MICROBIAL GENOMICS

RESEARCH ARTICLE

Brynildsrud *et al.*, *Microbial Genomics* 2019;5

DOI 10.1099/mgen.0.000290

MICROBIOLOGY SOCIETY

OPEN DATA OPEN MICROBIOLOGY

Gauging the epidemic potential of a widely circulating non-invasive meningococcal strain in Africa

Ola Brønstad Brynildsrud[1,2,*], Vegard Eldholm[1], Adelina Rakhimova[1], Paul A. Kristiansen[3] and Dominique A. Caugant[1,4,5]

Abstract

Neisseria meningitidis colonizes the human oropharynx and transmits mainly via asymptomatic carriage. Actual outbreaks of meningococcal meningitis are comparatively rare and occur when susceptible populations are exposed to hypervirulent clones, genetically distinct from the main carriage isolates. However, carriage isolates can evolve into pathogens through a limited number of recombination events. The present study examines the potential for the sequence type (ST)-192, by far the dominant clone recovered in recent meningococcal carriage studies in sub-Saharan Africa, to evolve into a pathogen. We used whole-genome sequencing on a collection of 478 meningococcal isolates sampled from 1- to 29-year-old healthy individuals in Arba Minch, southern Ethiopia in 2014. The ST-192 clone was identified in nearly 60% of the carriers. Using complementary short- and long-read techniques for whole-genome sequencing, we were able to completely resolve genomes and thereby identify genomic differences between the ST-192 carriage strain and known pathogenic clones with the highest possible resolution. We conclude that it is possible, but unlikely, that ST-192 could evolve into a significant pathogen, thus, becoming the major invasive meningococcus clone in the meningitis belt of Africa following upcoming mass vaccination with a polyvalent conjugate vaccine that targets the A, C, W, Y and X capsules.

DATA SUMMARY

Whole-genome sequencing data are available at the European Nucleotide Archive (ENA) (https://www.ebi.ac.uk/ena), under accession number PRJEB29233. Nanopore sequencing data for strains Nig8/13 and BL16188 are separately available under run accession numbers ERR2259086 and ERR2259087. All finished assemblies are publicly available from PubMLST (https://pubmlst.org/bigsdb?db=pubmlst_neisseria_isolates) under the accession numbers presented in Table S1 (available with the online version of this article).

INTRODUCTION

Neisseria meningitidis, the meningococcus, is a human commensal colonizing the mucosa of the upper respiratory tract that can be isolated from around 5–10% of the healthy population. The bacterium is transmitted among close-contacts through airborne droplets and the prevalence of carriage tends to be higher among children and adolescents, who serve as the main reservoirs [1].

Occasionally, *N. meningitidis* may invade the bloodstream and cause life-threatening disease such as meningitis and/or septicaemia [2]. Meningococcal disease occurs endemically throughout the world, but in a region designated as the meningitis belt in sub-Saharan Africa, outbreaks and epidemics occur every year in the dry season. During these epidemics, carriage prevalence of the outbreak strain can increase by a factor of 10 or more [3].

Except for very rare cases [4–6], disease-causing meningococci are encapsulated and 12 known serogroups are defined based on the structure of the capsular polysaccharide [7]. Six of these serogroups, A, B, C, W, X and Y, are responsible for nearly all cases of disease worldwide [8]. In contrast,

Received 03 June 2019; Accepted 02 August 2019; Published 27 August 2019

Author affiliations: [1]Division for Infection Control and Environmental Health, Norwegian Institute of Public Health, Oslo, Norway; [2]Department of Food Safety and Infection Biology, Faculty of Veterinary Science, Norwegian University of Life Science, Oslo, Norway; [3]Coalition for Epidemic Preparedness and Innovation, Oslo, Norway; [4]WHO Collaborating Centre for Reference and Research on Meningococci, Norwegian Institute of Public Health, Oslo, Norway; [5]Department of Community Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway.

***Correspondence:** Ola Brønstad Brynildsrud, olbb@fhi.no

Keywords: whole-genome sequencing; nanopore sequencing; ST-192 meningococci; capsule null; pathogen evolution; carriage; invasiveness.

Abbreviations: ANI, average nucleotide identity; bp, base pairs; cc, clonal complex; ENA, European Nucleotide Archive; LOS, lipooligosaccharide; MLST, multilocus sequence typing; rMLST, ribosomal multilocus sequence typing; ST, sequence type.

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Three supplementary tables are available with the online version of this article.

meningococci isolated from the oropharynx of healthy individuals are often unencapsulated [9]. Either down-regulation of capsule gene expression, phase variation in the capsule synthesis genes or inactivation of genes in the capsule gene cluster (*cps*) [10] might cause this. Finally, in some instances, the genetic inability of strains to produce a capsule is a consequence of a lack of regions of the capsule gene cluster. In these 'capsule null' strains, the *cps* operon is replaced by a noncoding region (*cnl* region), resembling genetic regions seen in other neisserial species, such as *Neisseria gonorrhoeae* and *Neisseria lactamica* [10, 11].

While *cnl* carrier isolates occur in various genetic lineages of meningococci [12], the few *cnl* isolates described as a cause of invasive disease belong to sequence types (STs) ST-198 [clonal complex (cc) 198], ST-845 (cc 845) and ST-192 (cc 192) [4, 13, 14]. The ST-192 cc is commonly identified in meningococcal carriage studies in Africa [15, 16], but rarely elsewhere in the world. In a carriage study performed in Ethiopia in 2014, we identified ST-192 isolates in nearly 60% of the healthy carriers [17]. Factors that could explain the success of ST-192 include more efficient adhesion and consequently entry, due to no masking of adhesins by a bacterial capsule [18], and the deletion of the *fetA* gene, which encodes an immunogenic surface protein that is a potential target for the host complement system [4].

We have previously shown how a benign meningococcal strain could evolve into a hypervirulent lineage through the acquisition of a few virulence genes, illustrating that minor genetic changes in a microbe can have major public-health consequences [19]. With the ongoing development of multivalent conjugate meningococcal vaccines to be implemented through large-scale campaigns in sub-Saharan Africa [20], it is essential to estimate the potential of *cnl* isolates to evolve into virulent and epidemic clones. To that purpose, we performed comparative genomic analyses on the isolates collected in the Ethiopian carriage study. We show that ST-192 isolates from healthy individuals lack several well-known virulence factors, including the full capsule operon, iron-acquisition machinery (including the deletion of *fetA*, encoding an enterobactin receptor) and the meningococcal disease-associated phage Φ (MDAΦ). These operons are only functional if acquired through homologous recombination of a complete and intact operon, rather than piecemeal horizontal transfer of smaller units. Based on the size of the capsule gene cluster missing from the Ethiopian ST-192 genomes, which exceeds the size of fragments typically acquired by this group, it seems unlikely that ST-192 could quickly evolve into an epidemic clone.

