequence of the 16S rRNA of *N. gonorrhoeae* (Rossau, Heyndrickx, and Van Heuverswyn 1988), was obtained from 44 strains. Eight pairs of strains had identical sequences, and the final data set contained 36 unique sequences. The identical sequences were *N. elongata* I8 and *N. elongata* var. *intermedia* I3, *N. elongata* I2 and I4, *N. cinerea* F3 and F10, *N. polysaccharea* P4 and NCTC11858, *N. polysaccharea* P1 and P7, *N. perflava* M7 and *N. polysaccharea* P8, *N. meningitidis* HF116 and HF46, and *N. meningitidis* HF130 and N94II. Figure 4 shows a maximum-likelihood tree of these data. The tree has a very unusual structure, in which the five groups, although visible, are often not monophyletic. This is clearly seen for group 4, which is obviously polyphyletic, and group 2, in which *N. subflava* (U37) occupies an outlying position. While this may in part be due to the low rate of nucleotide substitution in this gene, which will result in a lack of phylogenetic resolution among closely related taxa, it is also likely that recombination is in part responsible. This is most apparent with respect to the *N. cinerea* isolates, which cluster close to those from group 1, and those of *N. lactamica*, which are distantly related to the other group 1 isolates, in contrast to the picture seen for the other genes. Surprisingly, the neighbor-joining bootstrap trees of these data differed from those estimated using maximum-likelihood in that group monophyly was well supported for groups 3 and 4, and less so for group 2 (the
relevant bootstrap values are given in parentheses in fig. 4).

There was also a surprising amount of diversity within the seven *N. lactamica* 16S rRNA gene sequences (up to 3%, which was twice that found between any of the other species within group 1). The randomness of the distribution of polymorphic sites between the sequences of isolates from four *N. lactamica* sequences (L18, L20, L22, and NCTC10617), representative of the total amount of diversity observed, was analyzed with the maximum chi-square test (Maynard Smith 1992). A highly nonrandom clustering of polymorphic sites was found ($P < 0.001$) for all pairwise comparisons except L18 versus NCTC10617. Examination of the complete *N. lactamica* data set revealed that the 16S rRNA gene fragments from strains L13, L17, and NCTC10617 were diverged throughout their length compared with those of other species within group 1 (fig. 5a). *Neisseria lactamica* strain L20, on the other hand, is located on a unique and long branch in the rRNA gene tree and appears to be a recombinant between the diverged L17 sequence and that of isolates from within group 1; strain L20 was identical to strain L17 from nucleotide 236 to nucleotide 1355, but differed at 29 sites between nucleotides 1–235, and this part of the gene was very similar to that of isolates of group 1, particularly *N. polysaccharea* (Fig. 5b).

### Analysis of Gene Sequences by Split Decomposition

We also examined how recombination might affect the phylogenetic relationships among the *Neisseria* by analyzing sequences using split decomposition (Bandelt and Dress 1992). Because this method does not assume a tree-like process of divergence a priori, conflicting phylogenetic signal in the data, such as that caused by recombination, will generate an interconnected network of lineages.
Interspecies Recombination and Neisseria Phylogeny

**FIG. 4.**—Maximum-likelihood phylogenetic tree of 16S rRNA sequences. The groups identified in figure 1 are indicated, including the group 3 species and the *N. lactamica* strains that are divergent in argF. The neighbor-joining bootstrap trees differed from the maximum-likelihood trees in that they supported the monophyly of groups 3 and 4. These bootstrap values are given in parentheses next to the group names, as is that for group 2, although monophyly in this case received much less support. The seven *N. lactamica* strains were also found to be monophyletic, but with less bootstrap support. All branch lengths are drawn to scale.

**FIG. 5.**—Mosaic structure in the 16S rRNA gene of *N. lactamica*. Only the polymorphic sites are shown, and the positions of these sites in the sequenced fragment are shown above, in vertical format. The nucleotides are shown for the upper sequence, and in the other sequences, only those nucleotide positions that differ from the upper sequence are shown. Dots represent nucleotides that are identical to those in the upper sequence.

