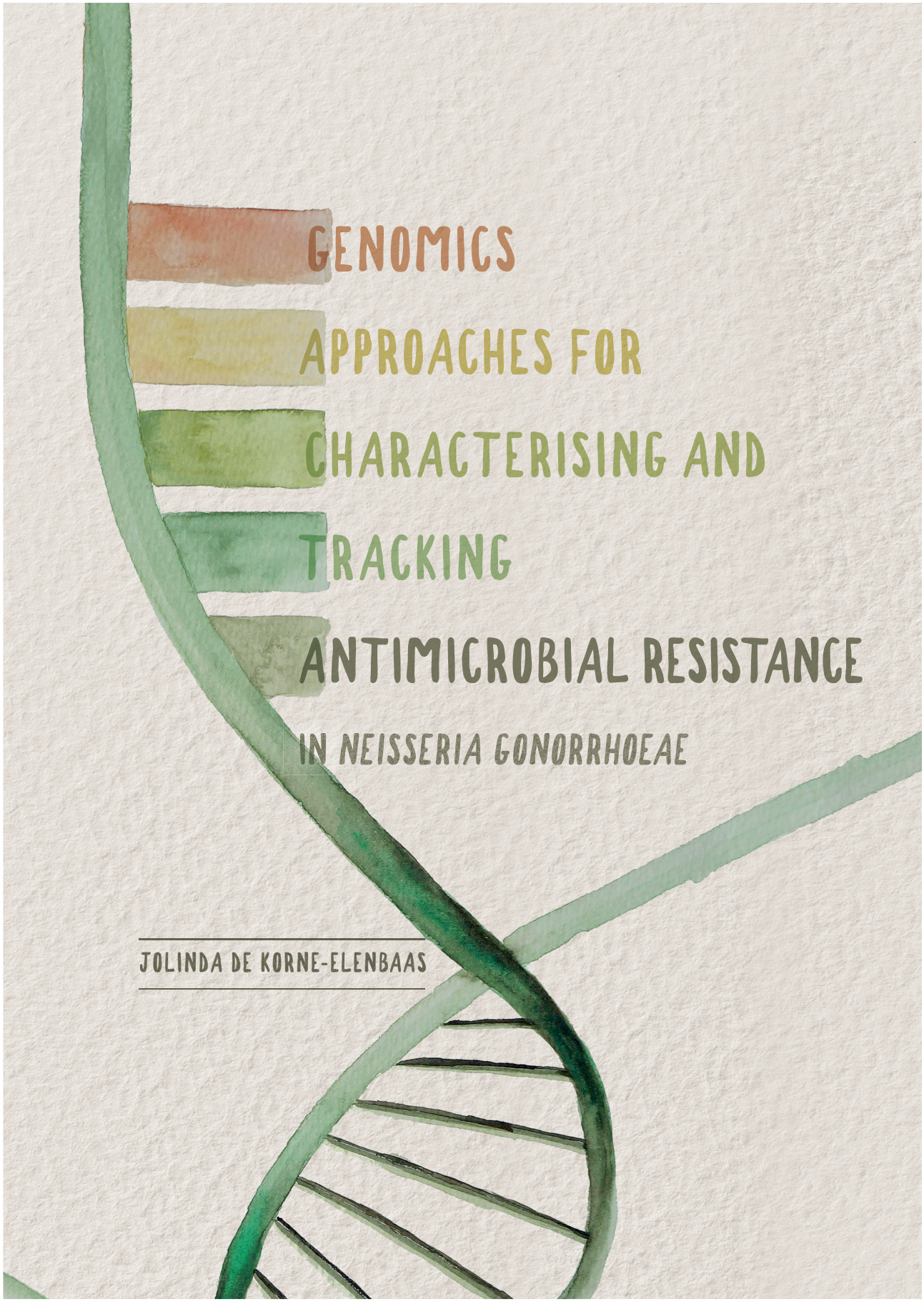


GENOMICS APPROACHES FOR CHARACTERISING AND TRACKING ANTIMICROBIAL RESISTANCE IN *NEISSERIA GONORRHOEAE*

JOLINDA DE KORNE-ELENBAAS

GENOMICS APPROACHES FOR CHARACTERISING AND TRACKING ANTIMICROBIAL RESISTANCE IN *NEISSERIA GONORRHOEAE*

JOLINDA DE KORNE-ELENBAAS

Genomics approaches for characterising and tracking antimicrobial resistance in *Neisseria gonorrhoeae*

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 18 januari 2023, te 11.00 uur
door Jacoba Wilhelmina Elenbaas
geboren te Goes

PROMOTIECOMMISSIE

Promotor:	prof. dr. H.J.C. de Vries	AMC-UvA
Copromotores:	dr. S.M. Bruisten	GGD Amsterdam
	dr. A.P. van Dam	AMC-UvA
Overige leden:	prof. dr. N.M. van Sorge	AMC-UvA
	prof. dr. C. Schultsz	AMC-UvA
	prof. dr. C.A. Russell	AMC-UvA
	prof. dr. W.J. Wiersinga	AMC-UvA
	prof. dr. P.H.M. Savelkoul	Universiteit Maastricht
	dr. C. Kenyon	Instituut voor Tropische Geneeskunde Antwerpen
	dr. O.B. Harrison	Oxford University

Faculteit der Geneeskunde

COLOPHON

Genomics approaches for characterising and tracking antimicrobial resistance in *Neisseria gonorrhoeae*
PhD thesis, University of Amsterdam, the Netherlands

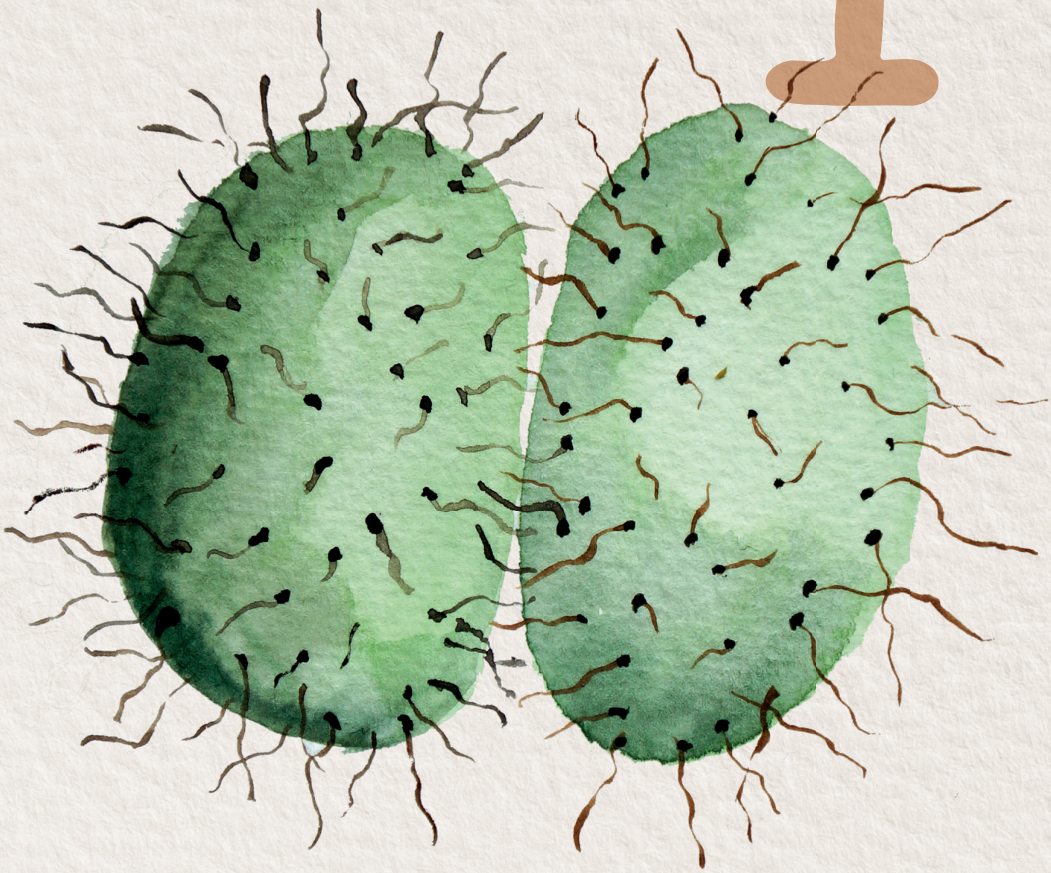
ISBN 978-94-6421-980-7

Watercolour art: Janine Hoekman
Printing: Ipskamp Printing | proefschriften.net
Layout and (cover) design: Leo Orth | persoonlijkproefschrift.nl

Financial support for printing of this thesis was kindly provided by the Public Health Service of Amsterdam, the Amsterdam UMC and NYtor B.V.

TABLE OF CONTENTS

1



GENERAL INTRODUCTION

ANTIMICROBIAL RESISTANCE AS MAJOR PUBLIC HEALTH THREAT

The discovery of antimicrobials was one of the greatest medical revolutions of the 20th century. With the introduction of antibiotics into clinical use, effective treatments for bacterial infections became available, which vastly reduced fatalities caused by common illnesses, surgery and childbirth. Alexander Fleming, who discovered penicillin in 1928, already warned about the risk of antimicrobial resistance (AMR) development through excessive use of antibiotics. Yet today, almost one century later, AMR is globally on the rise and poses a major public health threat. Worldwide, an estimated 1.27 million deaths were caused directly by AMR in 2019, meaning that those deaths would not have occurred when the causative pathogen would have been susceptible to treatment[1]. When taking deaths associated with AMR into account as well, this estimation would be more than twice as high. Moreover, these are most likely underestimations because of the lack of surveillance in low- and middle-income countries, which are the countries with the highest AMR burden[2]. The World Health Organization (WHO) has taken several steps to control AMR, including the publication of a list of multi-drug resistant bacteria for which treatment options are scarce and the development of new antibiotics is urgent[3]. The sexually transmittable bacterium *Neisseria gonorrhoeae* is one of the high priority pathogens on this list.

NEISSERIA GONORRHOEAE INFECTION AND DISEASE

PREVALENCE

N. gonorrhoeae causes the sexually transmitted infection (STI) gonorrhoea, of which the prevalence is globally on the rise[4]. The WHO estimated that 0.9% of the world's female population and 0.7% of the male population had urogenital gonorrhoea in 2016, corresponding to 30.6 million cases worldwide, with highest prevalence estimates in the African region and lowest estimates in Europe[5]. National data on case rates are only available for high-income countries that have public health agencies with comprehensive surveillance systems, such as the USA, Australia, Canada and multiple countries in Europe. For these countries, increasing gonorrhoea case rates have been observed over the last years with disproportionally high percentages of cases among men who have sex with men (MSM)[6–10].