Impact Statement

Meningococcal meningitis causes hundreds of thousands of deaths each year. However, the majority of those infected carry the bacterium in the nasopharynx without ever developing symptoms. Transmission occurs largely through these asymptomatic carriers. Whether an infection leads to clinical meningitis is thought to be determined almost exclusively by specific genetic properties harboured by certain lineages of meningococci. With a massive vaccination campaign targeting the major capsulated forms of the bacterium about to be rolled out, concern has been raised as to whether currently non-pathogenic strains can develop epidemic potential through genetic recombination. The current study evaluates the epidemic potential of sequence type (ST)-192, the most common non-pathogenic lineage in the meningitis belt of Africa. We conclude that it is unlikely that ST-192 will evolve into an epidemic clone. Our findings are of particular importance for surveillance and preparedness against meningitis, vaccine development and eradication campaigns.

METHODS

Sample collection and bacterial identification

Oropharyngeal swabs were collected from 7479 healthy volunteers (1–29-years-old), as described by Bårnes *et al.* [21]. Sampling took place in a 6 month period in 2014 in the district of Arba Minch (South Ethiopia) in four kebeles (the smallest administrative unit within a district): Genta Mechie, Zigiti Mechie, Gatse and Kolla Shelle. The 492 isolates putatively confirmed as *N. meningitidis* at the Norwegian Institute of Public Health were included in this analysis.

Illumina sequencing

Isolates were grown overnight on blood agar plates at 37 °C in an atmosphere of 5% CO_2. Genomic DNA was extracted using an automated MagNAPure isolation station and MagNAPure 96 DNA and Viral NA small volume kit (Roche), according to the manufacturer's instructions. DNA concentration was measured using a Qubit fluorometer (Invitrogen, Thermo Fisher Scientific). Sequencing libraries were generated using a KAPA HyperPlus kit (KAPA Biosystems, Roche) and NEXTflex DNA Barcodes (Bioo Scientific), as previously described [21]. Clean-up and size selection were performed using magnetic AMPureXP beads (Beckman Coulter). DNA was then quantified using a Qubit fluorometer and the size of fragments was measured using an Agilent 2100 Bioanalyzer system (Agilent Technologies). Library samples were pooled and sequenced on the Illumina MiSeq platform with MiSeq reagent kits v2 500-cycles (Illumina).

Post-sequencing analysis

Trimmomatic v0.36 [22] was used to trim and filter FASTQ files, with the KAPA adapters added to the filtering library, using the previously described settings [21]. Meningococcal genomes were assembled *de novo* using the software SPAdes v3.11.1 [23] in careful mode, using k-mer sizes 77, 99 and 127. Kraken 0.10.5-beta [24] was used to assign taxonomic labels. Fourteen isolates were excluded due to indicated contamination at this stage, leaving the total number of isolates included in the study as 478. Contigs shorter than 500 bp and contigs

with a mean k-mer-coverage ≤5 were additionally filtered out using an in-house script. The final assembly files consisted of a median number of 109 contigs/sample (range: 43 to 654 contigs) with a mean length of 18878 bp, covering the ~ 2.2 Mb of the *N. meningitidis* genome. The estimated mean coverage across all sequencing runs was 99.8 %.

Oxford Nanopore sequencing

Long-read sequencing was performed on two ST-192 isolates from Ethiopia, AR00441 and AR06000 (collected in 2014), and additionally on the following three capsulated strains collected as part of a 2010 carriage study in Burkina Faso: DL13591 (serogroup Y), DL20404 (serogroup W) and KL11168 (serogroup X). High-molecular-mass DNA was extracted from cultured bacteria using the genomic-tip 20/G kit and buffer set (Qiagen). Sequencing libraries were prepared employing the Oxford Nanopore rapid barcoding kit and 1D reads generated on the MinION platform using the R9.4 (FLO-MIN106) flow cell. The reads were demultiplexed and basecalled with Albacore version 2.3.4 (https://nanoporetech.com/community). Porechop version 0.2.3_seqan2.1.1 (https://github.com/rrwick/Porechop) was used to remove remaining adapters and split chimeric reads prior to assembly. Hybrid assembly employing both Oxford Nanopore and Illumina reads was carried out with Unicycler version 0.4.6 [25]. The 'normal' bridging mode was used. This resulted in the following completely resolved circular genome sizes: 2063454 base pairs (bp) (ST-192 strain AR00441), 2056963 bp (ST-192 strain AR06000), 2185934 bp (serogroup Y strain DL13591), 2168128 bp (serogroup W strain DL20404) and 2188654 bp (serogroup X strain KL11168). These genomes were compared with two complete genomes of ST-10217 isolates published previously, a non-groupable carrier strain BL16188 and a serogroup C invasive strain Nig8/13 [19]. All contigs were rotated so as to start with the *dnaA* (Uniprot Q9JXS7) gene.

Genome comparison and phylogenetic analyses

Due to differences in recombination patterns and synteny, we used different methods for phylogenetic analyses of the full dataset and of the ST-192 isolates only. For the full dataset, where we expected high levels of recombination, large differences in gene content and relatively low synteny, we used an alignment of concatenated alleles from the PubMLST *N. meningitidis* core genome multilocus sequence typing (MLST) v1.0 scheme. Missing alleles were filled with gap characters. Out of the 1605 loci included in this scheme, only 42 (2.6%) were present in <95% of isolates. For the subset of ST-192 genomes, where we expected a more clonal relation, we used the program parsnp v1.2 [26] to construct a whole-genome alignment from *de novo* assembled contigs, and the circularized genome of strain AR00441 was used as a reference. In this latter case, we filled all non-core regions with ambiguous ('N') sequence. Gubbins version 2.3.4 [27] with default parameters was used to identify and exclude recombination tracts from the whole-genome alignments prior to phylogenetic analyses. As outgroups in this analysis, we used *N. lactamica* (AR01252_T8 – BIGS ID no. 44922) for the full dataset and OL24904 (BIGS ID no. 56376) for the ST-192 subset. Within ST-192, we gathered the length of each recombination block from the Gubbins output GFF file. Phylogenetic trees were recreated using the program IQ-TREE version 1.6.7 [28] with 1000 bootstraps of the ultrafast bootstrap algorithm [29]. The evolutionary model used was the generalized time-reversible model with four gamma categories, as favoured by the program bModelTest [30]. Phandango [31] was used for visualization of inferred recombination hotspots. In this analysis, we additionally included the ribosomal MLST (rMLST) score as metadata.

