
Evolution of Sequence Type 4821 Clonal Complex Meningococcal Strains in China from Prequinolone to Quinolone Era, 1972–2013

Qinglan Guo,¹ Mustapha M. Mustapha,¹ Mingliang Chen, Di Qu, Xi Zhang, Min Chen, Yohei Doi, Minggui Wang, Lee H. Harrison

The expansion of hypervirulent sequence type 4821 clonal complex (CC4821) lineage *Neisseria meningitidis* bacteria has led to a shift in meningococcal disease epidemiology in China, from serogroup A (MenA) to MenC. Knowledge regarding the evolution and genetic origin of the emergent MenC strains is limited. In this study, we subjected 76 CC4821 isolates collected across China during 1972–1977 and 2005–2013 to phylogenetic analysis, traditional genotyping, or both. We show that successive recombination events within genes encoding surface antigens and acquisition of quinolone resistance mutations possibly played a role in the emergence of CC4821 as an epidemic clone in China. MenC and MenB CC4821 strains have spread across China and have been detected in several countries in different continents. Capsular switches involving serogroups B and C occurred among epidemic strains, raising concerns regarding possible increases in MenB disease, given that vaccines in use in China do not protect against MenB.

The incidence of meningococcal disease and *Neisseria meningitidis* strain distribution vary over time, within and between countries and regions (1). Six serogroups (A, B, C, W, X, and Y) account for nearly all cases of invasive meningococcal disease (IMD) globally (1). Serogroup C (MenC) cases were rare in China until 2003–2005, when several MenC outbreaks were reported in Anhui Province (2–4). These outbreaks were caused by a previously unreported hypervirulent clonal complex (CC) 4821 lineage (5). A pharyngeal carriage survey and national public health surveillance during 2004–2005 identified CC4821 among sporadic IMD case-patients and asymptomatic

carriers across 11 provinces, demonstrating the wide geographic distribution of CC4821 (2). During 2005–2012, MenC CC4821 became the leading cause of endemic meningococcal disease in China (6). Further analyses of historic isolate collections identified MenB and MenC CC4821 from carriage surveys in the 1970s (2,7). These studies demonstrated that CC4821 had been mostly associated with asymptomatic carriage over several decades before it emerged as a main cause of IMD (2,7,8). Also, our recent analyses of quinolone resistance among historic meningococcal isolate collections in China found a substantial temporal shift toward increased quinolone non-susceptibility from the prequinolone era (before ≈1985) to the quinolone era (none versus >70%), particularly within hypervirulent CC4821 and CC5 lineages (7). Such findings support the hypothesis that quinolone resistance could have played a role in the emergence of MenC CC4821 outbreaks in China.

Meningococci have a dynamic genome that evolves rapidly through point mutations and frequent recombination. Such genetic changes give rise to strains with novel capsular or other major surface antigens that evade existing population immunity (9). A study by Zhu et al. found extensive genomic variation among 22 CC4821 invasive and carriage isolates from 12 provinces in China during 2005–2011 (8). In that study, CC4821 isolates belonged to 2 distinct phylogenetic groups, and results indicated that group 1, containing the epidemic reference strain 053442, might be more invasive than group 2 and that MenB and MenC coexisted within both groups 1 and 2 (8).

Our study describes phylogenetic relationships within a collection of historic and current isolates from Shanghai in the prequinolone and quinolone eras and explores how the isolates fit into the larger genetic profile of CC4821 from China (8). We aimed to shed light on the genomic factors underlying the abrupt transition of this lineage from a minority strain to a leading cause of endemic disease and outbreaks.

Author affiliations: Fudan University Huashan Hospital, Shanghai, China (Q. Guo, M. Wang); University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA (M.M. Mustapha, Y. Doi, L.H. Harrison); Shanghai Municipal Center for Disease Control and Prevention, Shanghai (M. Chen, X. Zhang, M. Chen); Fudan University Shanghai Medical College, Shanghai (D. Qu)

¹These authors contributed equally to this article.

DOI: <https://doi.org/10.3201/eid2404.171744>

Materials and Methods

Strain Selection and Molecular Typing

A total of 374 meningococcal isolates were collected from IMD case-patients, close contacts, and asymptomatic carriers during pharyngeal carriage surveys and routine, laboratory-based public health surveillance conducted during 1965–1985 and 2005–2013 (7). The interruption during 1986–2004 was because of the decreased incidence of meningococcal cases and the increased ability to identify *N. meningitidis* in hospitals, necessitating fewer isolates to be referred (7). In all, 52 CC4821 isolates were identified and underwent molecular characterization using traditional PCR sequencing of the *porA*, *porB*, *fetA*, *fHbp*, *nadA*, *nhba*, and *gyrA* genes and pulsed-field gel electrophoresis as described previously (7,10,11). We selected 8 of these 52 isolates, representing strains from different periods, serogroups, or pulsed-field gel electrophoresis groups, for whole-genome sequencing (WGS) and in-depth phylogenetic analysis. We downloaded assembled contiguous genome sequences (contigs) for 24 additional genome sequences from previous studies of CC4821 in China from GenBank (NM11003, accession no. NZ_ANBU00000000) (5,8). Therefore, a total of 32 whole-genome sequences underwent core genome phylogenetic analyses for this study (online Technical Appendix Table 1, <https://wwwnc.cdc.gov/EID/article/24/4/17-1744-Techapp1.pdf>).

Genome Sequencing and Assembly

We performed single molecule real-time sequencing (PacBio; Pacific Biosciences, Menlo Park, CA, USA) on 4 CC4821 isolates (NM040, NM062, NM205, and NM323) and sequenced the remaining 4 isolates (NM001, NM050, NM193, and NM313) by using Illumina HiSeq paired-end sequencing (Illumina, San Diego, CA, USA). We assembled PacBio genomes by using HGAP 4.0 (<https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/HGAP>) and Illumina genomes by using SPAdes 3.7 (12). We annotated assembled contigs by using the Prokka v1.11 (13) pipeline and submitted them to the PubMLST *Neisseria* genome database (<http://pubmlst.org/neisseria>) with ID numbers 41414–41421, where allelic numbers were assigned to all identified genes (14).

Phylogenetic Analyses

We aligned assembled genomes ($n = 32$) and produced a core genome phylogenetic tree with 1,000 rapid bootstrap replicates by using RAxML and Mauve 2.3 as previously described (15–17). Serogroup A reference genome Z2491 was included as an outgroup (18). We aligned gene sequences selected for focused analyses with MEGA 5.2 (<http://www.megasoftware.net>) and constructed maximum-likelihood phylogenetic trees under the HKY model of evolution (19).

We included comparison sequences from a global collection of 133 genomes from the PubMLST *Neisseria* genome database representing all major invasive disease lineages in some of these phylogenetic analyses as references.

Recombination

We assessed recombination by using ClonalFrameML and Gubbins (20,21). We then mapped the recombination to the CC4821 reference genome 053442 (online Technical Appendix).

Gene Content

We assessed gene content by using the Roary 3.6 ortholog clustering program (22), which identifies presence or absence of orthologous gene sequences using a cutoff of 90% sequence identity ($-i = 90$). We defined core genes as genes present in $\geq 90\%$ of the genomes. When comparing the gene content of 2 groups of genomes, we defined a gene as specific to that group if it was present in $\geq 90\%$ of the genomes in the group and in $< 20\%$ of the genomes in the comparison group. We downloaded gene functional annotation from the COG database (23) and compared genes containing recombinant sequences with nonrecombinant ones based on the major COG classes (cellular processes and signaling, information storage and processing, metabolism, poorly characterized) using uncorrected χ^2 tests.

Results

Out of 52 Shanghai isolates, 18 (34%) were from 1972–1977 (3 IMD isolates and 15 asymptomatic carriage isolates), and the remaining 34 (23 IMD isolates, 6 isolates from close contacts, and 5 carriage isolates) were isolated during 2005–2013. Most (56%) Shanghai isolates were serogroup C, 42% were serogroup B, and 1 was nongroupable (online Technical Appendix Table 2).

Phylogenetic Analyses of CC4821 Isolates in China

We conducted comparative genome analyses for 8 isolates from Shanghai and 24 publicly available genomes from across China (Figure 1). Among the 8 newly sequenced isolates, 4 were from asymptomatic carriers from 1972–1977 (NM193, NM205, NM313, and NM323), 3 were IMD isolates from 2005–2011 (NM001, NM062, and NM040), and 1 was from an asymptomatic contact of a meningitis patient (NM050). The remaining 24 genome sequences included invasive and carriage CC4821 isolates from 12 provinces across China from 2004–2011, as characterized in previous studies (4,8). Overall, MenC represented 66% (21/32) of isolates that we analyzed, and the remaining 34% (11/32) belonged to MenB.

The core genome phylogenetic tree classified CC4821 into 2 distinct groups (Figure 1). Group 1 consisted of several isolates that were very closely related; this group con-

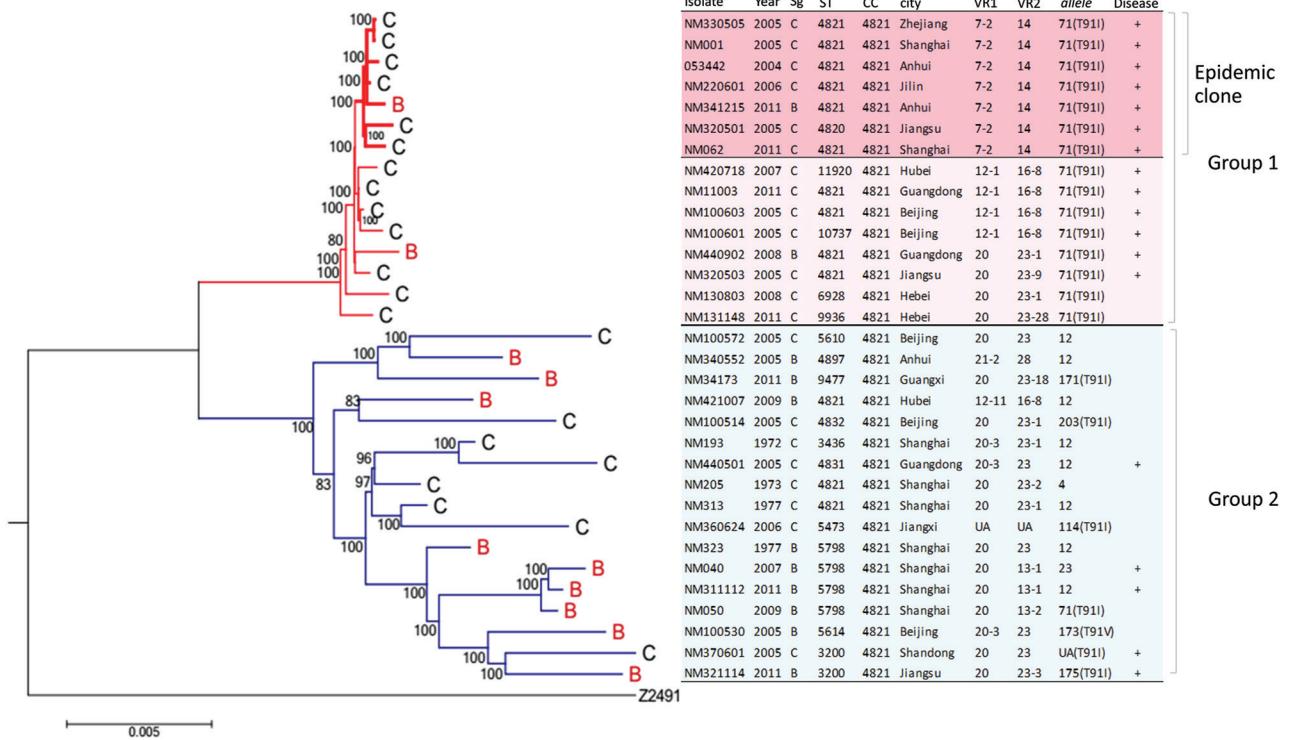


Figure 1. Core genome phylogenetic tree showing relationships between *Neisseria meningitidis* serogroups C and B CC4821 strains, China, 1972–2011. The strains cluster within 1 of 2 distinct phylogenetic groups, group 1 and group 2. Within group 1 is an antigenically distinct clonal group (epidemic clone) containing outbreak-associated strains. Tree is rooted using serogroup A reference strain (Z2491) as an outgroup. Maximum-likelihood phylogenetic trees of aligned core genome sequences were generated under a general time reversible model of evolution with gamma rate heterogeneity, with 1,000 rapid bootstrap replicates represented as a percentage. Only node labels with >80% bootstrap support are shown. Strain type and date and place of isolation are shown; + indicates strains isolated from invasive disease cases. Resistant point mutations on the T91 position of *gyrA* gene are shown alongside data on *gyrA* allele designation. Scale bar represents total substitutions per site. CC, clonal complex; PorA VR, outer membrane protein PorA variable regions; Sg, serogroup; ST, sequence type.

tained 15 isolates from 2004 to 2011, and most were from IMD case-patients. In addition, the epidemic reference strain 053442 clustered with a small group of highly similar group 1 isolates. This subgroup is defined as the epidemic clone based on core genome phylogenetic analyses and antigen gene profile (described hereafter). A second more diverse phylogenetic group (group 2) consisted of isolates from Shanghai during 1972–1977 and more recent ones collected from 9 provinces during 2005–2011. In contrast to group 1, only 29% of group 2 isolates were from IMD case-patients (Figure 1). MenB was found in both groups and was interspersed with MenC genomes. Two of 11 MenB isolates belonged to group 1, the remainder to group 2 (Figure 1; online Technical Appendix Table 1).

