Genetic diversity and characteristics of gyrA gene in *Neisseria* spp

Pan Zhao

Chinese Center for Disease Control and Prevention

Aiyu Zhang

Chinese Center for Disease Control and Prevention

Bingqing Zhu

Chinese Center for Disease Control and Prevention

Li Xu

Chinese Center for Disease Control and Prevention

zhujun shao (✉ shaozhujun@icdc.cn)

Chinese Center for Disease Control and Prevention

Research article

Keywords: DNA gyrase subunit A , Quinolone Resistance-Determining Region, *Neisseria meningitidis* Clonal Complex 4821, recombination.

Posted Date: August 4th, 2019

DOI: <https://doi.org/10.21203/rs.2.12407/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background *Neisseria meningitidis* bacteria characterized by clonal complex (CC4821) showed a high resistance rate to quinolones. The aim of this study was to assess whether the DNA gyrase A gene from *N.meningitidis* CC4821 strains collected in China featured any specific characteristics compared to other *Neisseria* species. Two hundred fifty two gyrase gene sequences were analyzed, among them 77 generated in this study from *N.meningitidis* CC4821 strains collected in China between 1978 and 2016. Results The quinolone resistance-related gene, coding for the DNA gyrase subunit A (GyrA) protein, was highly divergent within the *N.meningitidis* strains whereas *N. gonorrhoeae* and *N.lactamica* counterparts appeared well conserved. Only one position, 91 (83 in *E.coli* gyrA gene), was linked to quinolone resistance, all resistant strains featuring the substitution T91I. The *E.coli* position 87, which was mutated in quinolone-resistant strains, was also divergent (position 95) in some *Neisseria* resistant strains. Moreover, twenty eight additional putative resistance markers were identified. Finally, putative recombination events between *N.meningitidis* and either *N.subflava*, or *N.lactamica* or *N.cinerea* as well as between *N.meningitidis* strains were reported. Conclusions Analyzing the evolution of gyrA gene within *Neisseria* spp. is critical to monitor the quinolone resistance phenotype and the acquisition of new resistance markers. Such studies are necessary for the control of the meningococcal disease and the development of new drugs targeting DNA gyrase.

Background

Meningitis is an inflammation of the protective membranes covering the brain and the spinal cord (<https://www.cdc.gov/meningococcal>). This disease can have multiple causes like bacteria or virus but also fungus, parasite or even non-infectious agent like lupus. Bacterial meningitis can be caused by several types of bacteria including *Streptococcus pneumonia* or *Neisseria meningitidis* (*N.meningitidis*). In addition to *N.meningitidis*, the genus *Neisseria* contains at least 30 distinct species infecting humans, other mammals and even insects (1). *N.meningitidis* are classified through several different schemes, based on serological test (serogroup) or genetic tests (sequence type) (2, 3). Among the 12 described serogroups, which are based on the structure of the capsule polysaccharide (*cps*), 6 serogroups (A, B, C, X, Y and W) caused the majority of invasive meningococcal disease (IMD) globally (2). The strains, which reacted with, either more than one, or no serogroup-specific antiserum, were considered non-groupable (NG) (4). Strains lacking the operons required for either the synthesis, or the lipid modification and/or the transport of the capsule were identified as capsule null locus (*cnl*) (5). In addition, the strains can be grouped into different sequence types (STs) based on the multilocus sequence typing (MLST) method on 7 genes (*abcZ*, *adk*, *aroE*, *fumC*, *gdh*, *pdhC*, *pgm*) (3). An ST is characterized by a different sequence nucleotide for at least one of the 7 reference genes. So far, 14556 STs have been described. Furthermore, strains that are sharing 4 or more STs (identified as number for convenience) could be classified into the same clonal complex (CC) (3). So far, 9248 CCs have been described (<https://pubmlst.org/neisseria/>). The strains that could not be classified into an existing CC were called unassigned (UA).

Until 2003, the strains of *N. meningitidis* with either CC1 or CC5 of serogroup A were responsible for the majority of IMD cases in China (6). In 2003, an outbreak of a new serogroup C meningococcal disease caused by CC4821 was reported in Anhui province of China. This new hypervirulent clonal lineage did not belong to any of the previously reported sequence types and had not been reported in any other countries except one report from Japan (7) (8). Subsequently, CC4821 serogroup C became one of the leading lineages across China (6). Later on, CC4821 became a dominant lineage among serogroup B strains since they were first identified in 2005. In contrast to serogroup C strains, serogroup B strains were usually associated with sporadic infections (6). Our further analyses of historic isolate collections showed that CC4821 strains which included serogroup B and C strains were isolated as early as in 1978 and mostly associated with asymptomatic carrier (9).

Two main strategies have been developed to control the meningococcal disease. Vaccines specific to multiple serogroups have been generated and are globally used (10). As for other bacterial infections, antibiotics have been also used to control the infection. The most common, quinolone and its derivatives, is targeting the DNA gyrase A which is essential for DNA replication (11, 12). Quinolone interaction with DNA gyrase has been well studied thanks to the 3D structure of the *E.coli* protein (13). The mutation of critical sites lead to resistance and they are located in the so called Quinolone Resistant-Determining Region (QRDR) (14).