For the analysis of gene content among closed and circularized genomes, we created a custom pan-genome MLST scheme by including all loci that were present in at least one strain. The full gene list on PubMLST at the time of writing included 3048 loci, and 2453 loci were found in at least one of the closed genomes. The scheme was used in the genome comparator function of PubMLST [32]. The allelic calls were then used as external gene calls in the visualization software suite Anvi'o version 5.2 [33].

Data availability

Contigs and metadata used in this study are freely available at PubMLST under the BIGS accession numbers listed in Table S1. Table S1 also lists the accession numbers for the Illumina whole-genome sequencing data at the European Nucleotide Archive (ENA). All Illumina sequencing data are available at ENA under the project accession number PRJEB29233. The sequence data from the South African ST-192 samples are available at PubMLST under the BIGS identifiers 41961 and 41860. Nanopore sequencing data of the strains Nig8/13, BL16188, AR00441 and AR06000 are available from ENA under run accession numbers ERR2259086, ERR2259087, ERR3336386 and ERR3336387, respectively. The two former sequences are part of study accession number ERP106106, the two latter of PRJEB29233.

RESULTS

Overall genetic diversity

Among the 478 Ethiopian isolates included in the study, 346 (72.3%) were capsule null (*cnl*), i.e. completely missing the capsule operon, and 8 (1.7%) were non-groupable, i.e. missing various parts of the capsule operon. Genogroup X dominated among the remaining isolates (70; 14.6%), although a handful of these displayed indels in some of the capsule genes. (Two isolates had a truncated *csxA* gene, one isolate lacked a *csxA* gene, one had a truncated *csxC* gene and one isolate lacked the *ctrABCD* genes.) Other genogroups, in descending order, were W (28; 5.9%), Y (18; 3.8%), C (6; 1.3%) and B (2; 0.4%). Among the 354 isolates lacking a complete set of capsule genes (including non-groupable isolates), a clear majority, 278 (78.5%) belonged to ST-192, and 280 belonged to the ST-192 cc, which included ST-11598 and ST-11642. Remaining isolates with a *cnl* genotype belonged to ST-53 (36/10.2%, 46/13.0% in cc53, which included ST-2075, ST-7389 and

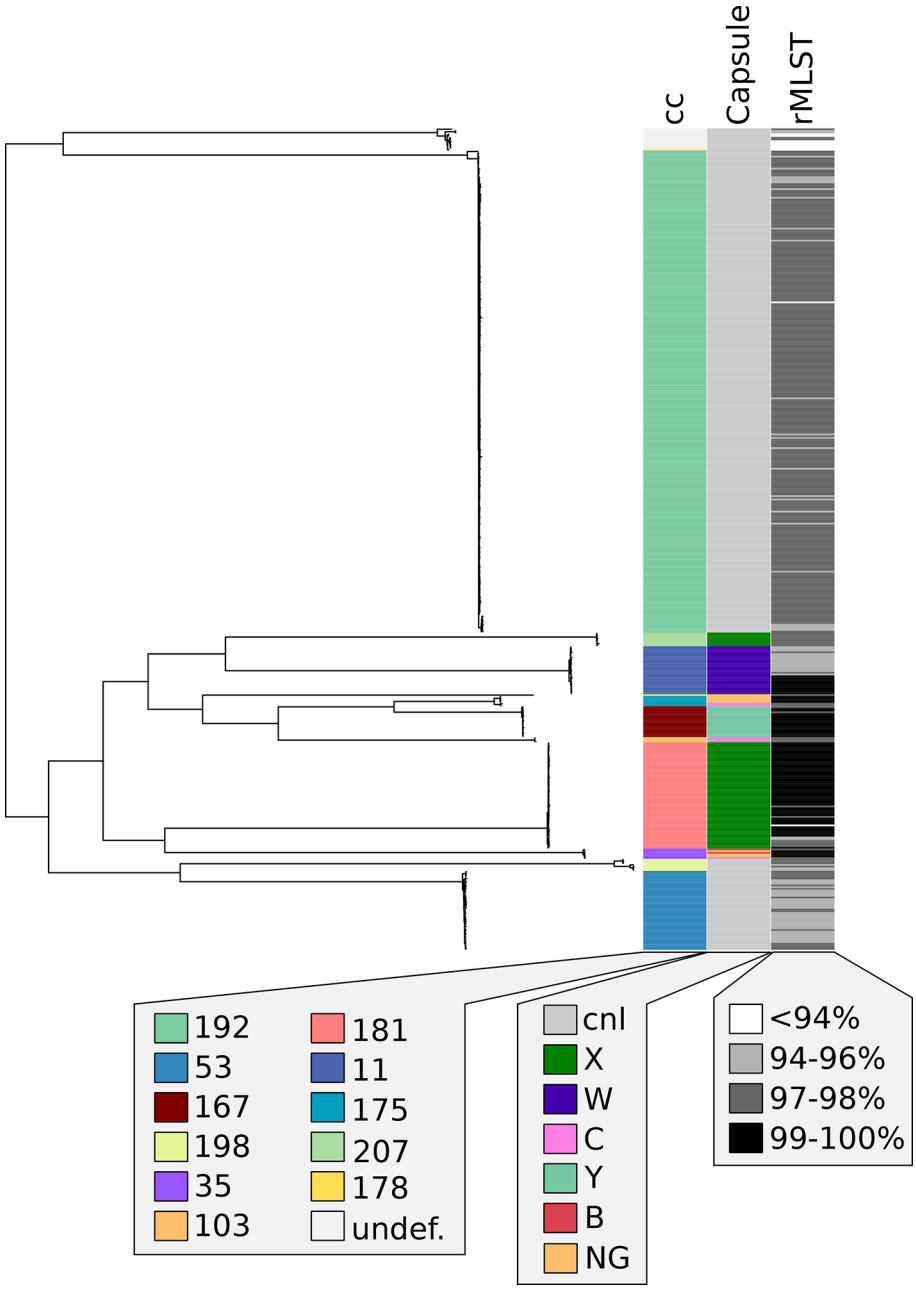

Fig. 1. Maximum-likelihood phylogenetic tree of all isolates sequenced in this study, with cc, capsule type and *N. meningitidis* rMLST identity annotated for each genome. Capsule cnl = capsule null; NG = non-groupable. Undef. = Clonal complex not defined.

ST-11594), ST-198 (4/1.1%, 7/2.0% in cc198, which included ST-11741), or to a series of closely related STs not assigned to a cc (STs 11587, 11593, 11595, 11597, 11617 and 11638, together 13/3.7%). The eight non-groupable capsule genotypes (missing or truncated in one or several genes of the operon) were ST-175 (4; 1.1%), ST-35 (3; 0.8%) and ST-178 (1; 0.3%).

