

Development and validation of a high throughput *Neisseria gonorrhoeae* genotyping method

Kohji KOMORI^{a†}, Kotaro AOKI^{b#†}, Yoshikazu ISHII^b, Ken SHIMUTA^c, Makoto
OHNISHI^c, Kazuhiro TATEDA^b

^a Department of Microbiology and Infectious Diseases, Toho University

Graduate School of Medicine, Tokyo, Japan

^b Department of Microbiology and Infectious Diseases, Toho University School

of Medicine, Tokyo, Japan

^c Department of Bacteriology I, National Institute of Infectious Diseases, Tokyo,

Japan

[†] These authors contributed equally to this work.

Running title: A high throughput method for *Gonococcus* genotyping

17 **# Address correspondence to Kotaro AOKI, Ph.D.**, Department of
18 Microbiology and Infectious Diseases, Toho University School of Medicine, 5-
19 21-16 Omori-nishi, Ota-ku, Tokyo 143-8540, Japan.
20 TEL: +81-03-3762-4151
21 kotaro.aoki@med.toho-u.ac.jp
22
23 **Key words:** *Neisseria gonorrhoeae*; multilocus sequence typing; *N.*
24 *gonorrhoeae* multi-antigen sequence typing; *N. gonorrhoeae* sequence typing
25 for antimicrobial resistance; whole-genome sequencing
26
27
28

Abstract

Background

Neisseria gonorrhoeae genotyping by whole-genome sequencing (WGS) is expensive for a large sample set, a less expensive and more efficient genotyping method is required. We developed a high-throughput genotyping method for *N. gonorrhoeae* to improve molecular epidemiological typing and antimicrobial-resistant identification in *N. gonorrhoeae* antimicrobial susceptibility surveillance.

Methods

We used multiplex-tailed PCR to amplify and sequence 15 alleles from multilocus sequence typing (MLST), *N. gonorrhoeae* multiantigen sequence typing (NG-MAST), and *N. gonorrhoeae* sequence typing for antimicrobial resistance (NG-STAR). After indexing-PCR, we sequenced the DNA library using the MiSeq platform (Illumina). Sequencing reads were *de novo* assembly or constructing consensus sequences of alleles, then assigned sequence type. We used 54 previously characterized strains of *N. gonorrhoeae* and WGS data to validate our method.

Results

The allele identification results of MLST and NG-STAR in all strains agreed with the draft WGS. However, in NG-MAST, only 35 strains agreed. Disagreement was found in the NG-MAST of *porB* in 15 strains and of *tbpB* in seven strains. QRDR analysis perfectly predicted levofloxacin resistance. But was less successful in predicting reduced susceptibility or resistance phenotype to penicillin G, cefixime, or ceftriaxone using *penA*, *porB*, *ponA*, or *mtrR* alleles.

Conclusions

The successful performance in MLST and NG-STAR of our method was validated in this study. This method may be useful for large-scale genotyping for *N. gonorrhoeae* surveillance in a cost- and labor-saving manner. Phenotypic prediction of antimicrobial susceptibility by combining multiple alleles may be necessary for other than fluoroquinolones.

Introduction

The World Health Organization (WHO) estimated that there were 87 million new global cases of gonorrhea in 2016¹. Furthermore, the increasing frequency of antimicrobial resistant (AMR) *Neisseria gonorrhoeae*, the bacterium that causes gonorrhea, is a serious public health concern^{2,3}. Historically, resistance to penicillin G, fluoroquinolones, oral cephalosporins, ceftriaxone, and azithromycin has been problematic in the treatment of gonorrhea⁴. Ceftriaxone, an extended-spectrum cephalosporin or combination with azithromycin, are used as a first-line option for treating gonorrhea^{5,6}. However, in low- or middle-income countries, empiric therapy with inappropriate spectrum or dose oral antimicrobial agents such as cefixime, levofloxacin, or azithromycin that are over-the-counter systems for symptom-based sexually transmitted infections may select for ceftriaxone- or azithromycin-resistant *N. gonorrhoeae*^{4,7–9}. Understanding the development and spreading manor of AMR *N. gonorrhoeae* and antimicrobial resistant mechanisms are important to guide the appropriate use of the treatment option of gonorrhea and to motivate the development of novel agents to circumvent resistance.

One of the predominant mechanisms leading to resistance to ceftriaxone is a reduced affinity of penicillin binding protein 2 that causes mosaic-like mutations to cephalosporins^{10,11}. In 2009, a high-level ceftriaxone-resistant *N. gonorrhoeae* strain reported in Japan called H041 and WHO X severely limited the ability to treat gonorrhea^{12,13}. The first failure with combination therapy with both ceftriaxone and azithromycin was reported in the United Kingdom in 2016¹⁴. Therefore, the clonal spread of high-level ceftriaxone resistance alone and combinational resistance to both ceftriaxone and azithromycin should be carefully monitored.