PATHOGENESIS

N. gonorrhoeae - the gonococcus - is a fastidious Gram-negative, oxidase positive diplococcus, belonging to the Neisseriaceae family. Although most species of this family are commensals residing in the human nasopharynx, *N. gonorrhoeae* and its close relative *Neisseria meningitidis* are human pathogenic species, with the latter being a major cause of bacterial meningitis[11,12]. *N. gonorrhoeae* is a mucosal coloniser, able to infect human urethral, endocervical, rectal, oropharyngeal and ocular mucosa. Transmission occurs between mucosa via direct contact

with infected secretions. Adherence to mucosal cells is mediated through distinct bacterial outer membrane structures, including opacity (Opa) proteins, the major outer membrane porin protein (porB) and type IV pili[13]. These pili are able to adhere to the human epithelial cell surface and can subsequently retract to bring the gonococcus closer to the cell surface. Type IV pili also play crucial roles in natural competence, motility and immune evasion through phase- and antigenic variation[14–16]. After cell adherence, the gonococci form microcolonies on the cell surface and invade the epithelium, thereby activating the innate immune response. Through activation of inflammatory transcription factors and the release of pro-inflammatory cytokines and chemokines, large numbers of neutrophils are recruited to the infection site. The neutrophils phagocytose the gonococci, followed by efflux of neutrophils from the epithelium which makes up a purulent exudate (reviewed in[17]). Since part of the phagocytosed gonococci survive and replicate inside neutrophils, this is arguably an important transmission route for *N. gonorrhoeae*[18].

SYMPTOMS

The word gonorrhoea comes from the translation of the Greek words *gonos* (seed) and *rhoe* (flow), as mentioned already in the Bible (Leviticus 15). The 'flow of seed' reflects the frequent manifestation of gonococcal colonisation of the male *tractus urogenitalis*, which in most cases results in urethritis with penile discharge as a common symptom (mimicking semen). Other symptoms of male urogenital *N. gonorrhoeae* infection can be dysuria, bleeding or swelling and pain of the foreskin and testicles. In women, urogenital *N. gonorrhoeae* infections often stay asymptomatic, however they can lead to vaginal discharge, dysuria, bleeding and abdominal pain. The same holds for male and female anal- or pharyngeal infections, which can come with burning or itching but often stay asymptomatic and therefore unnoticed. When left untreated, these infections provide an important reservoir for further *N. gonorrhoeae* transmission[19]. Untreated urogenital infections potentially lead to epididymitis in men and pelvic inflammatory disease in women. Rarely, gonorrhoea can evolve into a systemic infection (disseminated gonococcal infections) when *N. gonorrhoeae* enters the bloodstream, causing septic arthritis, endocarditis, or in very rare cases meningitis[4,20].

DIAGNOSTICS

The oldest method for the detection of *N. gonorrhoeae* is microscopy, used for the identification of Gram-negative diplococci in a Gram staining. Although multiple diagnostic tools with higher sensitivity have been developed over the last decade, microscopy is still used to detect gonococci in male urethral samples, due to the short time to result. Culturing has long been used for *N. gonorrhoeae* detection, and its sensitivity improved marginally after the introduction of selective gonococcal culture medium in the 1960s[21]. However, more sensitive methods that detect gonococcal genetic material became available from the 1990s onwards[22]. These methods entail nucleic acid amplification (NAAT) tests, such as the polymerase chain reaction (PCR) and the transcription mediated amplification (TMA) assay. The latter is currently the most sensitive

test and works by amplifying bacterial ribosomal RNA (rRNA), which occurs in a much higher starting amount than DNA in a bacterial cell. Culturing is still part of the diagnostic toolset albeit no longer for detection, but mainly for phenotypic characterisation. For this purpose, pure cultures are exposed to different concentrations of antibiotics, enabling the determination of the minimum inhibitory concentration (MIC) which reflects the susceptibility of the isolate to the antibiotic. In high-resource settings, antimicrobial susceptibility profiles are routinely determined to collect data for *N. gonorrhoeae* AMR surveillance[23].

THE HISTORY OF ANTIMICROBIAL THERAPY AND RESISTANCE IN *N. GONORRHOEAE*

SULFONAMIDES

The first antibiotic sulfanilamide was discovered in 1908 and multiple other antibiotics belonging to the group of sulfonamides were discovered afterwards. Sulfonamides were the first antibiotics used for gonorrhoea treatment in the 1930s[24]. This class of antibiotics has a similar chemical structure as *para*-aminobenzoic acid (PABA) and binds to the bacterial dihydropteroate synthase (DHPS) enzyme, thereby competing with PABA in binding to these enzymes. Antibiotic binding to DHPS enzymes inhibits synthesis of folic acid, which is needed for bacterial RNA and DNA synthesis[25]. Overproduction of PABA overrules antibiotic activity and causes resistance to sulfonamides in *N. gonorrhoeae*. Sulfonamide resistance in *N. gonorrhoeae* can also be mediated by point mutations or mosaicism in the *folP* gene, encoding the target DHPS enzyme (Figure 1)[26,27]. Resistance emerged rapidly with the majority of strains being resistant by the late 1940s[24,28].

PENICILLIN

Penicillin replaced sulfonamides for the treatment of gonorrhoea and stayed the preferred treatment until the mid-1970s, when doses had already been increased to a level causing high injection discomfort and penicillinase-producing strains were circulating[29,30]. Penicillin belongs to the group of β-lactam antibiotics, which have chemical structures containing a β-lactam ring. Its bactericidal activity occurs through binding of the β-lactam ring to transpeptidase enzymes (penicillin-binding proteins), which inhibits the peptidoglycan cross-linking in the bacterial cell wall. Resistance to penicillin in *N. gonorrhoeae* can be mediated through carriage of a plasmid encoding a TEM-1 type β-lactamase (Figure 1), also referred to as penicillinase, which inhibits β-lactam activity through hydrolyzing the cyclic amide bond in the β-lactam ring[20]. This so-called β-lactamase plasmid is presumably acquired from *Haemophilus influenzae* and has rapidly spread internationally after its introduction[31]. Different types of the plasmid and the β-lactamase enzyme have been evolving since, with the TEM-1 type β-lactamase still being widely distributed among gonococci carrying the plasmid[32].

Chromosomally-mediated penicillin resistance mechanisms include mutations that prevent target protein binding in interplay with mutations that lead to increased efflux- or decreased influx of penicillin[33]. The strongest decrease in penicillin susceptibility is caused by mutations in its main target, the *penA* encoded penicillin-binding protein-2 (Figure 1)[34]. More recently, in the context of extended-spectrum cephalosporin resistance, mosaic *penA* structures have been described with up to 70 amino acid changes that can also confer high-level resistance to penicillin[35]. A mutation in the second penicillin target, the *ponA* encoded penicillin-binding protein-1, is additionally found in high-level resistant strains, however its impact is less defined[36]. Penicillin susceptibility can decrease further through mutations in the outer membrane porin protein PorB (encoded by *porB1b,* also known as *penB)* which decrease the membrane permeability and thereby the penicillin influx, or through upregulated penicillin efflux, caused by overexpression of the multiple transferable resistance CDE (MtrCDE) efflux pump[37–40].

TETRACYCLINE

During the emergence of high-level resistance to penicillin, tetracycline was introduced as treatment for gonorrhoea. Tetracyclines inhibit protein synthesis by binding to the 30S ribosomal subunit and subsequently inhibiting the binding of aminoacyl-tRNA to the mRNA-ribosome complex. The susceptibility to tetracycline soon decreased after acquisition of the *tetM* gene on the gonococcal conjugative plasmid, which is the largest gonococcal plasmid that can be transferred between gonococcal strains by conjugation[41]. This *tetM* gene, described for different bacterial species[42], encodes the tetracycline resistance protein TetM, which competes with tetracycline for binding to the ribosome and subsequently restores the protein synthesis (Figure 1)[43]. Remarkably, gonococcal strains carrying the conjugative plasmid with *tetM* are significantly more prevalent in low- and middle-income countries, indicating that factors in these countries stimulate maintenance of this plasmid-mediated resistance[32]. Chromosomally-mediated resistance mechanisms involve mutations in the ribosomal protein target structure (*rpsJ*), reduced influx (porB*)* or increased efflux (MtrCDE)[44,45].

SPECTINOMYCIN

Another antibiotic used for gonorrhoea treatment when high-level penicillin resistance emerged was the aminocyclitol spectinomycin, commercialised for gonorrhoea treatment in the 1960s[46]. Spectinomycin targets the 30S ribosomal subunit by binding to 16S rRNA and subsequently inhibits bacterial protein synthesis[47]. Resistance to spectinomycin in *N. gonorrhoeae* is caused by blocking antibiotic binding to the ribosomal target, initially mediated by a missense mutation in the 16S rRNA target, and later by missense mutations in the *rpsE* gene encoding the 30S ribosomal protein S5 (Figure 1)[48,49]. Spectinomycin was abandoned as first-line gonorrhoea therapy once unacceptable levels of spectinomycin resistance were reached in the early 1980s[50]. Despite resistance currently being very rare, the fear for rapid resistance development and the

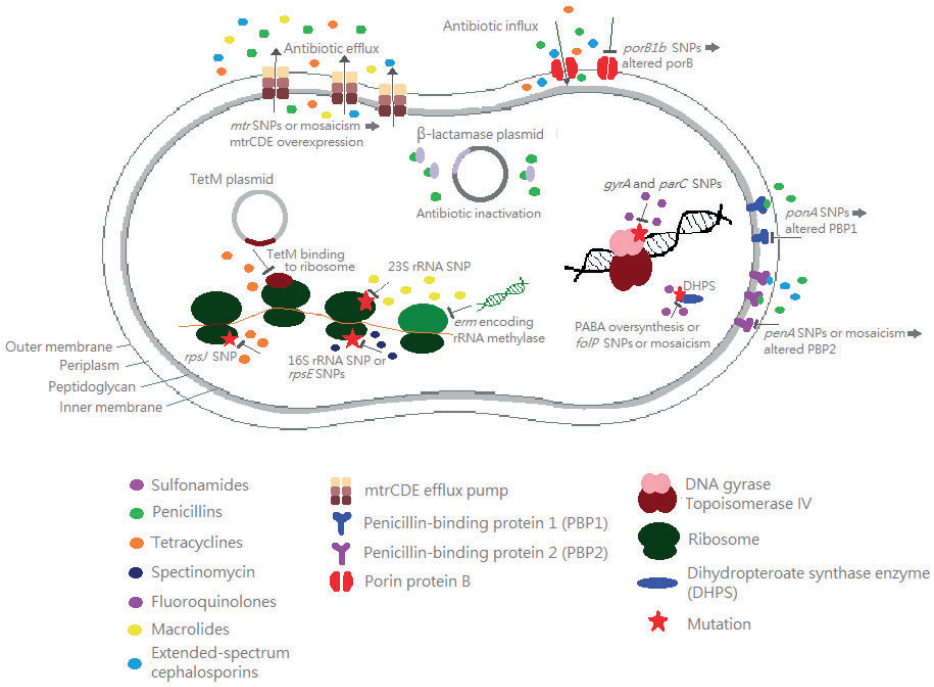


FIGURE 1. SCHEMATIC REPRESENTATION OF AMR MECHANISMS IN *N. GONORRHOEAE* TO FORMERLY- OR CURRENTLY USED ANTIBIOTICS FOR GONORRHOEA TREATMENT.

suboptimal treatment efficacy for pharyngeal gonorrhoea hampers its reintroduction as first-line gonorrhoea treatment[51].