Characterization of Antigen Gene Content

We examined 32 isolates that had undergone WGS to identify the major antigen gene (*PorA* VR1 and VR2, *porB*, *FetA*, *nhba*, and *fHbp*) alleles that corresponded to the epidemic strain, group 1 or group 2, as determined by the core

genome phylogenetic analysis (Figures 1, 2; online Technical Appendix Table 1, Figures 1–5). *nadA* was missing in all CC4821 study genomes. Group 1 genomes had diverse *fHbp* and *porA* alleles. All 15 group 1 isolates contained the *porB* 3–48 and *FetA* F3–3 alleles, and *nhba* allele 124. The epidemic clone contained a few highly related *porA* alleles that all encoded unique, conserved *porA* variable regions (*PorA* VR1 and VR2: P1.7–2, 14). Group 2 isolates had the most antigen gene diversity, containing 5–12 different alleles for each antigen-encoding gene at the nucleotide level and no clear predominance of any single allelic profile. We observed little overlap between the antigen gene allelic profiles in groups 1 and 2. Only 2 of 13 group 1 alleles were also found in group 2. None of the antigen gene alleles found in the epidemic clone was present in group 2, suggesting that the epidemic clone had a nonoverlapping repertoire of antigens compared with historic and current group 2 isolates (Figure 2; online Technical Appendix Table 1, Figures 1–5). Group 2 was predominantly associated with *PorA* P1.20 variants (82%); *PorB* 3–229 (29%);

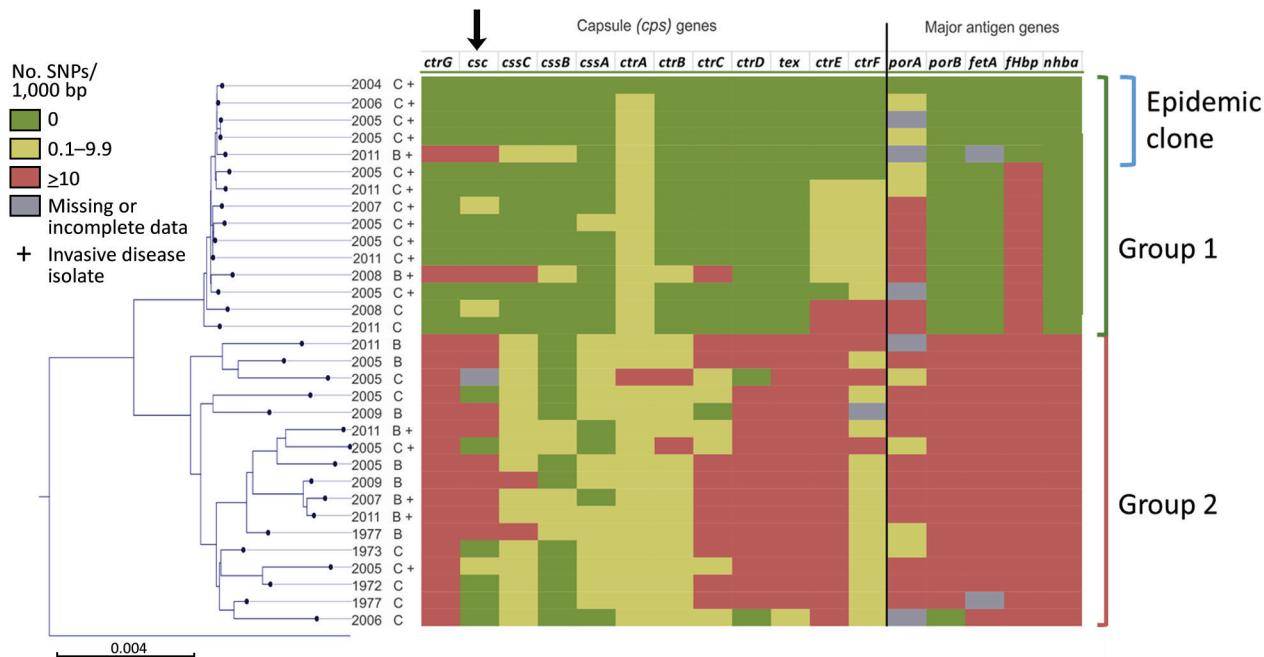


Figure 2. Genetic diversity within capsule and major antigen encoding genes among 32 *Neisseria meningitidis* clonal complex 4821 isolates, China, 1972–2011. Gene sequences from 32 clonal complex 4821 isolates were compared with the epidemic reference strain 053442 (the topmost isolate on the juxtaposed core genome phylogenetic tree). Scale bar represents total substitutions per site. SNP, single-nucleotide polymorphism.

FetA F3–9 (18%) and FetA F5–2 (18%); FHbp peptide 16 (65%); and *nhbA* 553 (41%) (online Technical Appendix Table 1).

Recombination Events Separating Group 1 and Group 2 Isolates

We identified extensive recombination within CC4821 lineage, as detected by ClonalFrameML (20). The estimated rate of recombination relative to mutation (R/θ) for the entire CC4821 core genome was 1.37 (95% CI 1.40–1.34), and the relative impact of recombination to mutation (r/m) was 20.83, indicating that ≈ 21 nt were acquired by recombination for every point mutation within the core genome of CC4821. Sublineage specific recombination rates (R/θ) were 1.79 (95% CI 1.53–2.05) for group 1, 1.47 (95% CI 1.15–1.79) for group 2, and 0.5 (95% CI 0.26–0.75) for the epidemic clone.

A total of 46 recombination fragments containing 120 genes were mapped to the node that separated group 1 from group 2, indicating that these recombination events could be linked to the divergence of group 1 and 2 sublineages (online Technical Appendix Table 3). Sequence alignment and phylogenetic analysis confirmed that the 46 genomic loci represented regions of marked sequence divergence, presumably caused by homologous recombination within a common ancestral genome. These 120 recombinant genes belonged to diverse functional categories; metabolism was the overrepresented functional assignment (51/120 [42.5%]), compared

with 28% of nonrecombinant genes ($p < 0.001$ by χ^2 test). Proportions of other functional groups were similar between recombinant and nonrecombinant genes.

Recombination Events Unique to the Epidemic Clone

The epidemic clone diverged from group 1 through several additional recombination events, affecting 24 genes across 7 genomic loci (Table). These genomic loci included capsule translocation genes (*ctrE* and *ctrF*); major outer membrane protein (OMP) antigen genes *porA*, *porB*, and *fhbP*; *pts* genes involved in carbohydrate transport and metabolism; and the *tkl* locus encoding DNA polymerase and several metabolic enzymes.

Gene Content

We identified a total of 3,292 unique genes, of which 1,730 core genes were shared by most CC4821 genomes. Eleven genes were present in $\geq 90\%$ of group 1 isolates but missing in most group 2 isolates. Genes that were predominantly found in group 1 included *lbpB* encoding lactoferrin binding protein B, *nhaP* encoding Na^+/K^+ antiporter, and genes encoding several putative enzymes and 4 hypothetical proteins whose functions are unknown (online Technical Appendix Table 4). These genes were spread across several locations on the reference genome 053442, suggesting that they were acquired separately rather than in a single event. The epidemic clone had no noticeable gene gain or loss compared with other group 1 isolates.

Table. Recombinant genes unique to the *Neisseria meningitidis* clonal complex 4821 epidemic clone, China, 1972–1977 and 2005–2013*

Gene ID	Gene name	Annotation
NMCC_0090	<i>ctrE</i>	Polysialic acid capsule modification protein LipA
NMCC_0091	<i>ctrF</i>	Polysialic acid capsule modification protein LipB
NMCC_0136		Putative RmuC-like protein
NMCC_0137		Putative metallo-dependent hydrolase
NMCC_0138		Putative periplasmic DNA ligase (polydeoxyribonucleotide synthase [ATP])
NMCC_0140	<i>ptsIIA</i>	Phosphotransferase system, enzyme IIA (protein IIA)
NMCC_0141	<i>ptsH</i>	Phosphocarrier protein HPr (phosphotransferase system, histidine-containing protein)
NMCC_0142	<i>ptsI</i>	Phosphoenolpyruvate-protein phosphotransferase (phosphotransferase system, enzyme I; protein I)
NMCC_0158	<i>porB</i>	Major outer-membrane protein P.IB (protein IB; PIB; porin)
NMCC_0350		Putative peptidase
NMCC_0351	<i>fHbp</i>	Factor H binding lipoprotein (lipoprotein GNA1870)
NMCC_0352	<i>fba</i>	Fructose-bisphosphate aldolase
NMCC_1338	<i>porA</i>	Major outer-membrane protein P.IA (protein IA; PIA; porin)
NMCC_1341	<i>greA</i>	Transcription elongation factor GreA (transcript cleavage factor GreA)
NMCC_1342	<i>aroA</i>	3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase; EPSP synthase; EPSPS)
NMCC_1343		Conserved hypothetical lipoprotein
NMCC_1363		Putative DnaQ-like exonuclease
NMCC_1364		Arg tRNA
NMCC_1365		Glu tRNA
NMCC_1366		Putative dioxygenase
NMCC_1367		Conserved hypothetical membrane protein
NMCC_1368		Putative ferredoxin
NMCC_1370	<i>tkt</i>	Transketolase
NMCC_2038	<i>fmt</i>	Methionyl-tRNA formyltransferase

*Gene IDs correspond to the clonal complex 4821 reference genome 053442. ATP, adenosine triphosphate; EPSP, 5-enolpyruvylshikimate-3-phosphate; EPSPS, EPSP synthase; ID, identifier.

Capsular Gene *cps* Cluster

The *cps* cluster of *N. meningitidis* consists of 6 regions (D-A-C-E-D'-B) required for capsule biosynthesis (region A), transport (region C), and translocation (region B) (24). The group 1 isolates were characterized by a novel capsule region A (*ctrG4-ctrE1-csc1-ctrC3-ctrB1-ctrA3*), associated with region C (*ctrA5-ctrB1-ctrC4-ctrD1*) and region E (*tex-orf1-orf2*) (online Technical Appendix Table 1). Nine of 15 regions A-C-E identified among group 1 isolates were identical, and another 4 were almost identical (1-bp variation in *csc*, *ctrA*, or *ctrB* over 14,489 bp). Only 2 diverged because of the allelic replacement containing the serogroup-specific polysialyltransferase gene (*csb*), resulting in capsular switches involving serogroups B and C (NM341215 and NM440902) (online Technical Appendix Table 2). Group 2 had substantial genetic diversity within the *cps* cluster, with no 2 isolates possessing an identical *cps* gene allelic profile (Figure 2).

