*Epidemiol. Infect.* (2015), **143**, 3001–3010. © Cambridge University Press 2015
doi:10.1017/S0950268815000473

# Multilocus variable-number tandem-repeat analysis of *Neisseria meningitidis* serogroup C in China

X. Y. SHAN[1,2,3,4], H. J. ZHOU[3,4], J. ZHANG[1], B. Q. ZHU[3,4], L. XU[3,4], Z. XU[3], G. C. HU[1], A. Y. BAI[1], Y. W. SHI[1], B. F. JIANG[2]* AND Z. J. SHAO[3,4]*

[1] *Jinan Municipal Centre for Disease Control and Prevention, Jinan City, Shandong Province, PR China*
[2] *Department of Epidemiology and Health Statistics, School of Public Health, Shandong University, Jinan City, Shandong Province, PR China*
[3] *State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, PR China*
[4] *Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, PR China*

*Received 21 November 2014; Final revision 18 February 2015; Accepted 23 February 2015; first published online 17 March 2015*

## SUMMARY

This study characterized *Neisseria meningitidis* serogroup C strains in China in order to establish their genetic relatedness and describe the use of multilocus variable-number tandem-repeat (VNTR) analysis (MLVA) to provide useful epidemiological information. A total of 215 *N. meningitidis* serogroup C strains, obtained from 2003 to 2012 in China, were characterized by MLVA with different published schemes as well as multilocus sequence typing. (i) Based on the MLVA scheme with a combination of five highly variable loci, 203 genotypes were identified; this level of discrimination supports its use for resolving closely related isolates. (ii) Based on a combination of ten low variable loci, clear phylogenetic relationships were established within sequence type complexes. In addition, there was evidence of microevolution of VNTR loci over the decade as strain lineages spread from Anhui to other provinces, the more distant the provinces from Anhui, the higher the genetic variation.

**Key words**: Bacterial typing, meningococcus, molecular epidemiology, *Neisseria meningitidis*, spread of disease.

## INTRODUCTION

*Neisseria meningitidis* is a Gram-negative, human-specific bacterium and despite effective antibiotics and partially effective vaccines, it remains one of the leading causes of bacterial meningitis worldwide. It can also cause sepsis, pneumonia, and other localized infections [1]. Based on the immunogenicity and structure of the capsule polysaccharide, *N. meningitidis* can be classified into 12 serogroups, and the majority of meningococcal cases are caused by members of serogroups A, B, C, X, Y and W [2]. The epidemiology of meningococcal disease varies substantially by geographical area [3]. In Africa, most meningococcal disease is associated with serogroup A, but serogroups C, X and W also occur, whereas in the Americas serogroups C and B predominate, the latter being the most prevalent in Australia and Europe. Serogroups Y and W have also been associated with a substantial proportion of infections in various countries [3].

In China, *N. meningitidis* serogroup A was responsible for most cases in the last century, while serogroups

* Author for correspondence: Professor Z. J. Shao, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, 102206, PR China
(Email: shaozhujun@icdc.cn) [Z. J. Shao]
(Email: bjiang@sdu.edu.cn) [B. F. Jiang ]

