GENOMIC ANALYSIS OF SEROGROUP Y *NEISSERIA MENINGITIDIS* ISOLATES REVEALS EXTENSIVE SIMILARITIES BETWEEN CARRIAGE AND DISEASE-ASSOCIATED ORGANISMS

Neil J. Oldfield,[1]# Odile B. Harrison,[2] Christopher D. Bayliss,[3] Martin C. J. Maiden,[2] Dlawer A. A. Ala'Aldeen,[1] and David P. J. Turner[1]

[1]School of Life Sciences, University of Nottingham, Nottingham NG7 2RD, UK

[2]Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

[3]Department of Genetics, University of Leicester, Leicester LE1 7RH, UK

Running Head: Genomic analysis of MenY isolates

Type of article: Major article

Abstract word count: 196

Text only word count: 3074

References: 37

Correspondence: Centre for Biomolecular Sciences, University of Nottingham, Nottingham NG7 2RD, UK. E-mail neil.oldfield@nottingham.ac.uk; Tel. (+44) 115748 6122; Fax (+44) 115846 8002

ABSTRACT

Background. *Neisseria meningitidis* is a frequent colonizer of the human nasopharynx with asymptomatic carriage providing the reservoir for invasive, disease-causing strains. Serogroup Y (MenY) strains are a major cause of meningococcal disease. High resolution genetic analyses of carriage and disease isolates can establish epidemiological relationships and identify potential virulence factors.

Methods. Whole genome sequence data were obtained from UK MenY carriage isolates from 1997-2010 (n=99). Sequences were compared to those from MenY invasive isolates from 2010 and 2011 (n=73) using a gene-by-gene approach.

Results. Comparisons across 1,605 core genes resolved 91% of isolates into one of eight clusters containing closely related disease and carriage isolates. Six clusters contained carried meningococci isolated in 1997-2001 suggesting temporal stability. One cluster of isolates, predominately sharing the designation Y: P1.5-1,10-1: F4-1: ST-1655 (cc23), was resolved into a sub-cluster with 86% carriage isolates and a second with 90% invasive isolates. These sub-clusters were defined by specific allelic differences in five core genes encoding glycerate kinase (*glxK*), valine-pyruvate transaminase (*avtA*), superoxide dismutase (*sodB*) and two hypothetical proteins.

Conclusions. High resolution genetic analyses detected long-term temporal stability and temporally-overlapping carriage and disease populations for MenY clones but also evidence of a disease-associated clone.

Keywords: *Neisseria meningitidis*; whole genome sequencing; carriage; serogroup Y; epidemiology

BACKGROUND

Neisseria meningitidis is carried in the nasopharynx of 10 to 30% of the population with carriage rates influenced by setting and generally higher in young adults and close contact populations [1, 2]. Occasionally, meningococci invade, resulting in meningococcemia and meningitis. Invasive meningococcal disease (IMD) results in substantial mortality and morbidity despite effective antibiotic treatment [3].

A key virulence factor is the polysaccharide capsule, which allows the bacterium to resist complement-mediated lysis and opsonophagocytosis [4]. Twelve serogroups are recognized [5], but only six, A, B, C, W, X and Y are responsible for the majority of disease worldwide [6]. DNA sequence-based approaches have been extensively applied to the analysis of the population structure of meningococci [7]. Multilocus sequence typing (MLST), using sequences of seven representative housekeeping genes, has demonstrated a highly structured population, with most strains belonging to groups of closely related genotypes referred to as clonal complexes (ccs) [8]. Some of these clonal complexes correspond to 'hyperinvasive lineages', which are responsible for most cases of disease worldwide [9, 10]. These clonal complexes are often associated with specific combinations of antigenic proteins, such as Porin A (PorA) and Ferric enterobactin transport protein A (FetA), and serogroups [11, 12].

Much of the IMD in Europe and North America is caused by a limited range of serogroup/genotype combinations, for example serogroup B (MenB) ST-41/44, ST-32 and ST-269 isolates and serogroup C (MenC) isolates from ST-11 and ST-8 complexes [6, 13]; however, in recent decades the incidence of IMD due to MenY organisms, often belonging to cc23, has increased in several countries, including the USA, Sweden, and the United Kingdom [14-18]. In the UK, several carriage studies performed between 2008 and 2012 detected evidence of recent

alterations in MenY carriage epidemiology in young adults [19-22]. For example, MenY meningococci were found in only 1-2% of participants and constituted only *ca.* 10% of recovered isolates when carriage was assessed in 1997-8 in first-year university students at the University of Nottingham, UK and during 1999-2001 in >48,000 15-17 year-old school students throughout the UK [23, 24]. In contrast, in 2008-9 and 2009-10, significantly higher rates of overall carriage, principally resulting from the high prevalence of MenY strains, were detected in university students in Nottingham [19, 20]. These observations were supported by subsequent multisite studies undertaken to investigate carriage in UK school and university students [21, 22]. Identification of isolates in the 2008-9 and 2009-10 Nottingham carriage studies relied on PCR amplification of capsule genes and, while some further typing information was generated for a subset of the 2008-9 isolates [19], only limited information was available on the numbers and genetic background of the different MenY-associated clonal complexes carried in 2009-10.

High resolution analyses of genome-wide genetic relationships can determine the prevalence of disease-causing isolates among collections of carriage isolates and detect specific disease-associated loci. The PubMLST.org/neisseria database, which employs the Bacterial Isolate Genome Sequence database (BIGSdb) platform, is a scalable, open-source web-accessible database, to identify, index and extract genetic variation data from whole genome sequence (WGS) data [25]. This approach was utilized to resolve an outbreak of ST-11 disease [26], to investigate the evolution and global spread of the ET-5/ST-32 lineage [27], and to describe MenY disease isolates in Sweden [28]. Additionally, a genealogical analysis of 108 representative meningococcal genomes led to the proposal of a new lineage nomenclature reflecting the increased resolution of WGS typing compared to MLST [29].

Here we investigated the population structure of MenY invasive and carriage isolates in the UK using WGS data generated from 99 carriage isolates obtained from school or university students (typically 16 to 20 years old) between 1997 and 2010 and compared these genomic data with 73 publically available genomes from invasive MenY strains isolated in 2010-11.

METHODS

Isolates

A total of 99 MenY isolates, all obtained from nasopharyngeal carriers in Nottingham (East Midlands), UK, were included in the WGS analysis (Supplementary Table 1). Of these, 77 were isolated from students attending the University of Nottingham in 2009 [20] and were chosen as follows: (i) 20 obtained in September 2009 from first-year students; (ii) 18 obtained in September 2009 from second-year students; (iii) 19 obtained in December 2009 from first-year students; (iv) 20 obtained in December 2009 from second-year students [20]. To provide context, 10 isolates were chosen randomly from a collection of MenY meningococci isolated from sixth-form school students in Nottingham in 1999-2001 [24] and six isolates were chosen from MenY carried isolates obtained from first-year students at the University of Nottingham during 1997-8 [23]. All of these isolates were chosen as known MenY organisms based on PCR or serological typing methods, without prior knowledge of their clonal complex. Six additional MenY carriage isolates were chosen as representative examples of the predominant MenY lineages circulating in a 2008-9 cohort of first-year students at the University of Nottingham [19].

Genomic DNA Extraction, Illumina Sequencing, Assembly and Deposition

Meningococci were grown overnight on heated horse-blood ('chocolate') agar (Oxoid) at 37°C in an atmosphere of air plus 5% CO_2 and genomic DNA extracted using the Wizard Genomic DNA Purification Kit (Promega). Genomic DNA was sequenced as described previously [29]. Short-read sequences were assembled using the VelvetOptimiser de novo short-read assembly program optimization script after which resultant contiguous sequences (contigs) were uploaded to the PubMLST.org/neisseria database. Sequence reads were deposited in the European Nucleotide Archive (Supplementary Table 1). Genome sequences of the 73 MenY disease isolates for the epidemiological year 2010-11 in England, Wales and Northern Ireland (Supplementary Table 2) were accessed via the Meningitis Research Foundation Meningococcus Genome Library database (http://pubmlst.org/perl/bigsdb/bigsdb.pl?db=pubmlst_neisseria_mrfgenomes; accessed September 2015).

Genomic Analyses

The genome assemblies deposited in the database are automatically curated and annotated for all loci currently defined in the database thus identifying alleles with ≥98% sequence identity. Over 2,600 loci were defined at the time of analysis. These have a 'NEIS' prefix and are organized into schemes which enables, for example, the rapid identification of isolate genogroup, clonal complex, and PorA and FetA antigen types. Further analysis was undertaken using the BIGSdb Genome Comparator tool implemented within the database using the *N. meningitidis* cgMLST v1.0 core genome scheme (1,605 loci) [29]. Output distance matrices (Nexus format) were used to generate NeighborNet graphs with SplitsTree4 (v4.13.1).