Whole-genome sequencing (WGS) using a massive parallel sequencer (MPS) can simultaneously analyze multiple typing methods¹⁵, such as multilocus sequence typing (MLST)^{16,17}, *N. gonorrhoeae* multi-antigen sequence typing (NG-MAST)¹⁸, and *N. gonorrhoeae* sequence typing for antimicrobial resistance (NG-STAR) scheme for 7 alleles associated with cephalosporins, fluoroquinolones, penicillins, and macrolides¹⁹. WGS using MPS is widely used and allows for a comprehensive analysis, but its cost per strain is too expensive

for large-scale surveillance^{15,18,20,21}. Hence, a less expensive and more

efficient AMR *N. gonorrhoeae* genotyping method is required.

To address this need, we developed and validated a high throughput method

that takes full advantage of the features of the MiSeq (Illumina, San Diego, CA,

USA) sequence length that is commonly used in the analysis of bacteria.

Materials and Methods

Primers and condition for multiplex-tailed PCR

The sequences of the two primer sets and the concentrations for multiplex-

tailed PCR that amplifies 15 alleles of MLST, NG-MAST, and NG-STAR are

shown in Table S1. The primer design was a consensus sequence referring the

genome sequence data of the 14 WHO *N. gonorrhoeae* reference strains²¹. We

prepared primers with a tail sequence added to the 5' end based on the 16S

rRNA amplicon sequence preparation protocol from Illumina²². If the target

region was longer than the MiSeq merged read (approximately 450 bp), it was

amplified with two pairs of primers (left and right side) per region (Figure 1).

Template DNA was extracted by boiling method that from the bacterial cell

suspension in sterilized water by heating at 95°C for 15 minutes and then centrifuged. Primer concentrations were adjusted to ensure equal amplification efficiency for each allele, and validation was performed by quantitative PCR with KAPA SYBR FAST qPCR Master Mix (2×) (Nippon Genetics Co., Ltd., Tokyo, Japan). We used TaKaRa Ex Premier DNA Polymerase (TaKaRa Bio, Shiga, Japan) for multiplex-tailed PCR with an initial incubation at 94°C for 5 min, followed by 30 cycles at 98°C for 10 s, 50°C for 15 s, 68°C for 30 s, and then 68°C for 1 min. The left and right multiplex PCR products were then mixed and purified using Agencourt AMPure XP (Beckman Coulter, Inc., CA, USA).

Indexing-PCR and sequencing by MiSeq

For indexing-PCR, barcodes for multiplex sequencing by MiSeq were added to each sample of the multiplex PCR amplicon using the Nextera XT Index Kit v2 (Illumina, Inc., CA, USA) and 2X KAPA HiFi HotStart ReadyMix (Nippon Genetics Co., Ltd., Tokyo, Japan). Indexing-PCR was conducted at 95°C for 3 min, followed by 8 cycles of 95°C for 30 s, 55°C for 30s, 72°C for 30s, and then 72°C for 5 min. After purification, the pooled library was sequenced with 350-

bp/250-bp paired-end reads on the MiSeq platform using the MiSeq reagent kit

v3 600 cycle kit (Illumina).

Allele identification from the multiplex PCR amplicon sequencing data

Figure 2 shows an internal analysis pipeline that was constructed from an open-

source software installed on our laboratory server machine. The template DNA-

specific sequence of each primer on read was trimmed using Trimmomatic

v.0.39²³. Before SPAdes v.3.15.3²⁴ for *de novo* assembly, aligned reads to each

allele database were determined by BBtools ([https://jgi.doe.gov/data-and-](https://jgi.doe.gov/data-and-tools/software-tools/bbtools/)

[tools/software-tools/bbtools/](https://jgi.doe.gov/data-and-tools/software-tools/bbtools/)). Subsequently, the identification allele number of

the MLST, NG-MAST, and NG-STAR alleles was performed by similarity

analysis using the Basic Local Alignment Search Tool (BLAST)²⁵. These allele

sequence databases were obtained from PubMLST

(<https://pubmlst.org/organisms/neisseria-spp>, September 2022). Short and

partial *gyrA* and *parC* sequences were obtained by mapping them to reference

sequences using the BWA sw option²⁶ and then determining the consensus

sequences using Samtools²⁷. quinolone-resistance determining region (QRDR)

analysis was performed by the PointFinder tool²⁸.

148

149 **Strains for validation**

150 To validate our novel typing method for *N. gonorrhoeae*, we used 54 of 55

151 strains characterized by a draft WGS (one strain was excluded because of

152 discrepancies between the published and our reanalyzed draft WGS data)²⁸.

