Article

Genomic Diversity and Chromosomal Rearrangements in *Neisseria Gonorrhoeae* and *Neisseria Meningitidis*

Boris Shaskolskiy *, Dmitry Kravtsov, Ilya Kandinov, Ekaterina Dementieva and Dmitry Gryadunov

Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 119991 Moscow, Russia

* Correspondence: Boris Shaskolskiy; b.shaskolskiy@biochip.ru

Abstract: Chromosomal rearrangements in *N. gonorrhoeae* and *N. meningitidis* were studied with the determination of mobile elements and their role in rearrangements. The results of whole-genome sequencing and *de novo* genome assembly for 50 *N. gonorrhoeae isolates* collected in Russia were compared with 96 genomes of *N. gonorrhoeae* and 138 genomes of *N. meningitidis* from the databases. Rearrangement events with the determination of the coordinates of syntenic blocks were analyzed using the SibeliaZ software, the minimum number of events that allow one genome to pass into another was calculated using the DCJ-indel model using the UniMoG program. Population-level analysis revealed a stronger correlation between changes in the gene order and phylogenetic proximity for *N. meningitidis* in contrast to *N. gonorrhoeae*. Mobile elements were identified, including Correa elements, Spencer-Smith elements (in *N. gonorrhoeae*), Neisserial intergenic mosaic elements, IS elements of IS5, IS30, IS110, IS1595 groups, Nf1-Nf3 prophages, NgoΦ1-NgoΦ9 prophages, Mu-like prophages Pnm1, Pnm2, MuMenB (in *N. meningitidis*). More than 44% of the observed rearrangements most likely occurred with the participation of mobile elements, including prophages. No differences were found between the Russian and global *N. gonorrhoeae* population both in terms of rearrangement events and in the number of transposable elements in genomes.

Keywords: *Neisseria gonorrhoeae*, *Neisseria meningitidis*, whole genome sequencing, chromosomal rearrangements, mobile elements

1. Introduction

Considerable part of modern approaches in the field of molecular biology are based on whole genome sequencing (WGS) technologies that establish the entire sequence of cellular DNA. The results of WGS and *de novo* genome assembly showed great variability in the structure of genomes even in closely related organisms. Genome rearrangements are characteristic of both eukaryotic and prokaryotic organisms [1] and changes in the order of genes can lead to changes in the properties of a microorganism [2,3]. As an example, bacteria *Escherichia coli* and representatives of the genus *Shigella* differ greatly in pathogenicity and are even assigned to different genera, but they have similar genome sequences [4–6]. *Yersinia pseudotuberculosis* and *Yersinia pestis* also have distinct pathogenicity, despite the fact that they have more than 97% identity in 75% of chromosomal genes [7]. Studies of the evolution of prokaryotes have shown that the order of genes in prokaryotes is relatively poorly preserved in the process of evolution and, as a rule, changes faster than their amino acid sequences [2,3,8,9].

Comparison of genome sequences of closely related species reveals the role of genomic rearrangements in the formation of macroscopic polymorphisms through various recombination events: insertions, deletions, inversions, and translocations. Previous studies have demonstrated the evolution of the bacterial genomes of *Yersinia spp.* [10–12], *Burkholderia spp.* [13], *Streptococcus spp.* [14,15], *Shigella spp.* [16] and the role of genomic rearrangements in the evolution and adaptation of bacteria to environmental conditions, the acquisition of resistance to antibiotics, and changes in pathogenicity. The frequency of

observed genomic rearrangements correlates with the number of mobile genetic elements and the state of repair/recombination systems in genomes [9,17]. Genomic rearrangements were studied even before the widespread use of sequencing methods using restrictases and electrophoretic methods [18]. The determination of whole genome sequences has given a new impetus to work in this direction.

The genus *Neisseria* includes two species of bacteria pathogenic to humans: *N. gonorrhoeae* and *N. meningitidis*, the causative agents of gonorrhea and meningococcal infection, respectively. Unlike gonococcus, meningococcus can be present in the human nasopharyngeal mucosa without causing invasive disease. The literature mainly describes rearrangements in two pathogenic of *Neisseria* species [19–25]; in several studies, however, a comparative whole genome analysis of pathogenic and commensal *Neisseria* was carried out [26,27], in particular, to identify genome regions associated with virulence.

As in other bacteria, changes in the order of genes in bacteria of the *Neisseria* genus can occur as a result of recombination which is divided into homologous, site-specific, and illegitimate. Homologous recombination is an exchange of aligned regions of homologous DNAs, and its frequency is determined primarily by the extent of homology. Site-specific recombination occurs between specific sequences with short homologous regions. Illegitimate recombination involves DNA sequences without homology. In *Neisseria*, as in many other species, the RecA protein, which has multiple activities related to DNA recombination and repair, plays the central role in the process of homologous recombination. At the same time, *N. gonorrhoeae* has both RecBCD and RecF DNA repair pathways [28].

Neisseria genome is characterized by the presence of various repetitive DNA fragments which serve as "hot spots" of recombination. Meningococcal genomes have a particular abundance and diversity of repetitive DNA, which contributes to both genome fluidity and physical variability. About 20% of its chromosome are made up of repetitive sequences of all kinds [22,29].

The most important sequences ("hot spots" of recombination) involved in genomic rearrangements in *Neisseria* include Neisserial intergenic mosaic elements (NIMEs), Correia repeat enclosed element (CREE), and various insertion sequences (ISs). NIMEs with a length of 70-200 bp include 50-150 bp repetitive sequences (RS elements) flanked by 20 bp inverted repeats ATTCCC(N)$_8$GGGAAT called duplicated repeat sequences (dRS3). dRS3 are present in about 200 copies in the *N. gonorrhoeae* genome and up to 700 copies in *N. meningitidis* and are the second most common repetitive sequences after the DNA uptake sequence (DUS). In addition to flanking RS elements, dRS3 can act as sites for phage integration [20,24,26,30,31]. For example, the *N. meningitidis* Z2491 strain has been found to contain 681 RS elements [30].

Correia elements consist of a characteristic core region that defines their length (104-108 or 153-157 bp), surrounded by 26 bp inverted repeats called Correia repeats, and may carry promoter sequences at their ends at position -10 and/or -35 (Black promoter, Snyder promoter). Some longer CREEs carry a recognition site for integration host factor (IHF) that affects the expression of neighboring genes [19,21,32]. *N. gonorrhoeae* contains 120-150 CREE copies dispersed throughout the genome [21,32], while *N. meningitidis* contains twice as many copies, about 250 [20]. Commensal *Neisseria*, with the exception of *N. sicca*, have a lower number of CREEs compared to pathogenic ones [26].

N. gonorrhoeae elements about 650 bp in length containing 19 bp inverted repeats 5'-CGTTTCAGACGGCATCGGG//CCCGATGCCGCCTGAAACG-3' were first described in [23]. These elements named Spencer-Smith Repeat Enclosed Elements (SSREEs) were found in different strains of *N. gonorrhoeae* (3 elements in each strain) at approximately the same positions in the genomic sequences. The inner SSREE segment has open reading frames, but they do not encode transposases. It was shown that SSREEs contributed to the appearance of inversions in strain FA1090, being a template for homologous recombination [23].

Bacteria contain various insertion sequences (ISs) which are small (from 700 to 2,500 bp) mobile DNA fragments that can move within the chromosome. IS elements have one

or two open reading frames encoding proteins responsible for functions involved in their motility (transposases). On the right and left sides, ISs are bounded by short terminal inverted repeats (IRs) [33]. IS elements are grouped into families, depending primarily on the type of transposase [34,35].

Analysis of three *N. gonorrhoeae* strains (FA1090, NCCP11945, and TCDC-NG08107), the complete genome sequences of which were determined in 2012, showed that the gonococcal genome contains at least 16 IS1016 elements about 700 bp in length [23]. In particular, IS1016 are found at both ends of DNA fragments that turned out to be inverted, indicating that these ISs are associated with large-scale chromosomal rearrangements. Moreover, shorter copies of IS1016 have been found around maf genomic islands.

To describe horizontal transfer and integration of genes, the model of minimal mobile elements (MMEs) was proposed [36]. MMEs are cassettes surrounded by conservative genes encoding proteins, the content of which differs between strains or species. Cassettes can be incorporated into the genome by homologous recombination. Comparative analysis of *Neisseria* genome sequences revealed more than 30 potential MME sites, many of which contain strain-specific genes that may have been acquired through horizontal transfer even from other genera of bacteria [36].