RESULTS

General Features of Sequenced MenY Carriage Genomes

After de novo assembly, the 100-bp paired Illumina reads from the 99 MenY carriage isolates produced contiguous sequences between 2,018,731 bp to 2,214,168 bp in size, consistent with expectations for meningococcal genomes (Supplementary Table 1). Genome assemblies were automatically annotated in a 'gene-by-gene' approach using the BIGSdb platform and strain designation data extracted (Supplementary Table 1). Isolates from cc23 predominated (58%), followed by cc174 (18%), cc167 (11%) and cc22 (7%). The most prevalent strain designations were Y: P1.5-1,10-1: F4-1: ST-1655 (cc23), Y: P1.5-1,2-2: F5-8: ST-23 (cc23) and Y: P1.21,16: F3-7: ST-1466 (cc174), which collectively accounted for 48% of the carriage isolates (Table 1). Of the 16 carriage strains isolated in 1997-2001, 11 shared identical strain designations with 2008-10 carriage isolates suggesting persistence of these strain designations over this 7-13 year time period (Table 1).

To investigate the occurrence of these carriage strain designations amongst invasive MenY isolates, identical typing information was extracted from the WGS data of 73 invasive UK MenY meningococci isolated during 2010-11 available via the MRF Meningococcus Genome Library database (Supplementary Table 2). Isolates from cc23 predominated (79%), followed by cc174 (10%), cc167 (5%) and cc22 (3%). The most prevalent strain designations among the invasive isolates matched those found in the carriage collection (Table 1). Ten designations were present in both carriage and invasive isolates: these designations accounted for 74% of carriage and 73% of invasive isolates, respectively (Table 1).

WGS Analysis of MenY Isolates Identifies Clusters of Highly Related Isolates

158 To allow higher resolution genealogical analyses, comparison of all 172 MenY genomes was
159 undertaken using the BIGSdb Genome Comparator tool, the principal output of which is a
160 distance matrix based on the number of variable loci within those loci selected for analysis, these
161 differences were then resolved into a network using standard algorithms [30]. Comparison of the
162 genomes using the core *N. meningitidis* cgMLST v1.0 scheme [29] identified 1,157 loci which
163 varied in at least one isolate and resolved isolates into two distinct groups comprising 56 and 116
164 isolates, respectively (Figure 1). Only thirteen loci were found to be identical between these two
165 groups: these included loci encoding ribosomal and hypothetical proteins. Within the two groups,
166 distinct clusters of isolates containing multiple examples of both carriage and invasive isolates
167 were evident. Group 1 comprised three clusters, containing isolates belonging to cc167, cc22 and
168 cc174. Group 2 contained only cc23 meningococci, which formed five distinct clusters of
169 carriage and invasive organisms (Figure 1). Overall 91% (157/172) of isolates localized to one of
170 these eight clusters.
171
172 **Relationships Between Invasive and Carriage MenY Isolates in Identified Clusters**
173 To visualize the relationships among closely related isolates, NeighborNet graphs were generated
174 for each cluster with color-coding of isolate names detailing provenance (Figures 2, 3 and 4).
175 Amongst the 25 isolates in the cc174 cluster (Figure 2*A*), evidence of extensive genetic
176 similarities between carriage isolates was apparent with, for example, only 6 allelic differences
177 distinguishing isolates 22014 and 23214. Highly-related 2009-10 carriage isolates were often
178 isolated from students in the same year group suggestive of intra-year group transmission. This
179 was also apparent in other clusters of isolates, such as cc22 (*e.g.* isolates 22667 and 21258; 8
180 allelic differences) (Figure 2*C*). Conversely, the cc22 cluster revealed highly related

meningococci isolated from individuals in different year groups suggestive of inter-year group transmission (*e.g.* isolates 23009 and 21513; 3 allelic differences) (Figure 2*C*).

The cc167 cluster (Figure 2*B*) and cc23 cluster 4 (Figure 3*D*) each resolved into distinct sub-clusters. The ST-767 cc167 sub-cluster (Figure 2*B*) contained carriage isolates from 2001, 2008 and 2009 and a 2011 invasive isolate (M11 240071), suggestive of a long-lived clone capable of causing disease. Only 27 allelic differences distinguished M11 240071 from N117.1; 62 differences distinguished the former from NO01020675 – a carriage isolate obtained in 2001 (Figure 2*B*).

In some cases, clusters containing isolates with identical designations could also be resolved into distinct sub-clusters on the basis of WGS analysis. Notably, cc23 cluster 1 could be resolved into two sub-clusters (Figure 3*A*). The first contained a carriage isolate from 2000 (NO0010442), five 2008-10 carriage isolates and two 2010-11 invasive isolates. Since NO0010442 is only 34 allelic differences apart from 21251 (a 2009 carriage isolate) and 42 from the invasive isolate M10 240732, this sub-cluster represents another persistent clone, capable of causing disease.

WGS Analysis Resolves cc23 Cluster 5 into Invasive- and Carriage-Associated Sub-clusters

The cc23 cluster 5 contained the largest number of MenY isolates analyzed. Despite predominantly sharing a common strain designation, WGS-based analysis resolved meningococci in this cluster into two sub-clusters (Figure 4): sub-cluster 1 with 18 carriage isolates and three invasive isolates; and sub-cluster 2 with three carriage and 27 invasive meningococci. A total of 997 loci were identical among all cc23 cluster 5 isolates. These sub-clusters were defined by specific allelic differences in five core genes, encoding glycerate kinase

(*glxK*), valine-pyruvate transaminase (*avtA*), superoxide dismutase (*sodB*) and two hypothetical proteins (Table 2).

DISCUSSION

Nucleotide sequence-based methods involving small numbers of genes have been invaluable in characterizing the population structure and antigenic repertoires of meningococci [31]. The advent of WGS has greatly enhanced resolution and has begun to provide improved insights into the genetic relationships among bacterial isolates [32]. Since carriage is directly relevant to the epidemiology of IMD, we undertook to resolve the genealogical relationships between carriage and invasive isolates. We focused on MenY lineages due to recent observations of fluctuations in MenY disease and carriage levels in the UK. Although meningococci of this serogroup have been less prevalent globally as causes of disease compared to serogroups A, B and C [33], the proportion of IMD attributable to MenY organisms, predominately those belonging to cc23, increased markedly, a trend first recognized in the mid-1990s in the USA [14, 34], and more recently in other countries including the UK [17, 18] and Sweden [15, 35]. The higher MenY IMD case load in the UK was concomitant with a significant increase in MenY carriage, as first detected in studies of nasopharyngeal carriage in students at the University of Nottingham undertaken from 2008 to 2010 [19, 20].

The automated extraction of strain designation information from WGS data demonstrated the similarity of MenY isolates from carriage and invasive disease. This similarity was confirmed by the enhanced discrimination afforded by core genome analysis of the WGS data which resolved most of the isolates into one of eight defined clusters. While most isolates in a particular cluster shared the same strain designation (*i.e.* ST, PorA and FetA types), each cluster

contained variants, demonstrating the enhanced discrimination afforded by WGS. A key finding was that every cluster contained both invasive and disease isolates, indicating that all MenY lineages have the ability to cause disease.

Bacterial populations are often viewed as unstable collections of rapidly evolving clones with frequent extinctions or replacement of older clones. Temporal shifts are potentially important components of IMD epidemiology. Thus, analysis of IMD cases indicated replacement of an 'early' cc23 MenY lineage in the USA by an antigenically and genetically distinct 'late' strain type [36, 37]. A parallel shift in carriage of these clones was assumed but not investigated. A significant finding from the present study was that six out of the eight MenY clusters contained historic carriage isolates *(i.e.* from 1997-2001). The stability of this association appears to be strong as it was detected with only sixteen historic genome sequences. Thus, these six MenY clusters are long-lived and have been present within the UK for a 7-13 year time period. The uneven distribution (*e.g.* cc167 and cc23 cluster 1) and apparent outlier position (*e.g.* cc174 and cc22) of historic isolates in some clusters is suggestive of within-cluster evolution over time. The exception to this generalization was cc23 cluster 5, which was the largest cluster and yet contained no historic strain types, potentially suggesting the arrival of a non-UK associated epidemic lineage or major alterations in the genetic structure of a long-lasting UK MenY clone. The presence of long-lasting clones indicates that the genetic structure of meningococcal clones is stable and that extinctions of clones are rare events. The presence of a long-lived host-adapted commensal population has importance as introduction of the MenACWY vaccine into the main carrier population has the potential to radically-perturb a long-lasting association with unknown consequences.