ST-192 formed a very distinct clone within the full phylogeny (Fig. 1), with an intra-ST average nucleotide identity (ANI) of 99.9%, versus a 97.3% ANI of ST-192 to other STs. The ST-192 cc clustered first with the six closely related genotypes of the 13 *cnl* isolates not assigned to a cc described above. We could only find evidence of 23 recombination events within ST-192, with a median size of 2803 nucleotides (range: 17–10213), versus a median fragment size of 834 nucleotides (range: 4–37541) in the full dataset (Fig. 2) (Table S2). At least three of these recombination events were found in all Ethiopian ST-192 and involved known virulence factor systems. (Lipopolysaccharide synthesis through *lptA*, iron uptake through *tbpA* and type IV pili through *pilDFG*.) The

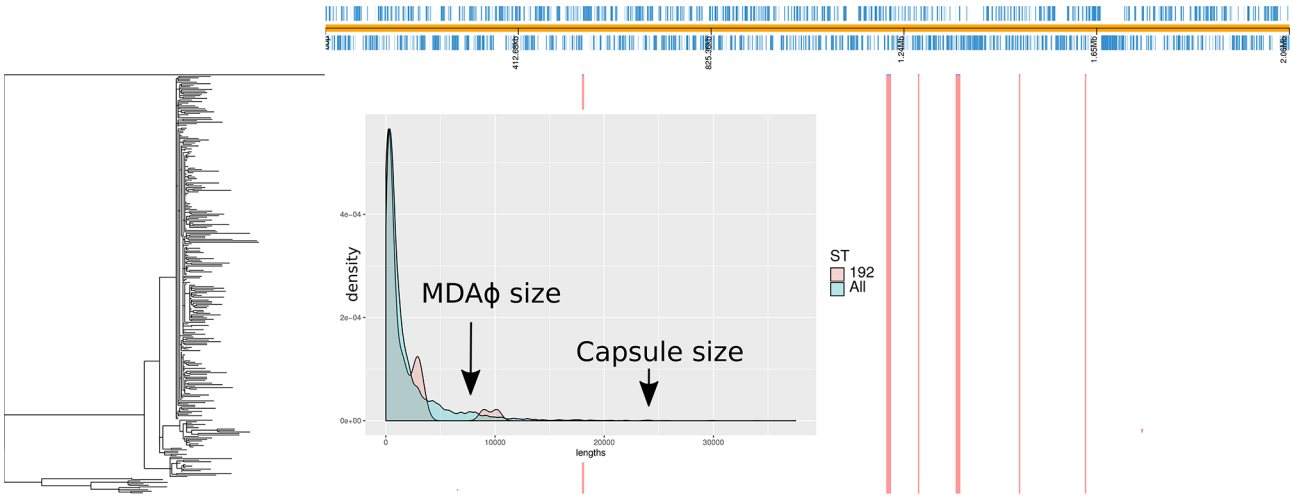

Fig. 2. Maximum-likelihood phylogenetic tree of all ST-192 sequenced in this study, with a Burkina Faso ST-192 genome (OL24904 – BIGS ID no. 56376) used as an outgroup. Recombinations across a whole-genome alignment of all genomes are visualized as vertical bars in red (multiple strains) and blue (single strain). In the middle of the figure, a kernel density plot of recombination fragment sizes for ST-192 and the full dataset is pictured. The size of the MDA Φ phage and of a full capsule operon is also shown on the horizontal (size) axis. The crop does not block any recombinations, but the vertical line visible on the top and bottom of it goes through all isolates.

phylogenetic tree topology makes it clear that the capsule locus must have been independently acquired (and lost) through horizontal transfer multiple times throughout the history of these meningococci, since both capsule X, C and non-capsule-expressing strains form polyphyletic groups in the tree.

There were some geographical patterns of ST distribution. For example, ST-11372 (49 isolates in total) was less frequent in Genta Mechie than expected ($P<1.0\times10^{-4}$, chi-square test) and ST-192 less frequent in Kolla Shelle than would be expected from the overall distribution pattern ($P<1.0\times10^{-3}$, chi-square test). However, after correcting for testing of multiple hypotheses, all significant patterns disappear. Thus, we assign no importance to this village-to-village difference in ST distribution, and consider all demographic variables to be uncorrelated to the genetic variation in our study population.

Next, we profiled the samples by means of rMLST, which assays 53 genes encoding ribosomal protein subunits [34]. This method has been shown to robustly characterize deeper evolutionary relationships as ribosomal genes are essential and under stabilizing selection. The rMLST score denotes the percentage of loci that were identical to that of the type strain. For example, an rMLST score of 98% would mean that 52/53 loci from the *rps* gene cluster were identical to that of the type strain. We found clear lineage-associated patterns of rMLST score to *N. meningitidis* as defined in PubMLST (Fig. 1). The STs with capsule X and capsule Y had the highest identity, with 46/49 of ST-11372 (X) and 6/6 of ST-11591 (also X) having 100% identity. A total of 8/8 ST-207 had 98% identity, and the final capsule X ST, ST-181, was the most variable with 95–97% identity across the 7 isolates. As for capsule Y, 16/18 had 100% identity, with the remaining 2 having a single ribosomal mutation, bringing them down to 98%.

Most ST-192 (251/278) had a 98% rMLST score, dispelling the notion that they differ wildly from more typical meningococci. We found the lowest ribosomal score within ST-53 (94–98%), as well as in the closely related atypical capsule null STs that cluster with ST-192 cc isolates (ST-11587, ST-11593, ST-11595, ST-11597, ST-11617 and ST-11638.) These were clearly atypical in that their ribosomal score varied between 90–96% and they differed a lot in MLST types despite relatively low genomic differences. In contrast to ST-192 cc isolates, all the above mentioned harboured both *porA* and *fetA* genes. Finally, capsule C and W strains exhibited highly variable rMLST scores ranging from 94–100%.

Genome comparison

With the advent of hybrid Illumina/nanopore sequencing, completely resolving genome assembly graphs (i.e. 'closing' genomes) has become much easier. Comparing closed genomes allows us to inspect the entire set of genetic elements shared and unique between invasive and non-invasive isolates. Armed with this technology, we set out to determine exactly how far off the carriage clone ST-192 is from evolving into a pathogen. Two previously closed genomes can contribute key information to this puzzle: Nig8/13 (invasive, serogroup C) (BIGS ID no. 56602) and BL16188 (non-invasive and capsule null) (BIGS ID no. 56603) [19]. BL16188 represents the closest known non-pathogenic relative to the Niger/Nigeria ST-10217 epidemic strain (represented by Nig8/13) and is, therefore, of great interest with respect to the minimum evolution needed to go from a carriage strain to invasiveness. It was previously demonstrated that conversion from a carriage to an epidemic phenotype was brought about in these two strains simply by the horizontal transfer of an intact capsule locus and the MDAΦ phage [19].