"Hot spots" of recombination are often associated with the locations of prophages, i.e., sets of phage genes integrated into chromosomal DNA. Prophage elements can include modules of genes for regulation, transposition, lysis, as well as phage head and tail proteins. In the study of the *N. gonorrhoeae* FA1090 genome sequence, five bacteriophages with double-stranded DNA (NgoΦ1-NgoΦ5) and four filamentous phages (NgoΦ6-NgoΦ9) have been identified [37]. Phages NgoΦ1 and NgoΦ2 have all the genes necessary for functioning, the rest of the double-stranded phages are incomplete. All four single-stranded phages are structurally similar and are associated with IS-like elements ISNgo2 and ISNgo3. It is believed that phages integrate into the genome using their own transposases. Meningococci do not have sequences of double-stranded phages present in *N. gonorrhoeae*; instead, they have Mu-like phages MuMenB and PNM1-PNM2 which are absent in *N. gonorrhoeae* [31,37].

When comparing *Neisseria* genomes *in silico*, four subtypes of filamentous prophages Nf (Neisserial filamentous phages) have been identified [38]. The genomes possessed intact copies of prophages, truncated variants, variants with internal deletions, and inverted sequences. Intact DNA of the Nf prophage, 8 kb long, is flanked by duplicated 5'-CT sequences similar in structure to the dRS3 sequence. At the right end, it has the gene encoding for the pivNM/irg transposase of the RNase H (retroviral integrase) superfamily.

Thus, studies have shown that during the evolution of *Neisseria spp.* numerous chromosomal rearrangements occur, and the structural elements which are most often involved in rearrangements are repetitive sequences and prophages. To determine the time frame during which changes in the order of genes can occur, the rearrangements in *N. gonorrhoeae* strains exposed to stress factors ((39-41)°C, nalidixic acid) has been determined in laboratory conditions [25,32]. Structural rearrangements associated with CREEs and prophage-associated elements have appeared after 8 weeks of growing the culture under stress and reseeding every three days.

The analysis of genomic rearrangements in bacteria of the genus *Neisseria* described in the literature was performed in several strains for which WGS data were available at the time of the study. Recently, there has been a significant increase in information about the structure of genomes based on WGS using long reads, which makes it possible to reliably determine the order of genes on a chromosome. In this work, we studied genomic rearrangements in *N. gonorrhoeae* and *N. meningitidis* at the population level using a wide range of genomes obtained from available databases. In our study, we performed the WGS for *N. gonorrhoeae* isolates collected in the Russian Federation and used the genome sequences of *N. gonorrhoeae* and *N. meningitidis* isolates sequenced *de novo* from the GenBank and BV-BRC (bv-brc.org) databases.

The aim of this work was to study the positions of rearrangements in the genomes of *N. gonorrhoeae* and *N. meningitidis*, changes in the order of genes with an analysis of the reasons for the transition to the current gene order, and to compare rearrangements in these two bacterial species with an assessment of the dependence of gene order on phylogenetic proximity.

2. Results

2.1. Whole genome sequencing of N. gonorrhoeae isolates collected in Russia

The results of WGS and genome assembly of 50 Russian *N. gonorrhoeae* isolates showed the presence of a 2.2 million bp circular chromosome in all strains. In all samples, 4,207 bp cryptic plasmids were identified. The isolates belonged to the G807 (NG-MAST 228, 807, 1544, 5941, 9570, 9576, 13054), G1993 (NG-MAST 1993, 5714, 14006, 14018) and G14942 (NG-MAST 9476, 9842, 14942) genogroups common in the Russian Federation [39]. The MLST typing showed the presence of eight MLST types: 1594, 1901, 11177, 13759, 14009, 14013, 15101, and 15664. A list of isolates and their characteristics are given in Table S1 of the Supplementary Materials.

Combining genome sequences generated using two different WGS platforms at the data processing stage allowed to obtain a high density of overlapping reads in the final assembly (hybrid assembly), which made it possible to process data without using a reference genome. *De novo* genome assembly using long reads enabled us to obtain data on the order of genes on the *N. gonorrhoeae* chromosome and to carry out an analysis of genomic rearrangements. As a control, we sequenced the DNA of the *N. gonorrhoeae* strain ATCC 49226 (Bioproject PRJNA768989, SAMN28870224) in our laboratory and compared the results with the sequence of this strain from the database. Nucleotide sequences were 99.97% identical without any changes in gene order.

2.2. Phylogenetic network of N. gonorrhoeae and N. meningitidis isolates

A comparative analysis was carried out using *de novo* assembled genomes of *N. gonorrhoeae* and *N. meningitidis* isolates. The sample of *N. gonorrhoeae* genomes consisted of 50 genomes of isolates collected in Russia and sequenced in this work and 96 genomes of isolates collected worldwide and available from public databases, including 15 WHO genomes, a total of 146 genomes. *N. gonorrhoeae* isolates belonged to 45 different MLST types, the most common of which were 1594, 7363 and 1901 (Table S1).

A sample of *N. meningitidis* genomes derived from 138 isolates contained genomes from 26 countries, including countries of the African meningitis belt (26 sub-Saharan countries from Senegal to Ethiopia). The isolates belonged to 38 MLST types, the most common of which were MLST 11, 23, 2881 (Table S1).

Based on the obtained sequences of the core genomes of *N. gonorrhoeae* and *N. meningitidis* isolates, we constructed a phylogenetic network that characterizes the relationship between isolates of both species (Figure 1a-c). As can be seen from Figure 1(c), the genetic diversity of *N. gonorrhoeae* and *N. meningitidis* isolates differs significantly: the population of *N. meningitidis* is characterized by a greater diversity.

The measure of genome diversity can be characterized by the value of π which determines the average number of pairwise nucleotide differences among genome sequences [40]. The calculated value of π for *N. gonorrhoeae* (2.9×10^{-3}) was about 6 times less than that for *N. meningitidis* (17.3×10^{-3}).
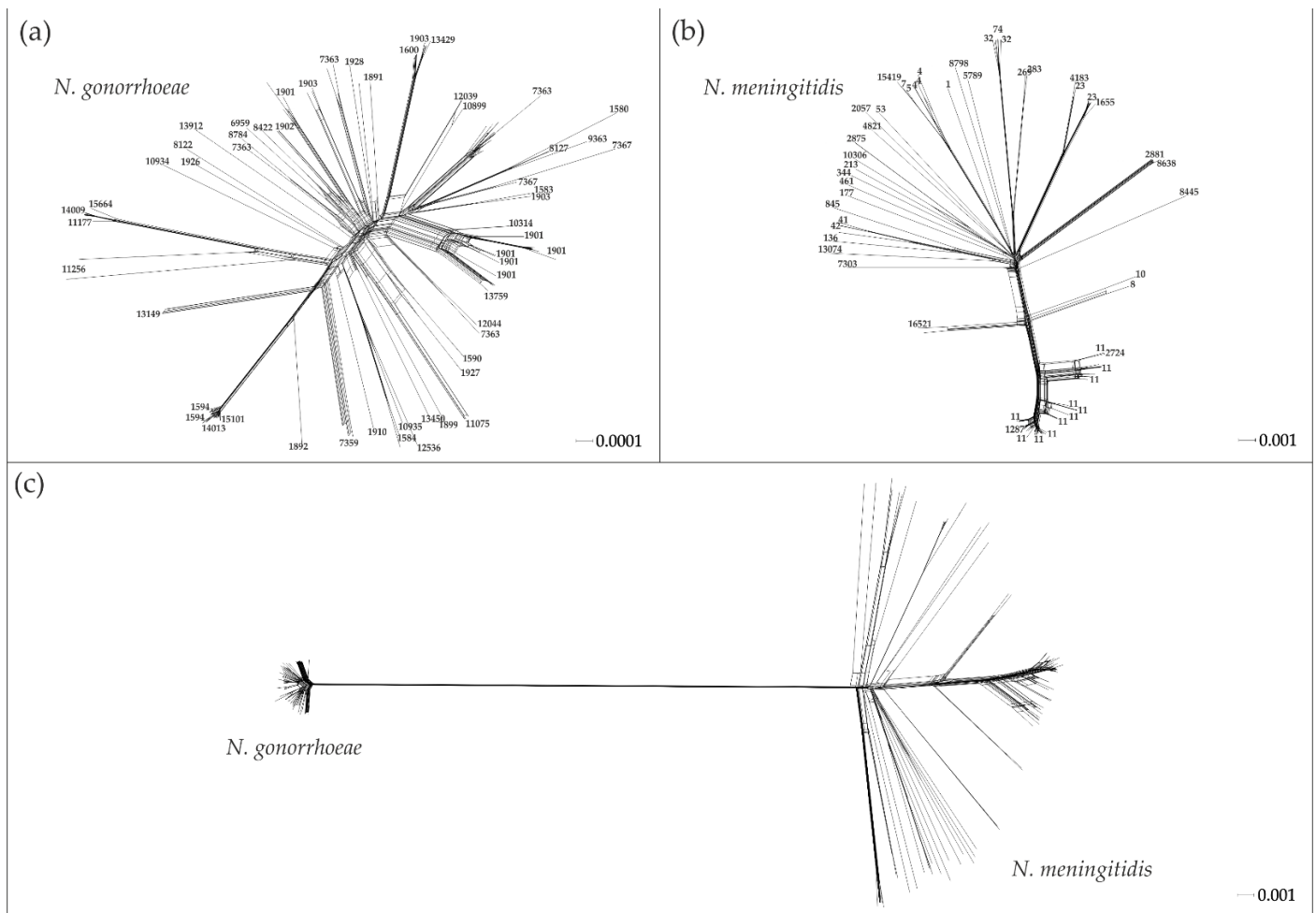
Figure 1. Phylogenetic network based on core genomes for isolates of *N. gonorrhoeae* (a), *N. meningitidis* (b), together for *N. gonorrhoeae* and *N. meningitidis* (c). Numbers denote MLST types.