FLUOROQUINOLONES

From mid- to late 1980s, the fluoroquinolone ciprofloxacin was widely used to treat gonorrhoea. However, resistance arose soon and in the mid-1990s already, the first countries abandoned ciprofloxacin as first-line therapy for gonorrhoea[52]. Many Asian and European countries followed in early- to mid-2000s due to the high prevalence of fluoroquinolone-resistant strains globally[53]. Fluoroquinolones inhibit DNA gyrase and topoisomerase IV, which are both type II topoisomerases that play important roles in bacterial DNA metabolism by breaking and joining double-stranded supercoiled DNA in order to relax its structure. Mutations in *gyrA* (encoding DNA gyrase subunit GyrA) and *parC* (encoding topoisomerase IV subunit ParC) genes, prevent ciprofloxacin from binding to the primary target proteins and inhibit its action (Figure 1)[54]. Specifically, missense mutations at positions S91 and D95 in *gyrA* cause ciprofloxacin resistance and co-occurrence of missense *parC* mutations even further increase ciprofloxacin resistance[55]. The global prevalence of ciprofloxacin resistant isolates is high, with measured prevalences of 35% in the American Gonococcal Isolate Surveillance Project (GISP) in 2019, 57%

in the European Gonococcal Antimicrobial Surveillance Project (Euro-GASP) in 2019 and 53% in the Dutch Gonococcal Resistance to Antibiotics Surveillance (GRAS) project in 2021[56–58]. The global WHO-GASP reported ciprofloxacin resistance in all countries for which data was available in 2018, even ranging to a prevalence of >70% in countries in Asia and South-America[59]. Despite the high resistance prevalence, British guidelines recommend treatment with ciprofloxacin in case susceptibility is known prior to treatment, since ciprofloxacin is orally administered and has relatively few side effects[60]. To enable rapid susceptibility testing, several NAAT assays for the detection of the predominant resistance determinant, the *gyrA*-S91F mutation, have been developed[61–67].

MACROLIDES

In many countries, the macrolide azithromycin has been used from the 1980s onwards for the treatment of bacterial STIs, including *Chlamydia trachomatis* infection and gonorrhoea. Macrolides inhibit protein synthesis by binding to the 50S ribosomal subunit and subsequently preventing peptidyltransferase from adding tRNA-bound amino acids to the peptide chain as well as blocking the peptide exit channel through interaction with the 23S rRNA[68]. Although a single oral dose of 2 grams azithromycin is effective against azithromycin-susceptible urogenital, rectal and oropharyngeal *N. gonorrhoeae* infections, empirical monotherapy with azithromycin has never been recommended as first-line treatment due to possible adverse effects and rapid resistance development[69,70]. Resistance to azithromycin arose from mid- to late 1990s in Latin America, where azithromycin treatment for STIs was introduced early[71–73]. Resistance spread globally and the prevalence of isolates with MICs above the epidemiological cut-off became highest in Europe, with a worrying increasing trend from 4% in 2016 to 9% in 2019[57,59]. This high prevalence can potentially be explained by the relationship between antimicrobial consumption and resistance prevalence, since Europe has the highest reported macrolide consumption of the world[74]. In the Netherlands, the prevalence increased even to 18% in 2021, despite macrolides not being used for gonorrhoea treatment[58].

In the past, genomic presence of erythromycin ribosome methylase *(erm)* genes encoding rRNA methylases caused low-level azithromycin resistance, by blocking binding to the 23S rRNA target gene (Figure 1)[75]. In more recent azithromycin-resistant isolates, resistance is either caused by mutations in the 23S rRNA target gene, or by an increased macrolide efflux[76,77]. The A2059G or C2611T mutations in 23S rRNA can cause respectively high- or low-level azithromycin resistance[78,79]. The level of resistance caused by 23S rRNA mutations is also dependent on the number of 23S rRNA alleles that have the mutation, as a single mutated allele does not lead to resistance[80]. Efflux of macrolides is predominantly mediated by the MtrCDE efflux pump system, as for other previously mentioned antimicrobials. *MtrCDE* gene expression is regulated by the MtrR repressor, which binds to the promoter of the *mtrCDE* operon[81]. The G45D or A39T mutations in the *mtrR* gene or the alanine deletion in its repetitive region inhibit MtrR activity,

resulting in overexpression of the MtrCDE efflux pump and a subsequent increased efflux of azithromycin[82]. *MtrR* mutations are often reported in azithromycin intermediate-resistant isolates. More recently, mosaic *mtr* gene sequences have been found in low-level azithromycin resistant isolates[83–85]. This mosaicism is a result of the *mtrR* and *mtrCDE* genes being hotspots for interspecies recombination, with parts of the *N. meningitidis* and *N. lactamica mtr* genes being incorporated in the *N. gonorrhoeae mtr* genes[86]. Strains with mosaic *mtr* genes are globally on the rise, reflecting the increasing trend in low-level azithromycin resistance over the last years[85,87–91].

Other efflux pumps have also been reported to affect the azithromycin susceptibility, although to a lower extent or only when co-occurring with *mtr* mutations. These include an overexpressed MacAB efflux pump or the presence of a *mef* gene encoded efflux pump[92,93]. Interestingly, a GC deletion in the repetitive region of the *mtrC* gene has been identified as a mutation that leads to increased azithromycin susceptibility[94]. However, the effect of these mutations is overshadowed by mutations with greater impact on azithromycin susceptibility.

EXTENDED-SPECTRUM CEPHALOSPORINS

Cephalosporins are a class of β-lactam antibiotics that consists of different groups with certain modifications of the β-lactam ring structure, referred to as generations. Like other β-lactam antibiotics, cephalosporins inhibit bacterial cell wall synthesis (see 'Penicillin'). Extended-spectrum cephalosporins have an increased affinity to the transpeptidase enzymes and thereby an enhanced activity against Gram-negative bacteria[95]. Different dosing regimens have been used since the introduction of oral extended-spectrum cephalosporins for gonorrhoea treatment in the late 1990s until early 2000s. In Japan, low-dose regimens of oral cefixime were frequently used and a remarkable decrease in susceptibility was reported in the early 2000s, probably a result of suboptimal antibiotic concentrations[96,97]. Cefixime resistance increased and treatment failures have been described worldwide[98–102]. Currently, a single intramuscular dose of ceftriaxone is recommended for gonorrhoea treatment, whether or not in combination with azithromycin. However, a global decrease in ceftriaxone susceptibility has been observed during the last decade and ceftriaxone-resistant isolates have already been reported, initially in Japan in 2009 and later in multiple other countries[35,59,103–108]. In 2015, the ceftriaxone resistant FC428 strain was identified in Japan and isolates belonging to this strain were later also sporadically found in Canada, Australia, several countries in Europe and Singapore[109–117]. Although sustained transmission of the FC428 strain is reported in China, this strain has not extensively spread so far and ceftriaxone resistance is globally still very rare[118,119].

The primary determinant for extended-spectrum cephalosporin resistance and reduced susceptibility is the mosaic *penA* gene, encoding the main target penicillin-binding protein 2 (Figure 1). Mosaic *penA* genes have up to 70 amino acid alterations as they contain parts of the *penA* gene of commensal *Neisseria* spp., obtained through interspecies horizontal gene

transfer[120]. Compared to mosaic *penA* genes of reduced susceptible isolates, mosaic *penA* genes of ceftriaxone-resistant isolates contained additional mutations conferring high-level ceftriaxone resistance[121]. Whereas most mutations in the mosaic *penA* gene need co-occurrence of other *penA* mutations to reduce ceftriaxone susceptibility, a change in the amino acid A501 in a non-mosaic *penA* gene can substantially reduce the susceptibility on its own[122–124]. Despite *penA* mutations being the primary extended-spectrum cephalosporins resistance determinants, a decreased influx- or increased efflux can further reduce the susceptibility in a stepwise manner. Decreased influx is mainly caused by amino acid changes at positions G120 and A121 in PorB (Figure 1)[37]. Previously described *mtrR* mutations are responsible for upregulated cephalosporin efflux[39]. More recently, mutations in the RNA polymerase encoding *rpoB* and *rpoD* genes were suggested to be involved in reduced susceptibility, although these are not widely found among reduced susceptible isolates[125]. Since mutations do not always result in the same phenotype, it is thought that the mechanisms of resistance or reduced susceptibility to extended-spectrum cephalosporin contain a still unknown 'factor X', implying that the mechanism is not completely resolved yet[20,40].

The development of cephalosporin resistance is a major public health concern, as cephalosporins are last resort antibiotics. During recent years it has been thought that resistance would arise soon, either through further spread of resistant strains with the mosaic *penA* gene or through the emergence of novel resistant strains. For instance, the β-lactamase TEM-1 gene, which is responsible for tetracycline resistance, differs from an extended-spectrum β-lactamase gene with only two single nucleotide polymorphisms (SNPs)[126]. Also, high recombination rates in *N. gonorrhoeae* indicate high probability of incorporation of resistance determinants in the gonococcal genome[127]. To prevent the spread of cephalosporin-resistant strains, empirical dual therapy with ceftriaxone and azithromycin is currently recommended by the WHO and the European International Union on STIs[128,129]. However, the added value of dual therapy on treatment efficacy is doubted and azithromycin resistance is globally on the rise. Several countries, including the Netherlands, therefore attempted to reduce the use of macrolides by changing the treatment recommendation to ceftriaxone monotherapy [60,130–132].

GENOMICS AND BIOINFORMATICS

The genome is the blueprint of life. Bacterial genomes contain information needed to understand emergence and spread of strains, mechanisms of AMR, infection- and transmission dynamics and bacterial population biology. To extract this information from the genome, several resources are available.