153 The strains contained eight MLST-ST, 38 NG-MAST-ST, 10 *penA* alleles, five

154 *mtrR* alleles, seven *porB* alleles, two *ponA* and 23S rRNA gene alleles, and

155 three QRDR mutation patterns detected in GyrA and ParC from translated *gyrA*

156 and *parC* (Tables 2 and 3).

157

158 **Counting average sequence reads depth**

159 The multiplex-PCR amplicon sequence reads were mapped using the BWA

160 “sw” option to each allele sequence obtained by *de novo* assembly. Then, the

161 sequence depth in the BAM file was determined using the pileup function in

162 BBtools (<https://jgi.doe.gov/data-and-tools/software-tools/bbtools/>).

163

164 **Nucleotide sequence accession numbers**

The MiSeq sequencing reads were deposited in the National Center for Biotechnology Information (NCBI) under BioProject number PRJNA901191. The Sequence Read Archive (SRA) accession number of fastq data of each sample were from SRR22270692 to SRR22270746 and are shown in Table S2.

Results

Evaluating sequencing read depth of each allele

The statistical range of the average sequence depth in all 15 alleles (*porB* for NG-MAST and NG-STAR was calculated as one) of each strain was as follows: mean = 425x-28,423x; median = 348x-28,320x; and standard deviation = 285x-17,901x. Sequence depth plots of each allele are shown in Figure 3.

Validating our method using a draft WGS as a reference

For the 54 strains of *N. gonorrhoeae* evaluated in our study, our novel typing method was in agreement with the draft WGS method using MiSeq in all alleles of MLST and NG-STAR (Tables 1 and 2). The NG-MAST results of 35 strains agreed with the draft WGS (Table 1). When evaluating the NG-MAST sequence

type (ST), only one strain (NG-MAST-ST6771) both agreed and disagreed with the draft WGS data (three agreements and one disagreement) (Table 1).

Predicting antimicrobial resistance

All strains harboring amino acid substitution patterns of S91F and D95G or S91F and D95N in GyrA were levofloxacin resistant (Table 2). Harboring a *penA*10.001 allele was associated with reduced susceptibility to cefixime but not to ceftriaxone. A strain with a *penA*169.001 allele was resistant to both cefixime and ceftriaxone. All strains harboring *penA*10.001, *penA*72.001, *penA*106.001, *penA*150.001, *penA*168.001, *penA*169.001, *mtrR*1, *mtrR*33, *porB*8, *porB*10, *porB*11, *porB*12, *porB*24, and *ponA*1 alleles showed resistance or reduced susceptibility to penicillin G. We did not observe an association between the 23S rRNA allele and azithromycin susceptibility.

Characterizing the disagreement with the draft WGS

We identified two causes of disagreement in *porB* allele detection (Table 3). First, for 13 of 15 strains, the BLAST analysis results for detecting an exact match allele were confounded by variable allele lengths in the NG-MAST

database (unlike the MLST database) leading to ambiguity. Second, we detected a single nucleotide difference at different positions in 2 of the 15 strains (TUM15755 and TUM15777).

We also determined two causes of disagreement in *tbpB* allele detection (Table 4). First, there were between two and seven mismatches in multiple *tbpB* PCR primers among five of the seven strains, which prevented the acquisition of *tbpB* contigs for these strains. Second, we found that the contig of *tbpB* was divided into two contigs in these seven strains.

Discussion

The successful performance of MLST and NG-STAR of our developed high-throughput typing method was validated in this study. By allowing the analysis of a large number of *N. gonorrhoeae* strains concurrently, this method is both cost-effective and efficient. Furthermore, this method can potentially be used for typing *N. gonorrhoeae* and predicting AMR phenotype without culture from gonorrhea patients because of the PCR amplification specificity of each allele.

217 Although this method is not intended to evaluate all aspects of *N. gonorrhoeae*,
218 it will likely be a valuable tool in large-scale AMR *N. gonorrhoeae* surveillance.
219
220 Even though allelic analysis of MLST and NG-STAR was an exact match, NG-
221 MAST did not produce the expected results with a small number of strains.
222 There are several reasons that may explain this discrepancy. First, false higher
223 than expected similarity scores for the allele were calculated by the BLAST
224 analysis, which was the primary cause of the poor performance for the *porB*
225 alleles in our study. These false higher similarity scores are likely due to the
226 inconsistent lengths of the sequences registered in the NG-MAST database. A
227 single nucleotide difference in a longer sequence will result in a higher similarity
228 score than an exact match of a shorter sequence. Second, the *porB*
229 disagreement in TUM15755 and TUM15777 is likely related to accidental
230 nucleotide amplification errors in the multiplex-tailed PCR. The sequencing
231 depth of *porB* was sufficient (more than 100×, Figure 3), so it is unlikely that
232 sequencing errors were the cause. There is an option to use PCR enzyme with
233 higher proofreading performance to prevent sequence amplification error in
234 PCR, but at the risk that a slight primer sequence mismatch could result in a