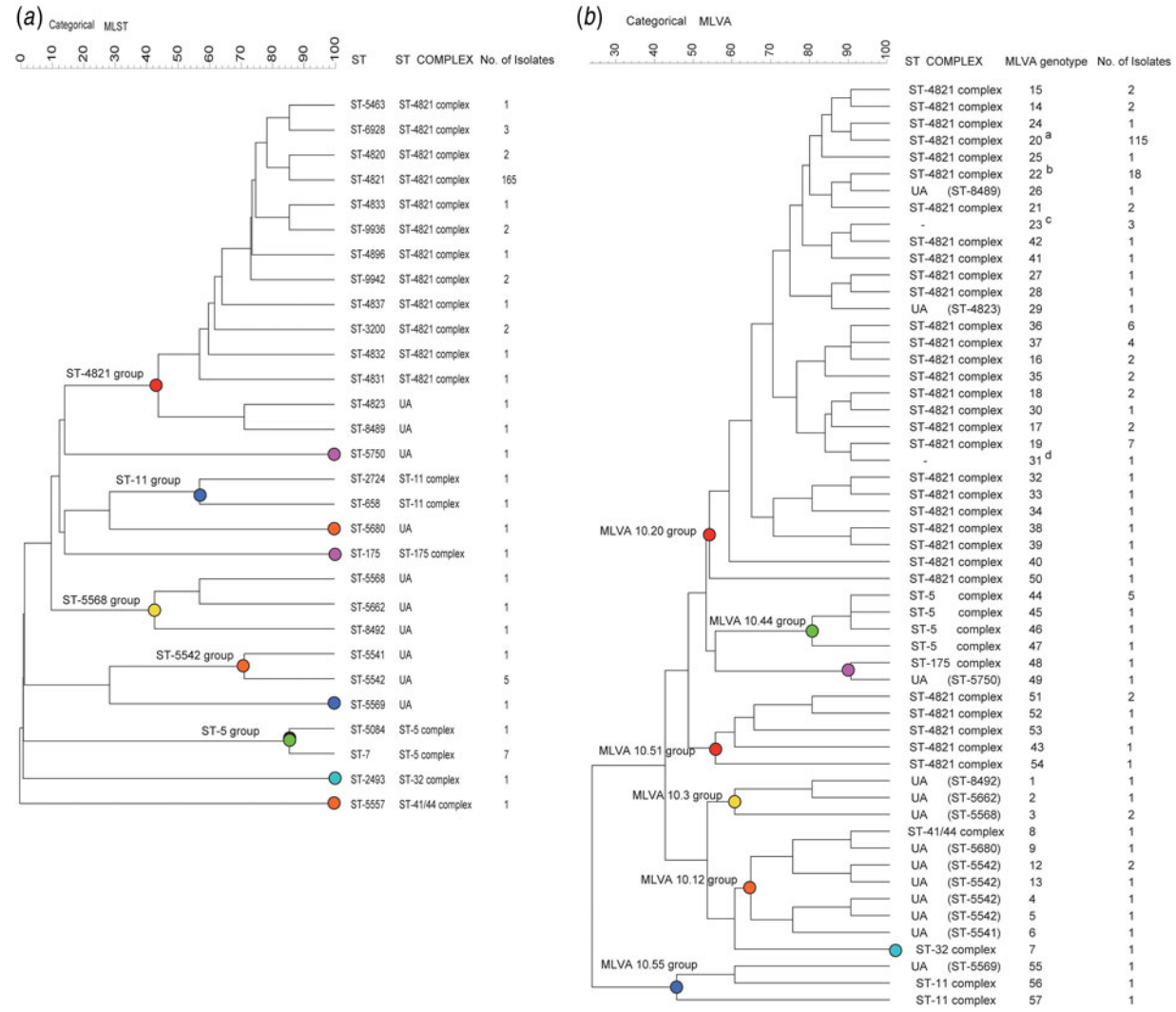


**Fig. 1.** Dendrogram constructed using multilocus sequence typing (MLST) and multilocus variable-number tandem-repeat (VNTR) analysis (MLVA). The dendrogram was generated by UPGMA, using the Dice-predicted similarity value of two patterns. The statistics program was provided by Bionumeric, v. 5.1. (*a*) Dendrogram constructed using sequence types (STs) in *N. meningitidis* serogroup C strains, excluding six isolates without a MLST result. ST groups were named when more than two STs clustered into one group. (*b*) Dendrogram constructed using MLVA scheme with 10 low variable VNTR loci (3, 6, 7, 8, 9, 11, 12, 13, 14, 15). Some STs are given in parentheses to indicate that isolates of UA have different STs. [a] Of 115 isolates, three have no MLST results. [b] Of 18 isolates, one has no MLST result. [c] Of three isolates, one has no MLST result. [d] This isolate has no MLST result. MLVA groups (*b*) were labelled with corresponding colours according to the groups that were established by MLST (*a*).

B and C occurred only sporadically [4]. However, during 2003 and 2005, a sudden increase in the number of cases with serogroup C strains occurred in Anhui province. Multilocus sequence typing (MLST) indicated that sequence type (ST) 4821 complex, a new hypervirulent lineage, was the cause of these outbreaks [5]. After 2003, *N. meningitidis* serogroup C spread nationwide through asymptomatic carriage and currently it has been isolated in more than 22 provinces in China.

*N. meningitidis* serogroup C is a monomorphic organism, but it can be divided into many clones by strain genotyping methods in order to establish its genetic relatedness in epidemiological and phylogenetic studies. Although widely adopted as a standardized method for investigation of meningococcal disease outbreaks [6], pulsed-field gel electrophoresis (PFGE) lacks the discriminatory power necessary to resolve clonal relationships in serogroup C strains that have evolved in China. MLST, has been widely applied for phylogenetic study of several bacterial species including *N. meningitidis* but has not proven to be sufficiently discriminatory for *N. meningitidis* since

https://doi.org/10.1017/S0950268815000473 Published online by Cambridge University Press

Table 1. *Characteristics of VNTR loci for 215 isolates and details at genomes of* N. meningitidis *strains FAM18 and 053442*

| VNTR locus | Repeat sequence | Size (bp) | Locus in FAM18 | | Locus in 053442 | | No. of alleles | Nei's Diversity index |
|---|---|---|---|---|---|---|---|---|
| | | | Genome coordinates | No. of repeats | Genome coordinates | No. of repeats | | |
| VNTR1 [15, 16] | CAAACAA | 7 | 601 072–601 274 | 29 | 597 047–597 213 | 24 | 40 | 0·96 |
| VNTR2 [16] | CATTTCT | 7 | 716 022–716 154 | 19 | 713 382–713 430 | 7 | 26 | 0·914 |
| VNTR3 [14, 16] | GCTTCAGTTACAGCTTCTTTG | 21 | 1 407 985–1 408 068 | 4 | 1 401 314–1 401 397 | 4 | 4 | 0·444 |
| VNTR4 [14, 16] | CAAG | 4 | 1 444 059–1 444 090 | 8 | 1 436 140–1 436 171 | 8 | 17 | 0·869 |
| VNTR5 [14–16] | GCCAAAGTT | 9 | 277 433–277 666 | 26 | 1 913 637–1 913 708 | 8 | 19 | 0·69 |
| VNTR6 [16] | CCGCTGCTACTGCCGCTGCTGAAGCACCTG | 30 | 932 818–932 907 | 3 | 901 179–901 238 | 2 | 3 | 0·265 |
| VNTR7 [16] | TACGGCTGCCGCGTCAAA | 18 | 1 191 565–1 191 582 | 1 | 1 181 818–1 181 853 | 2 | 3 | 0·205 |
| VNTR8 [16] | CGGATACGCTCTTGG | 15 | 1 250 095–1 250 139 | 3 | 1 241 966–1 242 010 | 3 | 4 | 0·172 |
| VNTR9 [16] | CAGATT | 6 | 1 824 619–1 824 596 | 4 | 1 811 289–1 811 336 | 8 | 12 | 0·462 |
| VNTR11 [14, 16] | GGGTAGCGG | 9 | 1 819 003–1 819 047 | 5 | 1 805 717–1 805 734 | 2 | 4 | 0·098 |
| VNTR12 [14, 16] | CGTATTTTCCCAT | 13 | 1 844 470–1 844 534 | 5 | 1 829 277–1 829 302 | 2 | 6 | 0·165 |
| VNTR13 [14] | TTTCCTG | 7 | 2 074 906–2 074 912 | 1 | 146 752–146 758 | 1 | 4 | 0·106 |
| VNTR14 [14] | TGTTTTC | 7 | 423 366–423 379 | 2 | 425 607–425 627 | 3 | 7 | 0·187 |
| VNTR15 [14] | GGC | 3 | 1 967 113–1 967 127 | 5 | 261 448–261 465 | 6 | 7 | 0·149 |
| VNTR18 [14] | AGCC | 4 | 1 297 888–1 297 987 | 25 | 1 288 655–1 288 730 | 19 | 45 | 0·959 |
| VNTR19 [14] | GCTT | 4 | 1 892 701–1 892 836 | 34 | 1 878 081–1 878 164 | 21 | 29 | 0·943 |

