MOLECULAR EPIDEMIOLOGY

A comparison of multilocus sequence typing and fluorescent fragment-length polymorphism analysis genotyping of clone complex and other strains of Neisseria meningitidis

J. V. HOOK and C. ARNOLD

Molecular Biology Unit, Sexually Transmitted and Blood Borne Virus Laboratory, Central Public Health Laboratory, 61 Colindale Avenue, London NW9 5HT

Five National Collection of Type Culture (NCTC) strains and 14 isolates of Neisseria meningitidis, representing 13 outbreak isolates from within the UK, were examined by multilocus sequence typing (MLST) for seven house-keeping genes. The results were compared with those of fluorescent amplified fragment-length polymorphism (FAFLP) analysis. Phylogenetic inferences were made from 3284-nucleotide lengths of sequence for the 19 isolates, by distance and parsimony methods. Two clusters of isolates were delineated. The larger, comprising eight isolates – S1, S3, Ironville, P9, ET-37 (M99-241951), P7, P10 and P60 – shared 100–99.2% similarity and varied in only 40 nucleotides (~1.22% variation) from the consensus sequence alignment. This cluster could be equated to the ET-37 complex because it had allelic signatures identical to MLST sequence types 11 and 50. These eight isolates were also assigned to one group by FAFLP. The reference ET-5 complex isolate ‘ET-5 (NG144/82)’ and an isolate (X9) from an outbreak in the north of England were also grouped together by MLST. They shared 99.2% similarity and differed within the avfE and famC genes by 4 and 17 nucleotides, respectively. Their MLST sequence types were 32 and 661 and, therefore, these two isolates could be equated to the ET-5 complex. They also grouped together by FAFLP. A comparison of the resources required to apply MLST to the 19 isolates examined with those needed to characterise them by FAFLP indicated that FAFLP (a fragment-based genotyping method) is more cost-effective than the partial sequencing approach, MLST.

Introduction

Neisseria meningitidis is an important cause of septicaemia and meningitis world-wide [1]. The predominant serogroups in Europe, America, Australia and Canada are B and C, whereas in Asia and Africa serogroup A predominates [2–4]. In England and Wales, 50% of confirmed cases have been serogroup B, 35% serogroup C and the remainder minor serogroups (W135, X, Y and 29E) or non-groupable [5]. Epidemic meningococcal disease has been attributed to a few clonal groupings including the electropherotype (ET) 37 complex and the ET-5 complex [2]. Timely recognition of cases due to such clonal groupings would raise awareness of the threat of epidemic disease and facilitate control measures.

Recently, the classical method of multilocus enzyme electrophoresis (MLEE [6]) has been superseded by multilocus sequence typing (MLST [7]). It has recently been proposed that MLST, based on DNA sequencing of ‘house-keeping’ genes, should be preferred for identifying clonal groups [8–10]. Concurrently, the application of fluorescent amplified fragment-length polymorphism (FAFLP) analysis to N. meningitidis has been reported and it has been established that some isolates from meningococcal disease outbreaks within the UK form distinct clusters [11]. However, in the absence of MLST data these clusters can be only tentatively equated with ET complexes. To investigate this further, representative strains and isolates from the earlier FAFLP study [11] have been chosen for analysis by MLST. The results are discussed in the context of the need to recognise clone complexes among disease-associated and other isolates of N. meningitidis quickly. Therefore, the aims of this study were to compare the genotypic groupings revealed by MLST and FAFLP.
and to assess the other features of, and resources deployed in, the two methods.

**Materials and methods**

**Bacterial culture and DNA extraction**

Five NCTC strains and 14 isolates of *N. meningitidis* were examined (Table 1). Bacteria were grown for 24 h on chocolate agar at 37°C in an atmosphere of CO₂ 5% in air. Chromosomal DNA was isolated as described by Jones [12], with modifications [13] to use cetyltrimethylammonium bromide.