significant decrease in amplification reaction efficiency, Third, two types of poor
contig formation at the *tbpB* allele may be due to primer mismatches against the
allele and its surrounding sequences in the affected strains that resulted in
inefficient PCR amplification. This indirectly demonstrates the low conservation
of the allele sequence. This is not surprising because although NG-MAST has a
higher discrimination power than MLST, it makes primer design difficult to
amplify all alleles. Recently, a novel AMR *N. gonorrhoeae* was reported without
NG-MAST analysis²⁹, suggesting that NG-MAST may no longer be necessary in
typing *N. gonorrhoeae*. We believe that WGS, not NG-MAST, is required for
high resolution gonococcal transmission analysis³⁰.

Because Japan has been an epicenter for AMR *N. gonorrhoeae*^{12,15,31–34}, it is
appropriate to use Japanese strains to validate our typing methodology.
However, since the validation of the primer region mismatch is insufficient, it
may be necessary to update multiplex-tailed PCR primers and that PCR
conditions while using this method in the future. We are in process of applying
and analyzing this method to a historical collection constructed by 220 strains

isolated in Japan between 1972 and 2005. The preliminary results have been

mostly favorable (manuscript in preparation).

Notably, this method of NG-STAR for *gyrA* and *parC* only targets a narrow

sequence around the QRDRs and thus does not allow for allele identification of

gyrA and *parC*. Because the goal of NG-STAR is to detect mutations that

contribute to AMR, the allele number does not necessarily have to be

identifiable. The idea behind this is to keep the number of PCR products as

small as possible to minimize the vulnerability of the multiplex-tailed PCR. If

long read sequencers become available, the vulnerability of this method would

be reduced as the multiplex-tailed PCR, which is currently divided into left and

right assays, could be integrated into a single assay.

In conclusion, we have developed a cost- and labor-saving high throughput *N.*

gonorrhoeae genotyping method to add molecular epidemiological information

to strains collected by AMR *N. gonorrhoeae* surveillance. By avoiding the

sequencing of unnecessary regions in the three schemes, this method has

succeeded in increasing the number of strains that can be analyzed in one

MiSeq assay by approximately 5-fold compared with that obtained by WGS.

This method will be valuable in AMR *N. gonorrhoeae* as it is suitable for initial analyses and can quickly identify new strains for subsequent WGS analysis.

Acknowledgment

This research was supported in part by AMED under Grant Numbers

JP15fk0108014h0001, JP18fk0108062j0001, and JP21fk0108605j0001.

Table 1. MLST and NG-MAST agreement in our novel typing method versus draft whole-genome sequencing

MLST, multilocus sequence typing; NG-MAST, *N. gonorrhoeae* multi-antigen sequence typing

Target	Sequence type	Novel typing method vs. MiSeq draft whole-genome sequencing	
		Agreement	Disagreement
MLST	1579	6	0
	1594	3	0
	1597	1	0
	1901	19	0
	7359	11	0
	7360	1	0
	7827	1	0
	15378	1	0
Total		54	0

NG-MAST	247	0	1
	1407	3	0
	1513	2	0
	2212	2	0
	2392	0	1
	2958	1	0
	3422	1	0
	3709	1	0
	3926	1	0
	4015	1	0
	4018	2	0
	4186	0	1
	5334	0	1
	5687	2	0
	5887	1	0
	6543	0	3
	6769	0	1
	6771	3	1
	6773	1	0
	6798	0	4
	6800	2	0
	7038	2	0
	10611	1	0
	13194	1	0
	14169	1	0
	14174	0	1
	14201	1	0
	14245	0	1
	15501	1	0
	15526	0	1
	15528	0	1
	18001	1	0
	19619	1	0
	19620	1	0
	19621	1	0
	19622	0	1

19623	0	1
19706	1	0
Total	35	19

Table 2. Validation of NG-STAR by our novel typing method versus draft whole-genome sequencing and antimicrobial resistant phenotype prediction

^a Relationship between alleles and antimicrobial resistant phenotype: *penA*, penicillin G reduced susceptibility (>0.25 mg/L) / cefixime reduced susceptibility (>0.125 mg/L) / ceftriaxone reduced susceptibility (>0.125 mg/L); 23S rRNA, azithromycin non-susceptible (>1 mg/L); *mtrR*, *porB* and *ponA*, penicillin G reduced susceptibility (>0.25 mg/L); GyrA and ParC, levofloxacin resistant (>1 mg/L)¹⁵. NG-STAR, *N. gonorrhoeae* sequence typing for antimicrobial resistance; a, according to NG-STAR metadata; Del, deletion; amino acid abbreviations: A, alanine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; K, lysine; L, leucine; N, asparagine; P, proline; S, serine; T, threonine; V, valine; Y, tyrosine.