Little evidence for networked evolution was observed within the complete strain sets sampled for each gene (data not shown). However, because of the long lengths of many of the branches (i.e., those linking groups), which might overwrite any networked evolution within groups, we also analyzed separately those sequences from group 1, for which most data are available (because of the short branch lengths and the separation of the *N. lactamica* strains from the other group 1 sequences, those strains from groups 2 and 3 were included in this analysis of the 16S rRNA gene). In every case, at least some of these sequences formed an interconnected network, which was most apparent in argF and the 16S rRNA gene (figs. 6 and 7, respectively). For example, in the rRNA splits graph, *N. lactamica* strain L20 has pathways both to the other *N. lactamica* strains and to strains of the other groups, as expected if this sequence is a recombinant. Another possible example of recombination concerns *N. subflava* (U37), which is connected to other members of group 2 as well as to the group 4 strains. The splits graphs for recA and rho are available from the authors on request. Overall, such complex evolutionary patterns constitute strong evidence for recombination.

**Discussion**

The primary purpose of this work was to establish if there was any consistency in the phylogenetic rela-
Fig. 6.—Splits graph of group 1 argF sequences. The complex network indicates the lack of a treelike relationship between these sequences. A fit parameter of 61.7% was obtained, indicating that all of the phylogenetic signal in the data is difficult to represent in the splits graph. All branch lengths are drawn to scale.

tionships of human nasopharyngeal *Neisseria* isolates reconstructed using a number of housekeeping genes. Our most notable observation is that *Neisseria* isolates can be divided into five internally consistent groups, supported by high bootstrap values, but that the relationships among groups vary depending on the gene analyzed. Although a strict definition of a group cannot be made, the presence of phylogenetically distinct clusters of sequences suggests that sequence variation is not continuous across the human commensal *Neisseria* species and that some degree of ecological isolation has occurred. For example, the Elongata group was distinct in all analyses, and the clustering of the *N. meningitidis*, *N. gonorrhoeae*, *N. polysaccharea*, and *N. lactamica* isolates that was obtained with all genes except the 16S rRNA gene agrees with many previous studies (Guibourdenche, Popoff, and Riou 1986; Rossau et al. 1989; Barrett and Sneath 1994). Indeed, these very closely related isolates are given species names to reflect their very different medical significance. It is interesting that microbiologists have little difficulty distinguishing among these “species,” although none of them was recognized as a consistently distinct group using sequence data.

That the gene trees are incongruent in places is taken as evidence for past recombination events. This is most clearly seen in two places. The first concerns the group 3 (Cinerea) sequences. For argF, the Cinerea group was closely associated with group 4 (the Pharyngis group), while for recA, it was closer to group 1 (the Meningitidis group), as it was for 16S rRNA, while for rho, these sequences were located deeper within group 1. The second example concerns *N. lactamica*. There are five strains of *N. lactamica* that clearly group together in the argF tree (fig. 1). The branch leading to this group has a high bootstrap value, and the group is obviously distinct from the remaining group 1 argF sequences. A group of distinct *N. lactamica* strains is also identified using recA, yet only two of the strains in the recA group (L13 and L18) are members of the divergent *N. lactamica* group identified by argF. Finally, in 16S rRNA, all the *N. lactamica* strains group together, but a long way from the other members of group 1.