In this analysis, we also included the genomes of two South African ST-192 isolates [14] that were highly unusual in that they caused invasive disease. These two genomes have not been closed, but they have been whole-genome sequenced with Illumina technology, which is indicative of resolution of the most important parts of the genome but incomplete contig ordering and with contig breaks in highly repetitive regions or at copy number variations.

In order to identify gene elements potentially important for invasiveness, we performed a pan-genome allelic all-versus-all comparison between the two invasive South African ST-192 isolates mentioned above [36718 (BIGS ID no. 41869) and 34526 (BIGS ID no. 41961)], the ST-10217 carrier BL16188 and its closely related hyperinvasive strain Nig8/13, as well as two non-invasive, capsule null (*cnl*) closed ST-192 genomes (AR00441 and AR06000). We binned the genome contents into seven groups based on presence/absence patterns. (A) Genes found in all invasive isolates, but not in any non-invasive isolates. (B) Genes that are exclusively found in the hyperinvasive/epidemic isolate Nig8/13. (C) Genes that are found in all ST-10217, but not in any ST-192 isolates. Although BL16188 displayed a carrier phenotype, it is possible that it was 'primed' with genetic elements that could more quickly turn it into a hyperinvasive pathogen. (D) Genes that are found in all ST-192 isolates, but not in any ST-10217 isolates. (E) Genes that are found in both the carrier ST-192 isolates, and not in any other isolates. (F) Genes that are found in all isolates except for the two South African ST-192. Finally, (G) core genes; genes present in all six isolates and, therefore, irrelevant to the virulence differences (Fig. 3).

We restrict our presentation here to known virulence factors across groups A–G, but a full complement is presented in

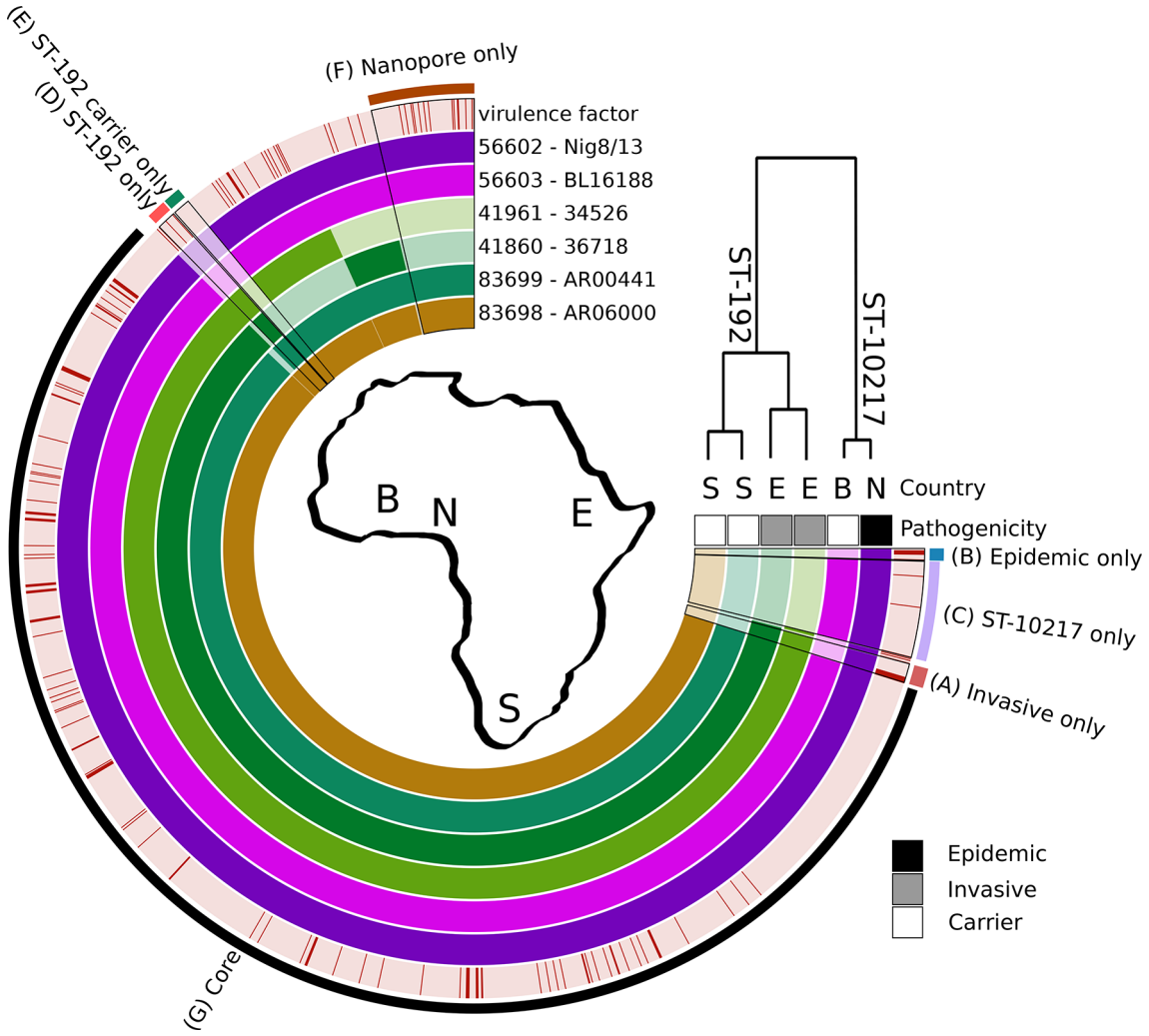

Fig. 3. Visualization of the pan-genomic gene-by-gene approach to analyse virulence factors across invasive and carrier ST-192 and ST-10217. Each gene in the full complement is organized in a radial concentric fashion. The six inner rings correspond to presence (dark colour) or absence (light colour) of a particular gene. The outer ring designates whether the gene is a known virulence factor (dark red). On the right is a phylogenetic tree of all six genomes. Groups A through G are annotated outside the outer ring. Also shown is the pathogenicity of each strain (carrier/invasive/epidemic) and the country of origin (E, Ethiopia; B, Burkina Faso; N, Nigeria; S, South Africa).

the supplementary material (Table S3). Virulence factors that were absent in all isolates (e.g. *nadA*) are not included. Note that although we only include the two closed ST-192 genomes in Fig. 3, the genome content of all 278 ST-192 isolates sequenced in this study are considered for the purposes of Table 1.