The aim of this study was to assess whether the gyrase A gene from CC4821 strains from China featured any specific characteristics. Phylogenetic analysis was performed in the context of other *Neisseria* species. In addition, genetic markers like ST, CC or serogroup, resistance markers as well potential recombination between strains have been studied.

Results

The aim of this study was to identify potential sites in GyrA protein linked to a specific character. In order to perform this analysis, it was necessary first to carefully analyze the dataset in order to identify any potential characteristics either genetic, temporal or geographical that could be linked to any GyrA amino acid changes.

Analysis of the dataset

The 252 *gyrA* sequences analyzed in this report concerned 15 *Neisseria* species, *N. meningitidis*, *N.gonorrhoeae* and *N.lactamica* being the most represented (Additional Figure 1a). Ten of these species have been reported to infect humans, *N. meningitidis* strains as well as *N.gonorrhoeae* strains being pathogenic for humans. This study focused on *N. meningitidis* and therefore 87% of sequences (218/252) concerned this species. The 252 strains featured in this analysis were isolated from 15 countries of 4 continents, and mostly in China (Additional Figure 1b). The analyzed sequences concerned strains collected from 1931 to 2017, with most of the strains collected after 2005 (Additional Figure 1c).

Eighty-eight STs were represented among the 218 *N. meningitidis* strains analyzed in this study (Additional Figure 2a-c). Thirty-nine of these STs belong to CC4821 and ST4821 was represented by 43 strains (Additional Figure 2a). ST7 and ST5, which belong to CC5, were also well represented with 28 and 11 strains respectively (Additional Figure 2b). Eleven serogroups of *N.meningitidis* were analyzed; B and C were represented by 60 strains, A by 47 and W by 12 (Additional Figure 2d). Based on STs, 19 clonal complexes were represented in the dataset. The majority was CC4821 (95/252) followed by CC5 (41/252) and unassigned (UA) (Additional Table 1).

Evolutionary analysis of 252 *gyrA* nucleotide sequences

A neighbor joining phylogenetic tree was constructed with 252 *gyrA* nucleotide sequences (Figure 1). The nucleotide sequences were significantly divergent with an overall p-distance of 0.052 (Table 1). An overview of the tree showed that the sequences from *N.meningitidis* were found across the tree demonstrating that *gyrA* gene was relatively divergent within that species. Moreover, no trend was found based on collection location, time or even genetic characteristics like CC or serogroup. For example, the sequences of CC4821 strains (in red in Figure 1) were found across the tree. Similarly, sequences of CC5 strains (in blue in Figure 1) were found in two lineages suggesting that the *gyrA* gene evolution was independent of the CC5 typing. Overall, these observations suggested that the evolution of *gyrA* sequences was independent of location, time as well genetic characteristics of the bacterial strains.

Among the 252 analyzed sequences, 48% (121) were identical (Table 1). We chose to maintain these sequences despite the sequence identity in order to keep valuable information in terms of geography, time and genetic characteristics. Nearly 60 % of the sequences (151) were found on the top of the tree, with no significant bootstrap value (Figure 1). These sequences were highly homogeneous, with a p-distance of 0.003 (Table 1). The remaining 101 sequences were more divergent, with a p-distance of 0.064 relative to sequences grouped on the top of the tree. Most of the nodes concerning these 101 sequences featured a bootstrap value > 70%. Moreover, the p-distance within this group of 101 sequences was 0.083 demonstrating that these sequences were highly divergent between each other. As major nodes of the tree featured a bootstrap value > 80%, we decided to arbitrarily assign sequences to 9 different genetic groups (Figure 1; Table 1).

Fifteen sequences were considered outliers. Even though they were from 2 genetically distinct groups at the bottom of the nucleotide tree, these sequences shared a common node with 90 % bootstrap suggesting that these sequences were divergent compared to the rest of the sequences. The lowest p-distance between the outlier group and the other 8 groups was 0.16. These sequences were from 11 species that were not found elsewhere in the tree. We could think that this might be the consequence of a sample bias. However, as these sequences shared 8 amino acid changes, it strongly suggested that these sequences were related and significantly divergent compared to the rest of the analyzed sequences (Table 1, Additional Figure 3). Group 1 consisted of 2 sequences, *N.meningitidis*-MK930374-100514-ST4832-CC4821-C-China-2005 and *N.subflava*-CP031251.1-M18660—2009. This group was supported