2.3. Estimation of the number of rearrangement events in genomes of N. gonorrhoeae and N. meningitidis

To determine the number of rearrangement events in genomes, we identified syntenic blocks using the SibeliaZ program. As the minimum length of syntenic block, we used the threshold value of 1248 since the length of 75% of the genes in the genomes of *N. gonorrhoeae* and *N. meningitidis* is less than 1,248 base pairs. A smaller threshold value introduces errors, showing the presence of changes within genes, while a larger value deals with syntenic blocks of a larger size, in which rearrangement events for individual genes are not taken into account.

Based on the above, syntenic blocks and their locations in *N. gonorrhoeae* and *N. meningitidis* genomes were determined. In 146 *N. gonorrhoeae* genomes, 220 syntenic blocks were identified. The number of common blocks, i.e., blocks found in all analyzed genomes, was 51. The total number of blocks per genome varied from 65 to 86, averaging 75. For 138 *N. meningitidis* genomes, 327 syntenic blocks were identified. The number of common blocks found in all analyzed genomes was 126. The total number of blocks per genome varied from 158 to 186, averaging 175. The list of identified syntenic blocks for *N. gonorrhoeae* and *N. meningitidis* is given in Table S2.

Next, we identified the minimum number of changes in the genome structure that allow a genome with one order of syntenic blocks to pass into a genome with a different order of syntenic blocks (the number of genome rearrangements). To do this, we used the "double-cut-and-join with indels" model which allows to compare genomes and identify insertions, deletions, inversions, translocations fusions and fissions of genome fragments

using the UniMoG program. For the studied sample of *N. gonorrhoeae* genomes, the number of changes in genomes varied from 0 to 28; for *N. meningitidis* genomes, from 0 to 43, with the minimum length of syntenic block set at 1248 bp. It should be noted that the number of rearrangement events in the studied *N. gonorrhoeae* population did not depend on the NG-MAST type and the MLST type of isolates, while in the *N. meningitidis* population it did not depend on the MLST type and the serogroup.

We examined the possibility of genome changes during the experiment, i.e., the possibility of the appearance of rearrangements during various manipulations with cell cultures. To do this, two *N. gonorrhoeae* DNA samples from isolates obtained in Russia were re-sequenced and re-analyzed (Bioproject PRJNA768989, SAMN31014989, SAMN22599488, SAMN31015000 and SAMN22600865). The analysis showed that, at least during the course of the experiment, the sequencing results were identical and no changes in synteny were observed.

Then we analyzed the relationship between the number of rearrangement events and phylogenetic proximity of genomes. Phylogenetic proximity or distance between genomes was assessed based on comparison of core genomes by calculating the number of base substitutions per site between sequences applying the Tamura-Nei model [41]. The relationship between the number of rearrangement events and the phylogenetic proximity of genomes for all studied genomes of *N. gonorrhoeae* (pairwise comparison of all 146 genomes) and *N. meningitidis* (pairwise comparison of all 138 genomes) is shown in Figure 2 a,b. For the plots shown in Figure 2, the calculated values of the Spearman ranking correlation coefficient were $\varrho = 0.329$ (p < 0.001) for *N. gonorrhoeae* and $\varrho = 0.624$ (p < 0.001) for *N. meningitidis*.

Thus, the correlation between the number of rearrangement events and the phylogenetic proximity of genomes were higher for the *N. meningitidis* population than for the *N. gonorrhoeae* population. It can be assumed that the dependence of the order of genes on phylogenetic proximity is observed only at a sufficient level of nucleotide diversity in the population, i.e., at sufficiently large phylogenetic distances between isolates. The higher nucleotide diversity of the *N. meningitidis* population compared to *N. gonorrhoeae* may in turn mean that *N. meningitidis* is more distant from their common ancestor.

2.4 Rearrangements in genomes of N. gonorrhoeae and N. meningitidis and coordinates of mobile elements

Since the aim of the study was to typify and characterize the recombination points that led to the formation of syntenic blocks and to determine the causes of recombination in the past, we considered not only whole mobile elements, but also their fragments. For instance, IS elements were taken into account if their length was at least 150 bp, phage elements – if their length was at least 25% of their reference sequence (see below).

When analyzing the genome sequences of our sample, we found Correia elements (CREEs) among the mobile elements that could contribute to genomic rearrangements: on average, 126 CREEs per genome for *N. gonorrhoeae* and 250 for *N. meningitidis* (Table 1), which corresponds to the literature data [20,21,26,32]. In *N. gonorrhoeae* isolates, we identified Spencer-Smith elements [23] with a length of about 650 bp long (three elements per genome). For *N. gonorrhoeae*, 80 NIMEs per genome were detected, for *N. meningitidis*, 288 NIMEs (Table 1).

We also identified the sequences of IS elements belonging to the IS5, IS30, IS110, and IS1595 groups (Table 1). The IS30 elements encoding the type DDE transposase (with a conserved amino acid triad, Asp-Asp-Glu, in the active site) working on the "copy-and-paste" principle [34,35] were found only in *N. meningitidis*.

The IS110 elements which encode a DEDD (Asp-Glu-Asp-Asp) type transposase were found in both *Neisseria* species. Although the chemical mechanisms of action of DEDD and DDE enzymes are similar, the mechanisms of operation of IS30 and IS110 elements are presumably different [34,35].
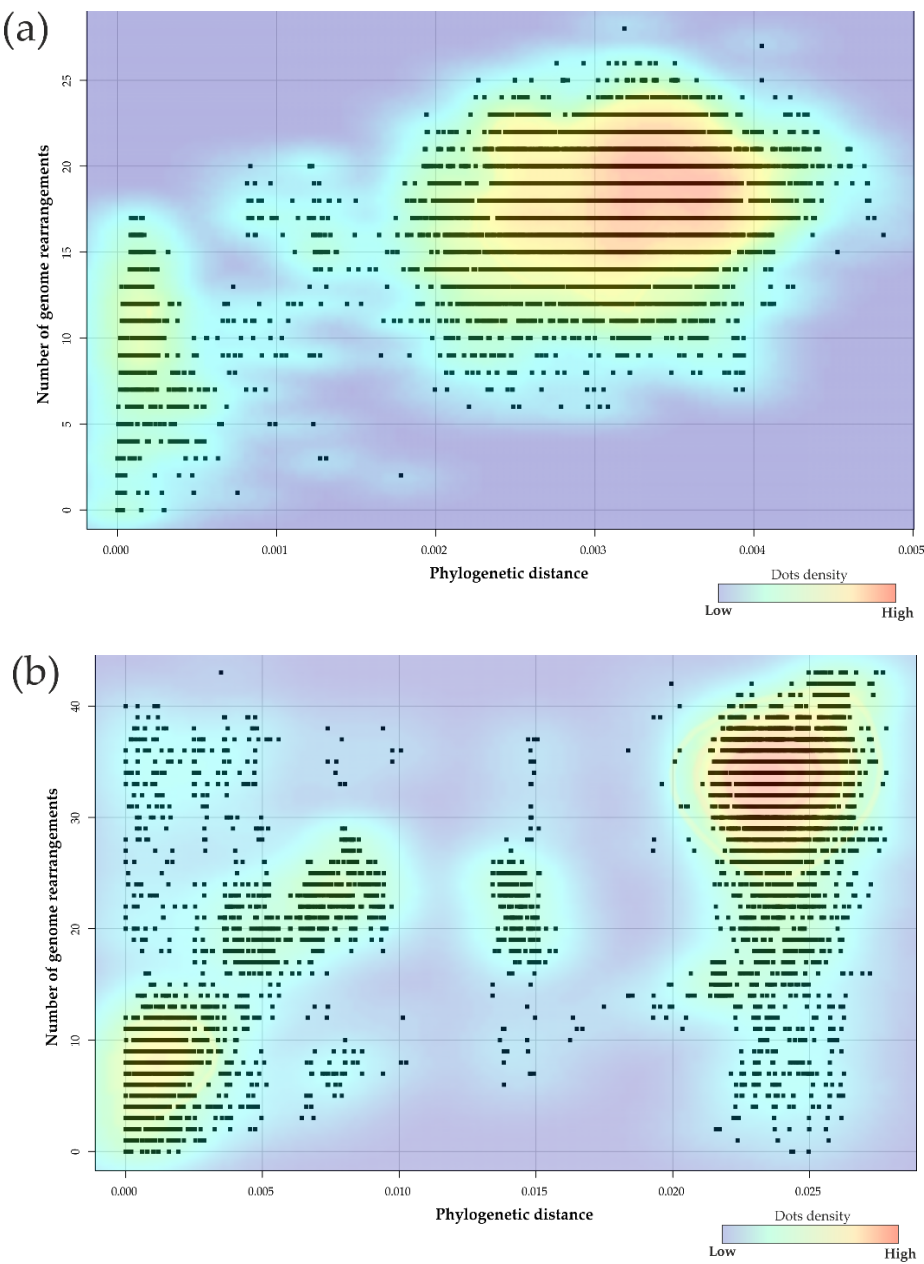
Figure 2. Relationship between phylogenetic distance estimated as the number of base substitutions per site between sequences in core genomes and the number of genome rearrangements for *N. gonorrhoeae* (a) and *N. meningitidis* (b).