Novel typing method vs MiSeq draft whole genome sequencing					
Allele (length)	Allele number	Characteristics of the translated allele to AMR ^a	Agreement	Disagreement	Number of strains with antimicrobial resistant phenotype ^a
<i>penA</i> (1,745 – 1,752 bp)	1.001	Type I Non- Mosaic	4	0	0 / 0 / 0
	2.002	Type II Non- Mosaic; A517G	11	0	4 / 0 / 0
	5.002	Type V Non- mosaic; A517G, G543S	2	0	2 / 0 / 0
	10.001	Type X Mosaic; N513Y	18	0	18 / 18 / 0
	34.001	Type XXXIV Mosaic; N513Y	5	0	5 / 3 / 0
	72.001	Type 72 Mosaic; N513Y	8	0	8 / 8 / 0
	106.001	Type 106 Non- Mosaic; A517G, G543S	1	0	1 / 0 / 0
	150.001	Type 150 Semi- Mosaic; A517G G543S	3	0	3 / 0 / 0
	168.001	Type 168 Mosaic; A501V, N513Y	1	0	1 / 1 / 0
	169.001	Type 169 Mosaic; A311V, N513Y	1	0	1 / 1 / 1
	Total		54	0	31 / 1
<i>mtrR</i>	1	-35A Del	25	0	25
(643 – 719 bp)	4	Wild type	5	0	1

	33	Wild type	12	0	12
	42	A39T	11	0	5
	304	A39T	1	0	0
	Total		54	0	43
<i>porB</i>	8	G120K, A121D	14	0	14
(30 bp)	10	G120K, A121D	1	0	1
	11	G120K, A121N	16	0	16
	12	G120K, A121G	1	0	1
	14	<i>porB1a</i>	1	0	0
	24	G120N, A121N	1	0	1
	100	Wild Type	20	0	10
	Total		54	0	43
<i>ponA</i>	1	L421P	36	0	36
(75 bp)	100	Wild type	18	0	7
	Total		54	0	43
23S rRNA gene	8	Wild type	1	0	0
(566 – 570 bp)	100	Wild type	53	0	1
	Total		54	0	1
GyrA	NA	S91F and D95G	27	0	27
	NA	S91F and D95N	12	0	12
	NA	Wild type	15	0	0
	Total		54	0	39
ParC	NA	S88P	11	0	11
	NA	E91G	1	0	1
	NA	Wild type	42	0	27
	Total		54	0	39

298

299

300 **Table 3. Cause of disagreement in NG-MAST of *porB* allele with draft**

301 **whole-genome sequence**

302

Strain	Expected <i>porB</i>	Detected <i>porB</i>	Detection results	
ID	allele (draft	allele (top hit by	for the top hit	Causes of false allele detection
(TUM)	WGS)	BLAST)	allele	
15752	182	11317	Single nucleotide mismatch	<i>porB</i> 182 is three bases shorter on the 3' end than <i>porB</i> 11317 in the NG-MAST database
15755	4028	4028	Single nucleotide mismatch	Likely due to PCR error
15762	206	11889	Single nucleotide mismatch	<i>porB</i> 206 is three bases shorter on the 3' end than <i>porB</i> 11889 in the NG-MAST database
15763	2569	12073	Single nucleotide mismatch	<i>porB</i> 2569 is three bases shorter on the 3' end than <i>porB</i> 12073 in the NG-MAST database
15767	3854	11337	Single nucleotide mismatch	<i>porB</i> 3854 is three bases shorter on the 3' end than <i>porB</i> 11337 in the NG-MAST database
15770	4033	11424	Single nucleotide mismatch	<i>porB</i> 4033 is three bases shorter on the 3' end than <i>porB</i> 11424 in the NG-MAST database
15775	3854	11337	Single nucleotide mismatch	<i>porB</i> 3854 is three bases shorter on the 3' end than <i>porB</i> 11337 in the NG-MAST database
15777	4061	9083	Single nucleotide mismatch	Likely due to PCR error
15778	3854	11337	Single nucleotide mismatch	<i>porB</i> 3854 is three bases shorter on the 3' end than <i>porB</i> 11337 in the NG-MAST database
15782	206	11889	Single nucleotide mismatch	<i>porB</i> 206 is three bases shorter on the 3' end than <i>porB</i> 11889 in the NG-MAST database
15790	4033	11424	Single nucleotide mismatch	<i>porB</i> 4033 is three bases shorter on the 3' end than <i>porB</i> 11424 in the NG-MAST database
15793	11399	4069	Single nucleotide mismatch	<i>porB</i> 11399 is six bases shorter on the 3' end than <i>porB</i> 4069 in the NG-MAST database