Group A – invasive versus non-invasive

Genes found in this group are potentially important for the invasiveness phenotype. Rather strikingly, the most obvious difference between the invasive and carrier isolates was that all the invasive isolates in this study had an intact MDAΦ phage, whereas the carrier isolates did not. Although the importance of this phage in ST-10217 has been previously described, it is noteworthy that the same change appears to play a major role in ST-192. In this case, the MDAΦ phage is the only relevant component that is present in the invasive and lacking in non-invasive strains. All other differences between invasive and carriage ST-192 isolates are either minor allelic changes in presumably unimportant genes, largely in genes encoding hypothetical proteins. It appears as if infection with the MDAΦ phage alone was sufficient to induce invasiveness in the South African ST-192. Notably, the presence of a capsule is not necessary for non-epidemic infections. However, the two ST-192 from South Africa were recovered 8 months apart from a single patient with a C6 complement deficiency. It is possible that a capsule just is not necessary for invasiveness in such patients.

Group B – epidemic versus non-epidemic

Two virulence factors were uniquely found in the hyper-invasive/epidemic ST-10217 (the group B comparison): an intact capsule (genes *ctrABCDG* and *csc*) and the NEIS0527 (*frpC*) gene. The *frpC* gene encodes an RTX family cytotoxin involved in adhesion, and is secreted by bacteria under iron-deprived conditions through a type I secretion system [35]. The capsule is involved in anti-phagocytosis and intracellular survival, but actually impedes colonization and mucosal adherence [18, 36]. Although not necessary for standalone invasive infections [4, 13], it is possible that the capsule is a necessary factor for a strain to become epidemic. The presence of a capsule seems to be uniformly present in all strains causing outbreaks and epidemics, whereas the opposite is not true; many strains have intact capsule operons without leading to outbreaks.

Groups C and D – ST-192 versus ST-10217

Some further differences between ST-192 and ST-10217 (comparisons C and D) were that ST-192 contains *pglB* and ST-10217 *pglB2*, which means that their O-linked glycosylation pathways differ and that the two STs produce functionally different lipooligosaccharides (LOSs) [37]. Furthermore, ST10217 has the ferric enterobactin receptor *fetA*, whereas it is lacking in ST-192. ST-192 does, however, have NEIS2529/NMB1449, which encodes a TonB-dependent receptor. It is not known whether *fetA* and NEIS2529 perform equivalent functions, although both are involved in iron acquisition. Also unique in ST-10217 was the capsule translocation system encoded by the genes *ctrEF*. Note that these genes were also found in the apathogenic BL16188, despite this isolate lacking a capsule. The *ctrEF* genes were located next to where the horizontally acquired capsule would be inserted, just upstream and on the opposite strand to *ctrG*. The fact that the apathogenic BL16188 had this system in place probably helped make the transformation to the capsule-expressing phenotype possible.

Group E – the ST-192 carrier phenotype

The carrier ST-192 (AR00441 and AR06000) did not contain any virulence factors that were absent in other groups. It is possible that they contain other genetic elements that suppress an invasive phenotype, such as for example phage defence systems, but we did not find any evidence for this in our presence/absence comparison (Table S3).

Group F – missing in invasive ST-192

These are the only two isolates for which closed genomes were not available, so it is possible that most genes missing under this label were due to technological artefacts, particularly concerning the sequencing of phase-variable genes. Many *N. meningitidis* genes have mono-/di-/oligonucleotide repeats that confound sequencing and assembly of short-read technologies. Indeed, most of the genes in this category are known phase-variable genes (*pglAEG*, *pilC*, *lgtAG*, *lbpB*) [38, 39] and are likely not missing in the invasive ST-192 at all. This does, however, illustrate some shortcomings of short-read technology used for pan-genomics.

Group G – core virulence factors

The majority of virulence factors were seen in all isolates, irrespective of whether or not they were invasive. This included systems to facilitate adherence, colonization, host invasion, immune system evasion, iron acquisition, efflux pumps, transcriptional regulators and various types of stress response proteins.

The current analysis reiterated that the difference between carriage and invasiveness can be minor from a genetic standpoint. In both ST-192 and ST-10217, the MDAΦ phage appears to be a sufficient condition for invasiveness, although we do not know if this result is generalizable to all STs of *N. meningitidis*.

DISCUSSION

Sub-Saharan Africa has suffered epidemic outbreaks of meningococcal meningitis every few years for at least a century [40]. Until now, only capsulated strains of serogroup A, B, C, W, X or Y have been capable of producing epidemics. This is likely due to the increased survivability of capsulated strains, mainly manifested through capsular protection against phagocytosis and antimicrobial peptides from the host immune system [18, 41]. Traditionally, serogroup A meningococci have dominated the meningitis

Table 1. Virulence factors in the seven different groups, as specified in Fig. 3

Group	Virulence factor	Function
A – invasive only	NEIS0023–NEIS0030 (including *tspB*)	MDAΦ phage
B – epidemic only	*ctrABCDG*, *cssABCE*, *csc*	Capsule
	frpC (NEIS0527)	Toxin
C – ST-10217 only	*pglB2*	Adhesion/pilin glycosylation
	ctrEF	Capsule translocation
	fetA	Iron uptake
D – ST-192 only	*pglB*	Adhesion/pilin glycosylation
	Ton system (NEIS2529)	Iron uptake
E – ST-192 carrier only	No known virulence factors	
F – nanopore only	*lgtAG*	Adhesion/LOS synthesis
	pilC	Adhesion/pili
	pglAEFG	Adhesion/pilin glycosylation
	lbpB, *nhba*, *tbpB*	Iron uptake
G – core	*hsf*, NEIS1772/NMB0394	Adhesion
	kdtA, *lgtBE*, *pgm*, *rfaCDEFK*	Adhesion/LOS synthesis
	pilBDEFGHIKT1T2UVXZ	Adhesion/pili
	pglCDHMOPQ	Adhesion/pilin glycosylation
	lctP, *nlpD*	Colonization
	farA, *mtrCDE*	Efflux pump
	fhbP, NEIS2103/NMB2127, *nspA*	Immune evasion
	mlp, NEIS1917/NMB1946, NEIS1933/NMB1961, *ompH*, *omp85*, *porAB*, *rmpM*, *tolC*	Invasion
	NEIS0982, NEIS1487	Invasion/infection potentiator
	bcp, *bfrAB*, *exbBD*, *fbpABC*, *fetB2*, *fur*, *hemH*, *hmbR*, NEIS0012, NEIS1282/NMB1346, NEIS1560, NEIS1658/NMB1738, NEIS1964-1966, *ton* (NEIS0338), *tonB* (NEIS1650)	Iron uptake
	fabZ, *lpxABCD*	Lipopolysaccharide synthesis
	NEIS0377	Membrane transporter
	misRS	Other/capsule expression
	NEIS1028/NMB1064, *purC* (NEIS0709)	Other/nucleotide metabolism
	nalP, *prc*	Other/protease
	NEIS0695/NMB0741	Other/protection
	mtrR	Other/transcriptional regulation
	katA	Stress/catalase
	nth	Stress/endonuclease
	NEIS1371–NEIS1373	Stress/iron-sulfur protein
	mntABC	Stress/manganese transport
	dsbA1, *dsbA2*, *dsbA3*	Stress/oxidoreductase
	sodB, *sodC*	Stress/superoxide dismutase
	NEIS0526/NMB1412, NEIS1805/NMB0365, NEIS1806/NMB0364	Toxins