Table 1. The number of mobile elements in genomes of *N. gonorrhoeae* and *N. meningitidis* (average number of elements per genome)

Mobile element		*N. gonorrhoeae*	*N. meningitidis*
Correia elements (CREEs)		126	250
Spencer-Smith elements (SSREEs)		3	0
Neisserial intergenic mosaic elements (NIMEs)		80	288
IS elements:	IS30	0	12
	IS110	10	6
	IS5 ssgr IS5	1	3
	IS5 ssgr IS427	0	7
	IS5 (all sub-groups)	6	22
	IS1595 ssgr IS1016	17	12

The identified elements of the IS1595 group belong only to the IS1016 subgroup encoding a DDEK type transposase. The detected elements of the IS5 group (DDE type transposase) belong to several subgroups: IS5, IS427, etc. These subgroups may differ in the expression of the transposase [34,35].

The presence of a number of prophages, both complete and fragments, was detected. Table 2 shows the sum of the lengths of phage elements when using different thresholds for their detection.

Table 2. Phage elements found in *N. gonorrhoeae* and *N. meningitidis* at different detection thresholds (cut-offs). Due to the high homology between filamentous phages, the Blast algorithm assigned the same fragment to several phages

Prophage	Sum of lengths of phage elements			
	N. gonorrhoeae	*N. meningitidis*	*N. gonorrhoeae*	*N. meningitidis*
Nf1	43%	98%	91%	113%
Nf1-MDA	42%	112%	86%	116%
Nf2	196%	72%	214%	94%
Nf3	0%	6%	0%	14%
NgoΦ1	64%	0%	130%	5%
NgoΦ2	40%	0%	138%	4%
NgoΦ3	88%	0%	165%	1%
NgoΦ3 (another fragment)	92%	1%	149%	17%
NgoΦ4	97%	0%	143%	19%
NgoΦ5	79%	2%	231%	82%
NgoΦ6	230%	5%	328%	48%
NgoΦ7	291%	43%	408%	80%
NgoΦ8	290%	43%	407%	80%
NgoΦ9	230%	44%	278%	51%
Pnm1 (Mu-like)	0%	7%	0%	16%
Pnm2 (Mu-like)	0%	8%	9%	67%
MuMenB (Mu-like)	0%	15%	11%	89%
Cut-off value	> 25%		>0.5%	

At a detection threshold of 25% or more, Neisserial filamentous phages Nf1, Nf1-MDA, Nf2, all five known phages with double-stranded DNA NgoΦ1-NgoΦ5, and all four known single-stranded (filamentous) phages NgoΦ6-NgoΦ9 were identified in the chromosomal DNA sequences of *N. gonorrhoeae*. The Nf1-MDA phage contains a meningococcal disease-associated (MDA) island, the presence of which correlates with hyperinvasive forms of gonorrhea in adults [31,37].

The chromosome of *N. meningitidis* was found to contain Nf1, Nf1-MDA, Nf2, and Nf3 single-stranded phages, NgoΦ7-NgoΦ9 single-stranded phages, and Mu-like Pnm1, Pnm2, and MuMenB phages. The Mu phage is a mobile element encoding its own transposase, as well as more traditional phage genes for head, tail, and lytic proteins [31,37].

The coordinates of syntenic blocks obtained using the SibeliaZ program were compared with the coordinates of mobile elements, thus determining which mobile elements could take part in the observed genomic rearrangements. The minimum distance from the block boundary to the beginning or end of the reading frame of the mobile element was considered to be no more than 700 bp for NIMEs, IS elements, prophages, and no more than 1800 bp for CREEs and SSREEs. The lists of *N. gonorrhoeae* and *N. meningitidis* syntenic blocks for which mobile elements flanking them were found are given in Table S3. Diagrams reflecting the number of different syntenic blocks flanked by mobile elements in *N. gonorrhoeae* and *N. meningitidis* are shown in Figure 3a and Figure 3b.
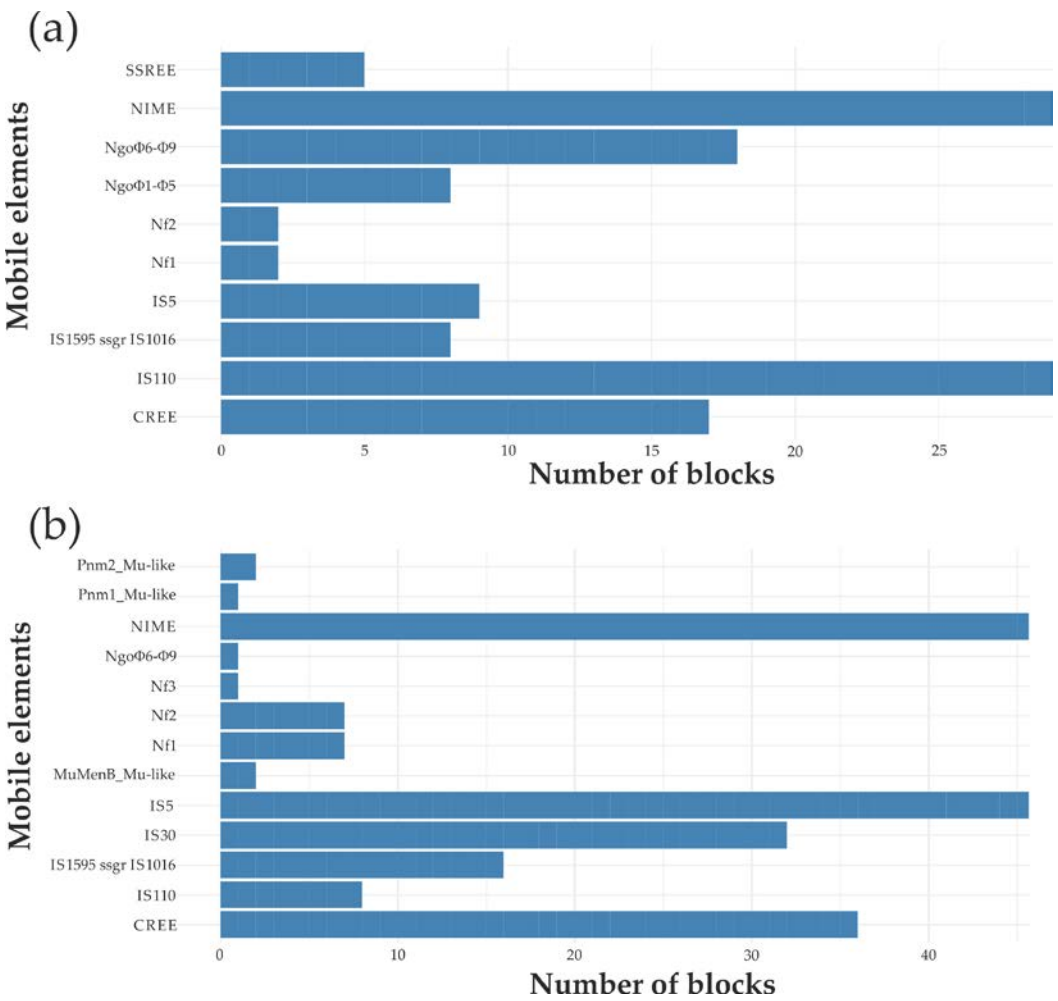
Figure 3. Syntenic blocks flanked by mobile elements in *N. gonorrhoeae* (a) and *N. meningitidis* (b).