There was evidence for antigenic shifts among members of the cc23 isolates, which distributed into five clusters distinguished by PorA type but not ST type. Four of the clusters differed in sequence for VR2 of PorA, a major target of bactericidal antibodies while the two clusters with identical PorA VR2 sequences had different PorA VR1 sequences, a variable target of bactericidal antibodies. The differences in the VR2 amino acid sequence are the increase in number of a three amino acid motif (NKQ) from one copy in P1.10-1, to two in P1.10-4 and three in P1.10-10: a rapid and minor change in protein structure. This was not a feature of all surface antigens, as there was limited variation in the FetA VR with four cc23 clusters having the same FetA variant. Further analysis of WGS data may indicate other antigenic variants or allelic variants of other genes that correlate with this segregation of cc23 isolates; nevertheless the PorA distribution suggests that minor differences in antigenicity may contribute to changes in population structure.

Geographic distribution of clones and potential sources of new clones was apparent from comparisons among WGS studies in different countries. Comparison of invasive cc23 isolates from Sweden, UK and USA identified three principal cc23 sub-lineages (designated 23.1, 23.2 and 23.3) with overlapping, but differentially prevalent repertoires in each country [28]. For example, the Swedish 'strain-type YI' which was largely responsible for the increase in Swedish MenY disease [16, 35], formed a cluster within the 23.1 sub-lineage, but very rarely caused disease in the UK [28]. Using the overlap in MenY WGS data analyzed, *i.e.* UK invasive cc23 strains isolated in 2010-11 examined previously [28], we further resolved the 23.1 sub-lineage into four sub-clusters (cc23 clusters 2-5) and found that cc23 cluster 1 corresponds to lineage 23.2. Cluster 5, which was responsible for most cases of UK IMD, was rarely observed in Swedish IMD isolates.

The resolution of cc23 cluster 5 into distinct carriage- and disease-associated sub-clusters, (1 and 2, respectively) was surprising as this cluster contained the highest number of MenY disease and carriage isolates. A confounding factor is that the sub-cluster 1 carriage isolates were all isolated in one geographical location, and hence may have a high level of one specific (highly transmissible) clone. Two of these isolates (20601 and 21619) were, however, isolated in the first week of term in September from first-year students who are presumed to have been colonized prior to arrival at the University. An alternative hypothesis is that the ability of sub-cluster 1 strains to cause disease is associated with rapid within host evolution into a sub-cluster 2 phenotype; however, sub-clusters were defined by differences in loci encoding proteins with hypothetical or core enzymatic functions not loci explicitly linked to adaptation to a systemic niche (*e.g.* survival in blood). A further possibility is that sub-cluster 1 has recently evolved from sub-cluster 2 into a highly transmissible carriage strain with a consequent reduction in virulence. It is unlikely that any of the five differences in sub-cluster 1 and 2 defining core loci are directly responsible for differences in virulence given the predicted functions of the five proteins, and that the allelic differences lead to, at most, only single amino acid changes which are unlikely to be functionally significant. Instead we hypothesize that differences outside the core alleles examined in this study, but which will co-segregate with the five core loci differences, are more likely to be responsible. A high-quality assembled cc23 genome is required to detect the effects on virulence mediated by non-core genes, and in order to determine how the transition between these sub-clusters has occurred.

In summary, high resolution genealogical relationships between MenY isolates highlighted the high degree of genetic similarity between carriage and invasive isolates and evidenced long-term stability of MenY clones. The detection and resolution of a highly prevalent

295 UK clone (Y: P1.5-1,10-1: F4-1: ST-1655 cc23) into invasive- and carriage-associated sub-clusters
296 exemplifies the improved precision of whole genome analysis for separating apparently identical
297 isolates.
298
299 **POTENTIAL CONFLICTS OF INTEREST**
300 All authors report no potential conflicts.
301
302 **FUNDING**
303 This work was supported by the Wellcome Trust [grant 087622/Z/08/2 to M.C.J.M].
304
305 **CORRESPONDING AUTHOR CONTACT INFORMATION**
306 Neil Oldfield, Centre for Biomolecular Sciences, University of Nottingham, Nottingham NG7
307 2RD, UK. E-mail: neil.oldfield@nottingham.ac.uk. Tel: (+44) (+44) 115748 6122. Fax: (+44)
308 115846 8002.
309
310 **ACKNOWLEDGMENT**
311 This publication made use of the Meningitis Research Foundation Meningococcus Genome
312 Library (http://www.meningitis.org/research/genome) developed by Public Health England, the
313 Wellcome Trust Sanger Institute and the University of Oxford as a collaboration and funded by
314 Meningitis Research Foundation.

REFERENCES

1. Christensen H, May M, Bowen L, Hickman M, Trotter CL. Meningococcal carriage by age: a systematic review and meta-analysis. Lancet Infect Dis **2010**; 10:853-61.

2. Soriano-Gabarró M, Wolter J, Hogea C, Vyse A. Carriage of *Neisseria meningitidis* in Europe: a review of studies undertaken in the region. Expert Rev Anti Infect Ther **2011**; 9:761-74.

3. Stephens DS, Greenwood B, Brandtzaeg P. Epidemic meningitis, meningococcaemia, and *Neisseria meningitidis*. Lancet **2007**; 369:2196-210.

4. Lewis LA, Ram S. Meningococcal disease and the complement system. Virulence **2014**; 5:98-126.

5. Harrison OB, Claus H, Jiang Y, et al. Description and nomenclature of *Neisseria meningitidis* capsule locus. Emerg Infect Dis **2013**; 19:566-73.

6. Harrison LH, Trotter CL, Ramsay ME. Global epidemiology of meningococcal disease. Vaccine **2009**; 27:B51-B63.

7. Jolley K, Brehony C, Maiden M. Molecular typing of meningococci: recommendations for target choice and nomenclature. FEMS Microbiol Rev **2007**; 31:89-96.

8. Maiden MC, Bygraves JA, Feil E, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A **1998**; 95:3140-5.

9. Yazdankhah SP, Kriz P, Tzanakaki G, et al. Distribution of serogroups and genotypes among disease-associated and carried isolates of *Neisseria meningitidis* from the Czech Republic, Greece, and Norway. J Clin Microbiol **2004**; 42:5146-53.

337 10. Caugant DA, Maiden MCJ. Meningococcal carriage and disease - population biology and
338 evolution. Vaccine **2009**; 27:B64-B70.
339 11. Russell JE, Jolley KA, Feavers IM, Maiden MC, Suker J. PorA variable regions of *Neisseria*
340 *meningitidis*. Emerg Infect Dis **2004**; 10:674-8.
341 12. Urwin R, Russell JE, Thompson EA, Holmes EC, Feavers IM, Maiden MC. Distribution of
342 surface protein variants among hyperinvasive meningococci: implications for vaccine design.
343 Infect Immun **2004**; 72:5955-62.
344 13. Trotter C, Chandra M, Cano R, et al. A surveillance network for meningococcal disease in
345 Europe. FEMS Microbiol Rev **2007**; 31:27-36.
346 14. Cohn AC, MacNeil JR, Harrison LH, et al. Changes in *Neisseria meningitidis* disease
347 epidemiology in the United States, 1998-2007: implications for prevention of meningococcal
348 disease. Clin Infect Dis **2010**; 50:184-91.
349 15. Bröker M, Bukovski S, Culic D, et al. Meningococcal serogroup Y emergence in Europe.
350 Hum Vaccin Immunother **2014**; 10:1725-8.
351 16. Hedberg ST, Toros B, Fredlund H, Olcen P, Molling P. Genetic characterisation of the
352 emerging invasive *Neisseria meningitidis* serogroup Y in Sweden, 2000 to 2010. Euro Surveill
353 **2011**; 16:19885.
354 17. Ladhani SN, Lucidarme J, Newbold LS, et al. Invasive meningococcal capsular group Y
355 disease, England and Wales, 2007-2009. Emerg Infect Dis **2012**; 18:63-70.
356 18. Ladhani SN, Flood JS, Ramsay ME, et al. Invasive meningococcal disease in England and
357 Wales: Implications for the introduction of new vaccines. Vaccine **2012**; 30:3710-6.

19. Bidmos FA, Neal KR, Oldfield NJ, Turner DPJ, Ala'Aldeen DAA, Bayliss CD. Rapid clonal expansion, persistence and clonal replacement of meningococcal carriage isolates in a 2008 university student cohort. J Clin Microbiol **2011**; 49:506-12.