by a bootstrap of 99% suggesting that these 2 sequences were very different from the rest of the sequences. The sequences shared 7 amino changes (Table 1, Additional Figure 3). However, the long branch corresponding to *N.subflava*-CP031251.1-M18660—2009 sequence suggested that this sequence was also significantly divergent from *N.meningitidis*-MK930374-100514-ST4832-CC4821-C-China-2005 and this was confirmed by a p-distance of 0.06 between these 2 sequences. Group 2 consisted of 2 sequences from *N.meningitidis* strains collected in Gansu and Guangxi province in China in 2007 and 2011, respectively, namely *N.meningitidis*-KF733178-G1-ST5636-UA-B-China-2007-R and *N.meningitidis*-MK930446-GX34173-ST9477-CC4821-B-China-2011. Both group 3 and 4 consisted of a single sequence, namely, *N.meningitidis*-MK930402-231003-ST12300-CC4821-B-China-2009 and *N.meningitidis*-KF733135-62-ST5751-UA-B-China-2006-R, respectively. The nodes corresponding to these sequences were supported by bootstrap values of 82 and 97% respectively. Furthermore, the lowest pairwise p-distances between these sequences and the other analyzed sequences were 0.032 and 0.025 respectively demonstrating that these sequences were significantly divergent compared to the rest of the analyzed sequences. Group 5 consisted of 26 sequences, mainly from *N.meningitidis*. However, a subgroup of 7 sequences supported by a bootstrap of 99% featured 5 sequences from *N.lactamica*. Group 6 consisted of 42 sequences from *N.meningitidis* at the exception of one sequence from *N. cinerea*, namely *N.cinerea*-LS483369.1-NCTC10294, which shared a node with *N.meningitidis*-KF733132-59-ST7962-CC77-NG-China-2009-R. However, the *N.cinerea* sequence featured a long branch suggesting a significant divergence compared to the *N.meningitidis* sequence. Twenty-five sequences, mainly from strains of ST7 and CC5 shared a common node with a bootstrap of 99% and a long branch. Finally, based on the available data, the sequences from group 6 appeared to be from strains that were resistant to quinolone. However, no amino acid substitution was shared by these strains suggesting that there was no common marker for the resistance phenotype (Additional Table 2). Group 7 was particular as it exclusively concerned sequences from *N. gonorrhoeae* species. These sequences were highly homogeneous with a p-distance of 0.002 and they shared a node with a bootstrap of 99% and a long branch. Moreover, these sequences shared 6 amino acid changes (Table 1, Additional Figure 3). Interestingly, the 12 analyzed sequences were from strains that were either susceptible or resistant to quinolone suggesting that this divergence was not related to a resistance phenotype. Group 8 consisted of most of *N.meningitidis* sequences analyzed in this report. The strains were collected in the last 90 years in 13 different countries from 4 continents. Despite the significant time span and geographic spread, these sequences were highly conserved with a p-distance of 0.003. Moreover, these sequences were from strains of 72 STs, 24 CCs, 10 serogroups, including the reference strain 053442. Overall, this showed that *gyrA* gene was highly conserved among most *N.meningitidis* strains despite different genetic characteristics, geographic locations or collection time.

Analysis of the divergence within the GyrA protein

The amino acid divergence within the GyrA protein was analyzed among 131 unique sequences (Additional Table 2). Two hundred fifty seven divergent positions were identified among the 931 amino

acid featured in the alignment (Figure 2). Even though these sites were found across the protein, the distribution of the divergence did not appear to be random. Two regions were highly conserved, from positions 530 to 620, and a smaller region between 300 and 330. According to the protein from *E.coli*, the first region corresponds to the end of the amino terminal domain and the beginning of the carboxy terminal domain. The second region corresponds to the tower domain of the protein based on the 3D model structure. Among the 257 divergent positions, 4 sites were highly divergent, namely 91, 417, 665 and 210. For example, more than 57% of the sequences featured a mutated residue at position 91 (Figure 2). Whereas 42% of the sequences had a T, 38% had an I, 14.7% had an S and the remaining sequences had either an F or a V (Additional Table 2). Overall, the divergence among the 257 variable sites was shared by 2-3% of the sequences (Additional Figure 4a). However, 52 divergent positions were only found in one sequence. These variable residues might be due to sequence artifact. Because these residues were not shared by other sequences, this meant that the variability at these positions had not been selected in the evolutionary process. Seventy-two positions were shared by at least 10 % of the sequences (14 out of 131), position 91 being the most variable as 75 sequences (57%) featured an amino acid substitution (Additional Figure 4b). Eighty-nine positions featured multiple amino acid residues (Additional Figure 4c-d). For example, the position 441 featured 6 amino residues (R, H, T, Q, G and N) in addition to the main residue E (Additional Table 2). This suggested that these positions were not only highly variable but also that these positions could allow different residues without affecting the protein function. The alignment of 131 sequences featured a few gaps, mainly located between positions 720-760 and the C terminus of the protein sequence.

The 257 divergent sites were analyzed among the complete dataset of 252 sequences in order to identify any potential correlation between divergence and geographic, species or genetic characteristics (Additional Table 2). As far as we can tell, no correlation was found between divergence and geographic location. Moreover, no temporal pattern was identified. However, a correlation between diversity and species was found for *N. gonorrhoeae* and *N. lactamica* (Table 2).