For *N. gonorrhoeae*, the total number of syntenic blocks was 220, out of which 97 (44.1%) were flanked by various mobile elements and prophages. For *N. meningitidis*, out of 327 syntenic blocks found, 158 (48.3%) were flanked by mobile elements and prophages. Thus, it was shown that more than 44% of genomic rearrangements in both bacterial species may have occurred with the participation of such mobile elements.

For both *N. gonorrhoeae* and *N. meningitidis*, the largest number of genomic rearrangements was associated with the presence of NIMEs; CREE and IS elements also played an important role (Figure 3). In *N. gonorrhoeae*, a significant proportion of rearrangements can be explained by the activity of single- and double-stranded phages NgoΦ1-NgoΦ9, the presence of large fragments (>25%) of which was found at the sites of rearrangements. The rearrangements in the *N. gonorrhoeae* genomes with the participation of Spencer-Smith elements (SSREE are absent in *N. meningitidis*) and rearrangements in *N. meningitidis* with the participation of Mu-like prophages (absent in *N. gonorrhoeae*) should also be noted. As for IS elements, the IS110 elements were the most important for the rearrangements in *N. gonorrhoeae*, and the IS5 elements in *N. meningitidis*.

Using the results of whole genome sequencing of 50 *N. gonorrhoeae* isolates collected in Russia, we analyzed the genomic rearrangements observed within the Russian *N. gonorrhoeae* population. The Wilcoxon rank-sum test allowed us to reveal that the average number of mobile elements per genome is the same both for the population as a whole and for the Russian isolates. The analysis showed that in terms of rearrangement events there is no difference between the Russian and global populations of *N. gonorrhoeae*.

3. Discussion

In this work, we analyzed chromosome rearrangements in *N. gonorrhoeae* and *N. meningitidis*, identified the mobile elements and determined their role in the formation of rearrangements. During the evolution of a microorganism, the nucleotide/amino acid sequence and the order of genes can change. It is generally accepted that the order of genes changes faster than the amino acid sequence [2,3,8,9], but the relationship between these processes has not been determined. In this work, we showed at the population level that for *N. gonorrhoeae* the statement about the interdependence of the order of genes and phylogenetic proximity is, in general, incorrect, in contrast to *N. meningitidis*, for which a stronger correlation between changes in the order of genes on the chromosome and phylogenetic proximity was found.

A strong positive correlation between rearrangement distance and amino acid distance for the genomes of 40 prokaryotes was noted by Novichkov et al. [9]. However, the distances were determined by somewhat different methods than those used by us. The distance of amino acids in [9] was characterized by the value of dN/dS, the ratio of the number of nonsynonymous (dN), i.e., leading to a change in the amino acid, and synonymous (dS) substitutions per site, while a quantitative measure of genome rearrangement was considered to be the synteny distance (dY) which, as noted in the article, is only a general characteristic of the evolutionary loss of the gene synteny and does not reflect specific events leading to rearrangement.

Based on the results of our work, in which we analyzed the observed changes in synteny in the genome and compared the populations of two different *Neisseria* species, it can be assumed that the correlation of the number of rearrangements in the genome correlates with the phylogenetic distance only at a high nucleotide diversity.

When analyzing the genome sequences of our sample, we identified mobile elements that could contribute to genomic rearrangements, including Correia elements, Spencer-Smith elements (in *N. gonorrhoeae*), Neisserial intergenic mosaic elements (NIMEs), a number of IS elements of the IS5, IS30, IS110, and IS1595 groups, Nf1-Nf3 prophages, NgoΦ1-NgoΦ9 single- and double-stranded prophages, Mu-like Pnm1, Pnm2, and MuMenB phages (in *N. meningitidis*). For the sample of *N. gonorrhoeae* genomes, 220 syntenic blocks were identified, out of which 97 (44.1%) were flanked by various mobile elements and prophages. For the sample of *N. meningitidis* genomes, 327 syntenic blocks were identified, out of which 158 (48.3%) were flanked by mobile elements and prophages. Thus, it was shown that more than 44% of the observed rearrangements in two species of *Neisseria* are most likely to have occurred with the participation of mobile elements, including prophages. The observed variability can provide pathogens with the ability to persist in the host for a long time and re-infect him.

It should be noted that homologous recombination can occur with the participation of a large number of sequences with high homology and it is not possible to trace all of them. In addition, it is known that in *N. gonorrhoeae* recombination processes are used for antigenic variation of the pilin protein PilE [42].

Based on the results of whole genome sequencing of the Russian *N. gonorrhoeae* isolates, we analyzed the genomic rearrangements observed within the Russian population of *N. gonorrhoeae*. Despite the fact that this population differs from the European population of *N. gonorrhoeae* [39,43,44] in terms of isolates belonging to NG-MAST and MLST types, as well as antibiotic resistance (in particular, isolates resistant to third-generation cephalosporins have not yet been found in Russia), we did not find any differences between the Russian and the global populations both in the number of rearrangement events and in the number of mobile elements in genomes.

Modern phylogeny methods rely not only on the amino acid or nucleotide sequence data using the distance-based methods, the maximum parsimony-based methods, and the maximum likelihood-based methods, but also on the gene order data [45]. Our results on the variability of the order of genes in the considered socially significant pathogens can become the basis for creating a more sophisticated pathogen genotyping method that

combine standard typing approaches, such as MLST, NG-MAST, and cgMLST, with syntenic rearrangements data.

4. Materials and Methods

4.1. Clinical isolates of N. gonorrhoeae collected in the Russian Federation

Clinical isolates of *N. gonorrhoeae* were collected at the State Research Center for Dermatovenerology and Cosmetology of the Ministry of Health of the Russian Federation. The samples were obtained by specialized medical organizations of the dermatovener-ological profile from clinical material (urethral preparations for men and cervical/urethral preparations for women) of patients diagnosed with primary symptomatic uncompli-cated gonorrhea, each sample from an individual patient. Sample collection, transporta-tion, cultivation, and storage were performed according to a previously described proto-col [43,44].

Fifty isolates obtained in 2015-2019 in 8 regions of the Russian Federation (Arkhan-gelsk, Astrakhan, Bryansk, Cheboksary, Kaluga, Omsk, Penza, Stavropol) were used. The isolates were grown under aseptic conditions on chocolate agar plates at 37 °C and 5% CO_2. As a control of sequencing, the *N. gonorrhoeae* ATCC 49226 strain was used.

4.2. Whole genome sequencing of N. gonorrhoeae isolates and de novo genome assembly

Isolation of genomic DNA for sequencing was performed from an overnight culture of gonococci using the Monarch Genomic DNA Purification Kit (New England Biolabs, UK). The resulting DNA preparations were further purified using the Agencourt AMPure XP beads (Beckman Coulter, USA). Final DNA concentrations were measured using a NanoDrop 2000 spectrophotometer and a Qubit 4 Fluorometer (Thermo Fisher Scientific, USA), the concentrations were in the range of 20-100 ng/µl.

Whole genome sequencing was performed in parallel on two platforms: FLO-MIN110 R9 and R10 sequencing cells in a MinION device (Oxford Nanopore Technolo-gies, UK) and a MiniSeq sequencing system (Illumina, USA). For the MinION cell, a li-brary of DNA fragments was prepared using the Oxford Nanopore and New England Biolabs reagent kits according to the Native barcoding genomic DNA protocol (with EXP-NBD104, EXP-NBD114, and SQK-LSK109). According to the manufacturer's protocol, the DNA ends were repaired with subsequent barcoding and adapter ligation. The final li-brary (5-50 fmol DNA) was loaded onto the flow cell. For sequencing on the Illumina platform, libraries were prepared using the DNA fragmentation method followed by PCR and indexing according to the Nextera XT DNA Library Prep Kit Reference Guide (Illu-mina, USA). After purification of the libraries, verification of the lengths and concentra-tions was carried out using automated capillary electrophoresis platforms TapeStation 4150 (Agilent, USA). The final library was normalized and denatured according to the MiniSeq system instructions. The cluster density averaged 170–250 K/mm^2 in all runs. Data output averaged 10.5 GB of data out of 12 GB theoretically possible on this device.