MOLECULAR TYPING

Molecular typing enables studying genetic diversity in bacterial populations in a reproducible and portable manner, based on sequences of genes or gene fragments. Each unique sequence is assigned an allele number and the combinations of allele numbers result in genetic barcodes, referred to as sequence types. Each typing scheme is developed for a defined purpose and has a certain resolution, depending on the genetic diversity of the typing genes and the number of genes included. The Multi-Locus Sequence Typing (MLST) scheme uses the sequences of fragments of housekeeping genes under stabilising selection[133]. For *N. gonorrhoeae*, the MLST scheme includes sequences of seven gene fragments (*abcZ, adk, aroE, fumC, gdh, pdhC* and *pgm*), which together have enough discriminatory power to separate *Neisseria* spp. and to infer phylogenetic relationships between isolates[134]. The *N. gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) scheme has been developed as a high-throughput method for the identification of recent *N. gonorrhoeae* transmission chains within a community[135]. This scheme includes fragments of two hypervariable genes: the outer membrane porin encoding *porB* gene, and the *tbpB* gene, encoding the surface-exposed part of the transferrin-binding protein. Both genes are under strong diversifying selection pressure by the human immune response, making NG-MAST highly discriminative. To track global emergence of antimicrobial resistant *N. gonorrhoeae* strains, the *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) scheme was set up, including gene fragments involved in resistance to β-lactam antibiotics, fluoroquinolones and macrolides (*penA, mtrR, porB1b, ponA, gyrA, parC* and 23S rRNA)(see 'The history of antimicrobial therapy and resistance in *N. gonorrhoeae*'). Multiple NG-STAR types have been associated with AMR, demonstrating the value of this typing scheme for AMR surveillance[136].

Before the era of whole-genome sequencing (WGS), MLST was the method of choice for bacterial typing. However, resolution improves dramatically by taking the whole genome into account, instead of only several gene fragments[137]. To increase resolution while retaining the reproducibility and portability of typing, a *N. gonorrhoeae* core genome MLST was developed, containing 1668 genes present in nearly all (>95%) gonococcal isolates[138]. Subsequent clustering of isolates that have allelic profile difference below a certain threshold constitute core genome groups, representing the gonococcal population structure. Interestingly, certain core genome groups are associated with AMR, which indicates predisposition to- or ancestral inheritance of AMR determinants in certain lineages.

WHOLE GENOME SEQUENCING

As a result of the improved cost-efficiency of WGS, this method has become the method of choice for genomic studies due to its high resolution[139]. Illumina platforms are widely used for high-throughput short-read sequencing, usually yielding sequence reads of 150 or 300 base pairs. Analysis pipelines including multiple bioinformatic tools are used to translate the raw

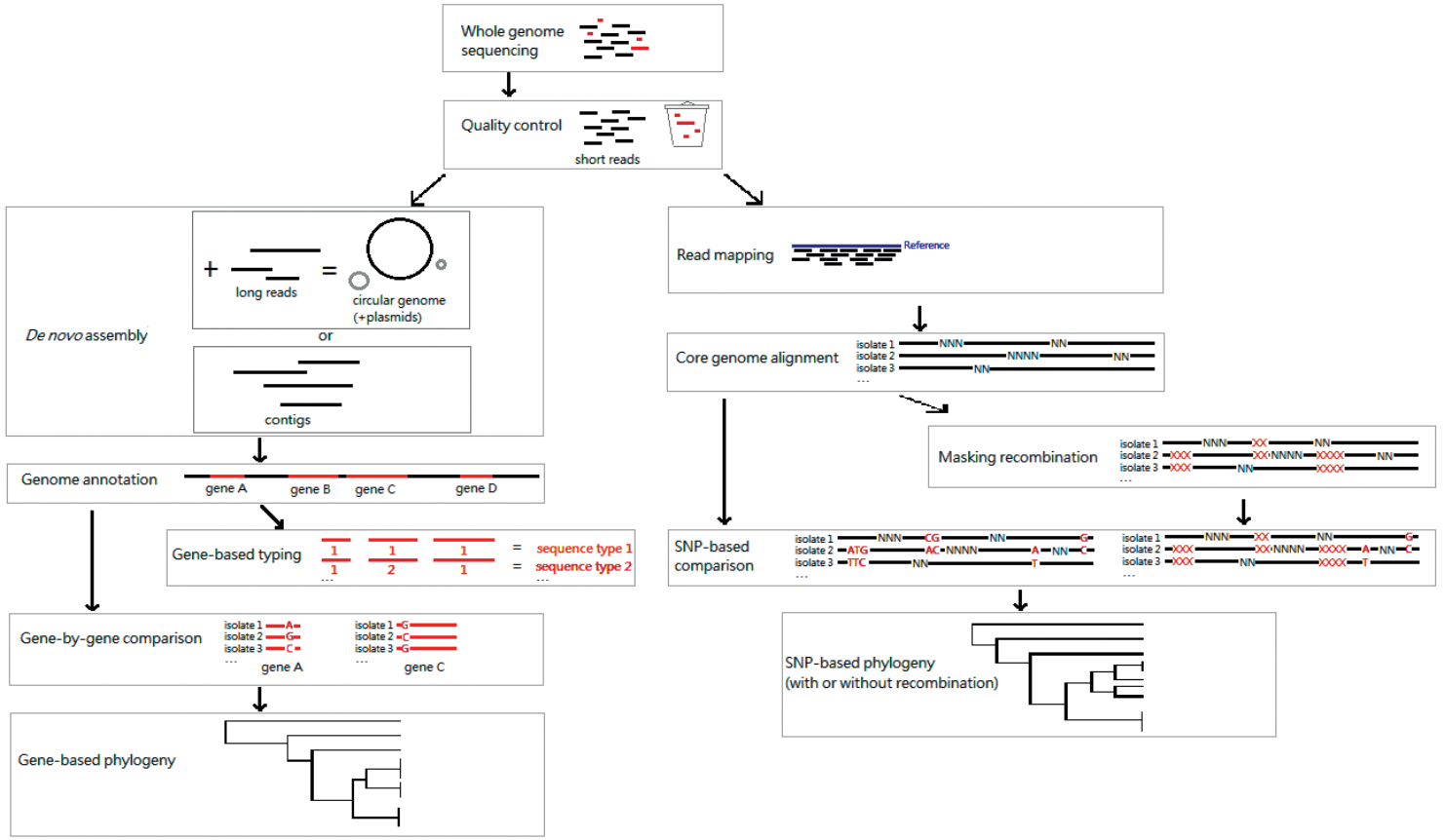


FIGURE 2. WORKFLOW WITH COMMON PRACTICES WHEN ANALYSING GENOMICS DATA.

sequences into interpretable information (several available tools are reviewed in[140]). Since the information that comes out is only as good as the data input, data analysis starts with a quality control step to remove too short- or bad quality reads from the raw data (Figure 2). The raw reads subsequently are *de novo* assembled to reconstruct the genome[141]. Due to repetitive regions of the genome being hard to assemble, this step yields a draft genome consisting of multiple segments, referred to as contigs[142]. Although draft genomes are sufficient for most analyses, complete genomes could be obtained by using additional long-read sequence data, produced by the MinION (Oxford Nanopore Technologies) or PacBio (Pacific Biosciences) platforms[143].

To extract gene information, genomes can be searched for open reading frames and these predicted protein sequences can be annotated by comparing them to large public protein databases[144]. One can either annotate the whole genome or, more specifically, only annotate typing genes or AMR genes. Multiple annotated genomes can be compared in a gene-by-gene manner, to study gene sequence diversity or to assess the prevalence of genes among a selection of isolates. The latter can be done to identify the core- and the accessory genome, with the core genome containing genes essential for survival and therefore present in nearly all isolates in a selection, and the accessory genome containing genes present in only a subset of isolates since these provide the genetic flexibility needed for phenotypic variation or environment adaptation[145]. The core- and accessory genome together form the pangenome, which is the complete gene pool of a set of isolates[146].

Besides gene-by-gene approaches for comparative analyses, SNP-based approaches are used to compare whole genomes including the intergenic regions (Figure 2)[147]. SNP-based comparisons between isolates can be done by mapping raw sequence reads of each isolate to the same reference genome. The genomic regions shared between all isolates and the reference genome can be compared and SNPs can be determined in an alignment of the shared regions. Afterwards, a phylogeny can be inferred from these SNPs to assess the genetic relatedness between the compared isolates. A consequence of this reference-based methodology is that regions that are absent in the reference genome are not included in the comparison. Therefore, the choice of reference genome strongly influences the results, which is especially challenging in comparative analyses of highly genetically diverse species, such as *Escherichia coli*[148]. For these species, a reference genome is selected that is genetically closest to all isolates in a certain set[149]. For *N. gonorrhoeae,* the genetic diversity is relatively limited and the reference genome FA1090 is widely used in comparative analyses, which also enables comparison between studies[150].

Recombination is an important factor to consider when doing comparative analyses, since recombination events can lead to high SNP counts between otherwise closely related isolates. Keeping recombination in the analyses could result in larger branch lengths in the phylogenetic

tree and to larger distances to common ancestors compared to a tree without recombination[151]. Since *N. gonorrhoeae* has high recombination rates, it is important to filter out recombination when inferring phylogenetic relationships or when studying gonococcal evolution (Figure 2). However, since masking recombination could lead to spurious clustering of isolates, this step should always be carefully considered and evaluated in the context of the research question[148].

USING GENOMICS IN AMR SURVEILLANCE

Genomics approaches can be applied in several ways to study AMR, with an important application being the identification of genetic AMR determinants by combining genotypic and phenotypic data. Defined AMR determinants can be tracked among populations, which forms part of AMR surveillance programmes such as Euro-GASP[85]. Genomic surveillance also elucidates gonococcal population dynamics and thereby enables the identification of emerging or vanishing lineages associated with AMR. Adding patient reported metadata further informs about gonorrhoea epidemiology and supports the identification of AMR key populations and transmission networks.

OUTLINE OF THIS THESIS

For this thesis, I have used genomics to perform AMR surveillance and to enhance the understanding of gonococcal biology. **Part I** of this thesis contains projects on genomic AMR surveillance on the national level in the Netherlands, which support global surveillance efforts. Routine determination of antimicrobial susceptibility profiles forms part of AMR surveillance. In **chapter 2**, a multiplex qPCR assay is presented that simultaneously detects *N. gonorrhoeae* and the fluoroquinolone resistance-mediating mutation *gyrA*-S91F to rapidly obtain fluoroquinolone susceptibility profiles, even for samples that can not be cultivated. Rapid fluoroquinolone susceptibility testing enables the reintroduction of oral ciprofloxacin for gonorrhoea treatment.