Characterization of CC4821 Isolates from Shanghai

Our results classified 52 CC4821 isolates from Shanghai into epidemic clone or groups 1 and 2 on the basis of associated *porA*, *porB*, *fetA*, *fHbp*, and *nhba* antigen genes defined by WGS (Figure 1; online Technical Appendix Table 2). A substantial proportion of strains in Shanghai (23/52 [44%]) belonged to group 1, with 0–1 antigen gene differences (Figure 1; online Technical Appendix Table 2, Figures 1–5). Shanghai group 1 isolates were exclusively

isolated during 2005–2013; were MenC (19/23), MenB (3/23), or nongroupable (1/23); and were isolated from IMD case-patients (18/23), close contacts of IMD case-patients (4/23), or an asymptomatic carrier (1/23). Within group 1, a total of 16 isolates (16/23 [70%]) contained *PorA* P1.7–2,14 and were consistent with the epidemic clone by all 5 antigen genes. Twenty-nine (29/52 [56%]) Shanghai isolates belonged to group 2 and differed from group 1 by 4–5 out of 5 antigen loci. Shanghai group 2 was dominated by carriage isolates (21/29 [72%]) from 1972–1977 and 2005–2013. MenB equaled or outnumbered MenC among historic (9/18 [50%]) and recent (10/11 [91%]) group 2 CC4821 isolates. MenB strains in group 1 possessed *fHbp* 498 or 22 encoding peptides 1.80 or 2.22 (subvariants belonging to *fHbp* variant group 1 or 2), as most (17/19 [89%]) of MenC epidemic clones did (online Technical Appendix Tables 1, 2).

All group 1 isolates contained fluoroquinolone-resistant *gyrA* allele 71 (corresponding to allele R1 in our previous study [7]), which was generated by a nonsynonymous mutation of fluoroquinolone-susceptible *gyrA* allele 12 (allele S1), creating an amino acid substitution of T91I. The *gyrA* allele 12/S1 was carried by 76% (22/29) of the group 2 strains (online Technical Appendix Table 2). In contrast, only 4 of 29 group 2 isolates from Shanghai and another 6 publicly available group 2 genomes contained resistant *gyrA* alleles, which were different from each other. Group 2 isolates with genotypic resistance to quinolones were

genetically highly diverse as evidenced by diverse sequence types, different antigen gene profiles, and expression of both MenB and MenC capsules. However, all quinolone-resistant isolates were from 2005–2013 (Figure 1; online Technical Appendix Tables 1, 2).

Discussion

This study presents detailed genomic analyses of the current major endemic meningococcal disease lineages in China, serogroups C and B CC4821. Phylogenetic analyses have suggested that the epidemic clone (corresponding to the epidemic China^{CC4821-R1-C/B} clone described in our previous work [7]) responsible for most recent CC4821 disease cases was nested within a distinct phylogenetic group (group 1), consistent with recent clonal expansion of a genetically distinct strain. In contrast, group 2 was temporally and genetically more diverse and accounted for a smaller proportion of IMD cases, as evidenced by the preponderance of asymptomatic carriage isolates within this group (8). Group 1 isolates were first identified during MenC outbreaks during 2003–2005, whereas group 2 isolates were mostly associated with carriage from the 1970s to 2013 (7).

This work adds to the description of CC4821 by a previous study (8) by demonstrating that, within group 1, a genetically distinct clone exists that shares the *porA* antigenic formula P1.7–2,14 and comprises strains associated with outbreaks and a large proportion of epidemic disease cases. This finding is in keeping with surveillance studies reporting MenC strains during 2005–2012 indicating that 55% of 238 confirmed meningococcal disease cases and 84% of 131 MenC strains belong to CC4821 with *porA* P1.7–2,14 (6), suggesting that the expansion of CC4821 was caused by clonal expansion of this antigenic type. Genomic analyses demonstrated that the epidemic CC4821 lineage had undergone 2 crucial genetic events compared with historic asymptomatic carriage isolates. First, CC4821 diverged into 2 major sublineages through extensive recombination events predominantly affecting genes involved in metabolic functions. Such extensive allelic exchanges might have enhanced the transmission fitness or the invasive potential within group 1. Second, a virulent, antigenically unique, epidemic strain emerged from within group 1 in a second set of more focused recombination events affecting major antigen genes, *porA*, *porB*, *fHbp*, capsule genes *ctrE* and *ctrF*, and several metabolic genes associated with oxidative phosphorylation and glycolytic processes. These genetic changes possibly account for the rapid dissemination and increased invasive potential of the epidemic clone.

This work also adds to the evidence that the emergence and persistence of virulent meningococcal strains occurs through the introduction of a novel antigenic variant in an immunologically naive population (9,15,25–27). In addition, even though a few major antigen gene repertoires

mediate the microevolution of virulent meningococcal strains, a larger and more assorted number of metabolic genes might be involved in the divergence of sublineages. Additional research is needed to elucidate the intricate interactions between various metabolic pathways in the fitness and virulence potential of meningococci.

Selection pressure of fluoroquinolones might also affect the evolution and adaptation of meningococcal strains, as evidenced by various resistant *gyrA* alleles recovered in many meningococcal lineages and singletons only in the quinolone era in China (7,28). All of the group 1 CC4821 strains contain *gyrA* allele 71/R1, in contrast to various resistant *gyrA* alleles in group 2, suggesting acquisition of this trait by group 1 at an early stage in evolution. The *gyrA* allele 71/R1 derived from *gyrA* allele 12/S1, which was possessed by most of the group 2 strains both in prequinolone and quinolone eras, indicating their common origin. The precise role of quinolone resistance in the emergence of the epidemic clone requires further study.

The relative rates and impact of recombination within the CC4821 lineage were considerably higher than previous estimates that examined recombination across 7 housekeeping genes (29–31). We identified 4,026 recombination events and 21 recombinant single-nucleotide polymorphisms (SNPs) for every point mutation within CC4821. Estimates of meningococcal recombination rates from whole-genome sequence data are limited. A study of MenA CC5 strains in Africa (27), an epidemic lineage notable for relatively low genetic diversity, found 34 recombinant sequences and 12 recombinant SNPs for every point mutation. A study of recombination rates within conserved multilocus sequence type loci among carriage isolates from the Czech Republic in 1993 found 6.2–16.8 SNPs caused by recombination for each point mutation (29). These data suggest that although homologous recombination is a shared mechanism for meningococcal strain emergence and persistence (29,31–33), the frequency and extent of recombination likely differs substantially between lineages.

High rates of recombination have led to multiple distinct capsular switch strains expressing group B and C capsule, as described in this study, and serogroup W (8,34). Both MenB and MenC were phylogenetically diverse and interspersed within both groups 1 and 2, suggesting multiple distinct capsular switches between these serogroups rather than clonal expansion of a single capsular switch strain (35,36). Moreover, recombination events that led to no apparent change in capsular phenotype were also prevalent.

Wide geographic and temporal spread of MenB strains in China is of concern given that vaccines currently in use do not protect against MenB disease. Furthermore, marked heterogeneity exists among both MenB and MenC strains within the gene encoding FHbp, a key component of MenB vaccines.

At the time of its emergence, CC4821 was confined to China and Taiwan (37). However, review of the PubMLST *Neisseria* genome database suggests a recent increase in the geographic spread of CC4821. CC4821 isolates were reported from a small number of asymptomatic carriers in Brazil (2014) (38,39) and Australia (2012–2017), as well as isolated disease cases and carriers from France (2009–2011), the United Kingdom (2011–2014), the United States (2007–2016), and India (2017) (14). A recent case report of a quinolone-resistant sequence type 4821 strain from Japan in a patient with no history of foreign travel also suggests local transmission of CC4821 in Japan (40). Global CC4821 carriage and IMD isolates on the PubMLST *Neisseria* database were antigenically diverse, containing PorA types associated with both group 1 and 2 strains, and expressed both serogroup C and B capsules. This finding suggests low-level dissemination of CC4821 strains with diverse virulence and antigenic types as opposed to clonal spread of a single epidemic strain. This pattern is in contrast to the pandemic spread of highly clonal serogroup A epidemic strains from China from the 1960s through the 1990s (25). Genome-based surveillance of these global CC4821 strains is needed to monitor the global spread of this clonal lineage.

Our study is limited by lack of isolates before the 1970s and during 1986–2004; isolates from those periods might have provided a clearer picture of the multiple evolutionary steps that led to the epidemic CC4821 clone. Also, the high frequency of recombination within capsular genes makes it difficult to accurately determine the direction of capsular switch.

In summary, we have presented detailed genomic analysis of a major hypervirulent meningococcal lineage associated with MenC and MenB in China and identified key genomic factors that might have led to the emergence and persistence of MenC in China. The potential emergence of MenB is of public health concern. Strengthened laboratory surveillance for disease cases and carefully planned carriage surveys are needed to monitor global trends, detect outbreaks, and inform immunization policies.

Acknowledgments

We thank Jane Marsh for providing thoughtful comments on an early draft of this manuscript.

This study made use of the PubMLST *Neisseria* genome database (<http://pubmlst.org/neisseria>), developed by Keith Jolley and hosted by the University of Oxford. The development of this site has been funded by the Wellcome Trust and European Union. This study also made use of the Meningitis Research Foundation Meningococcus Genome Library (<http://www.meningitis.org/research/genome>), developed by Public Health England, the Wellcome Trust Sanger Institute, and the

University of Oxford as a collaboration. The project is funded by the Meningitis Research Foundation.

This work was supported by grants from the National Natural Science Foundation of China (81673479, 81120108024, and 81601801); the Shanghai Pujiang Program (16PJD010); the Shanghai Rising-Star Program (17QA1403100); the Shanghai Medical Health Plan for Outstanding Young Talents (2017YQ039); and the 4th Three-Year Action Plan for Public Health of Shanghai (GWTD2015S01).

About the Author

Dr. Guo is an associate professor in the Institute of Antibiotics, Huashan Hospital, Fudan University. Her research interests include mechanisms of antimicrobial resistance in clinical isolates.

References

- Halperin SA, Bettinger JA, Greenwood B, Harrison LH, Jelfs J, Ladhani SN, et al. The changing and dynamic epidemiology of meningococcal disease. *Vaccine*. 2012;30(Suppl 2):B26–36. <http://dx.doi.org/10.1016/j.vaccine.2011.12.032>
- Zhang X, Shao Z, Yang E, Xu L, Xu X, Li M, et al. Molecular characterization of serogroup C *Neisseria meningitidis* isolated in China. *J Med Microbiol*. 2007;56:1224–9. <http://dx.doi.org/10.1099/jmm.0.47263-0>
- Ni JD, Jin YH, Dai B, Wang XP, Liu DQ, Chen X, et al. Recent epidemiological changes in meningococcal disease may be due to the displacement of serogroup A by serogroup C in Hefei City, China. *Postgrad Med J*. 2008;84:87–92. <http://dx.doi.org/10.1136/pgmj.2007.065680>
- Shao Z, Li W, Ren J, Liang X, Xu L, Diao B, et al. Identification of a new *Neisseria meningitidis* serogroup C clone from Anhui Province, China. *Lancet*. 2006;367:419–23. [http://dx.doi.org/10.1016/S0140-6736\(06\)68141-5](http://dx.doi.org/10.1016/S0140-6736(06)68141-5)
- Peng J, Yang L, Yang F, Yang J, Yan Y, Nie H, et al. Characterization of ST-4821 complex, a unique *Neisseria meningitidis* clone. *Genomics*. 2008;91:78–87. <http://dx.doi.org/10.1016/j.ygeno.2007.10.004>
- Zhou H, Shan X, Sun X, Xu L, Gao Y, Li M, et al. Clonal characteristics of invasive *Neisseria meningitidis* following initiation of an A + C vaccination program in China, 2005–2012. *J Infect*. 2015;70:37–43. <http://dx.doi.org/10.1016/j.jinf.2014.07.022>
- Chen M, Guo Q, Wang Y, Zou Y, Wang G, Zhang X, et al. Shifts in the antibiotic susceptibility, serogroups, and clonal complexes of *Neisseria meningitidis* in Shanghai, China: a time trend analysis of the pre-quinolone and quinolone eras. *PLoS Med*. 2015;12:e1001838. <http://dx.doi.org/10.1371/journal.pmed.1001838>
- Zhu B, Xu Z, Du P, Xu L, Sun X, Gao Y, et al. Sequence type 4821 clonal complex serogroup B *Neisseria meningitidis* in China, 1978–2013. *Emerg Infect Dis*. 2015;21:925–32. <http://dx.doi.org/10.3201/eid2106.140687>
- Harrison LH, Jolley KA, Shutt KA, Marsh JW, O’Leary M, Sanza LT, et al.; Maryland Emerging Infections Program. Antigenic shift and increased incidence of meningococcal disease. *J Infect Dis*. 2006;193:1266–74. <http://dx.doi.org/10.1086/501371>
- Thompson EA, Feavers IM, Maiden MC. Antigenic diversity of meningococcal enterobactin receptor FetA, a vaccine component. *Microbiology*. 2003;149:1849–58. <http://dx.doi.org/10.1099/mic.0.26131-0>
- Urwin R, Fox AJ, Musilek M, Kriz P, Maiden MC. Heterogeneity of the PorB protein in serotype 22 *Neisseria meningitidis*. *J Clin Microbiol*. 1998;36:3680–2.