Recombination may also explain why for each gene, very few distinct clusters of isolates were obtained within group 1, and, equally, why there were few clusters of isolates recognized by all four genes (*N. lactamica* strains L19 and L22 and *N. polysaccharea* strains NCT11858 and P4 were the only examples of isolates that clustered together with all four genes). This notion is supported by the split decomposition analysis, which reveals that many of the group 1 strains can be linked to each other (and, in the case of 16S rRNA, to other groups) by multiple pathways. Split decomposition has also been shown to give networks of relationships within housekeeping genes from different meningococcal isolates (Holmes, Urwin, and Maiden 1999), in agreement with previous studies that have detected frequent recombination within such genes (Zhou and Spratt 1992; Feil, Carpenter, and Spratt 1995; Zhou, Bowler, and Spratt 1997).
The sequences of 16S rRNA genes have been used frequency to define the phylogenetic relationships between bacterial species and might have been expected to be useful in this case. Yet the phylogenies we present using 16S rRNA sequences are also problematic. More specifically, although apparent, the five groups are not always monophyletic. A more obvious anomaly was the location of the *N. lactamica* isolates, which were placed far away from the remainder of the group 1 isolates, in contrast to the picture seen for other genes. This again suggests that recombination has occurred during the evolutionary history of these strains and also that there has been insufficient time in this slowly evolving gene for the groups to diverge from their ancestral sequences and thus appear distinct. Such observations illustrate the problems of using a single phylogenetic marker, even one as popular as 16S rRNA, in reconstructing the phylogenetic history of closely related bacteria that frequently recombine. We suggest that a wider sample of genes be taken in future studies of the systematics of bacteria in which recombination may be frequent.

Because it examines a large number of phenotypic characters and should be relatively buffered from occasional changes in characters that result from recombinational exchanges between species, numerical taxonomy might be expected to give a fairly good approximation of the relationships between the *Neisseria* species. However, the numerical taxonomic study of the *Neisseria* (Barrett and Sneath 1994) produced some clear anomalies. In our study, the trees produced for each gene placed all of the *N. elongata* strains in a closely related group that was only distantly related to the isolates of group 1, whereas the same strains were separated by numerical taxonomy into two groups, one that was closely allied with the species in group 1 and one that was only distantly related to the group 1 species. Similarly, groups 2 and 4 contained strains from several of the phenons described by Barrett and Sneath (1994).
The phenetic approach is based on biochemical tests, and our work emphasizes the fact that strains of bacteria that are closely related by sequence data (e.g., the *N. elongata* strains) may have dissimilar biochemical reactions.

Finally, the evidence we present for a history of recombination during the divergence of the five major groups in *Neisseria* may initially appear inconsistent with the finding that the same groups of isolates can, in general, be discerned using the different genes. One possible explanation is that recombinational replacements among distantly related *Neisseria* species are rare compared with those between closely related isolates. A rare recombinational event that introduces a diverged sequence into a group 1 strain, for example, results, for that sequence, in a large genetic distance between the isolate involved and the other members of the group. Occasionally, such exchanges may be fixed in the population, but further recombinational exchanges will tend to occur with isolates within group 1, as is evident from the split decomposition analysis. The introduced diverged sequence will then gradually become homogenized among the group 1 isolates by highly localized recombinational exchanges, although the introduction of multiple novel polymorphisms will shift the relationship of the group to the other groups.

In conclusion, our results suggest that reasonably consistent groups of *Neisseria* isolates can be recognized using gene sequences, even though the divergence between the groups appears not to have arisen by a bifurcating treelike process and has been strongly influenced by a history of interspecies recombinational exchanges.

**Acknowledgments**

This research was supported by the Wellcome Trust and the Royal Society. B.G.S. is a Wellcome Trust Principal Research Fellow. We thank Prof. John Maynard Smith and Dr. Martin Maiden for discussions, Prof. Peter Sneath for supplying strains, Prof. J.-Y. Riou and Dr. M. Guibourdencie for strain identification, and an anonymous referee for comments on an earlier version of the manuscript.

**LITERATURE CITED**


**COHAN, F. M. 1994. Genetic exchange and evolutionary divergence in prokaryotes. TREE 9:175–180.**


**HIGUCHI, R., and H. OCHMAN. 1989. Production of single-stranded DNA templates by exonuclease digestion following polymerase chain reaction. Nucleic Acid Res. 17:5865.**


Howard Ochman, reviewing editor

Accepted February 18, 1999