Indeed, 6 changes were only found in *N.gonorrhoeae*, D210K, E456K, E483K, V486I, N836S, D917G. Similarly, D907E was only found in *N.lactamica*. In contrast, no residue was specific to *N.meningitidis*. As the dataset was biased towards *N.meningitidis*, one explanation could be poor sampling. Only 12 *N.gonorrhoeae* sequences were analyzed compared to 120 *N.meningitidis*. However, the fact that the 6 residues were found in *N.gonorrhoeae* strains collected in different country and at different time suggested that these residues were indeed species specific. The search for a link between diversity and genetic marker like ST, CC or serogroup was unsuccessful suggesting that *gyrA* evolution was independent of the evolution of the rest of the genome. Finally, one position (91) appeared to be linked to quinolone resistance. All the resistant strains had an I or a F whereas all the sensitive strains had an S or a T. V was found in strains with undetermined resistance phenotype. Interestingly, *N.gonorrhoeae* appeared to have different markers, F for resistance and S for sensitivity.

Identification of potential resistance markers to quinolone

Among the 252 analyzed sequences, 61 were from strains tested for resistance to quinolone (Additional Table 1). These strains were further analyzed in order to identify additional potential markers for resistance. A change that would be found in resistant strains but not found in any sensitive strains would qualify. Thirty sites were identified (Additional Table 3). For example, H8N was found in 29 resistant strains but not featured in any sensitive strains. All the strains were analyzed for these 30 positions. Fifty-one different profiles had been identified, meaning that there were 51 combinations of these markers among all the analyzed strains (Table 3).

Among the 30 potential markers, position 91 and 103 were the most shared in the profiles. Some amino acid substitutions were only found in *N.gonorrhoeae* strains, like D95G or D95A. Finally, 24 of the 51 profiles concerned strains that were known to be resistant to quinolone. As resistance markers were initially described in *E.coli*, a comparison between *gyrA* sequences of *E.coli* and the reference strain *N.meningitidis* 53442 was necessary in order to check the position of these markers in *E.coli* sequence (Additional Table 4).

Recombination within *gyrA* gene between *N. spp.*

Among the 9 groups identified during the phylogenetic analysis, group 1 was of particular interest, concerning *N.meningitidis*-MK930374-100514-ST4832-CC4821-C-China-2005 and *N.subflava*-CP031251.1-M18660. These 2 sequences shared 7 residue changes, not seen in other strains. Furthermore, five of these changes were seen within 30 amino acids (Table 1). Finally, amino acid changes observed in one strain were not seen on other strains, like position 740 and 750. All these observations suggested a recombination between these 2 strains which was confirmed by a bootscan analysis (Figure 3). Other potential recombinations, with either cinerea, or lactamica strains and between meningitidis strains were also identified.

Discussion

The genus *Neisseria* has been found in more than 40 animal hosts mostly mammals but also insects (1). Nucleotide sequences for more than 30 different *Neisseria* species can be found in GenBank. The present study featured sequences from 14 of these species even though the dataset was highly biased towards *N.meningitidis*. Eleven species can be found in human, namely meningitidis, gonorrhoeae, lactamica, sicca, subflava, cinerea, elongata, flavescens, mucosa, pharyngis and polysaccharea (15, 16). Nine of these species were featured in the present analysis. The phylogenetic analysis of the *gyrA* gene presented in this report showed that *N.meningitidis* sequences were closely related to *N. gonorrhoeae* and *N.lactamica* sequences. To our knowledge, such phylogenetic analysis has never been realized before. However, several genes like *argF*, *recA*, *rho* as well as 16S *rRNA* have been used for evolutionary analyses (17, 18). The genetic relationship between the 3 species, meningitidis, gonorrhoeae and lactamica was

confirmed for the 3 genes *argF*, *recA*, *rho*. A genetic relationship between the 3 species was also confirmed at the genome level. However, *lactamica* appeared more divergent than *meningitidis* and *gonorrhoeae* (18). The present study featured 11 other species of *Neisseria*. However, as they were only represented by 1 or 2 sequences, it is likely that the poor sampling would compromise the reliability of the analysis. It is however striking to observe that 9 of these species shared 8 amino acid substitutions within GyrA protein.

The phylogenetic analysis on 218 sequences of *gyrA* gene from *N.meningitidis* strains showed that *gyrA* was highly divergent within the meningitidis strains. To our knowledge, only one other report featured a phylogenetic analysis of the *gyrA* gene from *N.meningitidis* but omitted to include other species therefore a comparison with other species was not possible (11). Whereas *gyrA* gene of meningitidis showed high divergence between the analyzed strains, *gyrA* gene of gonorrhoeae featured a remarkable conservation with 6 amino acid substitutions among the 12 analyzed sequences. A few hypotheses could be proposed to explain such discrepancy between species. First, a sampling artifact might be likely. Only 12 sequences have been analyzed whereas around 300 protein sequences are listed in GenBank. To our knowledge, no systematic phylogenetic analysis was performed on full length *gyrA* gene in *N.gonorrhoeae*. We briefly attempted to analyze the sequences available on the public domain and did not find significant any divergence among *N.gonorrhoeae* sequences thus confirming the present study. Another hypothesis could be that gonorrhoeae gyrase A might feature additional property that are not featured by meningitidis gyrase A. Gonorrhoeae strains do not have the same tropism as meningitidis strains and therefore additional property could be possible. Further studies would be necessary to assess the role of the 6 amino acid substitutions identified in the present study, namely D210K, E456K, E483K, V486I, N836S, D917G. These substitutions are shared by resistant (5 strains) as well as sensitive (4 strains) strains and are therefore very unlikely to be linked to a resistance phenotype.