12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455–77. <http://dx.doi.org/10.1089/cmb.2012.0021>
13. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9. <http://dx.doi.org/10.1093/bioinformatics/btu153>
14. Jolley KA, Maiden MC. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*. 2010;11:595. <http://dx.doi.org/10.1186/1471-2105-11-595>
15. Mustapha MM, Marsh JW, Krauland MG, Fernandez JO, de Lemos APS, Dunning Hotopp JC, et al. Genomic epidemiology of hypervirulent serogroup W, ST-11 *Neisseria meningitidis*. *EBioMedicine*. 2015;2:1447–55. <http://dx.doi.org/10.1016/j.ebiom.2015.09.007>
16. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30:1312–3. <http://dx.doi.org/10.1093/bioinformatics/btu033>
17. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*. 2010;5:e11147. <http://dx.doi.org/10.1371/journal.pone.0011147>
18. Parkhill J, Achtman M, James KD, Bentley SD, Churcher C, Klee SR, et al. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. *Nature*. 2000;404:502–6. <http://dx.doi.org/10.1038/35006655>
19. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011;28:2731–9. <http://dx.doi.org/10.1093/molbev/msr121>
20. Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLOS Comput Biol*. 2015;11:e1004041. <http://dx.doi.org/10.1371/journal.pcbi.1004041>
21. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res*. 2015;43:e15. <http://dx.doi.org/10.1093/nar/gku1196>
22. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31:3691–3. <http://dx.doi.org/10.1093/bioinformatics/btv421>
23. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. 2000;28:33–6. <http://dx.doi.org/10.1093/nar/28.1.33>
24. Harrison OB, Claus H, Jiang Y, Bennett JS, Bratcher HB, Jolley KA, et al. Description and nomenclature of *Neisseria meningitidis* capsule locus. *Emerg Infect Dis*. 2013;19:566–73. <http://dx.doi.org/10.3201/eid1904.111799>
25. Zhu P, van der Ende A, Falush D, Brieske N, Morelli G, Linz B, et al. Fit genotypes and escape variants of subgroup III *Neisseria meningitidis* during three pandemics of epidemic meningitis. *Proc Natl Acad Sci U S A*. 2001;98:5234–9. <http://dx.doi.org/10.1073/pnas.061386098>
26. Taha MK, Giorgini D, Ducos-Galand M, Alonso JM. Continuing diversification of *Neisseria meningitidis* W135 as a primary cause of meningococcal disease after emergence of the serogroup in 2000. *J Clin Microbiol*. 2004;42:4158–63. <http://dx.doi.org/10.1128/JCM.42.9.4158-4163.2004>
27. Lamelas A, Harris SR, Röltgen K, Dangy JP, Hauser J, Kingsley RA, et al. Emergence of a new epidemic *Neisseria meningitidis* serogroup A clone in the African meningitis belt: high-resolution picture of genomic changes that mediate immune evasion. *MBio*. 2014;5:e01974–14. <http://dx.doi.org/10.1128/mBio.01974-14>
28. Zhu B, Fan Y, Xu Z, Xu L, Du P, Gao Y, et al. Genetic diversity and clonal characteristics of ciprofloxacin-resistant meningococcal strains in China. *J Med Microbiol*. 2014;63:1411–8. <http://dx.doi.org/10.1099/jmm.0.078600-0>
29. Jolley KA, Wilson DJ, Kriz P, McVean G, Maiden MC. The influence of mutation, recombination, population history, and selection on patterns of genetic diversity in *Neisseria meningitidis*. *Mol Biol Evol*. 2005;22:562–9. <http://dx.doi.org/10.1093/molbev/msi041>
30. Didelot X, Maiden MC. Impact of recombination on bacterial evolution. *Trends Microbiol*. 2010;18:315–22. <http://dx.doi.org/10.1016/j.tim.2010.04.002>
31. Vos M, Didelot X. A comparison of homologous recombination rates in bacteria and archaea. *ISME J*. 2009;3:199–208. <http://dx.doi.org/10.1038/ismej.2008.93>
32. Budroni S, Siena E, Dunning Hotopp JC, Seib KL, Serruto D, Nofroni C, et al. *Neisseria meningitidis* is structured in clades associated with restriction modification systems that modulate homologous recombination. *Proc Natl Acad Sci U S A*. 2011;108:4494–9. <http://dx.doi.org/10.1073/pnas.1019751108>
33. Kong Y, Ma JH, Warren K, Tsang RS, Low DE, Jamieson FB, et al. Homologous recombination drives both sequence diversity and gene content variation in *Neisseria meningitidis*. *Genome Biol Evol*. 2013;5:1611–27. <http://dx.doi.org/10.1093/gbe/evt116>
34. He B, Jia Z, Zhou H, Wang Y, Jiang X, Ma H, et al. CC4821 serogroup W meningococcal disease in China. *Int J Infect Dis*. 2014;29:113–4. <http://dx.doi.org/10.1016/j.ijid.2014.08.022>
35. Lucidarme J, Lekshmi A, Parikh SR, Bray JE, Hill DM, Bratcher HB, et al. Frequent capsule switching in ‘ultra-virulent’ meningococci: are we ready for a serogroup B ST-11 complex outbreak? *J Infect*. 2017;75:95–103. <http://dx.doi.org/10.1016/j.jinf.2017.05.015>
36. Mustapha MM, Marsh JW, Krauland MG, Fernandez JO, de Lemos AP, Dunning Hotopp JC, et al. Genomic investigation reveals highly conserved, mosaic, recombination events associated with capsular switching among invasive *Neisseria meningitidis* serogroup W sequence type (ST)-11 strains. *Genome Biol Evol*. 2016;8:2065–75. <http://dx.doi.org/10.1093/gbe/evw122>
37. Chiou CS, Liao JC, Liao TL, Li CC, Chou CY, Chang HL, et al. Molecular epidemiology and emergence of worldwide epidemic clones of *Neisseria meningitidis* in Taiwan. *BMC Infect Dis*. 2006;6:25. <http://dx.doi.org/10.1186/1471-2334-6-25>
38. Nunes AM, Ribeiro GS, Ferreira IE, Moura AR, Felzemburgh RD, de Lemos AP, et al. Meningococcal carriage among adolescents after mass meningococcal c conjugate vaccination campaigns in Salvador, Brazil. *PLoS One*. 2016;11:e0166475. <http://dx.doi.org/10.1371/journal.pone.0166475>
39. Moura ARSS, Kretz CB, Ferreira IE, Nunes AMPB, de Moraes JC, Reis MG, et al. Molecular characterization of *Neisseria meningitidis* isolates recovered from 11–19-year-old meningococcal carriers in Salvador, Brazil. *PLoS One*. 2017;12:e0185038. <http://dx.doi.org/10.1371/journal.pone.0185038>
40. Kawasaki Y, Matsubara K, Takahashi H, Morita M, Ohmishi M, Hori M, et al. Invasive meningococcal disease due to ciprofloxacin-resistant *Neisseria meningitidis* sequence type 4821: the first case in Japan. *J Infect Chemother*. 2017 Dec 7:S1341–321X(17)30274-X [Epub ahead of print].

Address for correspondence: Min Chen, Shanghai Municipal Center for Disease Control and Prevention, 1380 W ZhongShan Rd, Shanghai, China; email: chenmin@scdc.sh.cn; Yohei Doi, Division of Infectious Diseases, University of Pittsburgh School of Medicine, S829 Scaife Hall, 3550 Terrace St, Pittsburgh, PA 15261, USA; email: yod4@pitt.edu; Minggui Wang, Institute of Antibiotics, Huashan Hospital, Fudan University, 12 M Wulumuqi Rd, Shanghai 200040, China; email: mgwang@fudan.edu.cn

Evolution of Sequence Type 4821 Clonal Complex Meningococcal Strains in China from Prequinolone to Quinolone Era, 1972–2013

Technical Appendix

Recombination

Recombination was assessed using maximum-likelihood inference under the ClonalFrame model of evolution implemented by ClonalFrameML (1). Aligned core genome sequences and a RAxML maximum-likelihood phylogenetic tree were used as input for ClonalFrameML under the standard model with 100 simulations (-emsim 100). Additional assessment of recombination was done using Gubbins (2) and compared to ClonalFrameML findings for consistency. The relative impact of recombination to mutation (r/m) was calculated from ClonalFrameML output using the formula $r/m = (R/\theta) \times \delta \times v$ where R/θ is the ratio of recombination to mutation rates, δ is the mean length of imports and v , the average distance of the imports as described (1).

The output from ClonalFrameML provided detailed breakdown of inferred recombinant sequences for every node of the core genome maximum-likelihood phylogenetic tree. Selected nodes, representing the divergence of group 1 from group 2 CC4821 lineages; and divergence of epidemic strains with P1.7–2 antigenic formula from within group 1 were selected for more in-depth analyses. Inferred recombinant sequences were mapped on to the reference genome 053442, using Blastn and custom Unix shell scripts to identify corresponding gene annotations.

Recombinations within these segments were further demonstrated using sequence alignments and focused phylogenetic analyses using corresponding sequences from several global reference strains for comparison.

References

1. Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. PLOS Comput Biol. 2015;11:e1004041. [PubMed](#)
<http://dx.doi.org/10.1371/journal.pcbi.1004041>
2. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015;43:e15. [PubMed](#) <http://dx.doi.org/10.1093/nar/gku1196>

Technical Appendix Table 1. Epidemiologic and genetic characteristics of 32 *N. meningitidis* isolates belonging to clonal complex 4821 lineage in China^{abc}