20. Ala'Aldeen DAA, Oldfield NJ, Bidmos FA, et al. Carriage of meningococci by university students, United Kingdom. Emerg Infect Dis **2011**; 19:1762-3.

21. Jeppesen CA, Snape MD, Robinson H, et al. Meningococcal carriage in adolescents in the United Kingdom to inform timing of an adolescent vaccination strategy. J Infect **2015**; 71:43-52.

22. Read RC, Baxter D, Chadwick DR, et al. Effect of a quadrivalent meningococcal ACWY glycoconjugate or a serogroup B meningococcal vaccine on meningococcal carriage: an observer-blind, phase 3 randomised clinical trial. Lancet **2014**; 384:2123-31.

23. Ala'Aldeen DA, Neal KR, Ait-Tahar K, et al. Dynamics of meningococcal long-term carriage among university students and their implications for mass vaccination. J Clin Microbiol **2000**; 38:2311-6.

24. Maiden M, Ibarz-Pavón AB, Urwin R, et al. Impact of meningococcal serogroup C conjugate vaccines on carriage and herd immunity. J Infect Dis **2008**; 197:737-43.

25. Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics **2010**; 11:595.

26. Jolley KA, Hill DM, Bratcher HB, et al. Resolution of a meningococcal disease outbreak from whole-genome sequence data with rapid Web-based analysis methods. J Clin Microbiol **2012**; 50:3046-53.

27. Harrison OB, Bray JE, Maiden MCJ, Caugant DA. Genomic analysis of the evolution and global spread of hyper-invasive meningococcal lineage 5. EBioMedicine **2015**; 2:234-43.

28. Törös B, Hedberg ST, Unemo M, et al. Genome-based characterization of emergent invasive *Neisseria meningitidis* serogroup Y in Sweden, 1995 to 2012. J Clin Microbiol **2015**; 53:2154-62.

29. Bratcher HB, Corton C, Jolley KA, Parkhill J, Maiden MC. A gene-by-gene population genomics platform: *de novo* assembly, annotation and genealogical analysis of 108 representative *Neisseria meningitidis* genomes. BMC Genomics **2014**; 15:1138.

30. Jolley KA, Maiden MC. Automated extraction of typing information for bacterial pathogens from whole genome sequence data: *Neisseria meningitidis* as an exemplar. Euro Surveill **2013**; 18:20379.

31. Brehony C, Jolley KA, Maiden MC. Multilocus sequence typing for global surveillance of meningococcal disease. FEMS Microbiol Rev **2007**; 31:15-26.

32. Jolley KA, Maiden MC. Using multilocus sequence typing to study bacterial variation: prospects in the genomic era. Future Microbiol **2014**; 9:623-30.

33. Stephens DS. Biology and pathogenesis of the evolutionarily successful, obligate human bacterium *Neisseria meningitidis*. Vaccine **2009**; 27:B71-7.

34. Baccarini C, Ternouth A, Wieffer H, Vyse A. The changing epidemiology of meningococcal disease in North America 1945–2010. Hum Vaccin Immunother **2013**; 9:162-71.

35. Törös B, Thulin Hedberg S, Jacobsson S, Fredlund H, Olcén P, Mölling P. Surveillance of invasive *Neisseria meningitidis* with a serogroup Y update, Sweden 2010 to 2012. Euro Surveill **2014**; 19:20940.

36. Harrison LH, Jolley KA, Shutt KA, et al. Antigenic shift and increased incidence of meningococcal disease. J Infect Dis **2006**; 193:1266-74.

37. Krauland MG, Dunning Hotopp JC, Riley DR, et al. Whole genome sequencing to investigate the emergence of clonal complex 23 *Neisseria meningitidis* serogroup Y disease in the United States. PLoS One **2012**; 7:e35699.

FIGURE LEGENDS

Figure 1. NeighborNet graph comparison of 172 UK MenY genome sequences analyzed using the BIGSdb Genome Comparator utilizing the *N. meningitidis* cgMLST v1.0 scheme. 91% of isolates analyzed localized to one of eight clusters. Strain designation(s) represent the most frequently occurring designation(s) in each cluster. Unlabeled nodes represent unassigned invasive (n=9) and carriage (n=6) isolates. Scale bar = number of allelic differences.

Figure 2. NeighborNet graphs comparing isolates in the (*A*) cc174, (*B*) cc167 and (*C*) cc22 clusters as defined in Figure 1. Sequences were analyzed using BIGSdb Genome Comparator tool utilizing the *N. meningitidis* cgMLST v1.0 scheme. Isolate names are color-coded as follows: 1997-2001 carriage isolates in fuchsia; 2008-9 carriage isolates in black; 2009-10 carriage isolates from first year students in green; 2009-10 carriage isolates from second year students in blue and invasive isolates from 2010-11 in red. Scale bar = number of allelic differences.

Figure 3. NeighborNet graphs comparing isolates in the cc23 cluster nos. 1, 2, 3 and 4 (panels *A-D*, respectively) as defined in Figure 1. Sequences were analyzed using BIGSdb Genome Comparator tool utilizing the *N. meningitidis* cgMLST v1.0 scheme. Isolate names are color-

424 coded according to the scheme described in the Figure 2 legend. Scale bar = number of allelic
425 differences.
426
427 **Figure 4.** NeighborNet graph comparison of isolates in the cc23 cluster 5 defined in Figure 1.
428 Sequences were analyzed using BIGSdb Genome Comparator tool utilizing the *N. meningitidis*
429 cgMLST v1.0 scheme. Isolate names are color-coded according to the scheme described in
430 Figure 2 legend. Scale bar = number of allelic differences.

Table 1. Frequency of Strain Designations in the MenY Carriage and Invasive Collections

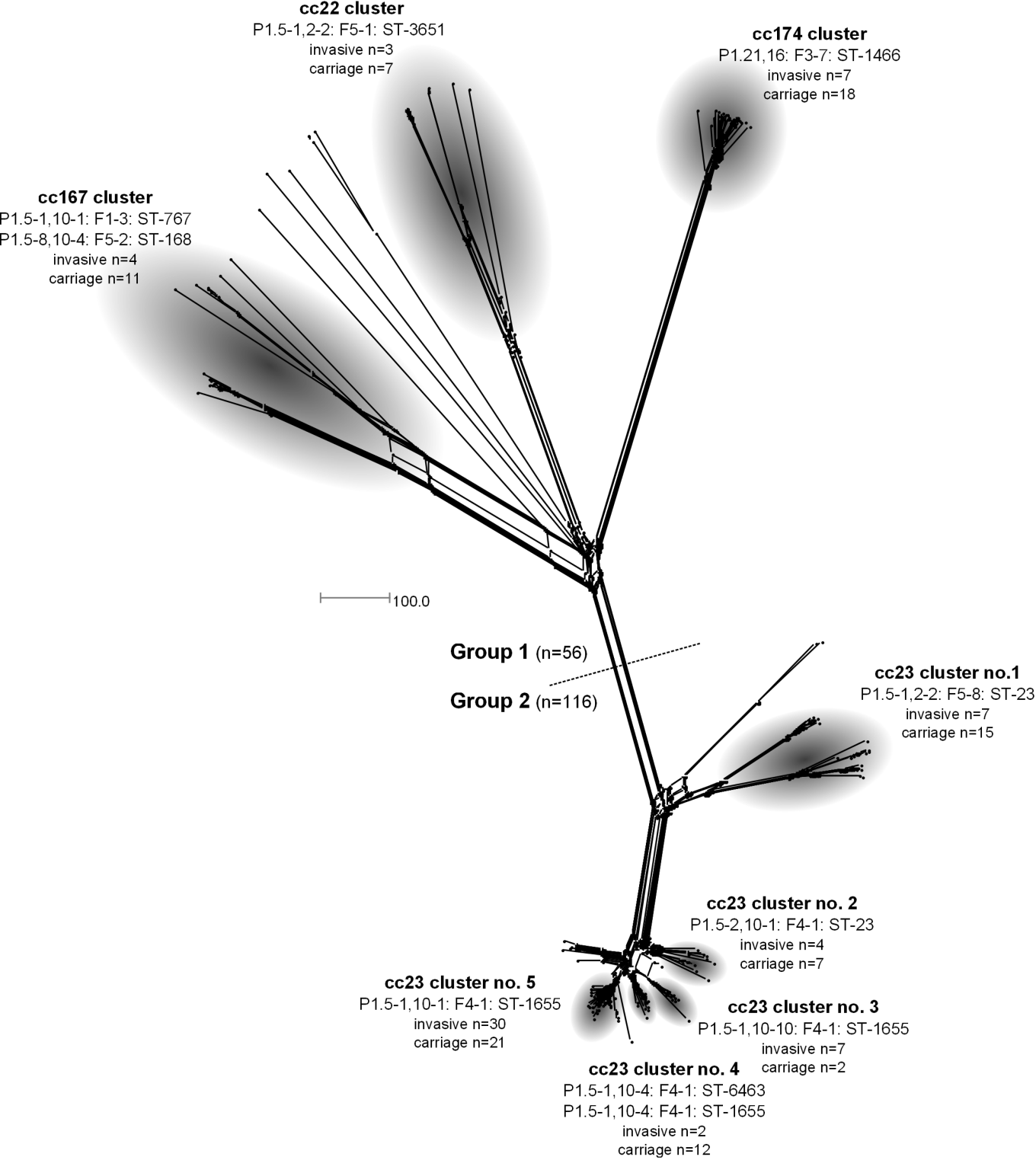
Strain designation	Carriage group 1997-2001 (*n*=16)	Carriage group 2008-10 (*n*=83)	Total carriage (*n*=99)	Invasive 2010-11 (*n*=73)	Total carriage and invasive (*n*=172)
Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	1	20	21	26	47
Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	5	10	15	5	20
Y: P1.21,16: F3-7: ST-1466 (cc174)	0	12	12	5	17
Y: P1.5-2,10-1: F4-1: ST-23 (cc23)	1	4	5	4	9
Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)	2	4	6	2	8
Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)	0	6	6	2	8
Y: P1.5-1,2-2: F5-1: ST-3651 (cc22)	0	4	4	2	6
Y: P1.5-1,10-10: F4-1: ST-1655 (cc23)	0	2	2	4	6
Y: P1.5-1,10-1: F1-3: ST-767 (cc167)	2	3	5	0	5
Y: P1.5-1,10-4: F4-1: ST-23 (cc23)	0	0	0	4	4
Y: P1.5-8,10-4: F5-2: ST-168 (cc167)	0	1	1	2	3
Y: P1.5-1,10-22: F5-1: ST-114 (cc22)	0	2	2	0	2
Y: P1.5-1,10-46: F3-9: ST-103 (cc103)	0	2	2	0	2
Y: P1.5-1,10-62: F1-3: ST-767 (cc167)	2	0	2	0	2
Y: P1.22,9: F3-7: ST-1466 (cc174)	0	1	1	1	2
Other[a]	3	12	15	16	31