**MLST**

Amplification and sequencing of the MLST housekeeping genes were as described previously [9, 10], with the addition of primers for the amplification of the *fumC* gene: PCR primers *fumC*-A1 (5'-CAC CGA ACA CGA CAC GAT GG-3) and *fumC*-A2 (5'-ACG ACC AGT TCG TCAAAAAC TC-3) and sequencing primers *fumC*-S1 (5'-TCG GCA CGG GGT TAA TGA ACA GC-3) and *fumC*-S2 (5'-CAA CGG TCG TTT CGC GAC-3). The PCR product was purified by agarose gel electrophoresis with a Qiagen gel extraction kit (no. 28706; Qiagen Ltd, Crawley, W. Sussex RH10 2AX). Sequencing reactions were performed with Big Dye Terminators (PE Biosystems, Warrington, Cheshire WA3 7PB) and the products were separated and detected with an ABI Prism 377 automated sequencer (PE Biosystems).

**DNA sequence manipulations and analysis**

The sequence data at each of the seven loci examined were submitted to the MLST database (The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, Oxford OX1 2PS).

The contigs were assembled and consensus sequences were edited with SEQMAN (a module of DNA-STAR’s lasergene; http://www.dnastar.com) and SEQED (Applied BioSystems, Warrington, Cheshire WA3 7PB), respectively. The sequences for each of the seven housekeeping genes were aligned with MEGA-LIGN (http://www.dnastar.com) and the ends were trimmed to that of a reference allele. The individual sequences were then concatenated into a single file and converted to PHYLYP format by READSEQ (Gilbert DG, Biology Department, Indiana University, Bloomington, IN, USA; gilbert@bio.indiana.edu) before phylogenetic analyses. The final data matrix comprised 19 samples each with 3284 nucleotides.

With the PHYLYP phylogenetic inference package (http://evolution.genetics.washington.edu/phylip.html), data were bootstrapped (×50) by SEQBOOT (see PHYLYP), distances were calculated in DNADIST [14], trees were computed with FITCH [15] and a majority-rule consensus tree was produced in CONSENSE (see PHYLYP). Also, the programme PUZZLE (http://www.zi.biologic.uni-muenchen.de/vstrimmer/puzzle.html) was used to construct phylogenetic trees from