PubMLST									<i>porA</i>	<i>porB</i>	<i>fetA</i>	<i>nhba</i>	FHbp	<i>gyrA</i>
ID	Isolate name	Year	Sg	PorA_VR1	PorA_VR2	<i>porB</i>	FetA_VR	(NEIS1364)	(NEIS2020)	(NEIS1963)	(NEIS2109)	<i>fhbp</i>	variant ^f	allele
57871	NM330505	2005	C	7-2	14	3-48	F3-3	ND	29	64	124	498	1.80/B	71 ^d
58133	NM001	2005	C	7-2	14	3-48	F3-3	56	29	64	124	498	1.80/B	71 ^d
12672	053442	2004	C	7-2	14	3-48	F3-3	7	29	64	124	498	1.80/B	71 ^d
57851	NM220601	2006	C	7-2	14	3-48	F3-3	1029	29	64	124	498	1.80/B	71 ^d
57872	NM341215	2011	B	7-2	14	3-48	F3-3	ND	29	ND	124	498	1.80/B	71 ^d
57869	NM320501	2005	C	7-2	14	3-48	F3-3	1030	29	64	124	323	1.276/B153	71 ^d
58134	NM062	2011	C	7-2	14	3-48	F3-3	56	29	64	124	22	2.22/A10	71 ^d
57862	NM420718	2007	C	12-1	16-8	3-48	F3-3	1031	29	64	124	489	1.419/B	71 ^d
41444	NM11003	2011	C	12-1	16-8	3-48	F3-3	1031	29	64	124	489	1.419/B	71 ^d
57853	NM100603	2005	C	12-1	16-8	3-48	F3-3	1031	29	64	124	489	1.419/B	71 ^d
57863	NM100601	2005	C	12-1	16-8	3-48	F3-3	1031	29	64	124	489	1.419/B	71 ^d
57852	NM440902	2008	B	20	23-1	3-48	F3-3	1026	29	64	124	22	2.22/A10	71 ^d
57870	NM320503	2005	C	20	23-9	3-48	F3-3	ND	29	64	124	489	1.419/B	71 ^d
57855	NM130803	2008	C	20	23-1	3-48	F3-3	1026	29	64	124	22	2.22/A10	71 ^d
57854	NM131148	2011	C	20	23-28	3-48	F3-3	1032	29	64	124	22	2.22/A10	71 ^d
57858	NM100572	2005	C	20	23	UA	F2-9	640	UA	2612	74	16	2.16/A19	12
57856	NM340552	2005	B	21-2	28	3-460	F2-9	UA	1275	2612	74	39	2.18/A	12
57857	NM34173	2011	B	20	23-18	3-81	F1-33	ND	43	2610	566	21	2.21/A07	171 ^d
57861	NM421007	2009	B	12-11	16-8	3-84	F5-7	1035	150	2191	553	16	2.16/A19	12
57860	NM100514	2005	C	20	23-1	3-458	UA	1034	UA	2620	UA	UA	2.16/A19	203 ^d
58135	NM193	1972	C	20-3	23-1	3-607	F1-5	1025	1273	2615	966	474	2.404/A	12
57859	NM440501	2005	C	20-3	23	UA	F5-70	1033	UA	2617	966	490	1.420/B	12
58136	NM205	1973	C	20	23-2	3-608	F5-135	643	1274	2616	553	16	2.16/A19	4
58137	NM313	1977	C	20	23-1	3-35	F14-1	1026	44	ND	553	16	2.16/A19	12
57864	NM360624	2006	C	UA	UA	3-48	F4-2	ND	29	255	UA	16	2.16/A19	114 ^d
58132	NM323	1977	B	20	23	3-229	F3-9	640	265	1465	553	16	2.16/A19	12
58130	NM040	2007	B	20	13-1	3-370	F5-2	1027	1270	2614	553	16	2.16/A19	23
57865	NM311112	2011	B	20	13-1	3-229	F5-2	1036	265	2614	553	16	2.16/A19	12
58131	NM050	2009	B	20	13-2	3-229	F5-2	1028	265	2614	553	16	2.16/A19	71 ^d
57866	NM100530	2005	B	20-3	23	3-229	F3-9	1033	265	1465	1346	16	2.16/A19	173 ^e
57868	NM370601	2005	C	20	23	UA	F3-9	640	UA	1465	UA	637	1.547/B	UA ^d
57867	NM321114	2011	B	20	23-3	3-229	F1-124	1037	265	2619	1368	677	1.547/B	175 ^d

^a *nadA* is missing in all of the 32 strains.

^b Light pink, lavender blush and light cyan indicated strains belonged to the epidemic clone, non-epidemic clones in phylogenetic group 1 and group 2, respectively.

^c Abbreviations: ND, not detected or incomplete sequence; Sg, serogroup; UA, unassigned; X, missing sequence data.

^d Amino acid mutation at position 91, T91I.

^e Amino acid mutation at position 91, T91V.

PubMLST ID	Isolate name	Year	Sg	PorA_VR1	PorA_VR2	'porB	FetA_VR	porA (NEIS1364)	porB (NEIS2020)	fetA (NEIS1963)	nhba (NEIS2109)	'fHbp variant ^f	FHbp allele	gyrA allele
------------	--------------	------	----	----------	----------	-------	---------	--------------------	--------------------	--------------------	--------------------	-------------------------------	----------------	----------------

^f FHbp variant: Novartis subvariant/Pfizer subfamily or subvariant. There are three major variants (variant 1, 2 and 3) in Novartis classification for FHbp proteins, named with a major variant number and a subvariant number (e.g., 1.80). Pfizer classifies FHbp into two subfamilies, A and B, with assigned number indicating subvariant (e.g., A10).

Technical Appendix Table 1 (continued)

PubMLST ID	Isolate name	Sg	ctrG	cssE	csc	csb ^e	cssC	cssB	cssA	ctrA	ctrB	ctrC	ctrD	tex(NEIS0059)	orf1(NEIS2910)	orf2(NEIS2854)	ctrE	ctrF
57871	NM330505	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	3	3
58133	NM001	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	3	3
12672	053442	C	4	1	1	X	3	1	3	1	1	4	1	4	1	1	3	3
57851	NM220601	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	3	3
57872	NM341215	B	13	X	X	325	2	13	3	263	1	4	1	4	1	1	3	3
57869	NM320501	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	3	3
58134	NM062	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	770	16
57862	NM420718	C	4	1	UA	X	3	1	3	5	1	4	1	4	1	1	770	16
41444	NM11003	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	770	16
57853	NM100603	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	770	16
57863	NM100601	C	4	1	1	X	3	1	613	5	1	4	1	4	1	1	770	16
57852	NM440902	B	5	X	X	1	205	3	3	271	18	7	1	4	1	1	770	16
57870	NM320503	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	770	16
57855	NM130803	C	4	1	9	X	3	1	3	5	1	4	1	4	1	1	898	533
57854	NM131148	C	4	1	1	X	3	1	3	5	1	4	1	4	1	1	827	532
57858	NM100572	C	12	1	UA	X	2	1	615	272	208	69	19	1087	1	1	151	535
57856	NM340552	B	5	X	X	1	2	1	148	5	1	32	225	1086	1	1	338	16
57857	NM34173	B	229	X	X	UA	2	1	614	5	1	4	1	1086	1	1	828	534
57861	NM421007	B	13	X	X	276	2	1	2	5	2	4	170	1090	1	UA	829	ND
57860	NM100514	C	173	1	1	X	26	1	616	10	209	3	1	1089	1	1	338	506
58135	NM193	C	173	1	1	X	2	1	612	5	2	7	18	1085	1	1	338	16
57859	NM440501	C	173	1	UA	X	2	1	12	5	39	13	45	1088	1	1	338	16
58136	NM205	C	173	1	1	X	203	1	95	5	2	7	18	62	1	1	338	16
58137	NM313	C	173	1	1	X	2	1	2	5	2	7	18	62	1	1	826	16
57864	NM360624	C	173	1	1	X	2	1	3	263	2	12	1	1091	87	1	830	536
58132	NM323	B	237	X	X	308	204	13	2	5	2	7	18	62	1	1	338	16
58130	NM040	B	5	X	X	108	2	13	3	5	2	7	18	62	1	1	338	16
57865	NM311112	B	5	X	X	320	2	13	2	5	2	7	18	62	1	1	338	16
58131	NM050	B	5	X	X	1	285	1	12	5	2	111	18	62	1	1	338	16
57866	NM100530	B	5	X	X	1	2	1	313	273	18	7	226	1092	1	1	338	16
57868	NM370601	C	173	11	1	X	2	13	3	2	210	3	7	1094	1	1	1175	537

PubMLST ID	Isolate name	Sg	ctrG	cssE	csc	csb ^e	cssC	cssB	cssA	ctrA	ctrB	ctrC	ctrD	tex(NEIS0059)	orf1(NEIS2910)	orf2(NEIS2854)	ctrE	ctrF
57867	NM321114	B	5	X	X	108	2	13	3	136	2	12	108	1093	1	1	338	16

^a *nadA* is missing in all of the 32 strains.

^b Light pink, lavender blush and light cyan indicated strains belonged to the epidemic clone, non-epidemic clones in phylogenetic group 1 and group 2, respectively.

^c Abbreviations: ND, not detected or incomplete sequence; Sg, serogroup; UA, unassigned; X, missing sequence data.

^d Amino acid mutation at position 91, T91I.

^e Amino acid mutation at position 91, T91V.

^f FHbp variant: Novartis subvariant/Pfizer subfamily or subvariant. There are three major variants (variant 1, 2 and 3) in Novartis classification for FHbp proteins, named with a major variant number and a subvariant number (e.g., 1.80). Pfizer classifies FHbp into two subfamilies, A and B, with assigned number indicating subvariant (e.g., A10).

Technical Appendix Table 2. Epidemiologic and genetic characteristics of 52 CC4821 *N. meningitidis* isolates in Shanghai^{abc}

Isolate															
name	Year	Source	Sg	ST	PFGE type	PorA_VR1	PorA_VR2	'porB	FetA_VR	NHBA-VR	nhba	fHbp	FHbp variant ^d	gyrA allele	
NM001	2005	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM002	2006	Contact	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM010	2006	Invasive	C	ST-9482	A	7-2	14	3-48	F3-3	503	124	512	1.437/B	71(R1)	
NM029	2007	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM033	2007	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM045	2008	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM048	2008	Contact	B	ST-9586	A	7-2	14	3-48	F3-3	503	124	22	2.22/A10	71(R1)	
NM053	2009	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM061	2011	Contact	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	22	2.22/A10	71(R1)	
NM062	2011	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	22	2.22/A10	71(R1)	
NM119	2011	Invasive	B	ST-9455	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM044	2008	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	22	2.22/A10	71(R1)	
NM055	2010	Invasive	C	ST-4821	A	7-2	14	3-48	F3-3	503	124	498	1.80/B	71(R1)	
NM015	2006	Invasive	C	ST-4821	A	7-2	14	3-48	F14-1	503	124	498	1.80/B	71(R1)	
NM032	2007	Contact	NG	ST-4821	A	7-2	14	3-48	NP	503	124	22	2.22/A10	71(R1)	
NM046	2008	Invasive	C	ST-4821	A	7-2	14	3-606	F3-3	503	124	498	1.80/B	71(R1)	
NM042	2007	Invasive	C	ST-9484	A	12-1	16-8	3-48	F3-3	503	124	636	1.546/B	71(R1)	

Isolate														
name	Year	Source	Sg	ST	PFGE type	PorA_VR1	PorA_VR2	'porB	FetA_VR	NHBA-VR	nhba	fHbp	FHbp variant ^d	gyrA allele
NM047	2008	Invasive	C	ST-4821	A	12-1	16-8	3-48	F3-3	503	124	489	1.419/B	71(R1)
NM049	2008	Invasive	C	ST-4821	A	12-1	16-8	3-48	F3-3	503	124	489	1.419/B	71(R1)
NM374	2013	Invasive	C	ST-4821	A	12-1	16-8	3-48	F3-3	503	124	489	1.419/B	71(R1)
NM014	2006	Invasive	C	ST-4821	A	20	23-7	3-48	F3-3	503	124	489	1.419/B	71(R1)
NM039	2007	Invasive	C	ST-4821	A	20	23-7	3-48	F3-3	503	124	489	1.419/B	71(R1)
NM025	2007	Carrier	B	ST-4821	A	20	23-2	3-48	F3-3	503	124	22	2.22/A10	71(R1)
NM023	2007	Carrier	B	ST-9585	B	18	23-6	3-1	F2-9	73	74	16	2.16/A19	12(S1)
NM323	1977	Carrier	B	ST-5798	C	20	23	3-229	F3-9	669	553	16	2.16/A19	12(S1)
NM040	2007	Invasive	B	ST-5798	C	20	13-1	3-370	F5-2	669	553	16	2.16/A19	23(S2)
NM059	2010	Carrier	B	ST-5798	C	20	13-1	3-229	F5-2	669	553	16	2.16/A19	12(S1)
NM050	2009	Contact	B	ST-5798	C	20	13-2	3-229	F5-2	669	553	16	2.16/A19	71(R1)
NM063	2011	Invasive	B	ST-5798	C	20	13-1	3-229	F5-2	669	553	16	2.16/A19	12(S1)
NM200	2011	Carrier	B	ST-5664	D	20	23-9	3-81	F1-112	910	967	16	2.16/A19	12(S1)
NM021	2007	Carrier	B	ST-5664	E	20	23-9	3-81	F1-91	688	965	1207	3.94/A	12(S1)
NM064	2011	Invasive	B	ST-9454	E	20	23-3	3-352	F1-91	688	965	322	1.275/B	113(R13)
NM312	1977	Carrier	C	ST-9591	F	20	23-3	3-35	F4-1	916	973	16	2.16/A19	12(S1)
NM313	1977	Carrier	C	ST-4821	F	20	23-1	3-35	F14-1	669	553	16	2.16/A19	12(S1)
NM325	1977	Carrier	C	ST-5081	F	20	23-7	3-460	F1-7	917	974	16	2.16/A19	12(S1)
NM322	1977	Carrier	C	ST-3436	F	20-3	23-7	3-460	F5-8	697	966	474	2.404/A	12(S1)
NM277	1977	Invasive	C	ST-3436	F	20-3	23-7	3-460	F1-140	697	966	474	2.404/A	12(S1)
NM278	1977	Invasive	C	ST-3436	F	20-3	23-1	3-460	F1-180	697	966	474	2.404/A	12(S1)
NM297	1977	Carrier	B	ST-4821	F	20-1	23-7	3-460	F5-138	669	553	16	2.16/A19	12(S1)
NM193	1972	Carrier	C	ST-3436	F	20-3	23-1	3-607	F1-5	697	966	474	2.404/A	12(S1)
NM213	1976	Carrier	B	ST-4821	F	20	23-1	3-460	F2-30	669	553	16	2.16/A19	12(S1)
NM220	1976	Carrier	B	ST-3436	F	20-3	23-7	3-460	F5-8	697	966	474	2.404/A	12(S1)
NM221	1976	Carrier	B	ST-3436	F	20-3	23-7	3-460	F5-8	697	966	474	2.404/A	12(S1)