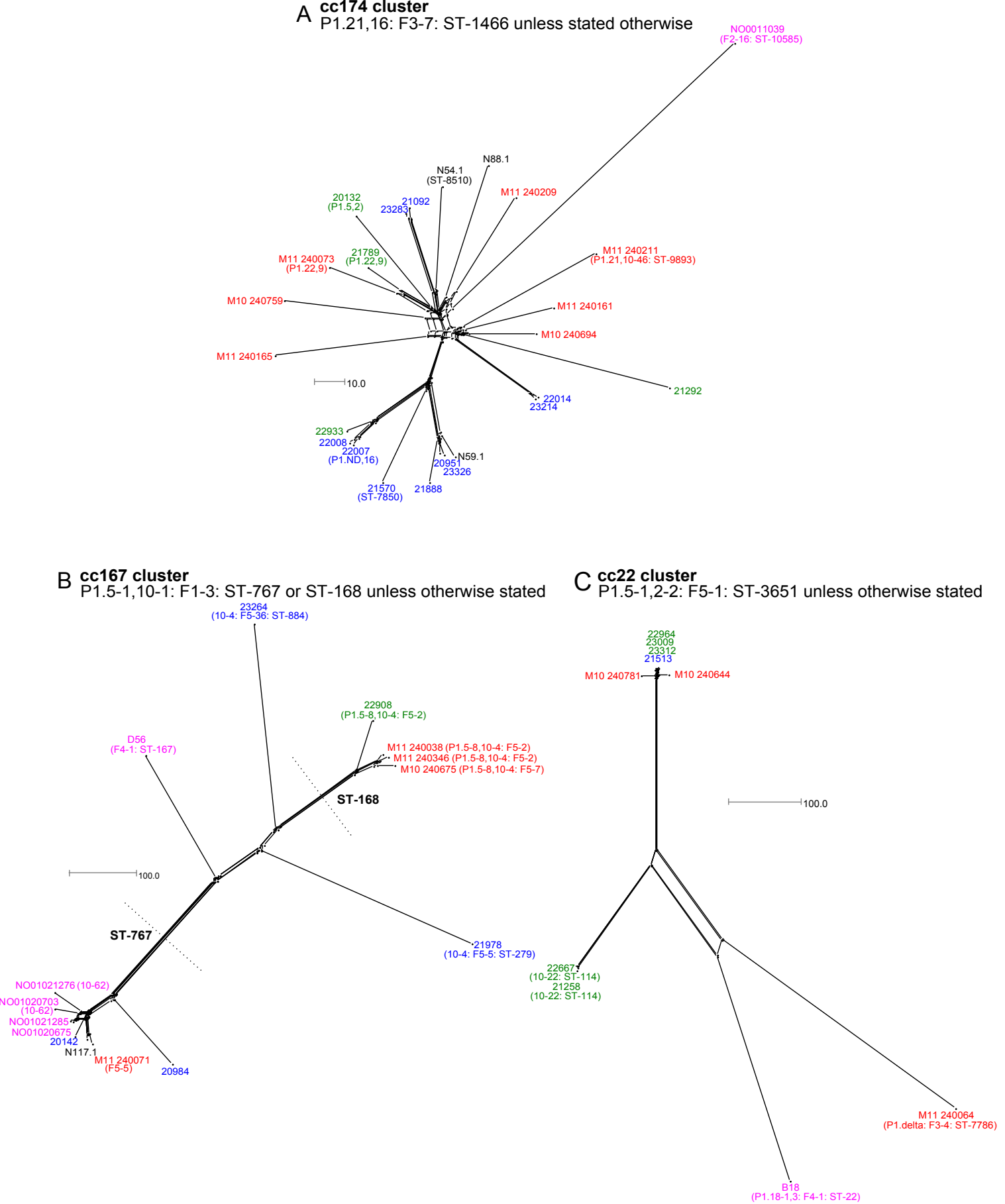
[a] Includes all strain designations occurring only once.

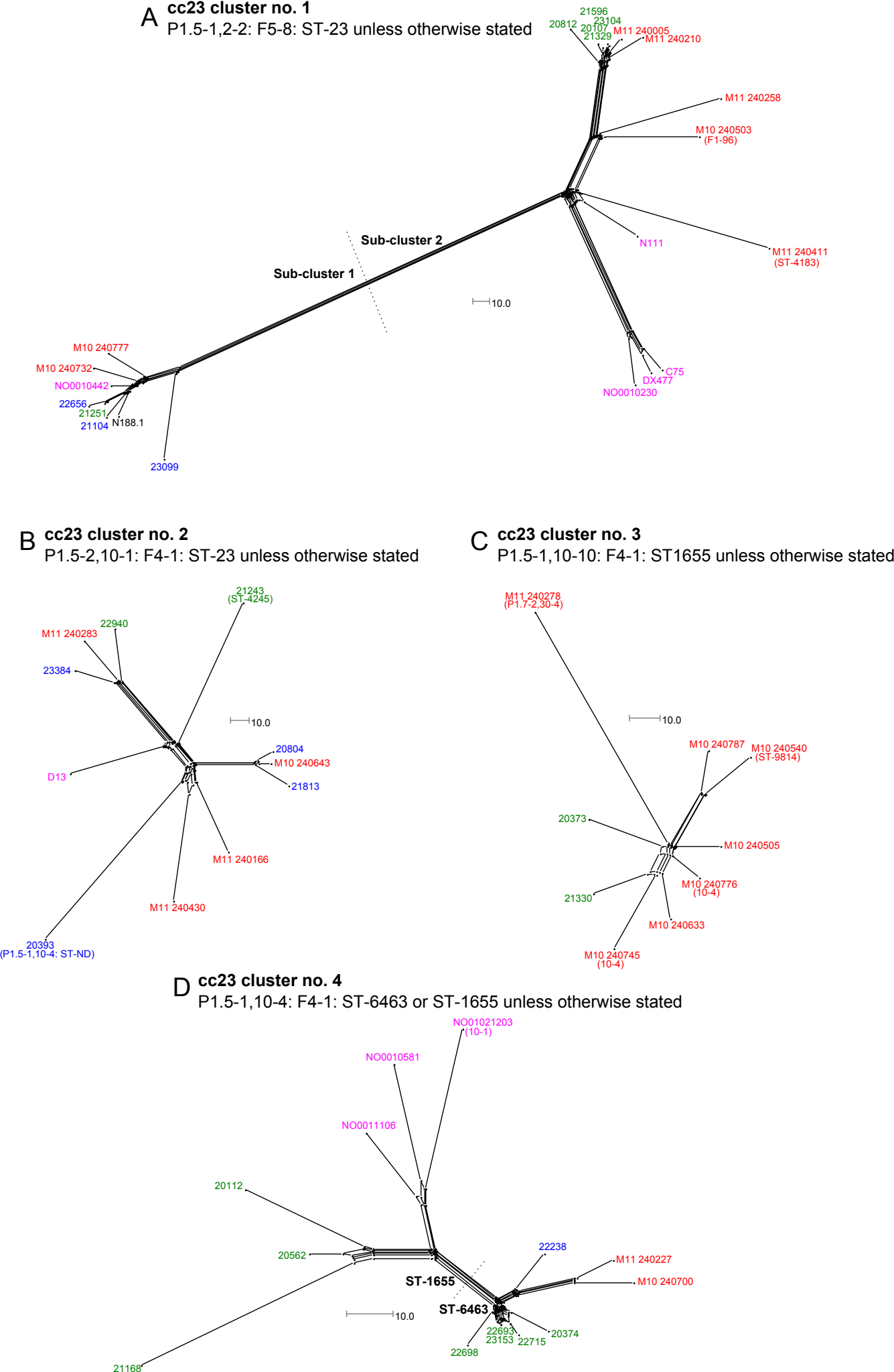
435 **Table 2. Loci with Allelic Differences between the Two Sub-clusters of cc23 Cluster 5**