In total, ~1 GB of raw data per sample in the fastQ format was generated for each platform. Hybrid *de novo* genome assembly after sequencing on two platforms was carried out using the Unicycler program (https://github.com/rrwick/Unicycler). All sequences were uploaded to the GenBank, Bioproject PRJNA768989 (www.ncbi.nlm.nih.gov/biopro-ject/?term=PRJNA768989), under assigned accession numbers (Table S1).

4.3. Selection of genomes from the databases

Genome sequences of isolates of the *Neisseria* genus were taken from the GenBank and BV-BRC databases (bv-brc.org). A total of 95 *N. gonorrhoeae* and 138 *N. meningitidis* genomes were selected, with particular attention paid to the fact that all genome se-quences were obtained by *de novo* assembly. Selected samples are listed in Table S1 of the Supplementary Materials.

4.4. Construction of phylogenetic network

The selected genomes of *Neisseria spp.* were processed using the Prokka 1.14.6 pro-gram (https://github.com/tseemann/prokka) to obtain gff files with annotated genomes.

The gff files were then analyzed by rapid large-scale prokaryotic pan-genome analysis using the Roary software [46]. After the Roary processing, core genomes containing aligned sequences of concatenated orthologs were obtained, and a phylogenetic network was constructed using the SplitsTree 5.0.0_alpha program [47].

4.5. Assessment of nucleotide diversity for N. gonorrhoeae and N. meningitidis

The nucleotide diversity of core genomes was assessed by calculating the π value [40]. Measurement of the Jukes-Cantor pairwise distances after realigning each pair of sequences, while ignoring the sites with gaps, was performed in the Matlab program.

4.6. Calculation of distance between genomes

The distance between genomes on the basis of core genomes, i.e. the number of base substitutions per site from between sequences were calculated for 146 nucleotide sequences of *N. gonorrhoeae* (95 genomes from the database and 50 genomes of the Russian isolates, that we obtained in this work) and 138 nucleotide sequences of *N. meningitidis*. Analyses were conducted using the Tamura-Nei model [41]. Evolutionary analyses were carried out using the MEGA11 software [48].

4.7. Determination of the minimum number of changes in the structure of one genome compared to other genomes

When studying genomic rearrangements, we determined the minimum number of changes in the genome structure that allow a genome with one order of syntenic blocks to turn into a genome with a different order of syntenic blocks due to such events as insertions, deletions, inversions, translocations, fusion and fission of blocks.

The chromosomal sequences of the *N. gonorrhoeae* and *N. meningitidis* isolates from our sample (*.gbk* files) were normalized in such a way that the sequences began with the *dnaA* gene and were converted to the *.fasta* format. At the next step, we used the SibeliaZ v1.2.4 algorithm (https://github.com/medvedevgroup/SibeliaZ) [49] to build locally collinear blocks. Default algorithm parameters were used, except for *k=15* (15 is recommended by the program developers for bacteria, 25 for mammals).

The resulting *.maf* file was then processed by the maf2synteny script available from the SibeliaZ repository. As an input parameter *b* which determines the minimum length of the syntenic block, we used the value of 1248 bp, which corresponds to the 3rd quartile (75%) of the distribution of all *N. gonorrhoeae* and *N. meningitidis* genes by length (gene sequences were taken from the PubMLST website). As a result, the files blocks_coords.txt and genomes_permutations.txt were obtained, with coordinates of syntenic blocks and permutation matrices, respectively.

Then we calculated the minimum number of events that allow one genome to turn into another using the DCJ-indel model in the UniMoG program [50] (https://bibiserv.ce-bitec.uni-bielefeld.de/dcj). Rearrangement events and calculation of DCJ-indel distance on the example of two *N. gonorrhoeae* genomes, ATCC 49226 and WHO U, are illustrated in Figure 4.

4.8. Determining coordinates of mobile elements

Correia elements. To identify Correia elements, we used an approach described earlier in [51]. Correia elements are flanked on the right and left sides by relatively conservative 26 bp sequences, the so-called inverted Correia repeats [23]. We used the fuzznuc utility from the EMBOSS package (https://www.ebi.ac.uk/Tools/emboss/), to the input of which four sequences were supplied,

5'-TATAGTGGGATTAACAAAAACCGGTACGG-3',
5'-TATAGTGGGATTAAATTTAAACCGGTACGG-3',
5'-TATAGTGGGATTAACAAAAATCAGGACAA-3',
5'-TATAGTGGGATTAAATTTAAATCAGGACAA-3',

with a maximum of three mismatches allowed. The found coordinates of the inverted repeats were matched with each other in Microsoft Excel, obtaining the coordinates of the Correia elements.
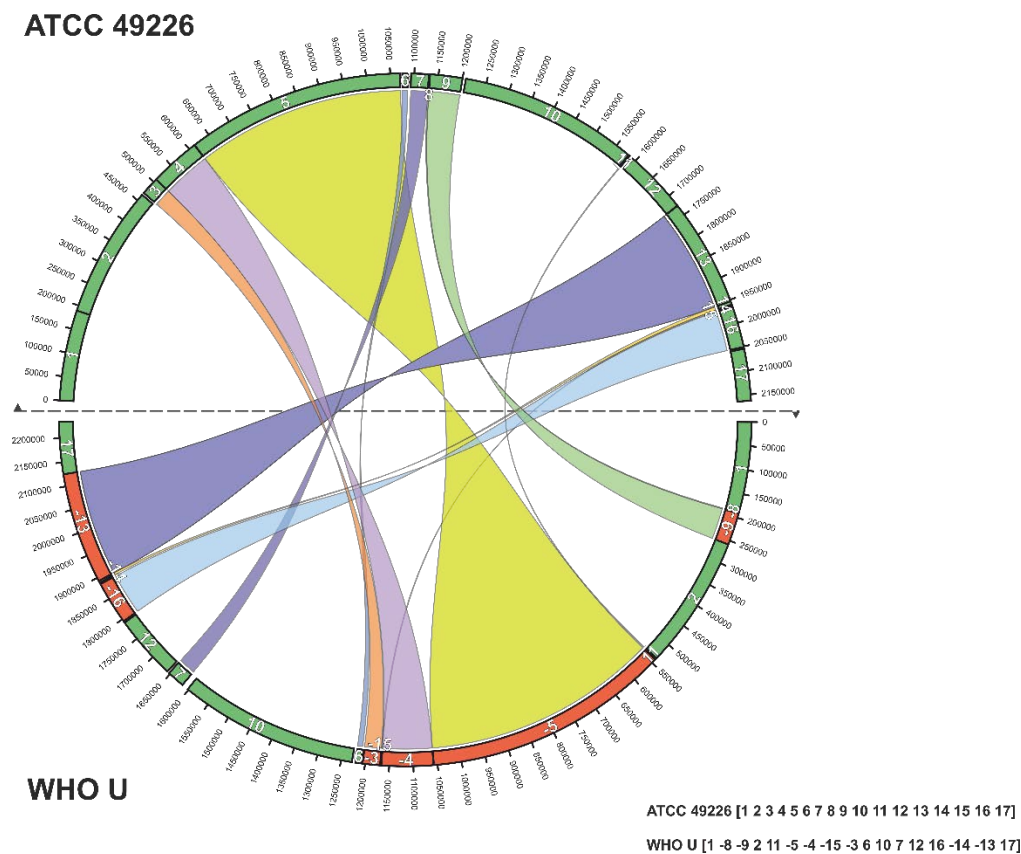
ATCC 49226 [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17]

WHO U [1 -8 -9 2 11 -5 -4 -15 -3 6 10 7 12 16 -14 -13 17]

Figure 4. Representation of rearrangements between genomes and calculation of the minimum number of rearrangement events on the example of two *N. gonorrhoeae genomes*, ATCC 49226 and WHO U. The results of pairwise comparison of genomes are presented as a string with the numbers of syntenic blocks and their order; the minus sign means a change in the direction of the genes (inversion). Blocks are highlighted in different colours, block numbers are marked with white numbers on the pie chart. Orange color reflects a change in the strand direction. The minimum number of rearrangement events that allow one genome to turn into another (DCJ distance) in this case is equal to 11.