Gene transfer or recombination between bacterial genomes has been well described (19, 20). The present study reported potential recombination events between *Neisseria* species. Recombination would mean that the bacteria have been replicating at the same time and same location. Among the reported recombinations, a recombination within *gyrA* gene between meningitidis and subflava was identified. Unfortunately, the metadata concerning the subflava strain were very limited, only the collection year was known after contacting the scientists who reported the genome sequence in GenBank. Knowing the country of collection would help to understand a potential recombination with *N.meningitidis* collected in China. Recombination between *Neisseria* species have been previously reported. For example, Wu et al. reported a likely recombination between *lactamica* and meningitidis (19). Recombination between species as well as intra-species is likely to allow the strains to acquire new resistant phenotype however further studies would be necessary to better assess the acquisition of the resistance phenotype.

The original aim of this study was to identify potential markers in *gyrA* gene that could be linked to either geographic, temporal or genetic characteristics. Only two criteria were found to be linked to *gyrA* mutations. One was the species; as discussed above, gonorrhoeae featured 6 unique amino acid substitutions. The other criteria was the position 91 which was exclusively an I or an F in quinolone

resistant strains and an S or a T in quinolone sensitive strains. The mutation at position 91 was already well described (19, 21-23). As quinolone resistance of gyrase A protein was originally described in *E.coli*, we felt that all positions related to quinolone resistance should be identified based on the *E.coli* sequence (Yoshida et al., 1990). Resistant markers were originally described as the Quinolone Resistance-Determining Region (QRDR), from position 67 to 106 in *E.coli* (75 to 114 in *Neisseria spp.*). The QRDR of *E.coli* featured 6 amino acid substitutions (A67S, G81C, S83W, A84P, D87N and Q106H). The corresponding region in *Neisseria spp.* featured 5 sites. However, the respective positions were different with only 2 positions conserved, 91 (83 in *E.coli*) and 95 (87 in *E.coli*). The 5 amino acid substitutions in *Neisseria spp.* were the following: T91I, D95G/A/N, N103D (95 in *E.coli*), A105S (97 in *E.coli*), I111V (103 in *E.coli*). The positions 67, 81, 84 and 106 which were divergent in *E.coli* were however conserved in *Neisseria spp.* at the respective positions 75, 89, 92 and 114. The alignment generated in the present study was in agreement with a previous report (24). A recent study identified 2 additional amino substitutions, located outside the QRDR *sensu stricto*, namely A51V and A119E. These 2 positions were however conserved in *Neisseria spp.* The discrepancy between *E.coli* and *Neisseria* suggests some difference in quinolone resistance. Unfortunately, to our knowledge, a comparison between *E.coli* and *Neisseria* resistance has not been performed and further studies would be necessary to assess whether *Neisseria spp.* are more or less resistant to quinolone than *E.coli*. The present study identified 28 positions that might be involved in a resistance phenotype. These positions are characterized by a substitution only found in resistant strains. Further analysis would be necessary to assess whether these substitutions have an effect on the resistance phenotype.

Conclusions

In summary, a phylogenetic analysis of *gyrA* gene from *N.meningitidis* in the context of other *Neisseria* species revealed the following points: *gyrA* evolution seems species specific for gonorrhoeae and lactamica. In contrast, the evolution of *gyrA* within meningitidis strains appears more divergent. Several potential recombination events have been identified that could explain why the meningitidis *gyrA* gene is less conserved. The consequence of recombination between species or within species remains to be studied. Only one position (91) appears to be linked to a quinolone resistance phenotype. This position corresponds to the position 83 in *E.coli* counterpart. Besides position 87 (95 in *Neisseria*), the other resistance markers identified in *E.coli* were not identified in *Neisseria spp.* However, additional putative resistance markers have been identified and would require additional work to assess their role in the resistance phenotype. Studying the evolution of *gyrA* within *Neisseria spp.* is critical to monitor the resistance to quinolone and the acquisition of new resistance markers. Such studies are necessary for the control of the meningococcal disease and the development of new drugs targeting DNA gyrase.