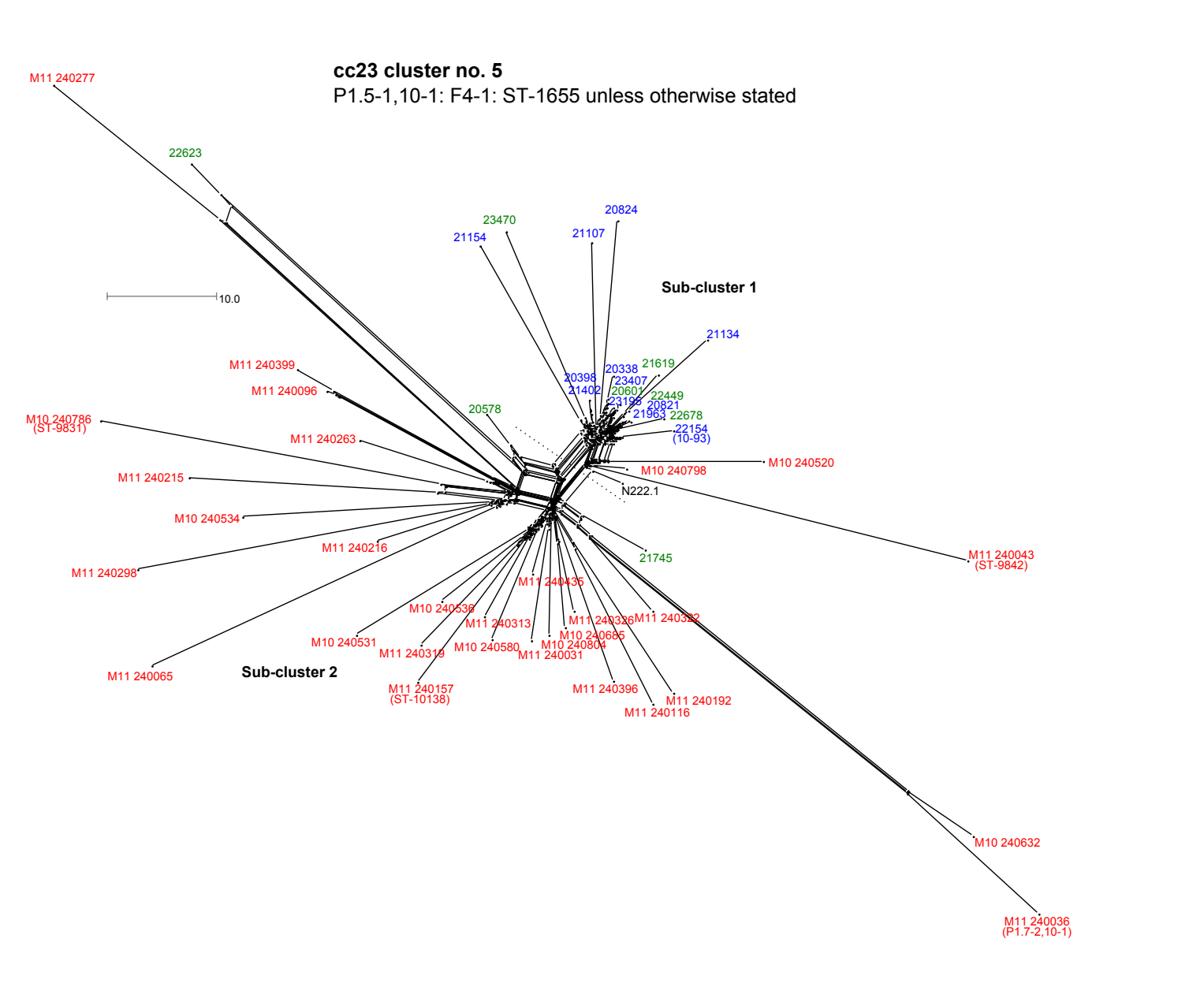
BIGSdb *Neisseria* locus identifier	Predicted protein/function (gene)	Allele number (%)		% nucleotide identity	Amino acid differences
		Sub-cluster 1	Sub-cluster 2		
NEIS0395	Valine-pyruvate transaminase (*avtA*)	112 (100)	113 (96.7)	99.9	1
NEIS0825	Superoxide dismutase (*sodB*)	155 (100)	22 (96.7)	99.8	1
NEIS0929	Hypothetical protein	42 (100)	3 (100)	99.6	0
NEIS1199	Glycerate kinase (*glxK*)	47 (100)	24 (100)	99.9	1
NEIS1568	Hypothetical protein	67 (100)	68 (96.7)	99.9	1

436


cc22 cluster
P1.5-1,2-2: F5-1: ST-3651
invasive n=3
carriage n=7
cc174 cluster
P1.21,16: F3-7: ST-1466
invasive n=7
carriage n=18
cc167 cluster
P1.5-1,10-1: F1-3: ST-767
P1.5-8,10-4: F5-2: ST-168
invasive n=4
carriage n=11
100.0
Group 1 (n=56)
Group 2 (n=116)
cc23 cluster no.1
P1.5-1,2-2: F5-8: ST-23
invasive n=7
carriage n=15
cc23 cluster no. 2
P1.5-2,10-1: F4-1: ST-23
invasive n=4
carriage n=7
cc23 cluster no. 5
P1.5-1,10-1: F4-1: ST-1655
invasive n=30
carriage n=21
cc23 cluster no. 3
P1.5-1,10-10: F4-1: ST-1655
invasive n=7
carriage n=2
cc23 cluster no. 4
P1.5-1,10-4: F4-1: ST-6463
P1.5-1,10-4: F4-1: ST-1655
invasive n=2
carriage n=12


A cc174 cluster
P1.21,16: F3-7: ST-1466 unless stated otherwise
NO0011039
(F2-16: ST-10585)
N88.1
N54.1
(ST-8510)
M11 240209
20132
(P1.5,2)
21092
23283
M11 240211
(P1.21,10-46: ST-9893)
M11 240073
(P1.22,9)
21789
(P1.22,9)
M10 240759
M11 240161
M10 240694
M11 240165
10.0
21292
22014
23214
22933
22008
22007
(P1.ND,16)
N59.1
20951
23326
21570
(ST-7850)
21888
B cc167 cluster
P1.5-1,10-1: F1-3: ST-767 or ST-168 unless otherwise stated
23264
(10-4: F5-36: ST-884)
22908
(P1.5-8,10-4: F5-2)
D56
(F4-1: ST-167)
M11 240038 (P1.5-8,10-4: F5-2)
M11 240346 (P1.5-8,10-4: F5-2)
M10 240675 (P1.5-8,10-4: F5-7)
ST-168
100.0
ST-767
21978
(10-4: F5-5: ST-279)
NO01021276 (10-62)
NO01020703
(10-62)
NO01021285
NO01020675
20142
N117.1
M11 240071
(F5-5)
20984
C cc22 cluster
P1.5-1,2-2: F5-1: ST-3651 unless otherwise stated
22964
23009
23312
21513
M10 240781
M10 240644
100.0
22667
(10-22: ST-114)
21258
(10-22: ST-114)
M11 240064
(P1.delta: F3-4: ST-7786)
B18
(P1.18-1,3: F4-1: ST-22)