15794	2547	11165	Single nucleotide mismatch	<i>porB</i> 2547 is three bases shorter on the 3' end than <i>porB</i> 11165 in the NG-MAST database
15796	4033	11424	Single nucleotide mismatch	<i>porB</i> 4033 is three bases shorter on the 3' end than <i>porB</i> 11424 in the NG-MAST database
15798	4033	11424	Single nucleotide mismatch	<i>porB</i> 4033 is three bases shorter on the 3' end than <i>porB</i> 11424 in the NG-MAST database

303

304

305 **Table 4. Primer mismatch disagreement of the *tbpB* allele between NG-**

306 **MAST and the draft whole-genome sequence**

307

Strain ID (TUM)	Expected <i>tbpB</i> allele (draft WGS)	Detection results	Number of mismatches on primers against MiSeq draft WGS data			
			left (forward)	left (reverse)	right (forward)	right (reverse)
15752	21	Contig divided into two	0	1	0	0
15756	1365	No contig generation in <i>tbpB</i>	0	6	7	5
15759	27	No contig generation in <i>tbpB</i>	0	0	2	5
15764	2532	No contig generation in <i>tbpB</i>	5	6	5	5

15782	2528	No contig generation in <i>tbpB</i>	0	2	5	0
15789	1127	No contig generation in <i>tbpB</i>	0	4	3	2
15794	21	Contig divided into two	0	1	0	0

308

309

310

311 Figure legends

312 **Figure 1: Model of MiSeq sequencing library preparation for MLST, NG-**
313 **MAST, and NG-STAR by multiplex PCR**

314 A. Multiplex-tailed PCR for target region amplification.

315 Design the primers for the left and right sides in the same arrangement for all
316 target alleles. There is a common tail sequence on the 5' side of the primers of
317 the multiplex PCR followed by indexing-PCR (the slanted part of the primer). To
318 avoid amplification by Left_Rv and Right_Fw primers, multiplex PCR for the left-
319 and right-side primers is performed separately.

B. Indexing-PCR for adding the index sequence by PCR.

Figure 2: Read data analysis workflow

MiSeq read data of multiplex PCR amplicon sequencing was analyzed by a *de novo* assembly-based workflow.

Figure 3. Sequence depth on each allele

Sequence depth of our novel method using MiSeq is shown as a box plot. The horizontal line in the box indicates the median, the upper end indicates the third quartile, and the lower end indicates the first quartile. The upper end of the whisker is drawn to the largest data point within 1.5× of the interquartile range (IQR) from the third quartile, while the lower end of the whisker is drawn to the smallest value within 1.5× of the IQR from the first quartile. This boxplot was generated by R ggplot2.

Reference

- 337 1. World Health Organization (WHO). Report on global sexually transmitted
338 infection surveillance, 2018. 2018. Available at:
339 <https://www.who.int/publications/i/item/9789241565691> [verified January 2023].
- 340 2. Barbaric J, Kuchukhidze G, Seguy N, *et al.* Surveillance and epidemiology of
341 syphilis, gonorrhoea and chlamydia in the non-European Union countries of the
342 World Health Organization European Region, 2015 to 2020. *Euro Surveill* 2022;
343 **27**: 2100197.
- 344 3. Unemo M, Lahra MM, Cole M, *et al.* World Health Organization Global
345 Gonococcal Antimicrobial Surveillance Program (WHO GASP): review of new
346 data and evidence to inform international collaborative actions and research
347 efforts. *Sex Health* 2019; **16**: 412–25.
- 348 4. Unemo M, Shafer WM. Antimicrobial resistance in *Neisseria gonorrhoeae* in
349 the 21st century: past, evolution, and future. *Clin Microbiol Rev* 2014; **27**: 587–
350 613.
- 351 5. Seña AC, Bachmann L, Johnston C, *et al.* Optimising treatments for sexually
352 transmitted infections: surveillance, pharmacokinetics and pharmacodynamics,

353 therapeutic strategies, and molecular resistance prediction. *Lancet Infect Dis*

354 2020; **20**: e181–91.

355 6. World Health Organization (WHO). WHO guidelines for the treatment of

356 *Neisseria gonorrhoeae*. 2016. Available at:

357 <https://www.who.int/publications/i/item/9789241549691> [verified January 2023].

358 7. Garrett NJ, McGrath N, Mindel A. Advancing STI care in low/middle-income

359 countries: has STI syndromic management reached its use-by date? *Sex*

360 *Transm Infect* 2017; **93**: 4–5.

361 8. World Health Organization (WHO). Guidelines for the management of

362 symptomatic sexually transmitted infections. 2021. Available at:

363 <https://www.who.int/publications/i/item/9789240024168> [verified January 2023].