Chapters 3, 4 and 5 describe genomic AMR surveillance projects that used isolates obtained from visitors of the Centre for Sexual Health in Amsterdam. **Chapter 3** contains a genomic population analysis of *N. gonorrhoeae* isolates obtained between 2014 and 2019 with a focus on ceftriaxone reduced susceptible isolates, of which the prevalence increased during that period in Amsterdam. The genotypes associated with reduced susceptibility and the corresponding AMR determinants were characterised. An update on the genomic AMR surveillance of *N. gonorrhoeae* is given in **chapter 4**, which investigated the dynamics of the *N. gonorrhoeae* population during the first COVID-19 lockdown in 2020 in Amsterdam. *N. gonorrhoeae* populations before- and during the lockdown were compared by phenotypic and genotypic characterisation of isolates obtained in a period before the lockdown and a period during the lockdown. In **chapter 5**, similar genomics methodologies are used for the surveillance of extended-spectrum β-lactamase-

producing *Escherichia coli* (ESBL-Ec), an opportunistic pathogen that can cause urinary tract infections or gastrointestinal symptoms, but is mainly asymptomatically carried in the intestines. ESBL genes confer resistance to extended-spectrum cephalosporins and can be horizontally transferred between different bacterial species, making them of great relevance to AMR surveillance. Increased carriage of ESBL-Ec was previously observed among MSM participating in the Amsterdam Cohort Studies, compared to the general Dutch population. This chapter describes the genomic characterisation of ESBL-Ec isolates obtained from MSM and the investigation on whether increased prevalence could be explained by sexual transmission between study participants.

In **part II** of this thesis, similar genomics approaches are used to understand gonococcal biology. **Chapter 6** describes genomic analyses of isolates obtained during the New AntiBiotic treatment Options for uncomplicated GOnorrhoea (NABOGO) trial, performed from 2017 to 2020 at the Centre for Sexual Health of Amsterdam. This trial assessed the efficacy of four different antibiotics for the treatment of gonorrhoea. Treatment failures were reflected by isolates from consecutive time points, whose genomes were genetically compared to identify genomic changes that occurred over the course of infection. The insights obtained in this chapter enhance the understanding of gonococcal infection dynamics. New insights into gonococcal population biology are obtained in **chapter 7**, by characterising the hitherto undefined chromosomally-encoded gonococcal accessory genome. For this purpose, a large comparative analysis was performed of over 8,000 *N. gonorrhoeae* genomes publicly available in the PubMLST database. Associations between the identified accessory genes and the core genome or AMR were also assessed.

In **chapter 8**, the insights obtained in this thesis are compared to recent literature. Avenues for future treatment and management of gonorrhoea are discussed in the context of AMR development and spread. Important considerations regarding the genomic and bioinformatic analyses used in this thesis are also described in this chapter.

REFERENCES

1. Murray CJ, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*. 2022;399(10325):629-655. doi:10.1016/S0140-6736(21)02724-0
2. Sulis G, Sayood S, Gandra S. Antimicrobial resistance in low- and middle-income countries: current status and future directions. *Expert Rev Anti Infect Ther*. 2022;20(2):147-160. doi:10.1080/14787210.2021.1951705
3. World Health Organization. 2019 Antibacterial agents in clinical development: an analysis of the antibacterial clinical development pipeline. Published online 2019. Accessed January 1, 2022. https://apps.who.int/iris/bitstream/handle/10665/330420/9789240000193-eng.pdf
4. Kirkcaldy RD, Weston E, Segurado AC, et al. Epidemiology of Gonorrhea: A Global Perspective. *Sex Health*. 2019;16(5):401-411. doi:10.1071/SH19061
5. Rowley J, Vander Hoorn S, Korenromp E, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bull World Health Organ*. 2019;97(8):548-562P. doi:10.2471/BLT.18.228486
6. Centers for Disease Control and Prevention. *Sexually Transmitted Disease Surveillance 2018*. Atlanta: U.S Department of Health and Human Services; 2018. doi:10.15620/cdc.79370
7. Kirby Institute. *HIV, Viral Hepatitis and Sexually Transmissible Infections in Australia: Annual Surveillance Report 2018*. Sydney Kirby Institute, UNSW Sydney; 2018. Accessed July 10, 2022. https://kirby.unsw.edu.au/sites/default/files/kirby/report/KI_Annual-Surveillance-Report-2018.pdf
8. Choudhri Y, Miller J, Sandhu J, et al. Gonorrhea in Canada, 2010-2015. *Can Commun Dis Rep Releve Mal Transm Au Can*. 2018;44(2):37-42. doi:10.14745/ccdr.v44i02a01
9. European Centre for Disease Prevention and Control. *Gonorrhoea: Annual Epidemiological Report for 2018*. Stockholm, ECDC; 2020. Accessed January 1, 2022. https://www.ecdc.europa.eu/sites/default/files/documents/gonorrhoea-annual-epidemiological-report-2018.pdf
10. Mohammed H, Mitchell H, Sile B, et al. Increase in Sexually Transmitted Infections among Men Who Have Sex with Men, England, 2014. *Emerg Infect Dis*. 2016;22(1):88-91. doi:10.3201/eid2201.151331
11. McSheffrey GG, Gray-Owen SD. *Neisseria gonorrhoeae*. In: Tang YW, Sussman M, Liu D, et al., eds. *Molecular Medical Microbiology (Second Edition)*. Academic Press; 2015:1471-1485. doi:10.1016/B978-0-12-397169-2.00082-2
12. Hoffman O, Weber RJ. Pathophysiology and Treatment of Bacterial Meningitis. *Ther Adv Neurol Disord*. 2009;2(6):1-7. doi:10.1177/1756285609337975
13. Pearce WA, Buchanan TM. Attachment role of gonococcal pili. Optimum conditions and quantitation of adherence of isolated pili to human cells in vitro. *J Clin Invest*. 1978;61(4):931-943.
14. Craig L, Pique ME, Tainer JA. Type IV pilus structure and bacterial pathogenicity. *Nat Rev Microbiol*. 2004;2(5):363-378. doi:10.1038/nrmicro885
15. Obergfell KP, Seifert HS. The Pilin N-terminal Domain Maintains *Neisseria gonorrhoeae* Transformation Competence during Pilus Phase Variation. *PLoS Genet*. 2016;12(5):e1006069. doi:10.1371/journal.pgen.1006069
16. Cahoon LA, Seifert HS. Transcription of a cis-acting, Noncoding, Small RNA Is Required for Pilin Antigenic Variation in *Neisseria gonorrhoeae*. *PLOS Pathog*. 2013;9(1):e1003074. doi:10.1371/journal.ppat.1003074
17. Quillin SJ, Seifert HS. *Neisseria gonorrhoeae* host adaptation and pathogenesis. *Nat Rev Microbiol*. 2018;16(4):226-240. doi:10.1038/nrmicro.2017.169
18. Johnson MB, Criss AK. Resistance of *Neisseria gonorrhoeae* to Neutrophils. *Front Microbiol*. 2011;2:77. doi:10.3389/fmicb.2011.00077
19. Hui B, Fairley CK, Chen M, et al. Oral and anal sex are key to sustaining gonorrhoea at endemic levels in MSM populations: a mathematical model. *Sex Transm Infect*. 2015;91(5):365-369. doi:10.1136/sextrans-2014-051760
20. Unemo M, Shafer WM. Antimicrobial Resistance in *Neisseria gonorrhoeae* in the 21st Century: Past, Evolution, and Future. *Clin Microbiol Rev*. 2014;27(3):587-613. doi:10.1128/CMR.00010-14
21. Thayer JD, Martin JE. A selective medium for the cultivation of *N. gonorrhoeae* and *N. meningitidis*. *Public Health Rep*. 1964;79(1):49-57. doi:10.2307/4592051

22. Hook EW, Kirkcaldy RD. A Brief History of Evolving Diagnostics and Therapy for Gonorrhea: Lessons Learned. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2018;67(8):1294-1299. doi:10.1093/cid/ciy271

23. World Health Organization. Global action plan to control the spread and impact of antimicrobial resistance in *Neisseria gonorrhoeae*. Published 2012. https://apps.who.int/iris/handle/10665/44863

24. Kampmeier RH. Introduction of sulfonamide therapy for gonorrhea. *Sex Transm Dis*. 1983;10(2):81-84. doi:10.1097/00007435-198304000-00007

25. Brown GM. The Biosynthesis of Folic Acid. *J Biol Chem*. 1962;237(2):536-540. doi:10.1016/S0021-9258(18)93957-8

26. Swedberg G, Fermér C, Sköld O. Point mutations in the dihydropteroate synthase gene causing sulfonamide resistance. *Adv Exp Med Biol*. 1993;338:555-558. doi:10.1007/978-1-4615-2960-6_113

27. Fermer C, Kristiansen BE, Sköld O, et al. Sulfonamide resistance in *Neisseria meningitidis* as defined by site-directed mutagenesis could have its origin in other species. *J Bacteriol*. 1995;177(16):4669-4675. doi:10.1128/jb.177.16.4669-4675.1995

28. Dunlop EMC. Gonorrhoea and the Sulphonamides. *Sex Transm Infect*. 1949;25(2):81-83. doi:10.1136/sti.25.2.81

29. Amies CR. Development of resistance of gonococci to penicillin: an eight-year study. *Can Med Assoc J*. 1967;96(1):33-35.