belt. However, the prevalence of serogroup A disease has been drastically reduced following large-scale vaccination with the conjugate serogroup A vaccine MenAfriVac since 2010. New cc types with non-A capsule types have since emerged, notably serogroup C, which had only rarely been seen in Africa in past decades, but which has recently re-emerged and caused massive yearly outbreaks since 2013. Now, as mass vaccination with a pentavalent ACWXY conjugate vaccine is being considered [20], the major question is whether non-targeted capsule types are set to take the stage. Assuming that the ACWXY vaccine is efficient against all incorporated capsule types, the only major capsule type left to acquire is B, which is commonly found in Europe and North America, but comparatively rare in Africa. Capsule homologues have also recently been identified in several species of non-pathogenic *Neisseria* [42]. However, within the phylogenetic lineage comprising *N. meningitidis*, *N. gonorrhoeae*, *N. lactamica*, *Neisseria cinerea*, *Neisseria polysaccharea* and '*Neisseria bergeri*', only *N. meningitidis* has a capsule [42], and only intra-species horizontal gene transfer of the capsule has been demonstrated. We consider it highly improbable that ST-192 could acquire a functional capsule operon from a non-pathogenic *Neisseria*.

The relationship between carrier strains of ST-192 and a hypothetical epidemic type B strain is similar to the relationship between carrier and invasive strains of ST-10217. The ST-10217 carrier strain represented by BL16188 is thought to have become hypervirulent after acquiring a type C capsule and the MDAΦ phage. We know from the present work and that of Ganesh *et al.* [14] that ST-192 can become invasive through the acquisition of this phage, so the major unknowns are if such a strain could also (1) acquire an intact type B capsule operon, and (2) that this would increase the survivability of the strain enough to induce hypervirulence. Horizontal transfer of intact whole capsule operons is rare. ST-10217 did not completely lack capsule genes: it harboured intact alleles of the *ctrEF* genes, which are responsible for capsule translocation. Comparatively, ST-192 completely lacks all capsule genes and would have to acquire a fragment of around 20 kb to be able to synthetize a capsule on its surface (region A, *cssABC*, *csb*, *ctrG*; region B, *ctrEF*; region C, *ctrABCD*). Furthermore, a capsule operon would have the highest probability of being functional if acquired through a single recombination event (as opposed to multiple, partly overlapping events). We have analysed the length of fragments acquired through homologous recombination within carrier strains in our collection from Ethiopia, as well as specifically within the ST-192 carrier strains. From these analyses, we conclude that the acquisition of a type B capsule operon by an ST-192 member is unlikely, as this will require the incorporation of a sequence tract exceeding the size of fragments typically horizontally acquired by this group (shown in this paper), or in *N. meningitidis* in general [43, 44].

Conclusion

The extremely successful ST-192 could theoretically evolve into an epidemic clone through the acquisition of a type B capsule, *fetA*, and infection with the MDAΦ phage, although this would require a series of highly unlikely events. We, therefore, propose that despite its ubiquity, ST-192 is not particularly poised to take over as a dominant invasive clone in the wake of mass vaccination with a polyvalent ACWXY vaccine.

Funding information
D. A. C. acknowledges funding from the Norwegian Research Council, grant no. 220829. O. B. B., V. E., A. R. and P. A. K. report no funding source other than their employer. These funders had no role in the study design, data collection, data analysis, data interpretation, writing of the report nor the decision to submit the paper for publication. The corresponding author had full access to the study data and had final responsibility for the decision to submit to publication.

Author contributions
Study design: all authors. Isolate collection: P. A. K. and D. A. C. Data generation: O. B. B., V. E. and A. R. Data interpretation: O. B. B., V. E. and D. A. C. Figures and analyses: O. B. B. Manuscript preparation: O. B. B. and D. A. C., with contributions from V. E., A. R. and P. A. K. All authors critically reviewed the manuscript and gave consent for publication.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
Ethical approval for the study was obtained from the Regional Committee for Medical Research Ethics, South-East Norway (2013/849), the World Health Organization, the Armauer Hansen Research Institute (AHRI)/ All-Africa Leprosy, Tuberculosis and Rehabilitation Training Centre (ALERT) Ethics Review Committee, and the National Research Ethics Review Committee, Ethiopia. Study information was given in the local language and written informed consent was obtained from all participants, or from parents in the case of children.

Data bibliography
1. Norwegian Institute of Public Health. European Nucleotide Archive, accession no. PRJEB29233 (2019).

2. Norwegian Institute of Public Health. European Nucleotide Archive, accession no. ERP106106 (2018).

References
1. Christensen H, May M, Bowen L, Hickman M, Trotter CL. Meningococcal carriage by age: a systematic review and meta-analysis. *Lancet Infect Dis* 2010;10:853–861.

2. Stephens DS, Greenwood B, Brandtzaeg P. Epidemic meningitis, meningococcaemia, and *Neisseria meningitidis*. *The Lancet* 2007;369:2196–2210.

3. Harrison LH, Trotter CL, Ramsay ME. Global epidemiology of meningococcal disease. *Vaccine* 2009;27 (Suppl. 2):B51–B63.

4. Findlow H, Vogel U, Mueller JE, Curry A, Njanpop-Lafourcade B-M *et al*. Three cases of invasive meningococcal disease caused by a capsule null locus strain circulating among healthy carriers in Burkina Faso. *J Infect Dis* 2007;195:1071–1077.

5. Johswich KO, Zhou J, Law DKS, St Michael F, McCaw SE *et al*. Invasive potential of nonencapsulated disease isolates of *Neisseria meningitidis*. *Infect Immun* 2012;80:2346–2353.