VNTR, Variable number tandem repeat.

Table 2. *Number of types and D values obtained by different combination of variable number tandem repeat (VNTR) loci in 215* N. meningitidis *strains*

| Combination of loci* | No. of type | *D* value | Combination of loci | No. of type | *D* value |
|---|---|---|---|---|---|
| All sixteen loci | 203 | 0·994 | 1 + 5 + 6 | 154 | 0·989 |
| 1 + 2 + 3 + 4 + 5 + 6 | 203 | 0·994 | 2 + 3 + 4 | 187 | 0·992 |
| 1 + 2 + 3 + 4 + 5 | 203 | 0·994 | 2 + 3 + 5 | 184 | 0·993 |
| 1 + 2 + 3 + 4 + 6 | 199 | 0·994 | 2 + 3 + 6 | 176 | 0·992 |
| 1 + 2 + 3 + 5 + 6 | 194 | 0·994 | 2 + 4 + 5 | 179 | 0·992 |
| 1 + 2 + 4 + 5 + 6 | 191 | 0·993 | 2 + 4 + 6 | 168 | 0·989 |
| 1 + 3 + 4 + 5 + 6 | 181 | 0·991 | 2 + 5 + 6 | 167 | 0·991 |
| 2 + 3 + 4 + 5 + 6 | 193 | 0·993 | 3 + 4 + 5 | 167 | 0·988 |
| 1 + 2 + 3 + 4 | 199 | 0·994 | 3 + 4 + 6 | 155 | 0·985 |
| 1 + 2 + 3 + 5 | 198 | 0·994 | 3 + 5 + 6 | 144 | 0·986 |
| 1 + 2 + 3 + 6 | 194 | 0·994 | 4 + 5 + 6 | 135 | 0·983 |
| 1 + 2 + 4 + 5 | 199 | 0·994 | 1 + 2 | 172 | 0·992 |
| 1 + 2 + 4 + 6 | 191 | 0·993 | 1 + 3 | 156 | 0·989 |
| 1 + 2 + 5 + 6 | 194 | 0·994 | 1 + 4 | 137 | 0·986 |
| 1 + 3 + 4 + 5 | 185 | 0·992 | 1 + 5 | 127 | 0·987 |
| 1 + 3 + 4 + 6 | 181 | 0·991 | 1 + 6 | 101 | 0·979 |
| 1 + 4 + 5 + 6 | 181 | 0·991 | 2 + 3 | 158 | 0·99 |
| 1 + 3 + 5 + 6 | 181 | 0·992 | 2 + 4 | 146 | 0·987 |
| 2 + 3 + 4 + 5 | 192 | 0·993 | 2 + 5 | 137 | 0·988 |
| 2 + 3 + 5 + 6 | 189 | 0·993 | 2 + 6 | 110 | 0·982 |
| 2 + 3 + 4 + 6 | 190 | 0·992 | 3 + 4 | 129 | 0·981 |
| 2 + 4 + 5 + 6 | 184 | 0·992 | 3 + 5 | 110 | 0·982 |
| 3 + 4 + 5 + 6 | 173 | 0·989 | 3 + 6 | 90 | 0·971 |
| 1 + 2 + 3 | 193 | 0·994 | 4 + 5 | 103 | 0·978 |
| 1 + 2 + 4 | 190 | 0·993 | 4 + 6 | 81 | 0·959 |
| 1 + 2 + 5 | 193 | 0·994 | 5 + 6 | 61 | 0·945 |
| 1 + 2 + 6 | 183 | 0·993 | 1 | 40 | 0·96 |
| 1 + 3 + 4 | 173 | 0·99 | 2 | 45 | 0·959 |
| 1 + 3 + 5 | 175 | 0·991 | 3 | 29 | 0·943 |
| 1 + 3 + 6 | 170 | 0·99 | 4 | 26 | 0·914 |
| 1 + 4 + 5 | 174 | 0·99 | 5 | 17 | 0·869 |
| 1 + 4 + 6 | 159 | 0·988 | 6 | 19 | 0·69 |

* 1, VNTR1; 2, VNTR18; 3, VNTR19; 4, VNTR2; 5, VNTR4; 6, VNTR5.

different serogroups often share common STs; for example some serogroup A and C strains fall into ST7, some serogroup C and W strains fall into ST11, and some serogroup B and C strains into ST4821 [7].