30. Williams RA. Penicillinase-producing *Neisseria gonorrhoeae*. *Am Fam Physician*. 1981;24(6):117-119.

31. Roberts M, Elwell LP, Falkow S. Molecular characterization of two beta-lactamase-specifying plasmids isolated from *Neisseria gonorrhoeae*. *J Bacteriol*. 1977;131(2):557-563. doi:10.1128/jb.131.2.557-563.1977

32. Cehovin A, Jolley KA, Maiden MCJ, et al. Association of *Neisseria gonorrhoeae* Plasmids With Distinct Lineages and The Economic Status of Their Country of Origin. *J Infect Dis*. 2020;222(11):1826-1836. doi:10.1093/infdis/jiaa003

33. Faruki H, Kohmescher RN, McKinney WP, et al. A community-based outbreak of infection with penicillin-resistant *Neisseria gonorrhoeae* not producing penicillinase (chromosomally mediated resistance). *N Engl J Med*. 1985;313(10):607-611. doi:10.1056/NEJM198509053131004

34. Powell AJ, Tomberg J, Deacon AM, et al. Crystal structures of penicillin-binding protein 2 from penicillin-susceptible and -resistant strains of *Neisseria gonorrhoeae* reveal an unexpectedly subtle mechanism for antibiotic resistance. *J Biol Chem*. 2009;284(2):1202-1212. doi:10.1074/jbc.M805761200

35. Ohnishi M, Golparian D, Shimuta K, et al. Is *Neisseria gonorrhoeae* initiating a future era of untreatable gonorrhea?: detailed characterization of the first strain with high-level resistance to ceftriaxone. *Antimicrob Agents Chemother*. 2011;55(7):3538-3545. doi:10.1128/AAC.00325-11

36. Ropp PA, Hu M, Olesky M, et al. Mutations in *ponA*, the Gene Encoding Penicillin-Binding Protein 1, and a Novel Locus, *penC*, Are Required for High-Level Chromosomally Mediated Penicillin Resistance in *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother*. 2002;46(3):769-777. doi:10.1128/AAC.46.3.769-777.2002

37. Olesky M, Zhao S, Rosenberg RL, et al. Porin-Mediated Antibiotic Resistance in *Neisseria gonorrhoeae*: Ion, Solute, and Antibiotic Permeation through PIB Proteins with *penB* Mutations. *J Bacteriol*. 2006;188(7):2300-2308. doi:10.1128/JB.188.7.2300-2308.2006

38. Olesky M, Hobbs M, Nicholas RA. Identification and analysis of amino acid mutations in porin IB that mediate intermediate-level resistance to penicillin and tetracycline in *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother*. 2002;46(9):2811-2820. doi:10.1128/AAC.46.9.2811-2820.2002

39. Ohneck EA, Zalucki YM, Johnson PJT, et al. A novel mechanism of high-level, broad-spectrum antibiotic resistance caused by a single base pair change in *Neisseria gonorrhoeae*. *mBio*. 2011;2(5):e00187-11. doi:10.1128/mBio.00187-11

40. Unemo M, Golparian D, Eyre DW. Antimicrobial Resistance in *Neisseria gonorrhoeae* and Treatment of Gonorrhea. *Methods Mol Biol*. 2019;1997:37-58. doi:10.1007/978-1-4939-9496-0_3

41. Morse SA, Johnson SR, Biddle JW, et al. High-level tetracycline resistance in *Neisseria gonorrhoeae* is result of acquisition of streptococcal tetM determinant. *Antimicrob Agents Chemother*. 1986;30(5):664-670. doi:10.1128/AAC.30.5.664

42. Roberts MC, Knapp JS. Host range of the conjugative 25.2-megadalton tetracycline resistance plasmid from *Neisseria gonorrhoeae* and related species. *Antimicrob Agents Chemother*. 1988;32(4):488-491. doi:10.1128/AAC.32.4.488

43. Chopra I, Roberts M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiol Mol Biol Rev*. 2001;65(2):232-260. doi:10.1128/MMBR.65.2.232-260.2001
44. Sparling PF, Sarubbi FA, Blackman E. Inheritance of low-level resistance to penicillin, tetracycline, and chloramphenicol in *Neisseria gonorrhoeae*. *J Bacteriol*. 1975;124(2):740-749. doi:10.1128/jb.124.2.740-749.1975
45. Hu M, Nandi S, Davies C, et al. High-Level Chromosomally Mediated Tetracycline Resistance in *Neisseria gonorrhoeae* Results from a Point Mutation in the rpsJ Gene Encoding Ribosomal Protein S10 in Combination with the mtrR and penB Resistance Determinants. *Antimicrob Agents Chemother*. 2005;49(10):4327-4334. doi:10.1128/AAC.49.10.4327-4334.2005
46. Holloway WJ. Spectinomycin. *Med Clin North Am*. 1982;66(1):169-173.
47. Carter AP, Clemons WM, Brodersen DE, et al. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*. 2000;407(6802):340-348. doi:10.1038/35030019
48. Galimand M, Gerbaud G, Courvalin P. Spectinomycin Resistance in *Neisseria* spp. Due to Mutations in 16S rRNA. *Antimicrob Agents Chemother*. 2000;44(5):1365-1366. doi:10.1128/aac.44.5.1365-1366.2000
49. Ilina E, Malakhova M, Bodoev I, et al. Mutation in ribosomal protein S5 leads to spectinomycin resistance in *Neisseria gonorrhoeae*. *Front Microbiol*. 2013;4. doi:10.3389/fmicb.2013.00186
50. Boslego JW, Tramont EC, Takafuji ET, et al. Effect of Spectinomycin Use on the Prevalence of Spectinomycin-Resistant and of Penicillinase-Producing *Neisseria gonorrhoeae*. *N Engl J Med*. 1987;317(5):272-278. doi:10.1056/NEJM198707303170504
51. Judson FN, Ehret JM, Handsfield HH. Comparative study of ceftriaxone and spectinomycin for treatment of pharyngeal and anorectal gonorrhea. *JAMA*. 1985;253(10):1417-1419.
52. Tanaka M, Kumazawa J, Matsumoto T, et al. High prevalence of *Neisseria gonorrhoeae* strains with reduced susceptibility to fluoroquinolones in Japan. *Genitourin Med*. 1994;70(2):90-93. doi:10.1136/sti.70.2.90
53. Unemo M, Shafer WM. Antibiotic resistance in *Neisseria gonorrhoeae*: origin, evolution, and lessons learned for the future. *Ann N Y Acad Sci*. 2011;1230:E19-E28. doi:10.1111/j.1749-6632.2011.06215.x
54. Lindbäck E, Rahman M, Jalal S, et al. Mutations in *gyrA, gyrB, parC*, and *parE* in quinolone-resistant strains of *Neisseria gonorrhoeae*. *APMIS Acta Pathol Microbiol Immunol Scand*. 2002;110(9):651-657. doi:10.1034/j.1600-0463.2002.1100909.x
55. Belland RJ, Morrison SG, Ison C, et al. *Neisseria gonorrhoeae* acquires mutations in analogous regions of *gyrA* and *parC* in fluoroquinolone-resistant isolates. *Mol Microbiol*. 1994;14(2):371-380. doi:10.1111/j.1365-2958.1994.tb01297.x
56. Centers for Disease Control and Prevention. National Overview - Sexually Transmitted Disease Surveillance, 2019. Published 2020. Accessed June 29, 2022. https://www.cdc.gov/std/statistics/2019/overview.htm
57. Day MJ, Jacobsson S, Spiteri G, et al. Significant increase in azithromycin "resistance" and susceptibility to ceftriaxone and cefixime in *Neisseria gonorrhoeae* isolates in 26 European countries, 2019. *BMC Infect Dis*. 2022;22(1):524. doi:10.1186/s12879-022-07509-w
58. Van Wees D, Visser M, van Aar F, et al. *Sexually Transmitted Infections in the Netherlands in 2021*. Rijksinstituut voor Volksgezondheid en Milieu RIVM; 2022. doi:10.21945/RIVM-2022-0023
59. Unemo M, Lahra MM, Escher M, et al. WHO global antimicrobial resistance surveillance for *Neisseria gonorrhoeae* 2017–18: a retrospective observational study. *Lancet Microbe*. 2021;2(11):e627-e636. doi:10.1016/S2666-5247(21)00171-3
60. Fifer H, Saunders J, Soni S, et al. 2018 UK national guideline for the management of infection with *Neisseria gonorrhoeae*. *Int J STD AIDS*. 2020;31(1):4-15. doi:10.1177/0956462419886775
61. Allan-Blitz LT, Wang X, Klausner JD. Wild-Type Gyrase A Genotype of *Neisseria gonorrhoeae* Predicts In Vitro Susceptibility to Ciprofloxacin: A Systematic Review of the Literature and Meta-Analysis. *Sex Transm Dis*. 2017;44(5):261-265. doi:10.1097/OLQ.0000000000000591
62. Siedner MJ, Pandori M, Castro L, et al. Real-Time PCR Assay for Detection of Quinolone-Resistant *Neisseria gonorrhoeae* in Urine Samples. *J Clin Microbiol*. 2007;45(4):1250-1254. doi:10.1128/JCM.01909-06

63. Hemarajata P, Yang S, Soge OO, et al. Performance and Verification of a Real-Time PCR Assay Targeting the *gyrA* Gene for Prediction of Ciprofloxacin Resistance in *Neisseria gonorrhoeae*. *J Clin Microbiol*. 2016;54(3):805-808. doi:10.1128/JCM.03032-15

64. Ellis O, Hemarajata P, Shahkolahi A, et al. A multisite implementation of a real-time polymerase chain reaction assay to predict ciprofloxacin susceptibility in *Neisseria gonorrhoeae*. *Diagn Microbiol Infect Dis*. 2019;94(3):213-217. doi:10.1016/j.diagmicrobio.2018.12.018

65. Perera SR, Khan NH, Martin I, et al. Multiplex Real-Time PCR Assay for Simultaneous Identification of *Neisseria gonorrhoeae* and Its Ciprofloxacin Susceptibility Status. *J Clin Microbiol*. 2017;55(11):3201-3209. doi:10.1128/JCM.00855-17

66. Ebeyan S, Windsor M, Bordin A, et al. Evaluation of the ResistancePlus GC (beta) assay: a commercial diagnostic test for the direct detection of ciprofloxacin susceptibility or resistance in *Neisseria gonorrhoeae*. *J Antimicrob Chemother*. 2019;74(7):1820-1824. doi:10.1093/jac/dkz036

67. de Korne-Elenbaas J, Pol A, Vet J, et al. Simultaneous Detection of *Neisseria gonorrhoeae* and Fluoroquinolone Resistance Mutations to Enable Rapid Prescription of Oral Antibiotics. *Sex Transm Dis*. 2020;47(4):238-242. doi:10.1097/OLQ.0000000000001141

68. Schlünzen F, Harms JM, Franceschi F, et al. Structural Basis for the Antibiotic Activity of Ketolides and Azalides. *Structure*. 2003;11(3):329-338. doi:10.1016/S0969-2126(03)00022-4

69. Handsfield HH, Dalu ZA, Martin DH, et al. Multicenter trial of single-dose azithromycin vs. ceftriaxone in the treatment of uncomplicated gonorrhea. Azithromycin Gonorrhea Study Group. *Sex Transm Dis*. 1994;21(2):107-111. doi:10.1097/00007435-199403000-00010

70. Bignell C, Garley J. Azithromycin in the treatment of infection with *Neisseria gonorrhoeae*. *Sex Transm Infect*. 2010;86(6):422-426. doi:10.1136/sti.2010.044586

71. Dillon JAR, Ruben M, Li H, et al. Challenges in the Control of Gonorrhea in South America and the Caribbean: Monitoring the Development of Resistance to Antibiotics. *Sex Transm Dis*. 2006;33(2):87-95. doi:10.1097/01.olq.0000187231.28812.29

72. Starnino S, Galarza P, Carvallo MET, et al. Retrospective Analysis of Antimicrobial Susceptibility Trends (2000–2009) in *Neisseria gonorrhoeae* Isolates from Countries in Latin America and the Caribbean Shows Evolving Resistance to Ciprofloxacin, Azithromycin and Decreased Susceptibility to Ceftriaxone. *Sex Transm Dis*. 2012;39(10):813-821. doi:10.1097/OLQ.0b013e3182631c9f