<table>
<thead>
<tr>
<th>Reference strain/isolate</th>
<th>Origin/year of isolation</th>
<th>Serogroup</th>
<th>MLST sequence type*</th>
<th>FAFLP cluster (group)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCTC 12083</td>
<td>Manchester, UK/1986</td>
<td>Serogroup C</td>
<td>(Unique) 735*</td>
<td>Unique</td>
</tr>
<tr>
<td>NCTC 10025</td>
<td>Chicago, USA/1937</td>
<td>Serogroup A</td>
<td>Unique</td>
<td>Unique</td>
</tr>
<tr>
<td>P4</td>
<td>Puntuppyllid UK/1999</td>
<td>C.1.Pi.15</td>
<td>Unique</td>
<td>Unique (VII)</td>
</tr>
<tr>
<td>P5</td>
<td>Puntuppyllid UK/1999</td>
<td>C.1.Pi.14</td>
<td>Unique</td>
<td>Unique (VII)</td>
</tr>
<tr>
<td>X17</td>
<td>Puntuppyllid UK/1999</td>
<td>Serogroup C</td>
<td>(Unique) 728†</td>
<td>3 (XII)</td>
</tr>
<tr>
<td>X6</td>
<td>Puntuppyllid UK/1999</td>
<td>Serogroup B</td>
<td>(Unique) 728†</td>
<td>3 (XII)</td>
</tr>
<tr>
<td>E5-5 (NG144820)†</td>
<td>Norway/1982</td>
<td>B.15.Pi.17.16</td>
<td>32</td>
<td>2 (X)</td>
</tr>
<tr>
<td>X9</td>
<td>Puntuppyllid UK/1999</td>
<td>B.15.Pi.17.16</td>
<td>661</td>
<td>2 (VIX)</td>
</tr>
<tr>
<td>E5-5 (NG144820)†</td>
<td>Norway/1982</td>
<td>B.15.Pi.17.16</td>
<td>661</td>
<td>2 (VIX)</td>
</tr>
<tr>
<td>NCTC 8554</td>
<td>Maryland, USA/1942</td>
<td>Serogroup C</td>
<td>Unique</td>
<td>Unique</td>
</tr>
<tr>
<td>NCTC 8569</td>
<td>Dordogne, France/1954</td>
<td>Serogroup C</td>
<td>Unique</td>
<td>Unique</td>
</tr>
<tr>
<td>NCTC 8249</td>
<td>Liverpool, UK/1952</td>
<td>Serogroup B</td>
<td>(Unique) 729†</td>
<td>Unique</td>
</tr>
<tr>
<td>S1</td>
<td>Southampton, UK/1998</td>
<td>C.1.Pi.5</td>
<td>50</td>
<td>(Ia)</td>
</tr>
<tr>
<td>S3</td>
<td>Southampton, UK/1998</td>
<td>C.1.Pi.5</td>
<td>50</td>
<td>(Ia)</td>
</tr>
<tr>
<td>Ironville (M97-252005)</td>
<td>Ironville, UK/1997</td>
<td>C.2a.Pi.5</td>
<td>11</td>
<td>(V)</td>
</tr>
<tr>
<td>P9</td>
<td>Puntuppyllid UK/1999</td>
<td>C.2a.Pi.5</td>
<td>11</td>
<td>(V)</td>
</tr>
<tr>
<td>P7</td>
<td>Puntuppyllid UK/1999</td>
<td>C.2a.Pi.5</td>
<td>11</td>
<td>(V)</td>
</tr>
<tr>
<td>P8</td>
<td>Puntuppyllid UK/1999</td>
<td>C.2a.Pi.5</td>
<td>11</td>
<td>(V)</td>
</tr>
<tr>
<td>P10</td>
<td>Puntuppyllid UK/1999</td>
<td>C.2a.Pi.5</td>
<td>11</td>
<td>(V)</td>
</tr>
</tbody>
</table>

*NK. not known.

*The MLST type was defined in Feavers et al. [9] and Maiden et al. [10] where sequence types 11 and 50 were equated to the ET-37 complex, and sequence type 32 to that of the ET-5 complex.

†The FAFLP cluster (group) was defined in Goulding et al. [11], and groups XIV to XII were recognised by extended analysis of the results.

*Type strain of *N. meningitidis*.

*Reference isolate of the ET-5 complex.

†New MLST type (K. Jolley, Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, Oxford OX1 2PS).

**Reference isolate of the ET-37 complex.
uncorrected (Hamming) distances by maximum like-
ilhood. The parsimony criterion in PAUP (http: //www.sinauer.com/Titles/1stwofford.html) was imple-
mented and the neighbour-joining method was used to
construct a consensus tree from the distance data. The
programme SPLITSTREE was also used to visualise
the data as split graphs generated by split decomposi-
tion analysis [16].