Multilocus variable-number tandem-repeat (VNTR) analysis (MLVA), has proved to be highly discriminatory and provide useful information on phylogenetic relationships between several bacterial species [8–13]. MLVA with various VNTR loci has also been applied for fine typing of meningococcus isolates with varying success in differentiating between and within some STs or ST complexes [14–16]. One research group [16] screened VNTR loci from the genomes of NM strains Z2491, MC58 and FAM18, but did not use any strains from China; we therefore tested these VNTR loci in *N. meningitidis* strains from China in order to determine a suitable scheme for such strains. Moreover, we have previously reported that MLVA further discriminated between isolates of *N. meningitidis* serogroup C ST4821 and identified 112 MLVA genotypes (GTs) in 118 isolates [17]. In the present study, we extended this work using MLST and MLVA, to characterize 215 serogroup C *N. meningitidis* isolates, including 182 of closely related ST4821 complex. In addition, we show that the use of different combinations of VNTR loci is highly discriminatory for strain typing and allows the resolution of phylogenetic relationships between strains of this serogroup that have evolved over different timescales in China.

## METHODS

### Bacterial strains and molecular typing

A total of 215 *N. meningitidis* serogroup C isolates, obtained from 22 provinces in China from 2003 to

https://doi.org/10.1017/S0950268815000473 Published online by Cambridge University Press

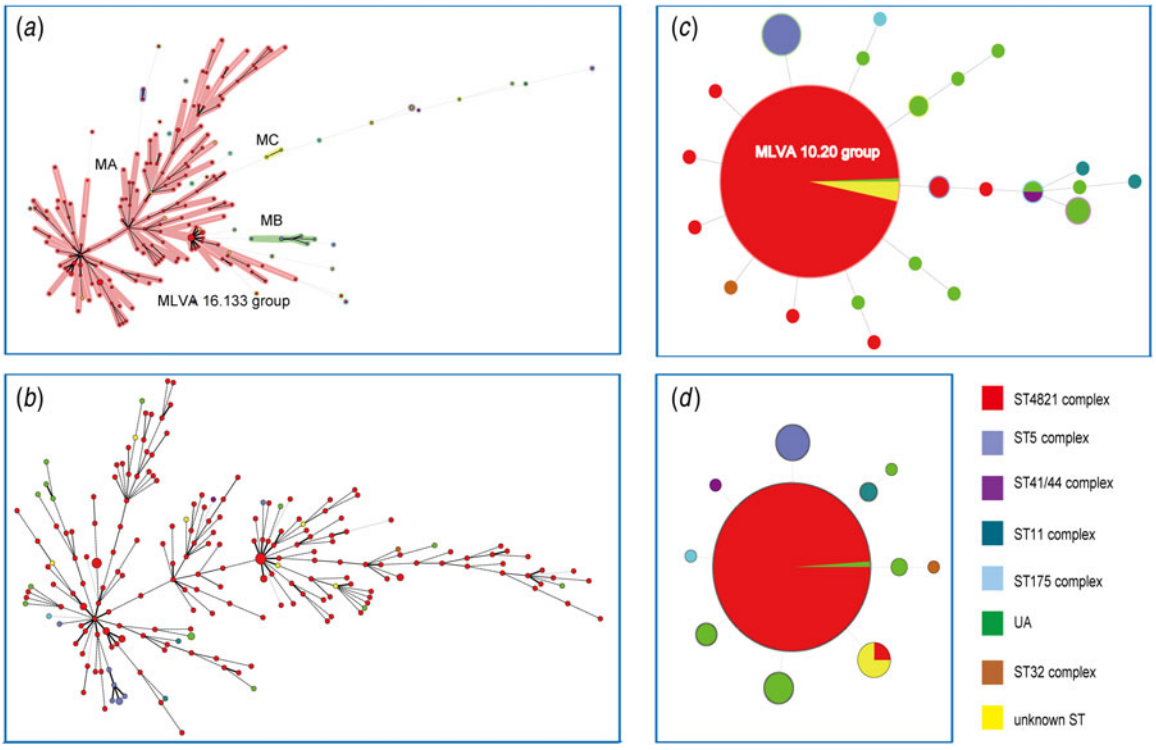


**Fig. 2.** Minimum spanning trees (MSTs) of the multilocus variable-number tandem-repeat (VNTR) analysis (MLVA) genotypes (GTs) and sequence types (STs) for 215 *N. meningitidis* serogroup C strains. The clustering was constructed by a MST algorithm. The circle size is proportional to the number of isolates belonging to the indicated MLVA genotype. The halos surrounding the circles denote different groups. Thick, solid lines represent single-locus variants; thin, solid lines represent double-locus variants. ST complexes of strains were assigned by various colours. (*a*) MLVA GTs were produced with all 16 VNTR loci used in this study, those differing by $\leqslant 5$ loci are regarded as a group. (*b*) MLVA GTs were produced with five highly variable VNTR loci (1, 2, 4, 18, 19). (*c*) MST was established with a similarity line that was drawn at 11/16 = 69% for MLVA. (*d*) MST was established with a similarity line that was drawn at 4/7 = 57% for MLST.