73. Dillon JAR, Trecker MA, Thakur SD, et al. Two decades of the gonococcal antimicrobial surveillance program in South America and the Caribbean: challenges and opportunities. *Sex Transm Infect*. 2013;89(Suppl 4):iv36-iv41. doi:10.1136/sextrans-2012-050905

74. Kenyon C, Buyze J, Wi T. Antimicrobial Consumption and Susceptibility of *Neisseria gonorrhoeae*: A Global Ecological Analysis. *Front Med*. 2018;5. doi:10.3389/fmed.2018.00329

75. Roberts MC, Chung WO, Roe D, et al. Erythromycin-Resistant *Neisseria gonorrhoeae* and Oral Commensal *Neisseria* spp. Carry Known rRNA Methylase Genes. *Antimicrob Agents Chemother*. 1999;43(6):1367-1372. doi:10.1128/aac.43.6.1367

76. Chisholm SA, Dave J, Ison CA. High-level azithromycin resistance occurs in *Neisseria gonorrhoeae* as a result of a single point mutation in the 23S rRNA genes. *Antimicrob Agents Chemother*. 2010;54(9):3812-3816. doi:10.1128/AAC.00309-10

77. Allen VG, Farrell DJ, Rebbapragada A, et al. Molecular analysis of antimicrobial resistance mechanisms in *Neisseria gonorrhoeae* isolates from Ontario, Canada. *Antimicrob Agents Chemother*. 2011;55(2):703-712. doi:10.1128/AAC.00788-10

78. Ng LK, Martin I, Liu G, et al. Mutation in 23S rRNA Associated with Macrolide Resistance in *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother*. 2002;46(9):3020-3025. doi:10.1128/AAC.46.9.3020-3025.2002

79. Galarza PG, Abad R, Canigia LF, et al. New Mutation in 23S rRNA Gene Associated with High Level of Azithromycin Resistance in *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother*. 2010;54(4):1652-1653. doi:10.1128/AAC.01506-09

80. Zhang J, van der Veen S. *Neisseria gonorrhoeae* 23S rRNA A2059G mutation is the only determinant necessary for high-level azithromycin resistance and improves in vivo biological fitness. *J Antimicrob Chemother*. 2019;74(2):407-415. doi:10.1093/jac/dky438

81. Hagman KE, Shafer WM. Transcriptional control of the mtr efflux system of *Neisseria gonorrhoeae*. *J Bacteriol*. 1995;177(14):4162-4165. doi:10.1128/jb.177.14.4162-4165.1995

82. Zarantonelli L, Borthagaray G, Lee EH, et al. Decreased azithromycin susceptibility of *Neisseria gonorrhoeae* due to *mtrR* mutations. *Antimicrob Agents Chemother*. 1999;43(10):2468-2472. doi:10.1128/AAC.43.10.2468

83. Thomas JC, Seby S, Abrams AJ, et al. Evidence of Recent Genomic Evolution in Gonococcal Strains With Decreased Susceptibility to Cephalosporins or Azithromycin in the United States, 2014-2016. *J Infect Dis*. 2019;220(2):294-305. doi:10.1093/infdis/jiz079

84. Golparian D, Bazzo ML, Golfetto L, et al. Genomic epidemiology of *Neisseria gonorrhoeae* elucidating the gonococcal antimicrobial resistance and lineages/sublineages across Brazil, 2015-16. *J Antimicrob Chemother*. 2020;75(11):3163-3172. doi:10.1093/jac/dkaa318

85. Sánchez-Busó L, Cole MJ, Spiteri G, et al. Europe-wide expansion and eradication of multidrug-resistant *Neisseria gonorrhoeae* lineages: a genomic surveillance study. *Lancet Microbe*. 2022;3(6):e452-e463. doi:10.1016/S2666-5247(22)00044-1

86. Wadsworth CB, Arnold BJ, Sater MRA, et al. Azithromycin Resistance through Interspecific Acquisition of an Epistasis-Dependent Efflux Pump Component and Transcriptional Regulator in *Neisseria gonorrhoeae*. *mBio*. 2018;9(4):e01419-18. doi:10.1128/mBio.01419-18

87. Sawatzky P, Demczuk W, Lefebvre B, et al. Increasing Azithromycin Resistance in *Neisseria gonorrhoeae* Due to NG-MAST 12302 Clonal Spread in Canada, 2015 to 2018. *Antimicrob Agents Chemother*. 2022;66(3):e0168821. doi:10.1128/AAC.01688-21

88. Banhart S, Selb R, Oehlmann S, et al. The Mosaic *mtr* Locus as Major Genetic Determinant of Azithromycin Resistance of *Neisseria gonorrhoeae*—Germany, 2018. *J Infect Dis*. 2021;224(8):1398-1404. doi:10.1093/infdis/jiab091

89. Gianecini RA, Poklepovich T, Golparian D, et al. Genomic Epidemiology of Azithromycin-Nonsusceptible *Neisseria gonorrhoeae*, Argentina, 2005-2019. *Emerg Infect Dis*. 2021;27(9):2369-2378. doi:10.3201/eid2709.204843

90. Sánchez-Busó L, Yeats CA, Taylor B, et al. A community-driven resource for genomic epidemiology and antimicrobial resistance prediction of *Neisseria gonorrhoeae* at Pathogenwatch. *Genome Med*. 2021;13(1):61. doi:10.1186/s13073-021-00858-2

91. Gernert KM, Seby S, Schmerer MW, et al. Azithromycin susceptibility of *Neisseria gonorrhoeae* in the USA in 2017: a genomic analysis of surveillance data. *Lancet Microbe*. 2020;1(4):e154-e164. doi:10.1016/S2666-5247(20)30059-8

92. Rouquette-Loughlin CE, Balthazar JT, Shafer WM. Characterization of the MacA–MacB efflux system in *Neisseria gonorrhoeae*. *J Antimicrob Chemother*. 2005;56(5):856-860. doi:10.1093/jac/dki333

93. Luna VA, Cousin S, Whittington WLH, et al. Identification of the Conjugative *mef* Gene in Clinical *Acinetobacter junii* and *Neisseria gonorrhoeae* Isolates. *Antimicrob Agents Chemother*. 2000;44(9):2503-2506. doi:10.1128/AAC.44.9.2503-2506.2000

94. Ma KC, Mortimer TD, Hicks AL, et al. *Increased Antibiotic Susceptibility in Neisseria gonorrhoeae through Adaptation to the Cervical Environment*. Microbiology; 2020. doi:10.1101/2020.01.07.896696

95. Klein NC, Cunha BA. Third-generation cephalosporins. *Med Clin North Am*. 1995;79(4):705-719. doi:10.1016/S0025-7125(16)30034-7

96. Ito M, Yasuda M, Yokoi S, et al. Remarkable Increase in Central Japan in 2001-2002 of *Neisseria gonorrhoeae* Isolates with Decreased Susceptibility to Penicillin, Tetracycline, Oral Cephalosporins, and Fluoroquinolones. *Antimicrob Agents Chemother*. 2004;48(8):3185-3187. doi:10.1128/AAC.48.8.3185-3187.2004

97. Yokoi S, Deguchi T, Ozawa T, et al. Threat to Cefixime Treatment for Gonorrhea. *Emerg Infect Dis*. 2007;13(8):1275-1277. doi:10.3201/eid1308.060948

98. Allen VG, Mitterni L, Seah C, et al. *Neisseria gonorrhoeae* treatment failure and susceptibility to cefixime in Toronto, Canada. *JAMA*. 2013;309(2):163-170. doi:10.1001/jama.2012.176575

99. Lewis DA, Sriruttan C, Müller EE, et al. Phenotypic and genetic characterization of the first two cases of extended-spectrum-cephalosporin-resistant *Neisseria gonorrhoeae* infection in South Africa and association with cefixime treatment failure. *J Antimicrob Chemother*. 2013;68(6):1267-1270. doi:10.1093/jac/dkt034

100. Ison CA, Hussey J, Sankar KN, et al. Gonorrhoea treatment failures to cefixime and azithromycin in England, 2010. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*. 2011;16(14):19833. doi:10.2807/ese.16.14.19833-en

101. Unemo M, Golparian D, Syversen G, et al. Two cases of verified clinical failures using internationally recommended first-line cefixime for gonorrhoea treatment, Norway, 2010. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*. 2010;15(47):19721. doi:10.2807/ese.15.47.19721-en

102. Unemo M, Golparian D, Stary A, et al. First *Neisseria gonorrhoeae* strain with resistance to cefixime causing gonorrhoea treatment failure in Austria, 2011. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*. 2011;16(43):19998. doi:10.2807/ese.16.43.19998-en

103. Unemo M, Golparian D, Nicholas R, et al. High-level cefixime- and ceftriaxone-resistant *Neisseria gonorrhoeae* in France: novel *penA* mosaic allele in a successful international clone causes treatment failure. *Antimicrob Agents Chemother*. 2012;56(3):1273-1280. doi:10.1128/AAC.05760-11

104. Camara J, Serra J, Ayats J, et al. Molecular characterization of two high-level ceftriaxone-resistant *Neisseria gonorrhoeae* isolates detected in Catalonia, Spain. *J Antimicrob Chemother*. 2012;67(8):1858-1860. doi:10.1093/jac/dks162

105. Lahra MM, Ryder N, Whiley DM. A New Multidrug-Resistant Strain of *Neisseria gonorrhoeae* in Australia. *N Engl J Med*. 2014;371(19):1850-1851. doi:10.1056/NEJMc1408109

106. Jeverica S, Golparian D, Mati i M, et al. Phenotypic and molecular characterization of *Neisseria gonorrhoeae* isolates from Slovenia, 2006-12: rise and fall of the multidrug-resistant NG-MAST genogroup 1407 clone? *J Antimicrob Chemother*. 2014;69(6):1517-1525. doi:10.1093/jac/dku026

107. Egli K, Roditscheff A, Flückiger U, et al. Molecular characterization of a ceftriaxone-resistant *Neisseria gonorrhoeae* strain found in Switzerland: a case report. *Ann Clin Microbiol Antimicrob*. 2021;20:52. doi:10.1186/s12941-021-00456-5

108. Pleininger S, Indra A, Golparian D, et al. Extensively drug-resistant (XDR) *Neisseria gonorrhoeae* causing possible gonorrhoea treatment failure with ceftriaxone plus azithromycin in Austria, April 2022. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*. 2022;27(24). doi:10.2807/1560-7917.ES.2022.27.24.2200455