Results
MLST of reference strains and isolates of N. meningitidis

Five N. meningitidis NCTC strains representing
serogroups A, B and C, and 14 isolates from disease
outbreaks, including eight previously grouped by
FAFLP as cluster I [11], were analysed by MLST for
seven house-keeping genes. The resulting sequence
data (3284 nucleotides for each specimen) were
aligned, and phylogenetic inferences were made by
means of methods for constructing distance and
parsimony trees. The consensus alignment showed that
there were 2896 invariant nucleotides. A consensus
distance tree based upon bootstrapped data (∗50
reiterations) summarised the evolutionary relationships
between the strains and isolates examined (Fig. 1a). A
comparison of phylogenies based on different models
(the models implemented were distance, parsimony,
split decomposition) indicated that the branching order
was reliable (data not shown). The rate of nucleotide
substitution (Kω) among those meningococci exam-
ined varied from 0 (100% similarity) to 0.0599 (94.2%
similarity). The NCTC strains and four isolates (P4, P5,
X17 and X6) that did not form an MLST cluster were
considered unique (Fig. 1a, Table 1). The isolates X6
and X17 from outbreaks unrelated in time and
geographical location were not dissimilar, and had a
unique sequence type (Table 1). Eight isolates – S1,
S2, Ironville, P9, ET-37 (M99-241951), P7, P6 and P10 –
fell into a MLST cluster, were related at 100–96.3%
similarity and had allelic signatures characteristic of
sequence types 11 and 50. The isolates Ironville, ET-37
(M99-241951) and those from the Pontypridd (South
Wales) outbreak (P9, P7, P10 and P6) were identical.
The two remaining isolates within this cluster (S1 and
S3) differed from the aligned consensus in 40
nucleotides within the aroE gene and were sequence
type 50 (Table 1). The isolates ET-5 (NG144/82) and
X9 formed a second MLST cluster, were 99.2% similar
and differed only in four and 17 nucleotides within the
aroE and fumC genes, respectively Their MLST types
are given in Table 1, and can be equated to the ET-5
complex.

Comparison of MLST and FAFLP

The MLST and FAFLP trees for NCTC strains and
isolates of N. meningitidis were compared (Fig. 1a
and b). The NCTC reference strains (10025, 8554 and
8569) and two isolates (P4 and P5) were each
considered unique by both methods (Table 1). Isolates
X6 and X17 had sequence type 728 and could be
assigned to groups XI and XII by FAFLP. Eight isolates
– S1, S3, Ironville, P9, ET-37 (M99-241951), P7, P10
and P6 – formed a cluster by both genotyping methods
(Fig. 1a and b, Table 1). The representatives ET-5
(NG144/82) and X9 of the ET-5 complex (MLST
sequence types 32 and 661) also formed a cluster by
FAFLP. Isolates 17 and X6 had sequence type 28 and
formed an FAFLP cluster, 3 (Table 1). Other compara-
tive features of MLST and FAFLP, such as investiga-
tive cost per sample, are shown in Table 2.

Discussion

Although MLEE was formerly used to study the
population structure of bacterial species, it has recently
been superseded by MLST. In MLEE, relatively
invariant genes are recognised by the electrophoretic
mobilities of the enzymes that they encode, whereas in
MLST corresponding differences in the DNA se-
quences of these genes are determined. With automated
sequencing, MLST is objective, reproducible and
portable, and both MLST and MLEE have been used to
demonstrate that epidemic virulent strains of N.
meningitidis fall, in the main, into two lineages known
as the ET-37 complex and the ET-5 complex, ir-
respective of serogroup [9, 10]. It has also been
proposed that these lineages could be recognised by

In this study, eight isolates were found to have the
allelic MLST sequence types 11 and 50, consistent
with them being part of the ET-37 complex [9, 10],
and similar clustering occurred when the same eight iso-
lates were typed by FAFLP [11]. Two representatives
of a second globally occurring hyper-invasive complex
(ET-5), comprising at least three subpopulations [8, 10],
were also studied. The reference isolate ET-5 (NG144/
82) had an allelic profile and sequence type character-
istic of the ET-5 complex, and isolate X9 clustered
with this strain by MLST. Although assigned a
different sequence type (661), X9 could be considered
a variant, or representative of a subpopulation, within
the ET-5 complex. Isolates X9 and ET-5 (NG144/82)
had also clustered by FAFLP. Therefore, comparison of
MLST with FAFLP – in this study and elsewhere
[9, 11] – shows that both are able to recognise
outbreaks of the ET-37 and ET-5 complexes. However,
there are significant differences in the resources
deployed in the methods, as summarised in Table 2;
in particular, FAFLP offers higher cost benefits.