2012, were used in this study; DNA from each of the isolates was prepared as published previously [17]. MLST was performed as per published protocols [18] and ST data for 135 of the 215 isolates was utilized from a recent study [7]. Sixteen VNTR loci and MLVA were as described previously [17].

### Data analysis

MLST and MLVA data were analysed using BioNumerics v. 5.1 software (Applied Maths, Belgium). The numbers of repeat units for each VNTR locus were saved as 'character type' data and then subjected to cluster analysis using the minimum spanning tree (MST) method. The discriminatory power of MLVA types were assessed by Simpson's index of diversity (*D*) calculated using the VNTR diversity and confidence extractor software (V-DICE) available at the HPA website (http://www.hpa-bioinfotools.org.uk/cgi-bin/DICI/DICI.pl). The polymorphism of each locus was represented by Nei's index [19]. Dendrograms were derived by clustering with the unweighted pair-group method with arithmetic averages (UPGMA), using 'categorical' character table values. All markers were given equal weight, irrespective of the number of repeats.

## RESULTS

### Phylogenetic patterns for the *N. meningitidis* serogroup C isolates determined by MLST

Twenty-nine STs were identified in the 215 *N. meningitidis* serogroup C isolates and a dendrogram constructed with the STs displayed 11 distinct clusters. Based on the major STs of each cluster, ten ST groups were established. One hundred and eighty-four isolates formed the ST4821 complex, and 165 of these were typed as ST4821, which is the most prevalent ST in *N. meningitidis* serogroup C in China (Fig. 1*a*).

https://doi.org/10.1017/S0950268815000473 Published online by Cambridge University Press