6. Xu Z, Zhu B, Xu L, Gao Y, Shao Z. First case of *Neisseria meningitidis* capsule null locus infection in China. *Infect Dis* 2015;47:591–592.

7. Harrison OB, Claus H, Jiang Y, Bennett JS, Bratcher HB *et al*. Description and nomenclature of *Neisseria meningitidis* capsule locus. *Emerg Infect Dis* 2013;19:566–573.

8. Jafri RZ, Ali A, Messonnier NE, Tevi-Benissan C, Durrheim D *et al*. Global epidemiology of invasive meningococcal disease. *Popul Health Metr* 2013;11:17.
9. Yazdankhah SP, Caugant DA. *Neisseria meningitidis*: an overview of the carriage state. *J Med Microbiol* 2004;53:821–832.
10. Tzeng YL, Thomas J, Stephens DS. Regulation of capsule in *Neisseria meningitidis*. *Crit Rev Microbiol* 2016;42:759–772.
11. Claus H, Maiden MCJ, Maag R, Frosch M, Vogel U. Many carried meningococci lack the genes required for capsule synthesis and transport. *Microbiology* 2002;148:1813–1819.
12. Neri A, Fazio C, Ambrosio L, Vacca P, Barbui A *et al*. Carriage meningococcal isolates with capsule null locus dominate among high school students in a non-endemic period, Italy, 2012-2013. *Int J Med Microbiol* 2019;309:182–188.
13. Hoang LMN, Thomas E, Tyler S, Pollard AJ, Stephens G *et al*. Rapid and fatal meningococcal disease due to a strain of *Neisseria meningitidis* containing the capsule null locus. *Clin Infect Dis* 2005;40:e38–e42.
14. Ganesh K, Allam M, Wolter N, Bratcher HB, Harrison OB *et al*. Molecular characterization of invasive capsule null *Neisseria meningitidis* in South Africa. *BMC Microbiol* 2017;17:40.
15. Nicolas P, Djibo S, Tenebray B, Castelli P, Stor R *et al*. Populations of pharyngeal meningococci in Niger. *Vaccine* 2007;25:A53–A57.
16. Greenwood BM, Aseffa A, Caugant DA, Diallo K, Kristiansen PA *et al*. Narrative review of methods and findings of recent studies on the carriage of meningococci and other *Neisseria* species in the African Meningitis Belt. *Trop Med Int Health* 2019;24:143–154.
17. Bårnes GK, Kristiansen PA, Beyene D, Workalemahu B, Fissiha P *et al*. Prevalence and epidemiology of meningococcal carriage in southern Ethiopia prior to implementation of MenAfriVac, a conjugate vaccine. *BMC Infect Dis* 2016;16:639.
18. Spinosa MR, Progida C, Talà A, Cogli L, Alifano P *et al*. The *Neisseria meningitidis* capsule is important for intracellular survival in human cells. *Infect Immun* 2007;75:3594–3603.
19. Brynildsrud OB, Eldholm V, Bohlin J, Uadiale K, Obaro S *et al*. Acquisition of virulence genes by a carrier strain gave rise to the ongoing epidemics of meningococcal disease in West Africa. *Proc Natl Acad Sci USA* 2018;115:5510–5515.
20. Chen WH, Neuzil KM, Boyce CR, Pasetti MF, Reymann MK *et al*. Safety and immunogenicity of a pentavalent meningococcal conjugate vaccine containing serogroups A, C, Y, W, and X in healthy adults: a phase 1, single-centre, double-blind, randomised, controlled study. *Lancet Infect Dis* 2018;18:1088–1096.
21. Bårnes GK, Brynildsrud OB, Børud B, Workalemahu B, Kristiansen PA *et al*. Whole genome sequencing reveals within-host genetic changes in paired meningococcal carriage isolates from Ethiopia. *BMC Genomics* 2017;18:407.
22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
23. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
24. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014;15:R46.
25. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 2017;13:e1005595.
26. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 2014;15:524.
27. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.
28. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.
29. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 2018;35:518–522.
30. Bouckaert RR, Drummond AJ. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evol Biol* 2017;17:42.
31. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM *et al*. Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics* 2018;34:292–293.
32. Maiden MCJ, Harrison OB. Population and functional genomics of *Neisseria* revealed with gene-by-gene approaches. *J Clin Microbiol* 2016;54:1949–1955.
33. Eren AM, Esen Özcan C, Quince C, Vineis JH, Morrison HG *et al*. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 2015;3:e1319.
34. Jolley KA, Bliss CM, Bennett JS, Bratcher HB, Brehony C *et al*. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology* 2012;158:1005–1015.
35. Linhartová I, Bumba L, Mašín J, Basler M, Osička R *et al*. Rtx proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol Rev* 2010;34:1076–1112.
36. Hammerschmidt S, Hilse R, van Putten JP, Gerardy-Schahn R, Unkmeir A *et al*. Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element. *Embo J* 1996;15:192–198.
37. Viburiene R, Vik A, Koomey M, Borud B. Allelic variation in a simple sequence repeat element of neisserial *pglB2* and its consequences for protein expression and protein glycosylation. *J Bacteriol* 2013;195:3476–3485.
38. Saunders NJ, Jeffries AC, Peden JF, Hood DW, Tettelin H *et al*. Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. *Mol Microbiol* 2000;37:207–215.
39. Zelewska MA, Pulijala M, Spencer-Smith R, Mahmood H-T-NA, Norman B *et al*. Phase variable DNA repeats in *Neisseria gonorrhoeae* influence transcription, translation, and protein sequence variation. *Microb Genom* 2016;2:e000078.
40. Greenwood B. Manson lecture: meningococcal meningitis in Africa. *Trans R Soc Trop Med Hyg* 1999;93:341–353.
41. Jarvis GA, Vedros NA. Sialic acid of group B *Neisseria meningitidis* regulates alternative complement pathway activation. *Infect Immun* 1987;55:174–180.
42. Clemence MEA, Maiden MCJ, Harrison OB. Characterization of capsule genes in non-pathogenic *Neisseria* species. *Microb Genom* 2018;4:e000208.
43. Linz B, Schenker M, Zhu P, Achtman M. Frequent interspecific genetic exchange between commensal neisseriae and *Neisseria meningitidis*. *Mol Microbiol* 2000;36:1049–1058.
44. Alfsnes K, Frye SA, Eriksson J, Eldholm V, Brynildsrud OB *et al*. A genomic view of experimental intraspecies and interspecies transformation of a rifampicin-resistance allele into *Neisseria meningitidis*. *Microb Genom* 2018;4:e000222.