We thank the Meningococcal Research Group (Division of Micro-
biology, University of Nottingham, Queen’s Medical Centre, Notting-
ham) for supplying isolates and serological data. We are grateful to P.
P. Mortimer and J. P. Clewley for critically reading this manuscript.
Fig. 1. (a) Unrooted consensus distance tree [14, 15] based on MLST and (b) unrooted consensus distance tree (PAUP) based on AFLP (data extracted from that in Goulding et al. [11]) showing the phylogenetic position of strains and isolates of *N. meningitidis*. Strains designated NCTC were obtained from the National Collection of Type Cultures (London). Isolates with the prefix P were from an outbreak within the community of Pontypridd (South Wales [17]), those with the prefix S were from an outbreak within the University of Southampton [9, 18]. Those labelled X were from outbreaks unrelated in time and geographical location. The Ironville isolate was from a community outbreak in the village of Ironville (Derbyshire). ET-5 (NG144/82) and ET-37 (M99-241951) were reference isolates within the hyper-invasive clone lineages ET-5 and ET-37 complexes, respectively. 1AFLP cluster 1, 2AFLP cluster 2 and 3AFLP cluster 3 were defined in Goulding et al. [11]. The bootstrap percentages quoted are the percentage times that a strain or isolate occurred at that node. The scale bars represent 0.01 substitutions per sequence position (a) and one fragment difference (b).
Table 2. Comparison of MLST and FAFLP

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>MLST</th>
<th>FAFLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruence with MLST</td>
<td>Strong evidence</td>
<td>Some evidence*</td>
</tr>
<tr>
<td>Major equipment required</td>
<td>DNA sequencer</td>
<td>DNA sequencer</td>
</tr>
<tr>
<td></td>
<td>PCR thermal cycler</td>
<td>PCR thermal cycler</td>
</tr>
<tr>
<td></td>
<td>Computer and software</td>
<td>Computer and software</td>
</tr>
<tr>
<td>Culture from specimen</td>
<td>Not essential</td>
<td>Required</td>
</tr>
<tr>
<td>Number of PCR reactions/specimen</td>
<td>14 (4 index sequencing)</td>
<td>1</td>
</tr>
<tr>
<td>Number of PCR primer pairs/specimen</td>
<td>~3.5 h</td>
<td>~3.5–5 h</td>
</tr>
<tr>
<td>Time to run reactions on DNA sequence</td>
<td>~0.015</td>
<td>~0.05</td>
</tr>
<tr>
<td>Percentage of genome sampled</td>
<td>DNA sequence</td>
<td>PCR fragments</td>
</tr>
<tr>
<td>Data type</td>
<td>Nucleotide variations</td>
<td>Presence-absence of sized bands</td>
</tr>
<tr>
<td>Data analysis</td>
<td>Intensive</td>
<td>Intensive</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>~100%</td>
<td>100% (± 1 bp)</td>
</tr>
<tr>
<td>Level of resolution</td>
<td>Moderate (7 loci)</td>
<td>High¹</td>
</tr>
<tr>
<td>Portability</td>
<td>Yes</td>
<td>Unproven</td>
</tr>
<tr>
<td>Potential for automation</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Start-up costs</td>
<td>Similar to FAFLP</td>
<td>Similar to MLST</td>
</tr>
<tr>
<td>Estimated cost/specimen</td>
<td>£300 (7 loci)</td>
<td>£15¹</td>
</tr>
<tr>
<td>Time to analyse 19 specimens (one person)</td>
<td>3–4 days</td>
<td>3–4 days</td>
</tr>
<tr>
<td>Potential applicability</td>
<td>All bacterial genomes</td>
<td>Most bacterial genomes</td>
</tr>
</tbody>
</table>

* C. Arnold (personal communication).
¹ Feavers et al. [10].
² Goulding et al. [11].
³ Olive and Bean [19]. sequencing costs have fallen since this publication appeared.

References