A
cc23 cluster no. 1
P1.5-1,2-2: F5-8: ST-23 unless otherwise stated
21596
23104
20812
20107
21329
M11 240005
M11 240210
M11 240258
M10 240503
(F1-96)
Sub-cluster 2
Sub-cluster 1
N111
M11 240411
(ST-4183)
10.0
M10 240777
M10 240732
NO0010442
22656
21251
21104
N188.1
23099
C75
DX477
NO0010230
B
cc23 cluster no. 2
P1.5-2,10-1: F4-1: ST-23 unless otherwise stated
21243
(ST-4245)
22940
M11 240283
23384
10.0
20804
M10 240643
D13
21813
M11 240166
M11 240430
20393
(P1.5-1,10-4: ST-ND)
C
cc23 cluster no. 3
P1.5-1,10-10: F4-1: ST1655 unless otherwise stated
M11 240278
(P1.7-2,30-4)
10.0
M10 240787
M10 240540
(ST-9814)
20373
M10 240505
M10 240776
(10-4)
21330
M10 240633
M10 240745
(10-4)
D
cc23 cluster no. 4
P1.5-1,10-4: F4-1: ST-6463 or ST-1655 unless otherwise stated
NO01021203
(10-1)
NO0010581
NO0011106
20112
20562
22238
M11 240227
M10 240700
ST-1655
ST-6463
10.0
22693
23153
22715
20374
22698
21168


cc23 cluster no. 5
P1.5-1,10-1: F4-1: ST-1655 unless otherwise stated
M11 240277
22623
23470
21154
20824
21107
Sub-cluster 1
10.0
21134
M11 240399
M11 240096
20398
21402
20338
23407
20601
23189
21619
22449
20821
21963
22678
22154
(10-93)
20578
M10 240786
(ST-9831)
M11 240263
M10 240520
M10 240798
M11 240215
N222.1
M10 240534
M11 240216
M11 240043
(ST-9842)
M11 240298
21745
M11 240435
M10 240536
M11 240313
M11 240326
M11 240322
M10 240685
M10 240531
M10 240804
M10 240580
M11 240319
M11 240031
M11 240065
Sub-cluster 2
M11 240157
(ST-10138)
M11 240396
M11 240192
M11 240116
M10 240632
M11 240036
(P1.7-2,10-1)

SUPPLEMENTARY TABLE 1. List of MenY carriage isolates used.

Isolate	Year of isolation	Month of isolation	Carrier[a]	Strain designation[b]	Number of contigs	Total length (bp)	N50 (bp)[c]	Mean contig size (bp)	ENA accession
B18	1997	October	1st year UoN	Y: P1.18-1,3: F4-1: ST-22 (cc22)	227	2214168	37942	9755	ERR351542
C75	1997	October	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	203	2117447	29278	10431	ERR351543
D13	1997	October	1st year UoN	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)	172	2106353	33068	12247	ERR351544
D56	1997	October	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-167 (cc167)	195	2121467	38802	10880	ERR351545
N111	1997	December	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	173	2113470	36801	12217	ERR351546
DX477	1997	November	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	168	2121286	36047	12627	ERR351547
NO0010230	2000	November	School	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	181	2116713	31598	11695	ERR351548
NO0010442	2000	November	School	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	190	2112482	30837	11119	ERR351549
NO0010581	2000	December	School	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)	167	2100396	37401	12578	ERR351550
NO0011039	2000	December	School	Y: P1.21,16: F2-16: ST-10585 (cc174)	128	2141415	48587	16730	ERR351551
NO0011106	2000	December	School	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)	173	2101749	37314	12149	ERR351552
NO01020675	2001	November	School	Y: P1.5-1,10-1: F1-3: ST-767 (cc167)	191	2137661	39769	11192	ERR351553
NO01020703	2001	November	School	Y: P1.5-1,10-62: F1-3: ST-767 (cc167)	193	2135549	35671	11066	ERR351554
NO01021203	2001	November	School	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	164	2105490	41413	12839	ERR351555
NO01021276	2001	November	School	Y: P1.5-1,10-62: F1-3: ST-767 (cc167)	186	2135528	36461	11482	ERR351556
NO01021285	2001	November	School	Y: P1.5-1,10-1: F1-3: ST-767 (cc167)	186	2139032	36184	11501	ERR351557
N54.1	2008	November	1st year UoN	Y: P1.21,16: F3-7: ST-8510 (cc174)	172	2129298	38858	12380	ERR144492
N59.1	2008	November	1st year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	153	2127669	45623	13907	ERR144494
N88.1	2008	November	1st year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	173	2130971	44359	12318	ERR144495
N117.1	2008	November	1st year UoN	Y: P1.5-1,10-1: F1-3: ST-767 (cc167)	221	2120273	28603	9594	ERR144498
N222.1	2008	November	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	386	2109526	14018	5466	ERR144504
N188.1	2008	November	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	242	2106309	24258	8704	ERR144511
20107	2009	September	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	179	2110265	32493	11790	ERR351558
20112	2009	September	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)	193	2103990	28482	10902	ERR351559
20132	2009	September	1st year UoN	Y: P1.5,2: F3-7: ST-1466 (cc174)	141	2135890	48590	15149	ERR351560
20142	2009	September	1st year UoN	Y: P1.5-1,10-1: F1-3: ST-767 (cc167)	215	2127862	34107	9898	ERR351561
20562	2009	September	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)	171	2100554	28487	12284	ERR351562
20578	2009	September	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	167	2097698	37308	12562	ERR351563
20588	2009	September	1st year UoN	Y: P1.18-1,34: F1-5: ST-6058 (cc41/44)	181	2125539	45273	11744	ERR351564
20601	2009	September	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	168	2099544	31372	12498	ERR351565
21168	2009	September	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)	211	2116387	31273	10031	ERR351566

21243	2009	September	1st year UoN	Y: P1.5-2,10-1: F4-1: ST-4245 (cc23)	170	2089142	28483	12290	ERR351567
21251	2009	September	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	197	2110345	28032	10713	ERR351568
21258	2009	September	1st year UoN	Y: P1.5-1,10-22: F5-1: ST-114 (cc22)	289	2202962	27749	7623	ERR351569
21265	2009	September	1st Year UoN	Y: P1.18-1,3: F1-5: ST-1768 ()	189	2165563	31318	11459	ERR351570
21292	2009	September	1st year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	147	2136683	43536	14536	ERR351571
21329	2009	September	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	169	2109115	32493	12480	ERR351572
21330	2009	September	1st year UoN	Y: P1.5-1,10-10: F4-1: ST-1655 (cc23)	162	2099958	37903	12963	ERR351573
21596	2009	September	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	175	2108527	31560	12049	ERR351574
21619	2009	September	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	166	2082731	31398	12547	ERR351575
21745	2009	September	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	196	2099659	30859	10713	ERR351576
21789	2009	September	1st year UoN	Y: P1.22,9: F3-7: ST-1466 (cc174)	143	2134285	48900	14926	ERR351577
20663	2009	September	2nd year UoN	Y: P1.5-1,10-46: F3-9: ST-103 (cc103)	178	2210571	41370	12419	ERR351578
20951	2009	September	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	137	2133694	43053	15575	ERR351580
20984	2009	September	2nd year UoN	Y: P1.5-1,10-1: F1-3: ST-767 (cc167)	223	2135313	34056	9576	ERR351581
21089	2009	September	2nd year UoN	Y: P1.5-1,10-7: F3-9: ST-4963 (cc103)	217	2199601	33028	10137	ERR351583
21092	2009	September	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	132	2133973	63093	16167	ERR351584
21104	2009	September	2nd year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	190	2111925	32293	11116	ERR351585
21107	2009	September	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	165	2097444	31424	12712	ERR351586
21134	2009	September	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	172	2094410	37671	12177	ERR351587
21154	2009	September	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	164	2098074	37322	12794	ERR351588
21402	2009	September	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	174	2095868	31368	12046	ERR351589
21813	2009	September	2nd year UoN	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)	181	2098661	32473	11595	ERR351590
21839	2009	September	2nd year UoN	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)	167	2100788	32658	12580	ERR351591
21888	2009	September	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	176	2129844	38468	12102	ERR351592
22007	2009	September	2nd year UoN	Y: P1.ND,16: F3-7: ST-1466 (cc174)	135	2132984	49331	15800	ERR351593
22008	2009	September	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	143	2135297	48935	14933	ERR351594
22014	2009	September	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	144	2138162	48898	14849	ERR351595
22154	2009	September	2nd year UoN	Y: P1.5-1,10-93: F4-1: ST-1655 (cc23)	222	2087396	30805	9403	ERR351596
22238	2009	September	2nd year UoN	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)	173	2098463	37765	12130	ERR351597
20373	2009	December	1st year UoN	Y: P1.5-1,10-10: F4-1: ST-1655 (cc23)	169	2093706	31633	12389	ERR351598
20374	2009	December	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)	170	2097407	38085	12338	ERR351599
20812	2009	December	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	185	2109647	32845	11404	ERR351600
22449	2009	December	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	164	2098687	31368	12797	ERR351602
22623	2009	December	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	174	2098169	32571	12059	ERR351603
22667	2009	December	1st year UoN	Y: P1.5-1,10-22: F5-1: ST-114 (cc22)	214	2208272	30081	10320	ERR351604