Methods

Bacterial Strains

More than 4,000 *N.meningitidis* isolates have been collected throughout China from IMD patients as well as close contacts and asymptomatic carriers by our laboratory since 1960s. The bacterial strains were propagated on single Petri dish containing Difco™ Columbia Blood Agar Base with 5% Sheep Blood in a 5% CO₂ atmosphere at 37 °C for 18 h. Single colonies were lysed and tested by PCR for the meningococcal-specific contact-regulated gene A (*crgA*) in order to identify bacterial species (25). Strain serogroups were determined by slide agglutination with rabbit antisera specific to each serogroup (BD Difco). Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. MLST was performed according to the description of Maiden *et al.* (3). Briefly, 7 genes (*abcZ*, *adk*, *aroE*, *fumC*, *gdh*, *pdhC* and *pgm*) were amplified and sequenced. STs and CCs were subsequently assigned by querying the sequence database available at <http://pubmlst.org/>. Seventy-seven CC4821 isolates were selected as representative isolates based on the typing results (ST, CC and serogroup), isolation year (1978-2016) and sample source (IMD patients, close contacts, or asymptomatic carriers) for in-depth phylogenetic analysis. The *gyrA* gene was sequenced by PCR using the primer pair *gyrA*-F, 5'-GTTCCGCGTCAAAATATGCT-3' and *gyrA*-2844R, 5'-GACTATAATCCGCTATATTGT-3' generating a 2905 bp long amplicon containing the 2751 nt of *gyrA* gene (11, 26). Seventy-seven complete gene sequences were generated and submitted to GenBank (MK930374-MK930450) (listed in Additional Table 1).

Dataset

The *gyrA* gene from the reference strain *N.meningitidis* 053442 (CP000381.1) was used to query the GenBank database and 556 additional *gyrA* gene sequences were selected from the BLAST output. Identical sequences were deleted unless the strains did not share the same CC, ST or serogroup. Sequences less than 90% of full length (2905 nt) were discarded. The remaining 175 sequences were combined with the 77 sequences generated in this study.

Sequence analysis

Two hundred fifty-two *gyrA* nucleotide sequences were aligned using Mega 6 and a neighbor joining phylogenetic tree was generated using the maximum composite likelihood nucleotide model (27, 28). Phylogenetic inference was tested with 1000 bootstraps (29). Nodes with bootstrap value > 70% were indicated. Average genetic p-distance within genetic groups as well as between genetic groups and pairwise distances were computed in Mega. Amino acid divergence was analyzed in Mega and processed in Excel. Potential recombination events were analyzed with the SimPlot software (30).

Abbreviations

IMD: Invasive Meningococcal Disease

gyrA: gyrase subunit A

CC: Clonal Complex

ST: Sequence Types

N. meningitidis: *Neisseria meningitidis*

QRDR: Quinolone Resistance-Determining Region

MLST: Multi-Locus Sequence Typing

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by the grants from the National Key Program for Infectious Disease of China (2018ZX10101002-001-014). National Natural Scientific Foundation (No. 81602903) from the Ministry of Science and Technology of the People's Republic of China, Sanming Project of Medicine in Shenzhen (No. SZSM201803081).

Authors' contributions

ZS designed and supervised the study. PZ generated and analyzed the data. PZ wrote the manuscript. All authors read the manuscript and approved of its submission.

Acknowledgements

The authors would like to thank the staff from the provincial and regional CDC Chinese offices. Assistance with the strain *N.subflava* M18660 from Dr. Nadav Topaz from US CDC, Atlanta, USA was appreciated.

References

1. Liu G, Tang CM, Exley RM. Non-pathogenic *Neisseria*: members of an abundant, multi-habitat, diverse genus. *Microbiology*. 2015;161(7):1297-312.
2. Harrison OB, Claus H, Jiang Y, Bennett JS, Bratcher HB, Jolley KA, et al. Description and nomenclature of *Neisseria meningitidis* capsule locus. *Emerging infectious diseases*. 2013;19(4):566-73.
3. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*. 1998;95(6):3140-5.
4. Dolan-Livengood JM, Miller YK, Martin LE, Urwin R, Stephens DS. Genetic basis for nongroupable *Neisseria meningitidis*. *The Journal of infectious diseases*. 2003;187(10):1616-28.
5. Claus H, Maiden MC, Maag R, Frosch M, Vogel U. Many carried meningococci lack the genes required for capsule synthesis and transport. *Microbiology*. 2002;148(Pt 6):1813-9.
6. Zhou H, Gao Y, Xu L, Li M, Li Q, Li Y, et al. Distribution of serogroups and sequence types in disease-associated and carrier strains of *Neisseria meningitidis* isolated in China between 2003 and 2008. *Epidemiology and infection*. 2012;140(7):1296-303.
7. Kawasaki Y, Matsubara K, Takahashi H, Morita M, Ohnishi M, Hori M, et al. Invasive meningococcal disease due to ciprofloxacin-resistant *Neisseria meningitidis* sequence type 4821: The first case in Japan. *Journal of infection and chemotherapy : official journal of the Japan Society of Chemotherapy*. 2018;24(4):305-8.
8. 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(9508):419-23.
9. Yang L, Shao Z, Zhang X, Xu L, Peng J, Xu X, et al. Genotypic characterization of *Neisseria meningitidis* serogroup B strains circulating in China. *The Journal of infection*. 2008;56(3):211-8.
10. Updated recommendations for use of meningococcal conjugate vaccines — Advisory Committee on Immunization Practices (ACIP), 2010. *MMWR Morbidity and mortality weekly report*. 2011;60(3):72-6.
11. 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. *Journal of medical microbiology*. 2014;63(Pt 11):1411-8.
12. Guo Q, Mustapha MM, Chen M, Qu D, Zhang X, Chen M, et al. Evolution of Sequence Type 4821 Clonal Complex Meningococcal Strains in China from Prequinolone to Quinolone Era, 1972-2013. *Emerging infectious diseases*. 2018;24(4):683-90.