Spencer-Smith elements. To identify Spencer-Smith elements, we used 19 bp repeats 5'-CGTTTCAGACGGCATCGGG-3' and 5'-CCCGATGCCGCCTGAAACG-3' [23]. In addition, to confirm the coordinates, we used the blast2seq utility from NCBI.

Neisserial intergenic mosaic elements (NIMEs). First, we used the fuzznuc utility to search for the flanking dRS3 sequences (5'-ATTCCCNNNNNNNGGGAAT-3'), then matched them in Excel so that they were ≤ 200 bp apart, and assembled a unified NIME sequence. Sequences located at a greater distance were discarded.

IS elements. To identify IS elements, we used the IS-Finder site (http://issaga.biotoul.fr); 284 files in the .xlsx format were obtained. The files were pooled and sorted by the "*similarity_pp1*" column, discarding all IS element predictions with less than 80% similarity and/or if they were too short (< 150 bp) and/or they were flagged as "*probable false positive*".

Prophage sequences were first searched using the *PhiSpy v.4.2.21* algorithm (https://github.com/linsalrob/PhiSpy) as described in [52], with *default* settings to find rough coordinates. To refine the coordinates, the *blastall* program from the NCBI was used, to the input of which the custom database we built was supplied. This database included the sequences of both known gonococcal and meningococcal phages: NgoΦ1-NgoΦ9 (*N. gonorrhoeae* strain FA1090), Nf1, Nf3 (*N. meningitidis* strain Z2491), Nf2 (*N. meningitidis* strain alpha710), Pnm1, Pnm2 (*N. meningitidis* strain Z2491), MuMenB (*N. gonorrhoeae* strain MC58).

The obtained coordinates of the mobile elements were compared with the coordinates of the syntenic blocks, determined as described in 4.7.

4.9. Statistcal data processing. Correlation analysis.

The Spearman's rank correlation test with precomputed null distribution was performed in the pspearman package version 0.3-1 of the R software. The Wilcoxon Rank Sum and Signed Rank Tests were performed in the R Package stats version 4.2.1.

References

1. Periwal, V.; Scaria, V. Insights into structural variations and genome rearrangements in prokaryotic genomes. *Bioinformatics* **2015**, *31*, 1–9. doi: 10.1093/bioinformatics/btu600.
2. Mushegian, A.R.; Koonin, E.V. Gene order is not conserved in bacterial evolution. *Trends Genet*. **1996**, *12*, 289–290. doi: 10.1016/0168-9525(96)20006-x.
3. Itoh, T.; Takemoto, K.; Mori, H.; Gojobori, T. Evolutionary instability of operon structures disclosed by sequence comparisons of complete microbial genomes. *Mol. Biol. Evol.* **1999**, *16*, 332–346. doi: 10.1093/oxfordjournals.molbev.a026114.
4. Jin, Q.; Yuan, Z.; Xu, J.; Wang, Y.; Shen, Y.; Lu, W.; Wang, J.; Liu, H.; Yang, J.; Yang, F.; et al. Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. *Nucleic Acids Res.* **2002**, *30*, 4432–4441. doi: 10.1093/nar/gkf566.
5. Lan, R.; Reeves, P.R. *Escherichia coli* in disguise: molecular origins of *Shigella*. *Microbes Infect*. **2002**, *4*, 1125–1132. doi: 10.1016/s1286-4579(02)01637-4.
6. Devanga Ragupathi, N.K.; Muthuirulandi Sethuvel, D.P.; Inbanathan, F.Y.; Veeraraghavan, B. Accurate differentiation of *Escherichia coli* and *Shigella* serogroups: challenges and strategies. *New Microbes New Infect.* **2017**, *21*, 58–62. doi: 10.1016/j.nmni.2017.09.003.
7. Martínez-Chavarría, L.C.; Vadyvaloo, V. *Yersinia pestis* and *Yersinia pseudotuberculosis* infection: a regulatory RNA perspective. *Front. Microbiol.* **2015**, *6*, 956. doi:10.3389/fmicb.2015.00956.
8. Wolf, Y.I.; Rogozin, I.B.; Kondrashov, A.S.; Koonin, E.V. Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. *Genome Res.* **2001**, *11*, 356–372. doi: 10.1101/gr.gr-1619r.
9. Novichkov, P.S.; Wolf, Y.I.; Dubchak, I.; Koonin, E.V. Trends in prokaryotic evolution revealed by comparison of closely related bacterial and archaeal genomes. *J. Bacteriol.* **2009**, *191*, 65–73. doi: 10.1128/JB.01237-08.
10. Liang, Y.; Hou, X.; Wang, Y.; Cui, Z.; Zhang, Z.; Zhu, X.; Xia, L.; Shen, X.; Cai, H.; Wang, J.; et al. Genome rearrangements of completely sequenced strains of *Yersinia pestis. J. Clin. Microbiol*. **2010**, *48*, 1619–1623. doi: 10.1128/JCM.01473-09.

11. Liang, Y.; Xie, F.; Tang, X.; Wang, M.; Zhang, E.; Zhang, Z.; Cai, H.; Wang, Y.; Shen, X.; Zhao, H.; et al. Chromosomal rearrangement features of *Yersinia pestis* strains from natural plague foci in China. *Am. J. Trop. Med. Hyg.* **2014**, *91*, 722–728. doi: 10.4269/ajtmh.13-0491.

12. Bochkareva, O.O.; Dranenko, N.O.; Ocheredko, E.S.; Kanevsky, G.M.; Lozinsky, Y.N.; Khalaycheva, V.A.; Artamonova. I.I.; Gelfand, M.S. Genome rearrangements and phylogeny reconstruction in *Yersinia pestis*. *PeerJ*. **2018**, *6*, e4545. doi: 10.7717/peerj.4545.

13. Bochkareva, O.O.; Moroz, E.V.; Davydov, I.I.; Gelfand, M.S. Genome rearrangements and selection in multi-chromosome bacteria *Burkholderia spp*. *BMC Genomics* **2018**, *19*, 965. doi: 10.1186/s12864-018-5245-1.

14. Andam, C.P.; Hanage, W.P. Mechanisms of genome evolution of *Streptococcus*. *Infect. Genet. Evol.* **2015**, *33*, 334–342. doi: 10.1016/j.meegid.2014.11.007.

15. Shelyakin, P.V.; Bochkareva, O.O.; Karan, A.A.; Gelfand, M.S. Micro-evolution of three *Streptococcus* species: selection, antigenic variation, and horizontal gene inflow. *BMC Evol. Biol.* **2019**, *19*, 83. doi: 10.1186/s12862-019-1403-6.

16. Seferbekova, Z.; Zabelkin, A.; Yakovleva, Y.; Afasizhev, R.; Dranenko, N.O.; Alexeev, N.; Gelfand, M.S.; Bochkareva, O.O. High rates of genome rearrangements and pathogenicity of *Shigella spp*. *Front. Microbiol.* **2021**, *12*, 628622. doi: 10.3389/fmicb.2021.628622.

17. Rocha, E.P. DNA repeats lead to the accelerated loss of gene order in bacteria. *Trends Genet*. **2003**, *19*, 600–603. doi: 10.1016/j.tig.2003.09.011.

18. Dempsey, J.A.; Wallace, A.B.; Cannon, J.G. The physical map of the chromosome of a serogroup A strain of *Neisseria meningitidis* shows complex rearrangements relative to the chromosomes of the two mapped strains of the closely related species *N. gonorrhoeae*. *J. Bacteriol.* **1995**, *177*, 6390–6400. doi: 10.1128/jb.177.22.6390-6400.1995.

19. Liu, S.V., Saunders, N.J.; Jeffries, A.; Rest, R.F. Genome analysis and strain comparison of correia repeats and correia repeat-enclosed elements in pathogenic *Neisseria*. *J. Bacteriol.* **2002**, *184*, 6163–6173. doi: 10.1128/JB.184.22.6163-6173.2002.

20. Bentley, S.D.; Vernikos, G.S.; Snyder, L.A.; Churcher, C.; Arrowsmith, C.; Chillingworth, T.; Cronin, A.; Davis, P.H.; Holroyd, N.E.; Jagels, K.; et al. Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS Genet.* **2007**, *3*, e23. doi: 10.1371/journal.pgen.0030023.

21. Snyder, L.A.; Cole, J.A.; Pallen, M.J. Comparative analysis of two *Neisseria gonorrhoeae* genome sequences reveals evidence of mobilization of Correia Repeat Enclosed Elements and their role in regulation. *BMC Genomics* **2009**, *10*, 70. doi: 10.1186/1471-2164-10-70.