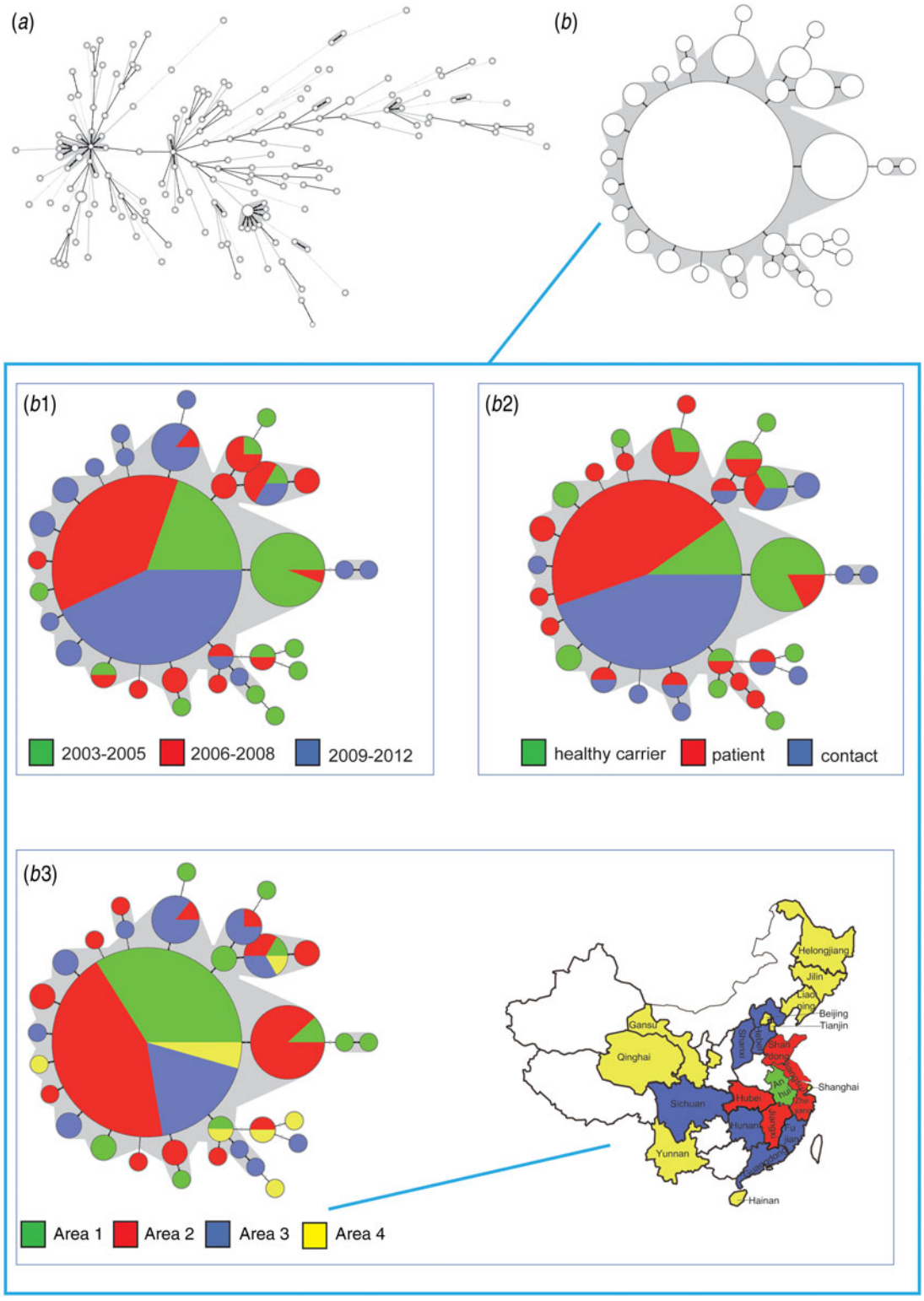


**Fig. 3.** Minimum spanning trees (MSTs) of the multilocus variable-number tandem-repeat (VNTR) analysis (MLVA) genotypes (GTs) for 182 *N. meningitidis* serogroup C ST4821 complex strains. The clustering was constructed by a MST algorithm. The circle size is proportional to the number of isolates belonging to the indicated MLVA genotype. The halos surrounding the circles denote different groups. Thick, solid lines represent single-locus variants; thin, solid lines represent double-locus variants. (*a*) MLVA GTs were generated with all the 16 VNTR loci. (*b*) MLVA GTs were generated with ten low variable VNTR loci (3, 6, 7, 8, 9, 11, 12, 13, 14, 15), those differing by single locus variation are regarded as a group. (*b*1) MST tree were constructed by year of isolation. (*b*2) MSTs were constructed by population. (*b*3) MSTs were

https://doi.org/10.1017/S0950268815000473 Published online by Cambridge University Press

Table 3. *Differences from the founder genotype MLVA10·20 for isolates of the ST4821 complex*

| No. of loci different | No. of isolates recovered in areas | | | | No. of isolates recovered in years | | | No. of isolates recovered in population | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Area 1 | Area 2 | Area 3 | Area 4 | 2003–2005 | 2006–2008 | 2009–2012 | Patient | Contact | Healthy carrier |
| 0 | 40 | 50 | 20 | 5 | 22 | 43 | 48 | 51 | 50 | 11 |
| 1 | 6 | 21 | 11 | 2 | 18 | 9 | 16 | 16 | 4 | 21 |
| 2 | 1 | 6 | 7 | 1 | 3 | 7 | 7 | 7 | 3 | 5 |
| 3 | 2 | 2 | 1 | 0 | 2 | 2 | 0 | 1 | 5 | 1 |
| 4 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 2 | 1 | 0 |
| 5 | 1 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 3 |
| Total | 51 | 80 | 40 | 11 | 49 | 62 | 71 | 77 | 64 | 41 |
| AD* | 0·45 | 0·54 | 0·83 | 1·64 | 0·69 | 0·53 | 0·42 | 0·53 | 0·53 | 1·20 |

* AD (average distance) means the average numbers of loci different from the founder genotype MLVA10·20, these values were calculated as follows: distance = $\sum$ [(no. of loci different) × (no. of isolates)]/total no. of isolates.

### Level of polymorphism and discriminatory power of MLVA loci

Of the 16 VNTR loci employed in MLVA of the 215 *N. meningitidis* serogroup C isolates, six loci (1, 2, 4, 5, 18, 19) displayed high variability, giving a Nei's diversity index >0·6. The remaining 10 loci (3, 6, 7, 8, 9, 11, 12, 13, 14, 15) were less variable (Nei's index <0·5) (Table 1).

### Phylogenetic patterns for the *N. meningitidis* serogroup C isolates determined by MLVA

Each MLVA result generates one digital combination, and based on these different combinations, GTs were identified and named starting with the numbers of VNTRs that were used in that scheme. Based on all 16 VNTR loci, 203 GTs were discriminated in the collection with a *D* value of 0·994 (Table 2, Fig. 2*a*), and these GTs were named from MLVA16·1 to 16·205 (without MLVA16·8 and 16·9). The discriminatory power of the combination of the six highly variable loci was equal to that observed for all 16 loci. We further compared the discriminatory power of all possible combinations based on the six highly variable loci and obtained *D* values between 0·945 and 0·994; 40 combinations gave a *D* value >0·99, and only one was <0·95 (Table 2). Interestingly, a combination of five highly variable loci (1, 2, 4, 18, 19) displayed the same discriminatory power as the six highly variable loci, and all of the 16 loci (Fig. 2*b*). As a consequence, these five highly variable loci were used to analyse the diversity of all isolates in the study.

MLVA using the 10 low variable loci identified 55 GTs which grouped into eight distinct clusters, on the dendrogram (Fig. 1*b*); the earliest recovered isolate was used to define the GT group and was considered to be the founder of that group. One hundred and eighty-four isolates fell into the MLVA10·20 group, being the largest GT group of *N. meningitidis* serogroup C strains in China. Most clusters differentiated by MLST tree shared the same set of isolates with the corresponding GT groups by MLVA. For example, isolates of ST4821 group were classified into MLVA10·20 and MLVA10·51 groups, ST5 group with MLVA10·44, ST11 with MLVA10·55, ST5568 with MLVA10·3 group, and ST5542 with MLVA10·12 (Fig. 1*b*).

The phylogenetic tree constructed using the MST algorithm with the 16 loci represented many distinct clusters. Cluster MA, named as MLVA16·133 group, comprised 175 strains of ST4821 complex and six isolates without MLST results, whereas cluster MB comprised seven strains of ST5 complex, and

constructed by geographical area. According to the neighbouring relationship between other provinces and Anhui, four areas were named and assigned by different colours: area 1 represents Anhui, the province where the founder strain MLVA10·20 was isolated; area 2 represents neighbouring provinces to Anhui, including Shandong, Jiangsu, Zhejiang, Jiangxi, and Hubei; area 3 represents provinces located at a one-province interval to Anhui, including Shanghai, Fujian, Guangdong, Hunan, Sichuan, Shanxi and Hebei; area 4 represents provinces located at more than two-province intervals to Anhui, including Beijing, Tianjin, Liaoning, Jilin, Heilongjiang, Gansu, Qinghai, Yunnan and Hainan.

https://doi.org/10.1017/S0950268815000473 Published online by Cambridge University Press

cluster MC four strains of ST5542 (Fig. 2*a*). The phylogenetic pattern established for the isolates with MLVA data was similar to that determined by MLST, but there were some discrepancies, for example, some isolates belonging to ST11 and ST4821 were separated in MLVA (Fig. 2*c*, *d*).

