Independent subsets of amplified fragments from the genome of *Neisseria meningitidis* identify the same invasive clones of ET37 and ET5

Jonathan N. Goulding, 1 John Stanley, 2 Will Olver, 3 Keith R. Neal, 4 Diawer A. A. Ala’Aldeen 3 and Catherine Arnold 2

1Department of Academic Paediatrics, Imperial College at St Mary’s Hospital, London, UK
2Molecular Biology Unit, Virus Reference Division, Central Public Health Laboratory, 61 Colindale Avenue, London NW9 5HT, UK
3,4Molecular Bacteriology and Immunology Group, Division of Microbiology 3 and Department of Public Health Medicine and Epidemiology 4, University of Nottingham, Queen’s Medical Centre, Nottingham, UK

To determine whether fluorescent amplified-fragment length polymorphism (FAFLP) analysis is an unbiased genome sampling technique, data were analysed from three different primer combinations, amplifying three independent fragment subsets from 123 isolates of *Neisseria meningitidis*. Using these data, dendrograms were generated with near-identical topologies that identified the same invasive clones of ET37 and ET5 and also identified the same outbreak clusters.

**INTRODUCTION**

The ability of *Neisseria meningitidis* to cause outbreaks with significant mortality is a source of considerable public health concern. All age groups, but especially children, may contract meningococcal septicaemia or meningitis and fatalities can occur in spite of antibiotic and supportive therapy. Five of the 13 serogroups, recognized on the basis of capsular polysaccharide antigens, are commonly associated with disease (A, B, C, Y and W-135), and antigenic diversity in the porB and porA outer-membrane proteins respectively defines serotypes and serosubtypes (Poolman et al., 1995).

Sensitive and reproducible meningococcal grouping and typing methods are required for population genetic and epidemiological investigations and for vaccine-related studies. However, the phenotypic methods used to characterize below the species and serogrouping level (e.g. serotyping) suffer from several defects including antigenic variability, poor expression or masking of surface antigens, the inability to type and subtype all isolates and the need continually to enlarge the reagent panel of monoclonal antibodies used for serotype and serotype characterization.

Molecular typing methods such as multilocus sequence typing (MLST) avoid these defects (Maiden et al., 1998). Genes determining two variable antigens (9-locus MLST), as well as slowly evolving ‘housekeeping’ genes, can be sequenced and identify both micro- and macrovariation between strains. The information from the variable genes can be used to identify the microvariation that distinguishes between strains circulating within a geographical location and the data from slowly accumulated variation in ‘housekeeping’ genes will reveal macrovariation and inform population genetics (Feavers et al., 1999).

Amplified-fragment length polymorphism (AFLP) analysis is a PCR-based genome sampling technique that generates a specific fragment profile for each bacterial clone (Vos et al., 1995). The technique is a convenient tool for the study of genetic diversity (Arnold et al., 1999; Dijkshoorn et al., 1996; Grady et al., 1996a, b; Janssen & Dijkshoorn, 1996; Janssen et al., 1996). For *N. meningitidis* outbreak strains, AFLP in its fluorescent format (FAFLP) yields groupings and a resolution between strains equal in precision to 9-locus MLST (Goulding et al., 2000).

FAFLP has been used successfully for genotyping *Streptococcus pyogenes* and *Staphylococcus aureus* as well as *N. meningitidis* (Grady et al., 1999; Goulding et al., 2000; Desai et al., 1998). In these studies, however, only one subset of amplified restriction fragments was analysed and the following two questions arise: (i) would different subsets of restriction fragments originating from completely different parts of the bacterial genome generate the same, or similar, dendrograms and (ii) would optimizing the selective primers result in higher discriminatory power?
METHODS

Strains. We examined the same 123 isolates of *N. meningitidis* described previously (Goulding et al., 2000). Thirty-three were disease-causing or disease-associated (from asymptomatic patient contacts). They included nine from the 1997 Southampton University outbreak in the south of England (S1 and S3–S10) and nine from the 1999 Pontypridd outbreak in Wales (P1 and P3–P10); 15 were from epidemiologically unrelated meningooccal disease (prefixed X) and were isolated between 1990 and 1993. Twenty-six strains, dating from 1934–1987, were from the National Collection of Type Cultures (NCTC; prefixed N). Fifty-eight carriage isolates were from throat swabs of healthy individuals from Ironville, a village in Derbyshire, UK, that experienced a protracted outbreak of invasive meningooccal disease (no prefix). Five other strains of known MLEE type 37 (ET37), including one from Ironville, and one known MLEE type 5 (ET5) strain were also analysed.