Isolate														
name	Year	Source	Sg	ST	PFGE type	PorA_VR1	PorA_VR2	'porB	FetA_VR	NHBA-VR	nhba	fHbp	FHbp variant ^d	gyrA allele
NM268	1977	Carrier	B	ST-4821	F	20	23-3	3-460	F5-172	915	972	16	2.16/A19	12(S1)
NM319	1977	Carrier	B	ST-3436	F	20-3	23-9	3-609	F5-8	697	966	474	2.404/A	12(S1)
NM205	1973	Carrier	C	ST-4821	F	20	23-2	3-608	F5-135	669	553	16	2.16/A19	4(S2)
NM267	1977	Carrier	B	ST-5798	F	20	23-1	3-48	F5-2	669	553	16	2.16/A19	12(S1)
NM290	1977	Invasive	C	ST-3436	F	20-3	23-3	3-229	F5-8	697	966	474	2.404/A	12(S1)
NM075	2013	Invasive	B	ST-10051	G	20	23-1	3-25	F1-91	9	7	5	1.5/B22	103(R10)
NM038	2007	Contact	B	ST-9502	H	18	25-11	3-229	F3-9	669	553	691	2.585/A	109(R11)
NM282	1977	Carrier	B	ST-3200	H	20	23	3-229	F3-9	669	553	16	2.16/A19	12(S1)
NM057	2005	Invasive	C	ST-4821	I	12-6	2-35	2-142	F1-8	20	808	39	2.18/A	4(S2)

^a *nadA* is missing in all of the CC4821 strains.

^b Light pink, lavender blush and light cyan indicated strains belonged to the epidemic clone, non-epidemic clones in phylogenetic group 1 and group 2, respectively, based on the similarity of antigen genes.

^c Abbreviations: NG, Nongroupable; NP, negative for PCR; Sg, serogroup; ST, sequence type.

^d FHbp variant: Novartis subvariant/Pfizer subfamily or subvariant. There are three major variants (variant 1, 2 and 3) in Novartis classification for FHbp proteins, named with a major variant number and a subvariant number (e.g., 1.80). Pfizer classifies FHbp into two subfamilies, A and B, with assigned number indicating subvariant (e.g., A10).

Technical Appendix Table 3. List of recombinant genes linked to the divergence of group 1 CC4821 sublineage.

Gene ID	Gene name	Annotation	COG pathway	COG process
NMCC_0117		putative FAD-dependent oxidoreductase	Amino acid transport and metabolism	Metabolism
NMCC_0585	<i>argH</i>	argininosuccinate lyase (arginosuccinase; ASAL)	Amino acid transport and metabolism	Metabolism
NMCC_0658	<i>trpB</i>	tryptophan synthase β chain	Amino acid transport and metabolism	Metabolism
NMCC_0883	<i>metX</i>	homoserine O-acetyltransferase (homoserine O-trans-acetylase; homoserine transacetylase; HTA)	Amino acid transport and metabolism	Metabolism
NMCC_0886	<i>metF</i>	5,10-methylenetetrahydrofolate reductase	Amino acid transport and metabolism	Metabolism
NMCC_0887	<i>metE</i>	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase (methionine synthase, vitamin-B12 independent isoz	Amino acid transport and metabolism	Metabolism
NMCC_1321	<i>pepN</i>	aminopeptidase N (α -aminoacylpeptide hydrolase)	Amino acid transport and metabolism	Metabolism
NMCC_1357	<i>aroD</i>	3-dehydroquinate dehydratase (3-dehydroquinase; type I DHQase)	Amino acid transport and metabolism	Metabolism
NMCC_1540	<i>serC</i>	phosphoserine aminotransferase (PSAT)	Amino acid transport and metabolism	Metabolism
NMCC_1547		putative amino-acid symporter	Amino acid transport and metabolism	Metabolism
NMCC_0140	<i>ptsIIA</i>	phosphotransferase system, enzyme IIA (protein IIA)	Carbohydrate transport and metabolism	Metabolism
NMCC_0141	<i>ptsH</i>	phosphocarrier protein HPr (phosphotransferase system, histidine-containing protein)	Carbohydrate transport and metabolism	Metabolism
NMCC_0403	<i>pglD</i>	pilin glycosylation protein PglD	Carbohydrate transport and metabolism	Metabolism
NMCC_0107	<i>folD</i>	FolD bifunctional protein (PPDC) [includes: methylenetetrahydrofolate dehydrogenase and methenyltetrahydrofolate cyclohy	Coenzyme transport and metabolism	Metabolism
NMCC_0114	<i>thiG</i>	thiazole biosynthesis protein ThiG	Coenzyme transport and metabolism	Metabolism
NMCC_0115		putative ThiS-like protein	Coenzyme transport and metabolism	Metabolism
NMCC_0116	<i>thiE</i>	thiamine-phosphate pyrophosphorylase (TMP pyrophosphorylase; TMP-PPase; thiamine-phosphate synthase)	Coenzyme transport and metabolism	Metabolism
NMCC_0125		putative NAD/FAD binding protein	Coenzyme transport and metabolism	Metabolism
NMCC_1271	<i>pdxH</i>	pyridoxamine 5'-phosphate oxidase (PNP/PMP oxidase; PNPOx)	Coenzyme transport and metabolism	Metabolism
NMCC_1453		lactoferrin binding protein B (LbpB)	Coenzyme transport and metabolism	Metabolism
NMCC_1618		truncated putative Tat-translocated NosX-like protein (pseudogene part 2)	Coenzyme transport and metabolism	Metabolism

Gene ID	Gene name	Annotation	COG pathway	COG process
NMCC_2063		putative 5-formyltetrahydrofolate cyclo-ligase (methenyl-THF synthetase)	Coenzyme transport and metabolism	Metabolism
NMCC_0124	<i>ppc</i>	phosphoenolpyruvate carboxylase (PEPCase; PEPC)	Energy production and conversion	Metabolism
NMCC_0126	<i>gpsA</i>	glycerol-3-phosphate dehydrogenase [NAD(P)+] ((NAD(P)H-dependent glycerol-3-phosphate dehydrogenase)	Energy production and conversion	Metabolism
NMCC_0133	<i>petA</i>	ubiquinol-cytochrome c reductase iron-sulfur subunit (Rieske iron-sulfur protein; RISP)	Energy production and conversion	Metabolism
NMCC_0134	<i>petB</i>	cytochrome b	Energy production and conversion	Metabolism
NMCC_0135	<i>petC</i>	cytochrome c1	Energy production and conversion	Metabolism
NMCC_0378		putative membrane-associated thioredoxin	Energy production and conversion	Metabolism
NMCC_0942		putative oxidoreductase	Energy production and conversion	Metabolism
NMCC_1270		putative iron/sulfur binding oxidoreductase	Energy production and conversion	Metabolism
NMCC_0035		conserved hypothetical integral membrane protein	Inorganic ion transport and metabolism	Metabolism
NMCC_0036		conserved hypothetical lipoprotein	Inorganic ion transport and metabolism	Metabolism
NMCC_0215	<i>fetB</i>	enterobactin uptake system binding lipoprotein FetB	Inorganic ion transport and metabolism	Metabolism
NMCC_0841		sulfate/thiosulfate import ATP binding protein CysA (sulfate-transporting ATPase)	Inorganic ion transport and metabolism	Metabolism
NMCC_0842		sulfate transport system permease protein CysW	Inorganic ion transport and metabolism	Metabolism
NMCC_1452	<i>lbpA</i>	lactoferrin binding protein A (LbpA)	Inorganic ion transport and metabolism	Metabolism
NMCC_1531	<i>norB</i>	nitric oxide reductase	Inorganic ion transport and metabolism	Metabolism
NMCC_0301	<i>fabD</i>	malonyl CoA-acyl carrier protein transacylase (MCT)	Lipid transport and metabolism	Metabolism
NMCC_0303	<i>fabH</i>	3-oxoacyl-[acyl-carrier-protein] synthase III (β -ketoacyl-ACP synthase III; KAS III)	Lipid transport and metabolism	Metabolism
NMCC_0307	<i>plsX</i>	fatty acid/phospholipid synthesis protein PlsX	Lipid transport and metabolism	Metabolism
NMCC_0361	<i>accB</i>	biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP)	Lipid transport and metabolism	Metabolism
NMCC_0386		putative inner-membrane acyltransferase	Lipid transport and metabolism	Metabolism
NMCC_1457	<i>fadD</i>	long-chain-fatty-acid-CoA-ligase (long-chain acyl-CoA synthetase)	Lipid transport and metabolism	Metabolism
NMCC_0118		putative transporter	Nucleotide transport and metabolism	Metabolism

Gene ID	Gene name	Annotation	COG pathway	COG process
NMCC_0299	<i>guaA</i>	GMP synthase [glutamine-hydrolyzing] (glutamine amidotransferase; GMP synthetase)	Nucleotide transport and metabolism	Metabolism
NMCC_0383	<i>fts</i>	formate-tetrahydrofolate ligase (formyltetrahydrofolate synthetase; FHS; FTHFS)	Nucleotide transport and metabolism	Metabolism
NMCC_1456	<i>pyrG</i>	CTP synthase (UTP-ammonia ligase; CTP synthetase)	Nucleotide transport and metabolism	Metabolism
NMCC_1620	<i>thyA</i>	thymidylate synthase (TS; TSase)	Nucleotide transport and metabolism	Metabolism
NMCC_2067	<i>pyrH</i>	uridylate kinase (UK; uridine monophosphate kinase; UMP kinase)	Nucleotide transport and metabolism	Metabolism
NMCC_0882		putative methionine biosynthesis protein MetW	Secondary metabolites biosynthesis, transport and catabolism	Metabolism
NMCC_1874		putative methyltransferase	Secondary metabolites biosynthesis, transport and catabolism	Metabolism
NMCC_0136		putative RmuC-like protein	Replication, recombination and repair	Information storage and processing
NMCC_0398	<i>dnaE</i>	DNA polymerase III α subunit	Replication, recombination and repair	Information storage and processing
NMCC_0619	<i>ligA</i>	DNA ligase (polydeoxyribonucleotide synthase [NAD+])	Replication, recombination and repair	Information storage and processing
NMCC_0657		putative DNA glycosylase	Replication, recombination and repair	Information storage and processing
NMCC_0688	<i>ihfA</i>	integration host factor α -subunit (IHF- α)	Replication, recombination and repair	Information storage and processing
NMCC_0735		putative deoxyribonuclease	Replication, recombination and repair	Information storage and processing
NMCC_1114	<i>recD</i>	exodeoxyribonuclease V α chain	Replication, recombination and repair	Information storage and processing
NMCC_1322		putative polynucleotidyl transferase	Replication, recombination and repair	Information storage and processing