### Phylogenetic patterns for the 182 ST4821 complex isolates determined by MLVA

A total 171 MLVA GTs were identified within the 182 ST4821 complex isolates based on all 16 loci (Fig. 3*a*). When the six highly variable loci were excluded, the phylogenetic pattern for the MLVA16·133 group appeared to be relatively monomorphic (Fig. 3*b*). MST analysis with the remaining loci showed that ST4821 complex strains fell into the MLVA10·20 group. Differences from the founder strain of MLVA10·20 were calculated according to geographical area, year of isolation, and population for all 182 ST4821 complex strains (Table 3), and corresponding MSTs were constructed (Fig. 3: *b*1, *b*2, *b*3). This showed that the average distance between clusters increased as the interval became larger; from 0·45 loci for isolates recovered in geographical area 1 to 1·64 loci for isolates from area 4. No higher variation was evident with number of years passed in the last decade and average distances from the founder genotype in populations differed from 0·53 loci in patients and contacts to 1·20 loci in healthy carriers.

## DISCUSSION

*N. meningitidis* serogroup C caused seven major outbreaks in China during 2003–2005, with over 200 isolates from meningitis cases as well as patient contacts, and healthy carriers. The majority of these isolates were assigned to ST4821 complex, which spread to most provinces in China. This epidemic spread has provided an opportunity to investigate bacterial microevolution and in this study we provide the first detailed insight into the molecular epidemiology of serogroup C in China based on MLVA, which was confirmed to be a suitable tool to examine the genomic heterogeneity of the species [20]. The *N. meningitidis* serogroup C isolates were assigned into 203 GTs by the MLVA scheme with 16 VNTR loci and from these data clonal relationships were determined using MST analysis. GTs matching at $\geqslant 5$ loci were regarded as clonally related and 23 distinct MLVA groups were identified with similar phylogenetic patterns as that revealed by MLST (Fig. 2*c*, *d*); these findings are consistent with a previous study [14].

During the last decade, MLVA16·133 group including most ST4821 complex isolates has been prevalent in China. Of the 182 ST4821 complex strains, 171 GTs were distinguished by the MLVA scheme with six highly variable loci and the total panel of 16 loci and these six loci were the primary contributors to diversification among the isolates. Therefore, loci with higher variability are suitable markers for resolving closely related isolates and are sufficient to discriminate *N. meningitidis* isolates of the same serogroup and ST. In addition, ten isolates of ST4821 complex, four of UA (ST5542), one of ST5; and two of ST11 complex did not cluster into a single group even when they were of the same ST complex. This suggests that other microevolutionary forces occurred over a range of gene loci in addition to VNTR.

Despite the general correspondence of MLST and MLVA groupings, isolates of some different STs fell into the same MLVA clone when based on ten loci, as exemplified by MLVA10·48 and MLVA10·49, and ST175 and ST5750 separately which share no alleles in the seven housekeeping genes used for MLST. The relationship between the two methods therefore warrants further study as MLST is clearly suitable to define the phylogeny of *N. meningitidis* serogroup C but has a lower strain-resolving power than MLVA, and is not sufficient to distinguish between closely related strains, as was achieved by MLVA with the panel of five highly variable loci. By contrast, the ten lower variable loci were not sufficient to discriminate closely related strains, but were able to clearly separate the three clonal groups (Fig. 2*c*).

Strains of the MLVA10·20 group caused a large outbreak of meningococcal disease at a school in 2003, and thereafter MLVA10·20 and its derivative strains circulated throughout the country for more than 10 years. Since these strains were derived from a common lineage, they were used to investigate the phylogenetic pattern by using combinations of loci with different variation values. Relatively little diversity was observed in the MLVA10·20 group but this approach may be helpful in assessing patterns of evolution of other *N. meningitidis* serogroup C strains over a long time scale. Chiou *et al*. [13] drew an analogy between MLVA for *Shigella sonnei* population trees, and suggested that VNTR loci with higher variability are useful markers for establishing the 'twig and leaves' of a phylogenetic tree, while loci with lower variability are more suited to define the 'trunk' [13]. This analogy

https://doi.org/10.1017/S0950268815000473 Published online by Cambridge University Press

is also appropriate for our data on *N. meningitidis* serogroup C strains in China.

In conclusion, *N. meningitidis* serogroup C infection was not frequently reported in China before 2003 when it became more common and spread to more than 22 provinces. The epidemic was caused mainly by strains of ST4821, which was first recovered from a patient in Anhui province. This isolate was indistinguishable by PFGE from the majority of ST4821 complex strains isolated in other provinces and thus points to Anhui province being the original source of the outbreak. Our evidence suggests that the further the provinces from Anhui, the higher the genetic variation that occurred. Moreover, there were significant differences found between the genetic distance from the founder in strains recovered from patients, contacts and healthy carriers. We speculate that these microevolutions of VNTR loci of *N. meningitidis* serogroup C may mediate phase variation and induce alterations in expression of surface antigens, therefore contributing to asymptomatic and persistent colonization of the upper respiratory tract of humans [21]. Whether similar microevolution exists in other meningococcal serogroups should be studied to explore the reasons for higher carriage rates in healthy persons.