22678	2009	December	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	166	2095958	30863	12627	ERR351605
22693	2009	December	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)	171	2097428	37784	12266	ERR351606
22698	2009	December	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)	186	2092600	31435	11251	ERR351607
22715	2009	December	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)	171	2096212	36414	12259	ERR351608
22908	2009	December	1st year UoN	Y: P1.5-8,10-4: F5-2: ST-168 (cc167)	159	2116817	40072	13314	ERR351609
22933	2009	December	1st year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	148	2134065	45033	14420	ERR351610
22940	2009	December	1st year UoN	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)	164	2095771	45309	12780	ERR351611
22964	2009	December	1st year UoN	Y: P1.5-1,2-2: F5-1: ST-3651 (cc22)	180	2189786	36166	12166	ERR351612
23009	2009	December	1st year UoN	Y: P1.5-1,2-2: F5-1: ST-3651 (cc22)	179	2190353	36561	12237	ERR351613
23104	2009	December	1st year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	193	2114323	28486	10956	ERR351614
23153	2009	December	1st year UoN	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)	174	2096746	31399	12051	ERR351615
23312	2009	December	1st year UoN	Y: P1.5-1,2-2: F5-1: ST-3651 (cc22)	217	2190866	28642	10097	ERR351616
23470	2009	December	1st year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	170	2096997	30857	12336	ERR351617
20338	2009	December	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	167	2098041	37322	12564	ERR351618
20393	2009	December	2nd year UoN	Y: P1.5-1,10-4: F4-1: ST-ND (-)	209	2101112	35760	10054	ERR351619
20398	2009	December	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	174	2097705	31424	12056	ERR351620
20411	2009	December	2nd year UoN	Y: P1.5-1,10-46: F3-9: ST-103 (cc103)	180	2165299	48063	12030	ERR351621
20804	2009	December	2nd year UoN	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)	177	2102522	37345	11879	ERR351622
20821	2009	December	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	178	2096794	31368	11780	ERR351623
20824	2009	December	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	166	2099157	37322	12646	ERR351624
21513	2009	December	2nd year UoN	Y: P1.5-1,2-2: F5-1: ST-3651 (cc22)	228	2190302	29668	9607	ERR351625
21570	2009	December	2nd year UoN	Y: P1.21,16: F3-7: ST-7850 (cc174)	148	2135897	46196	14432	ERR351626
21963	2009	December	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	174	2098738	31368	12062	ERR351627
21978	2009	December	2nd year UoN	Y: P1.5-1,10-4: F5-5: ST-279 (cc167)	190	2135482	35613	11240	ERR351628
22656	2009	December	2nd year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	202	2105759	28670	10425	ERR351629
23099	2009	December	2nd year UoN	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)	210	2119374	27893	10093	ERR351630
23195	2009	December	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	177	2097926	28482	11853	ERR351631
23214	2009	December	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	149	2136713	46325	14341	ERR351632
23264	2009	December	2nd Year UoN	Y: P1.5-1,10-4: F5-36: ST-884 (cc167)	177	2128184	37280	12024	ERR351633
23283	2009	December	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	141	2134836	49117	15141	ERR351634
23326	2009	December	2nd year UoN	Y: P1.21,16: F3-7: ST-1466 (cc174)	132	2128806	47571	16128	ERR351635
23384	2009	December	2nd year UoN	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)	174	2099241	35310	12065	ERR351636
23407	2009	December	2nd year UoN	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)	168	2091936	31267	12452	ERR351637

[a] UoN = University of Nottingham

[b] Derived from genome sequence data.

[c] Weighted median statistic indicating that 50% of the entire assembly is contained in contigs equal to, or larger than, this value.

SUPPLEMENTARY TABLE 2. Details of the 73 MenY invasive isolates used in this study. All were isolated during the epidemiological year 2010-11 in England, Wales and Northern Ireland. Further details can be found in the Meningitis Research Foundation Meningococcus Genome Library database.

Isolate	Strain designation
M10 240503	Y: P1.5-1,2-2: F1-96: ST-23 (cc23)
M10 240505	Y: P1.5-1,10-10: F4-1: ST-1655 (cc23)
M10 240507	Y: P1.21,16-5: Δ: ST-183 (cc23)
M10 240520	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240530	Y: P1.5-2,10-2: F3-1: ST-9813 (cc23)
M10 240531	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240534	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240536	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240540	Y: P1.5-1,10-10: F4-1: ST-9814 (cc23)
M10 240580	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240590	Y: P1.5-1,10-4: F4-1: ST-23 (cc23)
M10 240632	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240633	Y: P1.5-1,10-10: F4-1: ST-1655 (cc23)
M10 240643	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)
M10 240644	Y: P1.5-1,2-2: F5-1: ST-3651 (cc22)
M10 240675	Y: P1.5-8,10-4: F5-7: ST-168 (cc167)
M10 240685	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240694	Y: P1.21,16: F3-7: ST-1466 (cc174)
M10 240700	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)
M10 240732	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)
M10 240745	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)
M10 240759	Y: P1.21,16: F3-7: ST-1466 (cc174)
M10 240776	Y: P1.5-1,10-4: F4-1: ST-1655 (cc23)
M10 240777	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)
M10 240781	Y: P1.5-1,2-2: F5-1: ST-3651 (cc22)
M10 240786	Y: P1.5-1,10-1: F4-1: ST-9831 (cc23)
M10 240787	Y: P1.5-1,10-10: F4-1: ST-1655 (cc23)
M10 240798	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M10 240804	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240000	Y: P1.5-1,10-4: F4-1: ST-23 (cc23)
M11 240005	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)
M11 240007	Y: P1.5-1,10-10: F4-1: ST-1655 (cc23)
M11 240031	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240036	Y: P1.7-2,10-1: F4-1: ST-1655 (cc23)
M11 240038	Y: P1.5-8,10-4: F5-2: ST-168 (cc167)
M11 240043	Y: P1.5-1,10-1: F4-1: ST-9842 (cc23)
M11 240064	Y: P1.ND,ND: F3-4: ST-7786 ()
M11 240065	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240071	Y: P1.5-1,10-1: F5-5: ST-767 (cc167)
M11 240073	Y: P1.22,9: F3-7: ST-1466 (cc174)

M11 240096	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240116	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240126	Y: P1.5-1,10-4: F4-1: ST-23 (cc23)
M11 240157	Y: P1.5-1,10-1: F4-1: ST-10138 (cc23)
M11 240161	Y: P1.21,16: F3-7: ST-1466 (cc174)
M11 240165	Y: P1.21,16: F3-7: ST-1466 (cc174)
M11 240166	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)
M11 240192	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240209	Y: P1.21,16: F3-7: ST-1466 (cc174)
M11 240210	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)
M11 240211	Y: P1.21,10-46: F3-7: ST-9893 (cc174)
M11 240215	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240216	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240227	Y: P1.5-1,10-4: F4-1: ST-6463 (cc23)
M11 240258	Y: P1.5-1,2-2: F5-8: ST-23 (cc23)
M11 240263	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240277	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240278	Y: P1.7-2,30-4: F4-1: ST-1655 (cc23)
M11 240283	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)
M11 240298	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240312	Y: P1.5-1,10-4: F4-1: ST-23 (cc23)
M11 240313	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240319	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240322	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240326	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240346	Y: P1.5-8,10-4: F5-2: ST-168 (cc167)
M11 240396	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240399	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240411	Y: P1.5-1,2-2: F5-8: ST-4183 (cc23)
M11 240430	Y: P1.5-2,10-1: F4-1: ST-23 (cc23)
M11 240435	Y: P1.5-1,10-1: F4-1: ST-1655 (cc23)
M11 240437	Y: P1.5-1,10-4: F4-3: ST-784 (cc92)
M11 240442	Y: P1.5-1,10-1: F4-1: ST-23 (cc23)