22. Schoen, C.; Tettelin, H.; Parkhill, J.; Frosch, M. Genome flexibility in *Neisseria meningitidis*. *Vaccine* **2009**, *27 Suppl 2*, B103–111. doi: 10.1016/j.vaccine.2009.04.064.

23. Spencer-Smith, R.; Varkey, E.M.; Fielder, M.D.; Snyder, L.A. Sequence features contributing to chromosomal rearrangements in *Neisseria gonorrhoeae*. *PLoS One* **2012**, *7*, e46023. doi: 10.1371/journal.pone.0046023.

24. Cehovin, A.; Lewis, S.B. Mobile genetic elements in *Neisseria gonorrhoeae*: movement for change. *Pathog. Dis.* **2017**, *75*. doi: 10.1093/femspd/ftx071.

25. Spencer-Smith, R.; Gould, S.W.; Pulijala, M.; Snyder, L.A.S. Investigating potential chromosomal rearrangements during laboratory culture of *Neisseria gonorrhoeae*. *Microorganisms* **2018**, *6*, 10. doi: 10.3390/microorganisms6010010.

26. Marri, P.R.; Paniscus, M.; Weyand N.J.; Rendón, M.A.; Calton, C.M.; Hernández, D.R.; Higashi, D.L.; Sodergren, E.; Weinstock, G.M.; Rounsley, S.D.; So, M. Genome sequencing reveals widespread virulence gene exchange among human *Neisseria* species. *PLoS One* **2010**, *5*, e11835. doi: 10.1371/journal.pone.0011835.

27. Calder, A.; Menkitim C.J.; Çağdaş, A.; Lisboa Santos, J.; Streich, R.; Wong, A.; Avini, A.H.; Bojang, E.; Yogamanoharan, K.; Sivanesan, N.; et al. Virulence genes and previously unexplored gene clusters in four commensal *Neisseria spp*. isolated from the human throat expand the neisserial gene repertoire. *Microb. Genom.* **2020**, *6*, mgen000423. doi: 10.1099/mgen.0.000423.

28. Stohl, E.A.; Seifert, H.S. *Neisseria gonorrhoeae* DNA recombination and repair enzymes protect against oxidative damage caused by hydrogen peroxide. *J. Bacteriol.* **2006**, *188*, 7645–7651. doi:10.1128/JB.00801-06].

29. Achaz, G.; Rocha, E.P.; Netter, P.; Coissac, E. Origin and fate of repeats in bacteria. *Nucleic Acids Res.* **2002**, *30*, 2987–2994. doi: 10.1093/nar/gkf391.

30. Parkhill, J.; Achtman, M.; James K.D.; Bentley, S.D.; Churcher, C.; Klee S.R.; Morelli, G.; Basham, D.; Brown, D.; Chillingworth, T.; et al. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. *Nature* **2000**, *404*, 502–506. doi: 10.1038/35006655.

31. Rotman, E.; Seifert, H.S. The genetics of *Neisseria* species. *Annu. Rev. Genet.* **2014**, *48*, 405–431. doi: 10.1146/annurev-genet-120213-092007.

32. Elbeyioglu, F.; Roberts, S.B.; Spencer-Smith, R.; Pulijala, M.; Zelewska, M.A.; Nebel, J.C.; Snyder, L.A.S. Inversion of Correia repeat enclosed elements in *Neisseria gonorrhoeae*. *Microbiology (Reading)* **2017**, *163*, 31–36. doi: 10.1099/mic.0.000394.

33. Darmon, E.; Leach, D.R. Bacterial genome instability. *Microbiol, Mol, Biol, Rev*. **2014**, *78*, 1–39. doi: 10.1128/MMBR.00035-13.

34. Siguier, P.; Gourbeyre, E.; Chandler, M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev*. **2014**, *38*, 865–991. doi: 10.1111/1574-6976.12067.

35. Siguier, P.; Gourbeyre, E.; Varani, A.; Ton-Hoang, B.; Chandler, M. Everyman's guide to bacterial insertion sequences. *Microbiol. Spectr*. **2015**, *3*, MDNA3-0030-2014. doi: 10.1128/microbiolspec.MDNA3-0030-2014.

36. Snyder, L.A.; McGowan, S.; Rogers, M.; Duro, E.; O'Farrell, E.; Saunders, N.J. The repertoire of minimal mobile elements in the *Neisseria* species and evidence that these are involved in horizontal gene transfer in other bacteria. *Mol. Biol. Evol.* **2007**, *24*, 2802–2815. doi: 10.1093/molbev/msm215.

37. Piekarowicz, A.; Kłyz, A.; Majchrzak, M.; Adamczyk-Popławska, M.; Maugel, T.K.; Stein, D.C. Characterization of the dsDNA prophage sequences in the genome of *Neisseria gonorrhoeae* and visualization of productive bacteriophage. *BMC Microbiol.* **2007**, *7*, 66. doi: 10.1186/1471-2180-7-66.

38. Kawai, M.; Uchiyama, I.; Kobayashi, I. Genome comparison in silico in *Neisseria* suggests integration of filamentous bacteriophages by their own transposase. *DNA Res.* **2005**, *12*, 389–401. doi: 10.1093/dnares/dsi021.

39. Shaskolskiy, B.; Dementieva, E.; Kandinov, I.; Chestkov, A.; Kubanov, A.; Deryabin, D.; Gryadunov, D. Genetic diversity of *Neisseria gonorrhoeae* multi-antigen sequence types in Russia and Europe. *Int. J. Infect. Dis.* **2020**, *93*, 1–8. doi: 10.1016/j.ijid.2020.01.020.

40. Cutter, A.D. *A primer of molecular population genetics*. Oxford University Press: Oxford, UK, 2019.

41. Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526. doi: 10.1093/oxfordjournals.molbev.a040023.

42. Mehr, I.J.; Seifert, H.S. Differential roles of homologous recombination pathways in *Neisseria gonorrhoeae* pilin antigenic variation, DNA transformation and DNA repair. *Mol. Microbiol.* **1998**, *30*, 697–710, doi: 10.1046/j.1365-2958.1998.01089.x.

43. Shaskolskiy, B.; Dementieva, E.; Kandinov, I.; Filippova, M.; Petrova, N.; Plakhova, X.; Chestkov, A.; Kubanov, A.; Deryabin, D.; Gryadunov, D. Resistance of *Neisseria gonorrhoeae* isolates to beta-lactams (benzylpenicillin and ceftriaxone) in Russia, 2015-2017. *PloS One* **2019**, *14*, e0220339. doi: 10.1371/journal.pone.0220339.

44. Kubanov, A.; Solomka, V.; Plakhova, X.; Chestkov, A.; Petrova, N.; Shaskolskiy, B.; Dementieva, E.; Leinsoo, A.; Gryadunov, D.; Deryabin, D. Summary and trends of the Russian Gonococcal Antimicrobial Surveillance Programme, 2005 to 2016. *J. Clin. Microbiol.* **2019**, *57*, e02024-18. doi: 10.1128/JCM.02024-18.

45. Drillon, G.; Champeimont, R.; Oteri, F.; Fischer, G.; Carbone, A. Phylogenetic reconstruction based on synteny block and gene adjacencies. *Mol. Biol. Evol.* **2020**1; *37*, 2747–2762. doi: 10.1093/molbev/msaa114.

46. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.; Fookes M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. doi: 10.1093/bioinformatics/btv421.

47. Huson, D.H.; Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **2006,** *23*, 254–267. doi: 10.1093/molbev/msj030.

48. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. doi: 10.1093/molbev/msab120.

49. Minkin, I.; Medvedev, P. Scalable multiple whole-genome alignment and locally collinear block construction with SibeliaZ. *Nat. Commun.* **2020**, *11*, 6327. doi: 10.1038/s41467-020-19777-8.]

50. Braga, M.D.; Willing, E.; Stoye, J. Double cut and join with insertions and deletions. *J. Comput. Biol.* **2011**, *18*, 1167–1184. doi: 10.1089/cmb.2011.0118.

51. Roberts, S.B.; Spencer-Smith, R.; Shah, M.; Nebel, J.C.; Cook, R.T.; Snyder, L.A. Correia Repeat Enclosed Elements and non-coding RNAs in the *Neisseria* species. *Microorganisms* **2016**, *4*, 31. doi: 10.3390/microorganisms4030031.

52. Akhter, S.; Aziz, R.K.; Edwards, R.A. PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res.* **2012**, *40*, e126. doi: 10.1093/nar/gks406.