Gene ID	Gene name	Annotation	COG pathway	COG process
NMCC_1325	<i>ruvC</i>	crossover junction endodeoxyribonuclease RuvC (Holliday junction nuclease RuvC; Holliday junction resolvase RuvC)	Replication, recombination and repair	Information storage and processing
NMCC_1871	<i>recQ</i>	ATP-dependent DNA helicase RecQ	Replication, recombination and repair	Information storage and processing
NMCC_1326		putative Fis-like DNA binding protein	Transcription	Information storage and processing
NMCC_0003	<i>gltX</i>	glutamyl-tRNA synthetase (Glutamate-tRNA ligase; GluRS)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0031	<i>metG</i>	methionyl-tRNA synthetase (methionine-tRNA ligase; MetRS)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0363	<i>queA</i>	S-adenosylmethionine:tRNA ribosyltransferase-isomerase (queuosine biosynthesis protein QueA)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0384		putative GTP-dependent nucleic acid binding protein EngD	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0387	<i>tyrS</i>	tyrosyl-tRNA synthetase (tyrosine-tRNA ligase; TyrRS)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0656	<i>ksgA</i>	dimethyladenosine transferase (S-adenosylmethionine-6-N', N'- adenosyl(rRNA) dimethyltransferase; 16S rRNA dimethylase; h	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0687	<i>pheT</i>	phenylalanyl-tRNA synthetase β chain (phenylalanine-tRNA ligase β chain; PheRS)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0880	<i>efp</i>	elongation factor P (EF-P)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0884		50S ribosomal protein L36	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0885	<i>rpmE</i>	50S ribosomal protein L31 type B	Translation, ribosomal structure and biogenesis	Information storage and processing

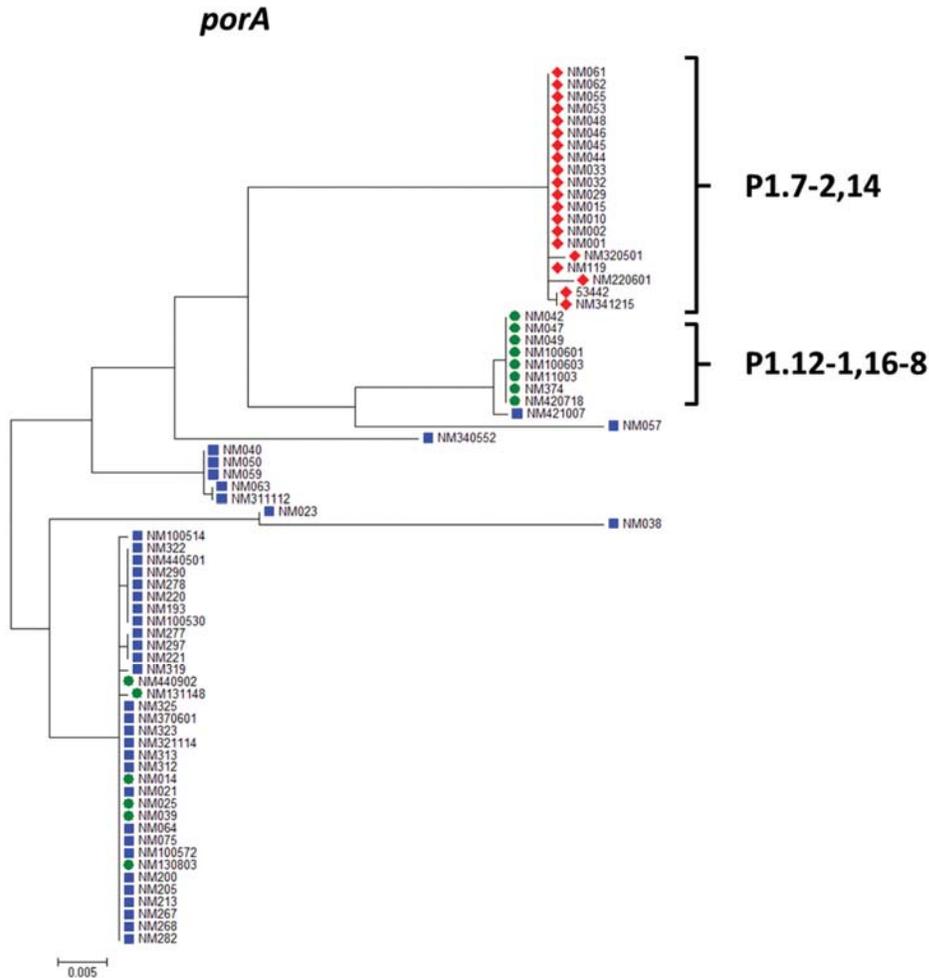
Gene ID	Gene name	Annotation	COG pathway	COG process
NMCC_1267	<i>gatA</i>	aspartyl-tRNA(Asn) amidotransferase subunit A (Asp-ADT subunit A)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_1269	<i>gatB</i>	aspartyl-tRNA(Asn) amidotransferase subunit B (Asp-ADT subunit B)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_1327	<i>dusB</i>	tRNA-dihydrouridine synthase B	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_1458	<i>trmU</i>	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_2065	<i>rpsB</i>	30S ribosomal protein S2	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_2066	<i>tsf</i>	elongation factor Ts (EF-Ts)	Translation, ribosomal structure and biogenesis	Information storage and processing
NMCC_0376		putative ATP binding protein	Cell cycle control, cell division, chromosome partitioning	Cellular processes and signaling
NMCC_0659		IgA-specific serine endopeptidase (IgA protease)	Cell cycle control, cell division, chromosome partitioning	Cellular processes and signaling
NMCC_0985	<i>ftsK</i>	DNA translocase FtsK2	Cell cycle control, cell division, chromosome partitioning	Cellular processes and signaling
NMCC_0417	<i>pilP</i>	type IV pilus biogenesis lipoprotein PilP	Cell motility	Cellular processes and signaling
NMCC_0418	<i>pilO</i>	type IV pilus biogenesis protein PilO	Cell motility	Cellular processes and signaling
NMCC_0419	<i>pilN</i>	type IV pilus biogenesis protein PilN	Cell motility	Cellular processes and signaling
NMCC_0032	<i>glmS</i>	glucosamine-fructose-6-phosphate aminotransferase [isomerizing] (hexosephosphate aminotransferase; D-fructose-6-phosph	Cell wall/membrane/envelope biogenesis	Cellular processes and signaling

Gene ID	Gene name	Annotation	COG pathway	COG process
NMCC_0034	<i>mItA</i>	membrane-bound lytic murein transglycosylase A (murein hydrolase A; GNA33)	Cell wall/membrane/envelope biogenesis	Cellular processes and signaling
NMCC_1324		putative lipid A biosynthesis (KDO)2-(lauroyl)-lipid IVA acyltransferase	Cell wall/membrane/envelope biogenesis	Cellular processes and signaling
NMCC_1616	<i>rfaK</i>	α 1,2 N-acetylglucosamine transferase	Cell wall/membrane/envelope biogenesis	Cellular processes and signaling
NMCC_1685	<i>murl</i>	glutamate racemase	Cell wall/membrane/envelope biogenesis	Cellular processes and signaling
NMCC_0300	<i>msbA</i>	lipid A export ATP binding/permease protein MsbA	Defense mechanisms	Cellular processes and signaling
NMCC_1115	<i>lolD</i>	lipoprotein-releasing system ATP binding protein LolD	Defense mechanisms	Cellular processes and signaling
NMCC_0448	<i>tpsB</i>	TpsA activation/secretion protein TpsB	Intracellular trafficking, secretion, and vesicular transport	Cellular processes and signaling
NMCC_0350		putative peptidase	Posttranslational modification, protein turnover, chaperones	Cellular processes and signaling
NMCC_0843	<i>cysT</i>	sulfate transport system permease protein CysT	Posttranslational modification, protein turnover, chaperones	Cellular processes and signaling
NEIMB_0895		hypothetical protein		
NMCC_0033		hypothetical protein		
NMCC_0304		hypothetical protein		
NMCC_0305				
NMCC_0306		hypothetical protein		
NMCC_0442		Cupin superfamily protein		
NMCC_0686		hypothetical protein		
NMCC_0689		hypothetical protein		
NMCC_0878		hypothetical protein		

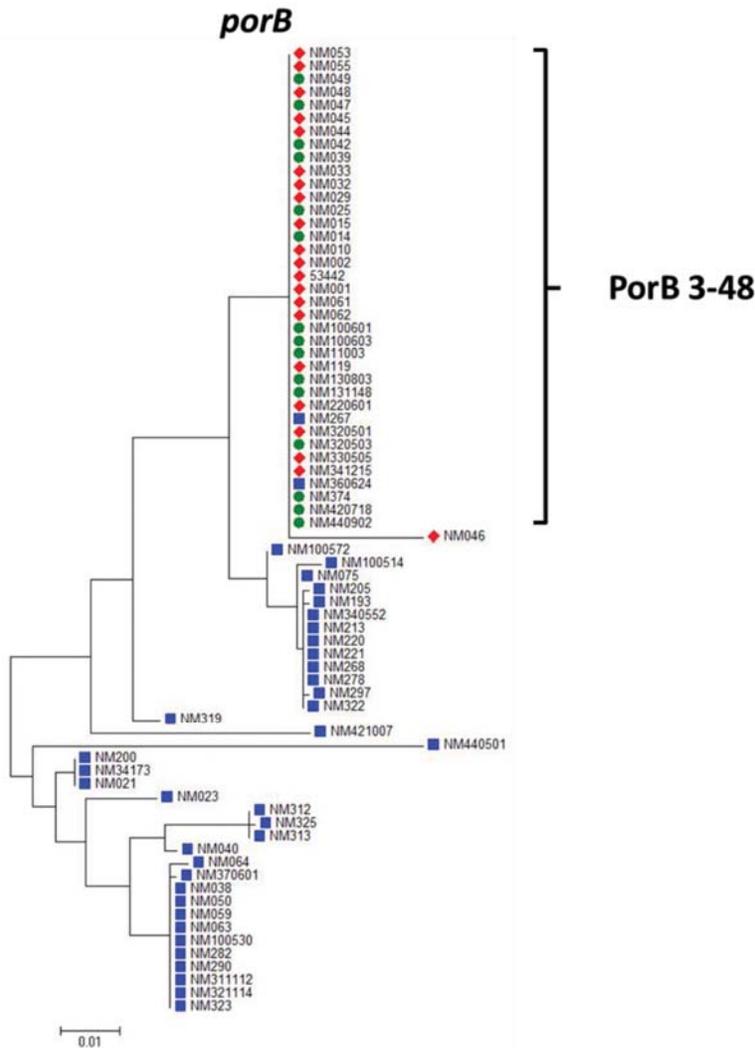
Gene ID	Gene name	Annotation	COG pathway	COG process
NMCC_0881		hypothetical protein		
NMCC_1539		hypothetical protein		
NMCC_1619		hypothetical protein		
NMCC_1872		hypothetical protein		
NMCC_1873		hypothetical integral membrane protein		
NMCC_2064		hypothetical protein		
NMCC_0302		conserved hypothetical integral membrane protein	Function unknown	Poorly characterized
NMCC_0879		conserved hypothetical protein	Function unknown	Poorly characterized
NMCC_1268		conserved hypothetical protein	Function unknown	Poorly characterized
NMCC_1541		conserved hypothetical protein	Function unknown	Poorly characterized
NMCC_1548		conserved hypothetical protein	Function unknown	Poorly characterized
NMCC_0369		putative phosphoribosyltransferase	General function prediction only	Poorly characterized
NMCC_0690		putative FxsA-like protein	General function prediction only	Poorly characterized
NMCC_1538		putative sulfatase	General function prediction only	Poorly characterized
NMCC_1617		putative sodium-dependent transporter	General function prediction only	Poorly characterized
NMCC_1875	<i>bioH</i>	carboxylesterase BioH (biotin synthesis protein BioH)	General function prediction only	Poorly characterized
NMCC_1876		putative GntX-like protein	General function prediction only	Poorly characterized

Technical Appendix Table 4. List of genes found in majority of group 1 CC4821 strains but not in group 2