364 9. Sarenje KL, Ngalamika O, Maimbolwa MC, Siame A, Munsaka SM, Kwenda

365 G. Antimicrobial resistance of *Neisseria gonorrhoeae* isolated from patients

366 attending sexually transmitted infection clinics in Urban Hospitals, Lusaka,

367 Zambia. *BMC Infect Dis* 2022; **22**: 688.

- 368 10. Ohnishi M, Watanabe Y, Ono E, *et al.* Spread of a chromosomal cefixime-
369 resistant *penA* gene among different *Neisseria gonorrhoeae* lineages.
370 *Antimicrob Agents Chemother* 2010; **54**: 1060–7.
- 371 11. Takahata S, Senju N, Osaki Y, Yoshida T, Ida T. Amino acid substitutions in
372 mosaic penicillin-binding protein 2 associated with reduced susceptibility to
373 cefixime in clinical isolates of *Neisseria gonorrhoeae*. *Antimicrob Agents*
374 *Chemother* 2006; **50**: 3638–45.
- 375 12. Ohnishi M, Golparian D, Shimuta K, *et al.* Is *Neisseria gonorrhoeae* initiating
376 a future era of untreatable gonorrhea?: detailed characterization of the first
377 strain with high-level resistance to ceftriaxone. *Antimicrob Agents Chemother*
378 2011; **55**: 3538–45.
- 379 13. Unemo M, Golparian D, Sánchez-Busó L, *et al.* The novel 2016 WHO
380 *Neisseria gonorrhoeae* reference strains for global quality assurance of
381 laboratory investigations: phenotypic, genetic and reference genome
382 characterization. *J Antimicrob Chemother* 2016; **71**: 3096–108.
- 383 14. Fifer H, Natarajan U, Jones L, *et al.* Failure of dual antimicrobial therapy in
384 treatment of gonorrhea. *N Engl J Med* 2016; **374**: 2504–6.

- 385 15. Hanao M, Aoki K, Ishii Y, Shimuta K, Ohnishi M, Tateda K. Molecular
386 characterization of *Neisseria gonorrhoeae* isolates collected through a national
387 surveillance programme in Japan, 2013: evidence of the emergence of a
388 ceftriaxone-resistant strain from a ceftriaxone-susceptible lineage. *J Antimicrob*
389 *Chemother* 2021; **76**: 1769–75.
- 390 16. Viscidi RP, Demma JC. Genetic diversity of *Neisseria gonorrhoeae*
391 housekeeping genes. *J Clin Microbiol* 2003; **41**: 197–204.
- 392 17. Bennett JS, Jolley KA, Sparling PF, *et al.* Species status of *Neisseria*
393 *gonorrhoeae*: evolutionary and epidemiological inferences from multilocus
394 sequence typing. *BMC Biol* 2007; **5**: 35.
- 395 18. Martin IMC, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Rapid
396 sequence-based identification of gonococcal transmission clusters in a large
397 metropolitan area. *J Infect Dis* 2004; **189**: 1497–505.
- 398 19. Demczuk W, Sidhu S, Unemo M, *et al.* *Neisseria gonorrhoeae* Sequence
399 Typing for Antimicrobial Resistance, a novel antimicrobial resistance multilocus
400 typing scheme for tracking global dissemination of *N. gonorrhoeae* strains. *J*
401 *Clin Microbiol* 2017; **55**: 1454–68.

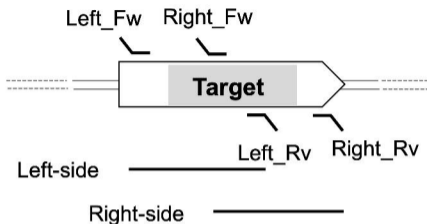
- 402 20. Unemo M, Lahra MM, Escher M, *et al.* WHO global antimicrobial resistance
403 surveillance for *Neisseria gonorrhoeae* 2017–18: a retrospective observational
404 study. *The Lancet Microbe* 2021; **2**: e627–36.
- 405 21. Wang F, Liu J-W, Li Y-Z, *et al.* Surveillance and molecular epidemiology of
406 *Neisseria gonorrhoeae* isolates in Shenzhen, China, 2010–2017. *Journal of*
407 *Global Antimicrobial Resistance* 2020; **23**: 269–74.
- 408 22. Illumina. 16S Metagenomic Sequencing Library Preparation. 2013.
409 Available at:
410 [https://support.illumina.com/documents/documentation/chemistry_documentation/](https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf)
411 [n/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf](https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf) [verified January
412 2023].
- 413 23. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
414 sequence data. *Bioinformatics* 2014; **30**: 2114–20.
- 415 24. Bankevich A, Nurk S, Antipov D, *et al.* SPAdes: a new genome assembly
416 algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**:
417 455–77.