FAFLP analysis. DNA extraction and FAFLP analysis were carried out as described previously (Goulding et al., 2000). Briefly, extracted DNAs were digested with *Mse*I and *Eco*RI, followed by ligation of oligonucleotide adaptors (linkers) specific to the cut ends of the digested fragments. PCR was performed using primers directed to the adaptor sequences. The difference between this study and that described previously was that here, the *Eco*RI-directed primer, *Eco*X (5′-GACTGCGTACCAATTCX-3′), had a different selective base at the 3′ end, where X is either A, C or T (i.e. Eco-A, Eco-C or Eco-T), in each of three experiments so that the PCR would select fragments from completely different parts of the genome in each experiment. Following electrophoretic separation of the fluorescently labelled fragments, binary data (i.e. presence/absence of a fragment of a particular size) were exported into a custom-designed software application that calculated Dice coefficients between pairs of isolates. These data were exported into NEIGHBOR, a part of the PHYLIP package which uses the unweighted pair group method with arithmetic averages (UPGMA) to construct a tree by successive clustering using an average-linked method of clustering.

RESULTS

The number of fragments found in each subset differed slightly. For the Eco-A subset, there were 26–65 fragments generated per isolate, none of which was common to all of the 103 isolates and 26 of which were unique to individual isolates. For the Eco-C subset, there were 21–53 fragments generated per isolate, three of which were common to all and 32 of which were unique to one isolate. For the Eco-T subset, there were 21–39 fragments generated per isolate, three of which were common to all and 28 of which were unique to one isolate. An *in silico* digest of one of the two sequenced strains of *N. meningitidis* (Parkhill et al., 2000) (not included in the 123 isolates studied) predicted that, in total, 184 fragments of between 100 and 400 bp would be amplified from this isolate by FAFLP using non-selective *Eco*RI and

**Fig. 1.** Dendrogram representing FAFLP data using an ‘A’ selective base on the *Eco*RI-directed primer (*Eco*-A) for 123 *N. meningitidis* isolates. Two groups of confirmed invasive disease-causing isolates are circled.
Mel primers: Eco-A, 51 fragments; Eco-C, 36 fragments; Eco-T, 28 fragments; and Eco-G, 69 fragments. This is within the range found experimentally and also indicates the level of complexity required for analysis when compared with a technique such as PFGE that mostly looks at only 10 to 20 fragments.

Dendrograms derived from our results all showed similar topology. Those derived from selective primer Eco-A (Fig. 1), Eco-C or Eco-T (not shown) and from pooled data (Fig. 2) all displayed two disease-causing groups, each of an inferred hypervirulent lineage. Such a group was defined as a collection of disease-causing strains associating together with a disease isolate(s) of known hypervirulent lineage on a dendrogram. Group 1 was the presumed ET37 complex group, and all three subsets (Eco-A, Eco-C or Eco-T) contained the same 28 isolates previously defined as being part of this complex. Only three of the presumed ET37 isolates did not appear in all three subsets. In the Eco-A subset, isolate 563 was located within group 1, but not in the other two subsets. Isolate N6455 was located on the edge of group 1 with the Eco-C and Eco-T subsets, but not with the Eco-A subset. With subsets Eco-A and Eco-C, isolate N11884 was found within group 1, but lay outside it with subset Eco-T. Group 2 isolates were defined previously as being part of the ET5 complex group, and the same six isolates of this complex grouped together in all three subsets. Pooled data (Fig. 2) confirmed these groupings and defined the relationship of the three outlying isolates described above more clearly, placing them within group 1 (ET37/ST-11-like). Seventy of the isolates were amplified at least twice, and these duplicates gave identical fragments. Gel and amplification controls also gave identical fragment patterns.

DISCUSSION

Data from this study indicate that the hypervirulent lineage-associated groupings are almost entirely consistent within the three fragment subsets analysed. Because, when using the FAFLP method, every fragment from each of the three different subsets is derived from a different part of the genome, it can be assumed that the sampling of the N.
meningitis genome by FAFLP using any single selective base on one primer is unweighted and gives an accurate as well as convenient indication of whether an isolate belongs to a hypervirulent lineage. Also, the data shown here indicate that FAFLP samples both non-variable and variable regions, as it can measure microvariation (outbreak analysis) as well as macrovariation (hypervirulent lineage detection). Compared with MLST, which also looks at a mixture of non-variable and variable genes by sequencing and is the ‘gold standard’ for defining hypervirulent lineages, FAFLP samples different sites throughout the genome and would be suitable as a complementary method to MLST. A disadvantage of the FAFLP method, however, is that PCR-based techniques such as MLST can be optimized to be performed on a clinical sample, whereas an isolate is required for FAFLP. This requirement means that, including analysis, FAFLP can take 3–4 days to perform for an outbreak.

Selective primers are frequently employed in FAFLP to limit the number of amplified genomic fragments within the preferred size boundaries to a convenient range (50–100 fragments for the current laser readers on slab and capillary gel sequencers). Our findings provide reassurance that genome sampling by FAFLP is unbiased and generates a robust and reproducible result, no matter which selective primer is employed.

ACKNOWLEDGEMENTS

We are grateful to Dr Philip Mortimer, Central Public Health Laboratory, for advice and encouragement.

REFERENCES