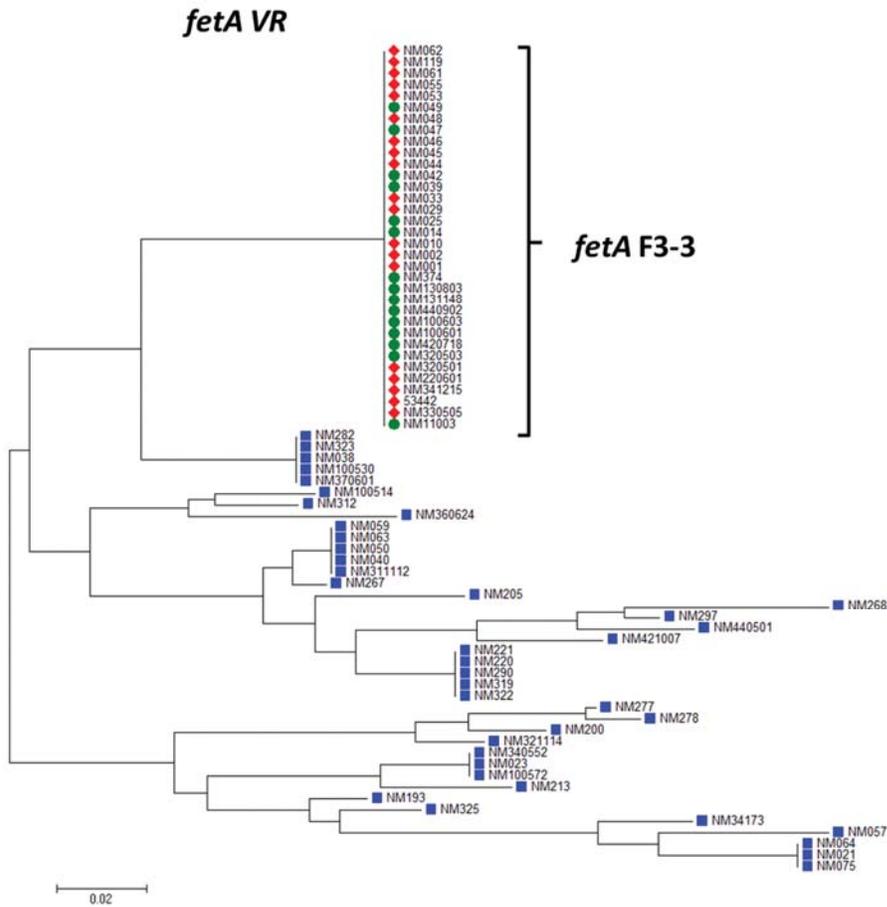
Gene ID	Gene name	Annotation
GNM001_00395		Cupin superfamily protein
GNM001_01225		Hypothetical protein
NEIMB_0683	<i>nhaP</i>	K(+)/H(+) antiporter NhaP
NMCC_0368		Hypothetical protein
NMCC_0369		Phosphoribosyl transferase domain protein
NMCC_0638		Hypothetical protein
NMCC_0790		Putative type I restriction enzyme P M protein
NMCC_1121	<i>mgsR</i>	Regulatory protein MgsR
NMCC_1206		Hypothetical protein
NMCC_1453	<i>lbpB</i>	Lactoferrin binding protein B (LbpB)
NMCC_2037	<i>ymdB</i>	O-acetyl-ADP-ribose deacetylase



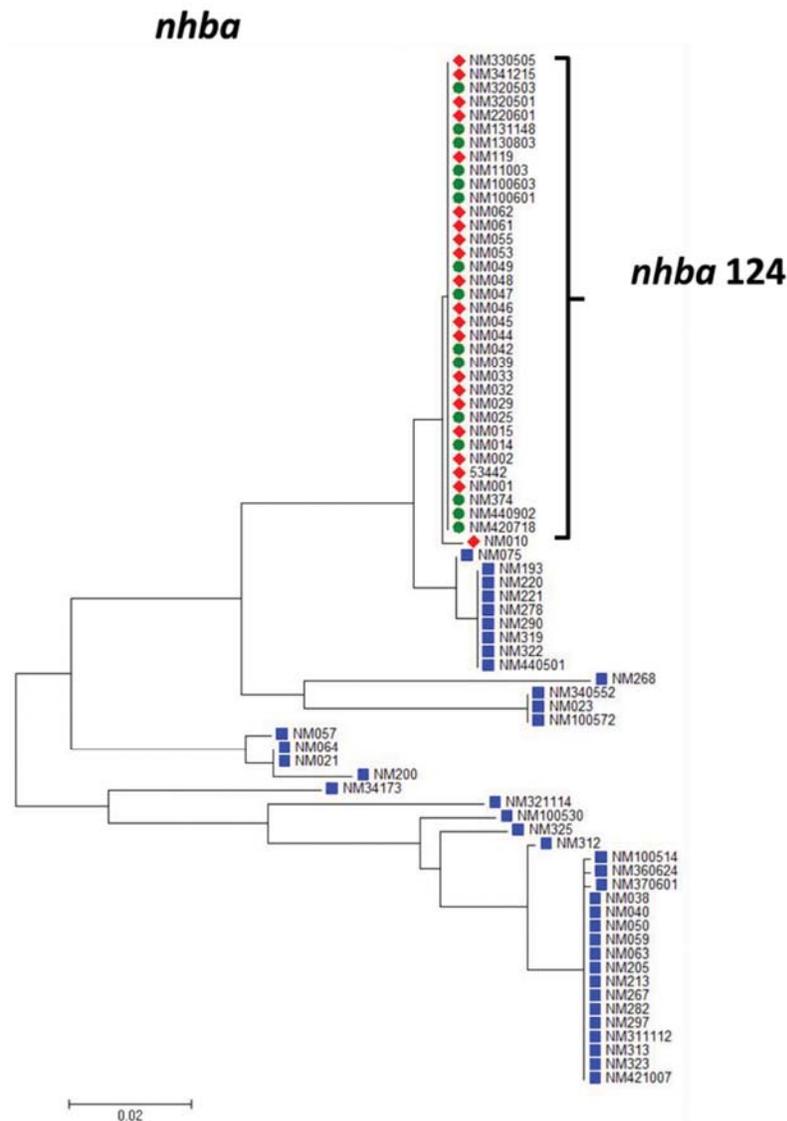
Technical Appendix Figure 1. Phylogenetic analysis of antigen gene sequence variation within the *Neisseria meningitidis* clonal complex 4821 lineage, based on nucleotide sequences of the *porA* gene. Phylogenetic trees of aligned antigen gene sequences for clonal complex 4821 isolates show extensive sequence diversity across all 5 antigen genes, most prominently among group 2 isolates (blue squares) compared with group 1 isolates (red diamonds and green dots). Group 1 isolates have unique *nada* and *fetA* alleles, whereas the epidemic clone (red diamonds) has acquired a genetically distinct *porA* genotype P1.7–2, 14 compared with remaining group 1 isolates (green dots). Scale bars represent total substitutions per site.



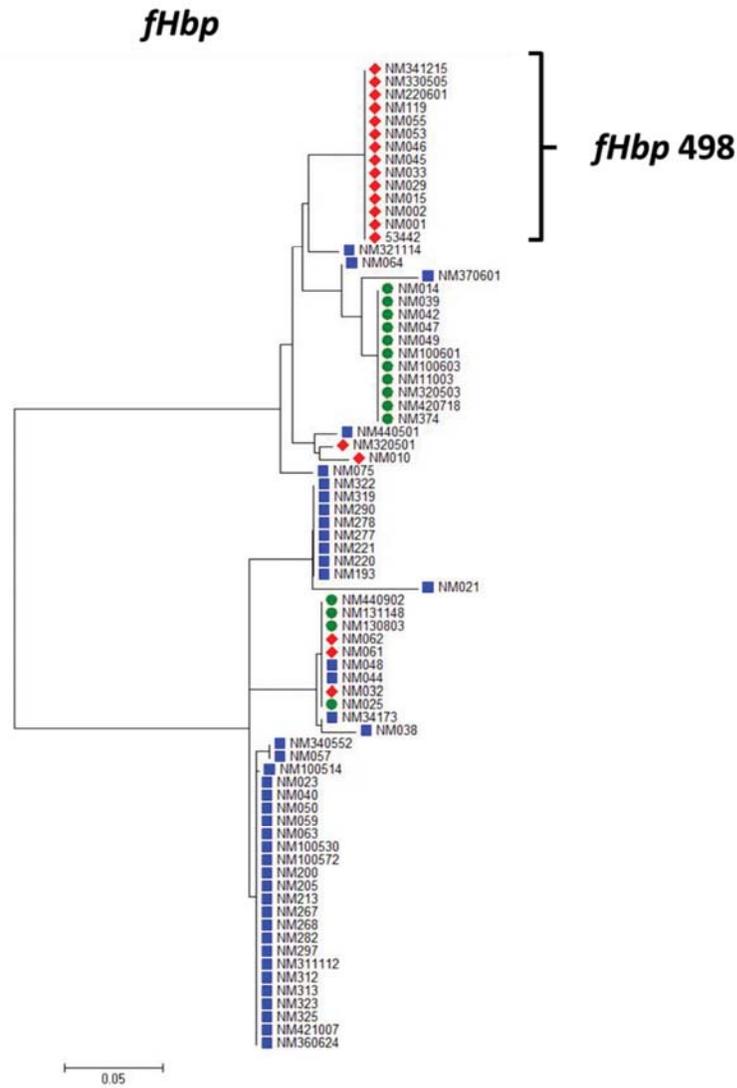
Technical Appendix Figure 2. Phylogenetic analysis of antigen gene sequence variation within the *Neisseria meningitidis* clonal complex 4821 lineage, based on nucleotide sequences of the *porB* gene. Phylogenetic trees of aligned antigen gene sequences for clonal complex 4821 isolates show extensive sequence diversity across all 5 antigen genes, most prominently among group 2 isolates (blue squares) compared with group 1 isolates (red diamonds and green dots). Group 1 isolates have unique *nada* and *fetA* alleles, whereas the epidemic clone (red diamonds) has acquired a genetically distinct *porA* genotype P1.7–2, 14 compared with remaining group 1 isolates (green dots). Scale bars represent total substitutions per site.



Technical Appendix Figure 3. Phylogenetic analysis of antigen gene sequence variation within the *Neisseria meningitidis* clonal complex 4821 lineage, based on nucleotide sequences of the *fetA* gene. Phylogenetic trees of aligned antigen gene sequences for clonal complex 4821 isolates show extensive sequence diversity across all 5 antigen genes, most prominently among group 2 isolates (blue squares) compared with group 1 isolates (red diamonds and green dots). Group 1 isolates have unique *nada* and *fetA* alleles, whereas the epidemic clone (red diamonds) has acquired a genetically distinct *porA* genotype P1.7–2, 14 compared with remaining group 1 isolates (green dots). Scale bars represent total substitutions per site.



Technical Appendix Figure 4. Phylogenetic analysis of antigen gene sequence variation within the *Neisseria meningitidis* clonal complex 4821 lineage, based on nucleotide sequences of the *nhba* gene. Phylogenetic trees of aligned antigen gene sequences for clonal complex 4821 isolates show extensive sequence diversity across all 5 antigen genes, most prominently among group 2 isolates (blue squares) compared with group 1 isolates (red diamonds and green dots). Group 1 isolates have unique *nada* and *fetA* alleles, whereas the epidemic clone (red diamonds) has acquired a genetically distinct *porA* genotype P1.7–2, 14 compared with remaining group 1 isolates (green dots). Scale bars represent total substitutions per site.



Technical Appendix Figure 5. Phylogenetic analysis of antigen gene sequence variation within the *Neisseria meningitidis* clonal complex 4821 lineage, based on nucleotide sequences of the *fHbp* gene. Phylogenetic trees of aligned antigen gene sequences for clonal complex 4821 isolates show extensive sequence diversity across all 5 antigen genes, most prominently among group 2 isolates (blue squares) compared with group 1 isolates (red diamonds and green dots). Group 1 isolates have unique *nada* and *fetA* alleles, whereas the epidemic clone (red diamonds) has acquired a genetically distinct *porA* genotype P1.7–2, 14 compared with remaining group 1 isolates (green dots). Scale bars represent total substitutions per site.