- 418 25. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment
419 search tool. *J Mol Biol* 1990; **215**: 403–10.
- 420 26. Li H, Durbin R. Fast and accurate short read alignment with Burrows-
421 Wheeler transform. *Bioinformatics* 2009; **25**: 1754–60.
- 422 27. Li H, Handsaker B, Wysoker A, *et al.* The Sequence Alignment/Map format
423 and SAMtools. *Bioinformatics* 2009; **25**: 2078–9.
- 424 28. Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM.
425 PointFinder: a novel web tool for WGS-based detection of antimicrobial
426 resistance associated with chromosomal point mutations in bacterial pathogens.
427 *J Antimicrob Chemother* 2017; **72**: 2764–8.
- 428 29. Berçot B, Caméléna F, Mérimèche M, *et al.* Ceftriaxone-resistant, multidrug-
429 resistant *Neisseria gonorrhoeae* with a novel mosaic *penA*-237.001 gene,
430 France, June 2022. *Eurosurveillance* 2022; **27**: 2200899.
- 431 30. Quainoo S, Coolen JPM, van Hijum SAFT, *et al.* Whole-genome sequencing
432 of bacterial pathogens: the future of nosocomial outbreak analysis. *Clin*
433 *Microbiol Rev* 2017; **30**: 1015–63.

- 434 31. Ameyama S, Onodera S, Takahata M, *et al.* Mosaic-like structure of
435 penicillin-binding protein 2 gene (*penA*) in clinical isolates of *Neisseria*
436 *gonorrhoeae* with reduced susceptibility to cefixime. *Antimicrob Agents*
437 *Chemother* 2002; **46**: 3744–9.
- 438 32. Ito M, Deguchi T, Mizutani K-S, *et al.* Emergence and spread of *Neisseria*
439 *gonorrhoeae* clinical isolates harboring mosaic-like structure of penicillin-binding
440 protein 2 in Central Japan. *Antimicrob Agents Chemother* 2005; **49**: 137–43.
- 441 33. Shimuta K, Watanabe Y, Nakayama S-I, *et al.* Emergence and evolution of
442 internationally disseminated cephalosporin-resistant *Neisseria gonorrhoeae*
443 clones from 1995 to 2005 in Japan. *BMC Infect Dis* 2015; **15**: 378.
- 444 34. Nakayama S-I, Shimuta K, Furubayashi K-I, Kawahata T, Unemo M,
445 Ohnishi M. New ceftriaxone- and multidrug-resistant *Neisseria gonorrhoeae*
446 strain with a novel mosaic *penA* gene isolated in Japan. *Antimicrob Agents*
447 *Chemother* 2016; **60**: 4339–41.

Figure 1

A. Multiplex-tailed PCR for target region amplification



B. Indexing-PCR

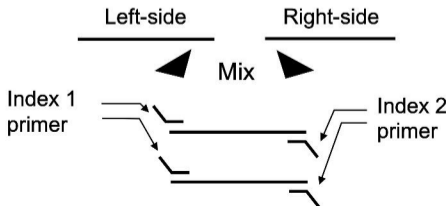


Figure 2

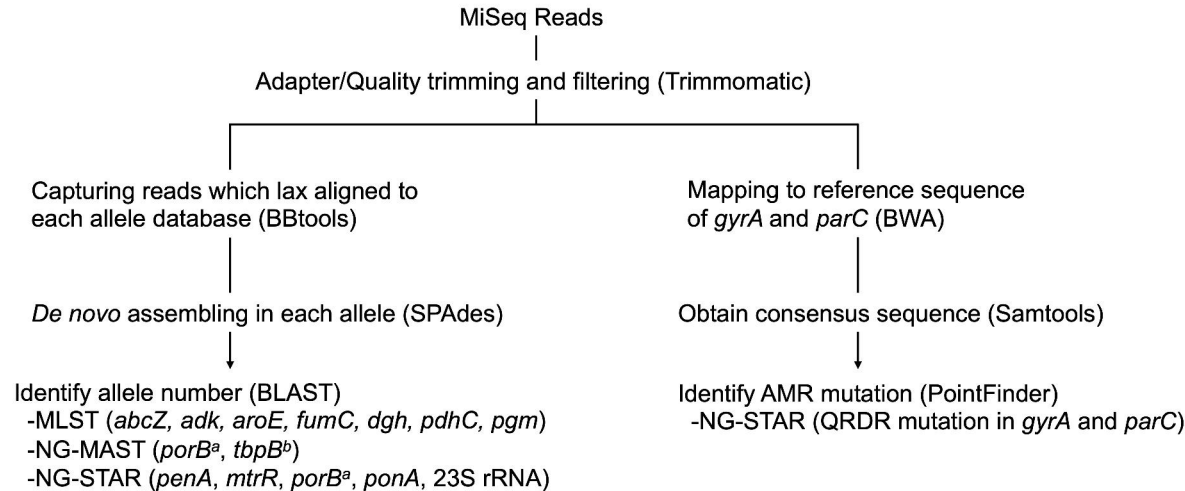


Figure 3