13. Malik M, Mustaev A, Schwanz HA, Luan G, Shah N, Oppegard LM, et al. Suppression of gyrase-mediated resistance by C7 aryl fluoroquinolones. *Nucleic acids research*. 2016;44(7):3304-16.
14. Yoshida H, Bogaki M, Nakamura M, Nakamura S. Quinolone resistance-determining region in the DNA gyrase *gyrA* gene of *Escherichia coli*. *Antimicrobial agents and chemotherapy*. 1990;34(6):1271-2.
15. Hung MC, Christodoulides M. The biology of *Neisseria* adhesins. *Biology*. 2013;2(3):1054-109.
16. Weyand NJ. *Neisseria* models of infection and persistence in the upper respiratory tract. *Pathogens and disease*. 2017;75(3).
17. Bratcher HB, Bennett JS, Maiden MC. Evolutionary and genomic insights into meningococcal biology. *Future microbiology*. 2012;7(7):873-85.
18. Smith NH, Holmes EC, Donovan GM, Carpenter GA, Spratt BG. Networks and groups within the genus *Neisseria*: analysis of *argF*, *recA*, *rho*, and 16S rRNA sequences from human *Neisseria* species. *Molecular biology and evolution*. 1999;16(6):773-83.
19. Wu HM, Harcourt BH, Hatcher CP, Wei SC, Novak RT, Wang X, et al. Emergence of ciprofloxacin-resistant *Neisseria meningitidis* in North America. *The New England journal of medicine*. 2009;360(9):886-92.
20. Arreaza L, Salcedo C, Alcalá B, Vazquez JA. What about antibiotic resistance in *Neisseria lactamica*? *The Journal of antimicrobial chemotherapy*. 2002;49(3):545-7.
21. Unemo M, Shafer WM. Antimicrobial resistance in *Neisseria gonorrhoeae* in the 21st century: past, evolution, and future. *Clinical microbiology reviews*. 2014;27(3):587-613.
22. Tanaka M, Nakayama H, Haraoka M, Saika T, Kobayashi I, Naito S. Susceptibilities of *Neisseria gonorrhoeae* isolates containing amino acid substitutions in *GyrA*, with or without substitutions in *ParC*, to newer fluoroquinolones and other antibiotics. *Antimicrobial agents and chemotherapy*. 2000;44(1):192-5.
23. Alcalá B, Salcedo C, de la Fuente L, Arreaza L, Uria MJ, Abad R, et al. *Neisseria meningitidis* showing decreased susceptibility to ciprofloxacin: first report in Spain. *The Journal of antimicrobial chemotherapy*. 2004;53(2):409.
24. Deguchi T, Yasuda M, Asano M, Tada K, Iwata H, Komeda H, et al. DNA gyrase mutations in quinolone-resistant clinical isolates of *Neisseria gonorrhoeae*. *Antimicrobial agents and chemotherapy*. 1995;39(2):561-3.
25. Taha MK. Simultaneous approach for nonculture PCR-based identification and serogroup prediction of *Neisseria meningitidis*. *J Clin Microbiol*. 2000;38(2):855-7.
26. Strahilevitz J, Adler A, Smollan G, Temper V, Keller N, Block C. Serogroup A *Neisseria meningitidis* with reduced susceptibility to ciprofloxacin. *Emerging infectious diseases*. 2008;14(10):1667-9.
27. Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution*. 2013;30(12):2725-9.

28. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101(30):11030-5.
29. Felsenstein J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution; international journal of organic evolution*. 1985;39(4):783-91.
30. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of virology*. 1999;73(1):152-60.
31. Morais Cabral JH, Jackson AP, Smith CV, Shikotra N, Maxwell A, Liddington RC. Crystal structure of the breakage-reunion domain of DNA gyrase. *Nature*. 1997;388(6645):903-6.
32. Reece RJ, Maxwell A. Probing the limits of the DNA breakage-reunion domain of the *Escherichia coli* DNA gyrase A protein. *The Journal of biological chemistry*. 1991;266(6):3540-6.

Additional Files

Additional Figure 1: Distribution of species, collection country and year of collection for the 252 *gyrA* gene sequences analyzed in this study. A. Distribution of species. The number of strains for each species is plotted. The dataset contains 218 *N.meningitidis* strains and 12 *N.gonorrhoeae* strains. The species are organized based on their host, human specific first. *Eikenella* belongs to the *Neisseria* genus. B. Distribution based on collection country. The dataset contains 172 strains from China. C. Distribution based on collection year.