## ACKNOWLEDGEMENTS

We acknowledge the staff of the Respiratory Bacterial Laboratory of the National Institute for Communicable Disease Control and Prevention, China. This work was supported by grants from the National Key Program for Infectious Disease of China (Contract No. 2013ZX10004221), the National Basic Research Program of China (973 Program, 2011CB504900) and the National Natural Science Foundation of China (No. 81201332).

This work is attributed to State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention.

## DECLARATION OF INTEREST

None.

## REFERENCES

1. **Rosenstein NE, *et al.*** Meningococcal disease. *New England Journal of Medicine* 2001; **344**: 1378–1388.
2. **Jafri RZ, *et al.*** Global epidemiology of invasive meningococcal disease. *Population Health Metrics*. Published online: 10 September 2013. doi:10.1186/1478-7954-11-17.
3. **Harrison LH1, Trotter CL, Ramsay ME.** Global epidemiology of meningococcal disease. *Vaccine* 2009; **27** (Suppl. 2): B51–B63.
4. **Zhang X, *et al.*** Molecular characterization of serogroup C *Neisseria meningitidis* isolated in China. *Journal of Medical Microbiology* 2007; **56**: 1224–1229.
5. **Shao Z, *et al.*** Identification of a new *Neisseria meningitidis* serogroup C clone from Anhui province, China. *Lancet* 2006; **367**: 419–423.
6. **Chiou CS, *et al.*** Molecular epidemiology and emergence of worldwide epidemic clones of *Neisseria meningitidis* in Taiwan. *BMC Infectious Diseases* 2006; **6**: 25.
7. **Zhou H, *et al.*** Clonal characteristics of invasive *Neisseria meningitidis* following initiation of an A + C vaccination program in China, 2005–2012. *Journal of Infection*. Published online: 5 August 2014. doi:10.1016/j.jinf.2014.07.022.
8. **Li Y, *et al.*** Features of variable number of tandem repeats in *Yersinia pestis* and the development of a hierarchical genotyping scheme. *PLoS ONE*. Published online: 21 June 2013. doi:10.1371/journal.pone.0066567.
9. **Thierry S, *et al.*** Genotyping of French *Bacillus anthracis* strains based on 31-loci multi locus VNTR analysis: epidemiology, marker evaluation, and update of the internet genotype database. *PLoS ONE*. Published online: 5 June 2014. doi:10.1371/journal.pone.0095131.
10. **Kahlisch L, *et al.*** High-resolution in situ genotyping of *Legionella pneumophila* populations in drinking water by multiple-locus variable-number tandem-repeat analysis using environmental DNA. *Applied and Environmental Microbiology* 2010; **76**: 6186–6195.
11. **Dimovski K, *et al.*** Analysis of *Salmonella enterica* serovar typhimurium variable-number tandem-repeat data for public health investigation based on measured mutation rates and whole-genome sequence comparisons. *Journal of Bacteriology* 2014; **196**: 3036–3044.
12. **González J, *et al.*** Comparison of 2 proposed MLVA protocols for subtyping non-O157:H7 verotoxigenic *Escherichia coli*. *Diagnostic Microbiology and Infectious Disease* 2014; **78**: 328–332.
13. **Chiou CS, *et al.*** Utility of multilocus variable-number tandem-repeat analysis as a molecular tool for phylogenetic analysis of *Shigella sonnei*. *Journal of Clinical Microbiology* 2009; **47**: 1149–1154.
14. **Schouls LM, *et al.*** Multiple-locus variable-number tandem repeat analysis of *Neisseria meningitidis* yields groupings similar to those obtained by multilocus sequence typing. *Journal of Clinical Microbiology* 2006; **44**: 1509–1518.
15. **Yazdankhah SP, Lindstedt BA, Caugant DA.** Use of variable-number tandem repeats to examine genetic diversity of *Neisseria meningitidis*. *Journal of Clinical Microbiology* 2005; **43**: 1699–1705.
16. **Liao JC, Li CC, Chiou CS.** Use of a multilocus variable-number tandem repeat analysis method for molecular subtyping and phylogenetic analysis of *Neisseria meningitidis* isolates. *BMC Microbiology* 2006; **6**: 44.

17. **Shan X, *et al.*** Genetic diversity of *Neisseria meningitidis* serogroup C ST-4821 in China based on multiple-locus variable number tandem repeat analysis. *PLoS ONE*. Published online: 6 November 2014. doi:10.1371/journal.pone.0111866.
18. **Zhang J, *et al.*** Molecular characteristics of *Neisseria meningitidis*isolated during an outbreak in a jail: association with the spread and distribution of ST-4821 complex serogroup C clone in China. *Biomedical and Environmental Sciences* 2013; **5**: 331–337.
19. **Malorny B, Helmuth R.** Multi-locus variable-number tandem repeat analysis for outbreak studies of *Salmonella enterica* serotype enteritidis. *BMC Microbiology* 2008; **8**: 84.
20. **Törös B, *et al.*** Evaluation of molecular typing methods for identification of outbreak-associated *Neisseria meningitidis* isolates. *APMIS* 2013; **121**: 503–510.
21. **Alamro M, *et al.*** Phase variation mediates reductions in expression of surface proteins during persistent meningococcal carriage. *Infection and Immunity* 2014; **82**: 2472–2484.