Additional Figure 2: Distribution of ST, CC, serogroup as well as quinolone resistance phenotype of the 218 *N.meningitidis* strains analyzed in this study. A. Distribution of ST among the 95 strains of CC4821. B. Distributions of CC and ST of 105 strains which are not of CC4821 and are represented by 2 or more ST in the dataset. ND; Not determined. Strains with no ST are indicated with "-". C. Distributions of CC and ST of 18 strains represented by only 1 ST in the dataset. D. Distribution of serogroup among the 218 *N.meningitidis* strains analyzed in this study. Serogroup information was not available for 26 strains.

Additional Figure 3: Neighbor phylogenetic tree of 252 GyrA protein sequences. The tree was annotated according to Figure 1.

Additional Figure 4: Distribution of amino acid divergence within the GyrA protein based on the 131 unique sequences analyzed in this study. A. Number of sequences sharing divergent positions. For example, 52 positions were only divergent in one sequence whereas 75 sequences were divergent at the same position (91). B. List of the divergent positions shared by more than 10% of the sequences (14 or more). The number of sequences and variable positions are shown as well as the residue position in *N. meningitidis* and the corresponding position in *E.coli*. For example, at position 487 in *N. meningitidis* (476 in *E.coli*), 16 sequences are divergent and 4 different substitutions can be found. C. Distribution of

number of possible changes among variable positions. For example, 162 positions have only one possible substitutions whereas 1 position can feature 6 different substitutions. D. List of the divergent positions with 4 or more possible of changes. Number of variable positions as well the residue positions in *N.meningitidis* and the corresponding in *E.coli* are shown.

Additional Table 1: List of 252 analyzed strains

Additional Table 2: Amino acid difference among the 252 GyrA protein sequences.

Additional Table 3: Potential quinolone resistance markers among the 252 GyrA protein sequences analyzed in this study.

Additional Table 4: Alignment of GyrA protein sequence from *E.coli* K12 and *N.meningitidis* reference strain 053442.

Figures

Figure 1

Figure 1

Neighbor phylogenetic tree of 252 *gyrA* gene sequences. Strain name is indicated as follows: species name-GB ID-strain ID-ST-CC-Serogroup-country of collection-year of collection-quinolone resistance phenotype. Missing information is indicated by an empty space. For example, *N.meningitidis*-AM889136.1-alpha14-ST53-CC53-cnl-Germany-1999 S; *Eikenella corrodens*-CP034670.1-KCOM3110—South_Korea-2017. Sequence name from *N.meningitidis* strains of CC4821 and CC5 are indicated in red

and blue font respectively. The 77 sequences generated in this study are underlined. The sequences of 9 reference strains are indicated by a black dot. Bootstrap values > 70% are indicated. Bootstrap values < 70% are indicated in parenthesis when necessary. The 9 genetic groups identified in this study are indicated as bracketed vertical lines.

Figure 2

Amino acid divergence among the GyrA protein based on 131 sequences. The positions within the 932 amino acid long alignment are indicated on the X axis. The percentage of sequences featuring a particular divergent position is indicated on the Y axis (left side). For example, 58% of the sequences feature a mutation at position 91. The divergence plot was generated from the amino acid difference table shown in Additional Table 2. The alignment features 10 gaps and the number of sequences featuring gaps is indicated on the right axis and shown as a black square. A map of E.coli GyrA protein featuring structural and functional domains is shown at the bottom for comparison. The map was generated based on the following references (13, 31, 32). The GyrA protein sequence of E.coli and N.meningitidis reference strain 53442 (GB-ID CP000381) were compared (Additional Table 4). The known quinolone-resistant site in E.coli are shown and corresponding positions in N.meningitidis are indicated with a blue line. It is worth noting that among the 8 resistant sites reported in E.coli, only positions 83 and 87 are divergent in N.meningitidis (sites 91 and 95).

Figure 3

Figure 3

Potential recombination events inter and intra species. A. Recombination between *N. subflava* and *N. meningitidis*. A bootscan plot was generated in Simplot with *N. meningitidis*-MK930374-100514-ST4832-CC4821-C-China-2005 as a query. B. Recombination between *N. cinerea* and *N. meningitidis*. A bootscan plot was generated in Simplot with *N. meningitidis*-KF733132-59-ST7962-CC77-NG-China-2009-R as a query. C. Recombination between *N. lactamica* and *N. meningitidis*. A bootscan plot was generated

in Simplot with N.meningitidis-MK930398-140901-ST8241-CC4821-B-China-2009 as a query. D. Recombination between N.meningitidis strains. A bootscan plot was generated in Simplot with N.meningitidis-MK930446-GX34173-ST9477-CC4821-B-China-2011 as a query. The strain that is predicted to contribute the most in the recombination process (backbone) is shown as a red line. The N.meningitidis reference strain 053442 (GB ID CP000381) is shown as a black line. A functional map for the reference strain 053442 is shown above based on the E.coli functional map shown in Figure 2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement1.jpg](#)
- [supplement2.jpg](#)
- [supplement2.jpg](#)
- [supplement4.xlsx](#)
- [supplement5.xlsx](#)
- [supplement6.xlsx](#)
- [supplement7.xlsx](#)
- [supplement8.jpg](#)
- [supplement8.xlsx](#)
- [supplement10.jpg](#)
- [supplement11.xlsx](#)