109. Eyre DW, Town K, Street T, et al. Detection in the United Kingdom of the *Neisseria gonorrhoeae* FC428 clone, with ceftriaxone resistance and intermediate resistance to azithromycin, October to December 2018. *Euro Surveill*. 2019;24(10):1900147. doi:10.2807/1560-7917.ES.2019.24.10.1900147

110. Ko KKK, Chio MT, Goh SS, et al. First Case of Ceftriaxone-Resistant Multidrug-Resistant *Neisseria gonorrhoeae* in Singapore. *Antimicrob Agents Chemother*. 2019;63(5). doi:10.1128/aac.02624-18

111. Nakayama S ichi, Shimuta K, Furubayashi K ichi, et al. New Ceftriaxone- and Multidrug-Resistant *Neisseria gonorrhoeae* Strain with a Novel Mosaic *penA* Gene Isolated in Japan. *Antimicrob Agents Chemother*. 2016;60(7):4339-4341. doi:10.1128/AAC.00504-16

112. Lefebvre B, Martin I, Demczuk W, et al. Ceftriaxone-Resistant *Neisseria gonorrhoeae*, Canada, 2017. *Emerg Infect Dis*. 2018;24(2). doi:10.3201/eid2402.171756

113. Lahra MM, Martin I, Demczuk W, et al. Cooperative Recognition of Internationally Disseminated Ceftriaxone-Resistant *Neisseria gonorrhoeae* Strain. *Emerg Infect Dis*. 2018;24(4):735-740. doi:10.3201/eid2404.171873

114. Terkelsen D, Tolstrup J, Johnsen CH, et al. Multidrug-resistant *Neisseria gonorrhoeae* infection with ceftriaxone resistance and intermediate resistance to azithromycin, Denmark, 2017. *Eurosurveillance*. 2017;22(42):17-00659. doi:10.2807/1560-7917.ES.2017.22.42.17-00659

115. Poncin T, Fouere S, Braille A, et al. Multidrug-resistant *Neisseria gonorrhoeae* failing treatment with ceftriaxone and doxycycline in France, November 2017. *Eurosurveillance*. 2018;23(21):1800264. doi:10.2807/1560-7917.ES.2018.23.21.1800264

116. Golparian D, Rose L, Lynam A, et al. Multidrug-resistant *Neisseria gonorrhoeae* isolate, belonging to the internationally spreading Japanese FC428 clone, with ceftriaxone resistance and intermediate resistance to azithromycin, Ireland, August 2018. *Eurosurveillance*. 2018;23(47):1800617. doi:10.2807/1560-7917.ES.2018.23.47.1800617

117. Poncin T, Merimeche M, Braille A, et al. Two cases of multidrug-resistant *Neisseria gonorrhoeae* related to travel in south-eastern Asia, France, June 2019. *Eurosurveillance*. 2019;24(36):1900528. doi:10.2807/1560-7917.ES.2019.24.36.1900528

118. Chen SC, Yuan LF, Zhu XY, et al. Sustained transmission of the ceftriaxone-resistant *Neisseria gonorrhoeae* FC428 clone in China. *J Antimicrob Chemother*. 2020;75(9):2499-2502. doi:10.1093/jac/dkaa196

119. Lin X, Chen W, Xie Q, et al. Dissemination and genome analysis of high-level ceftriaxone-resistant *penA* 60.001 *Neisseria gonorrhoeae* strains from the Guangdong Gonococcal antibiotics susceptibility Programme (GD-GASP), 2016-2019. *Emerg Microbes Infect*. 2022;11(1):344-350. doi:10.1080/22221751.2021.2011618

120. Ameyama S, Onodera S, Takahata M, et al. Mosaic-Like Structure of Penicillin-Binding Protein 2 Gene (*penA*) in Clinical Isolates of *Neisseria gonorrhoeae* with Reduced Susceptibility to Cefixime. *Antimicrob Agents Chemother*. 2002;46(12):3744-3749. doi:10.1128/AAC.46.12.3744-3749.2002

121. Tomberg J, Unemo M, Ohnishi M, et al. Identification of amino acids conferring high-level resistance to expanded-spectrum cephalosporins in the *penA* gene from *Neisseria gonorrhoeae* strain H041. *Antimicrob Agents Chemother*. 2013;57(7):3029-3036. doi:10.1128/AAC.00093-13

122. Tomberg J, Unemo M, Davies C, et al. Molecular and structural analysis of mosaic variants of penicillin-binding protein 2 conferring decreased susceptibility to expanded-spectrum cephalosporins in *Neisseria gonorrhoeae*: role of epistatic mutations. *Biochemistry*. 2010;49(37):8062-8070. doi:10.1021/bi101167x

123. Lee SG, Lee H, Jeong SH, et al. Various *penA* mutations together with *mtrR*, *porB* and *ponA* mutations in *Neisseria gonorrhoeae* isolates with reduced susceptibility to cefixime or ceftriaxone. *J Antimicrob Chemother*. 2010;65(4):669-675. doi:10.1093/jac/dkp505

124. Tomberg J, Fedarovich A, Vincent LR, et al. Alanine-501 Mutations in Penicillin-Binding Protein 2 from *Neisseria gonorrhoeae*: Structure, Mechanism, and Effects on Cephalosporin Resistance and Biological Fitness. *Biochemistry*. 2017;56(8):1140-1150. doi:10.1021/acs.biochem.6b01030

125. Palace SG, Wang Y, Rubin DH, et al. RNA polymerase mutations cause cephalosporin resistance in clinical *Neisseria gonorrhoeae* isolates. *eLife*. 2020;9:e51407. doi:10.7554/eLife.51407

126. Salverda MLM, De Visser JAGM, Barlow M. Natural evolution of TEM-1 β-lactamase: experimental reconstruction and clinical relevance. *FEMS Microbiol Rev*. 2010;34(6):1015-1036. doi:10.1111/j.1574-6976.2010.00222.x

127. Hamilton HL, Dillard JP. Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol Microbiol*. 2006;59(2):376-385. doi:10.1111/j.1365-2958.2005.04964.x

128. M Unemo, JDC Ross, AB Serwin, et al. 2020 European guideline for the diagnosis and treatment of gonorrhoea in adults. *Int J STD AIDS*. Published online 2020. doi:10.1177/0956462420949126

129. World Health Organization, Reproductive Health and Research. *WHO Guidelines for the Treatment of Neisseria gonorrhoeae.*; 2016. Accessed August 5, 2019. http://www.ncbi.nlm.nih.gov/books/NBK379221/

130. Fouéré S, Cazanave C, Hélary M, et al. Update on French recommendations for the treatment of uncomplicated *Neisseria gonorrhoeae* infections. *Int J STD AIDS*. 2021;32(11):1081-1083. doi:10.1177/09564624211023025

131. Cyr, SS, Barbee L, Workowski KA, et al. Update to CDC's Treatment Guidelines for Gonococcal Infection, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69. doi:10.15585/mmwr.mm6950a6

132. Rijksinstituut voor Volksgezondheid en Milieu. Gonorroe | LCI richtlijnen. Published 2020. Accessed July 13, 2022. http://lci.rivm.nl/richtlijnen/gonorroe

133. Maiden MCJ, Bygraves JA, Feil E, et al. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A*. 1998;95(6):3140-3145. doi:10.1073/pnas.95.6.3140

134. Bennett JS, Jolley KA, Sparling PF, et al. Species status of *Neisseria gonorrhoeae*: evolutionary and epidemiological inferences from multilocus sequence typing. *BMC Biol*. 2007;5:35. doi:10.1186/1741-7007-5-35

135. Martin IMC, Ison CA, Aanensen DM, et al. Rapid sequence-based identification of gonococcal transmission clusters in a large metropolitan area. *J Infect Dis*. 2004;189(8):1497-1505. doi:10.1086/383047

136. Demczuk W, Sidhu S, Unemo M, et al. *Neisseria Gonorrhoeae* Sequence Typing for Antimicrobial Resistance, a Novel Antimicrobial Resistance Multilocus Typing Scheme for Tracking Global Dissemination of *N. gonorrhoeae* Strains. *J Clin Microbiol*. 2017;55(5):1454-1468. doi:10.1128/jcm.00100-17

137. Maiden MCJ, van Rensburg MJJ, Bray JE, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol*. 2013;11(10):728-736. doi:10.1038/nrmicro3093

138. Harrison OB, Cehovin A, Skett J, et al. *Neisseria gonorrhoeae* Population Genomics: Use of the Gonococcal Core Genome to Improve Surveillance of Antimicrobial Resistance. *J Infect Dis*. 2020;222(11):1816-1825. doi:10.1093/infdis/jiaa002

139. Didelot X, Bowden R, Wilson DJ, et al. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet*. 2012;13(9):601-612. doi:10.1038/nrg3226

140. Carriço JA, Rossi M, Moran-Gilad J, et al. A primer on microbial bioinformatics for nonbioinformaticians. *Clin Microbiol Infect*. 2018;24(4):342-349. doi:10.1016/j.cmi.2017.12.015

141. Earl D, Bradnam K, John JS, et al. Assemblathon 1: A competitive assessment of *de novo* short read assembly methods. *Genome Res*. 2011;21(12):2224-2241. doi:10.1101/gr.126599.111

142. Kingsford C, Schatz MC, Pop M. Assembly complexity of prokaryotic genomes using short reads. *BMC Bioinformatics*. 2010;11(1):21. doi:10.1186/1471-2105-11-21

143. Wick RR, Judd LM, Gorrie CL, et al. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genomics*. 2017;3(10):e000132. doi:10.1099/mgen.0.000132

144. Richardson EJ, Watson M. The automatic annotation of bacterial genomes. *Brief Bioinform*. 2013;14(1):1-12. doi:10.1093/bib/bbs007

145. Brockhurst MA, Harrison E, Hall JPJ, et al. The Ecology and Evolution of Pangenomes. *Curr Biol*. 2019;29(20):R1094-R1103. doi:10.1016/j.cub.2019.08.012

146. Medini D, Donati C, Tettelin H, et al. The microbial pan-genome. *Curr Opin Genet Dev*. 2005;15(6):589-594. doi:10.1016/j.gde.2005.09.006

147. Schürch AC, Arredondo-Alonso S, Willems RJL, et al. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. *Clin Microbiol Infect*. 2018;24(4):350-354. doi:10.1016/j.cmi.2017.12.016

148. Gorrie CL, Silva AGD, Ingle DJ, et al. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. *Lancet Microbe*. Published online August 6, 2021. doi:10.1016/S2666-5247(21)00149-X

149. Schwengers O, Hain T, Chakraborty T, et al. ReferenceSeeker: rapid determination of appropriate reference genomes. *J Open Source Softw*. 2020;5(46):1994. doi:10.21105/joss.01994

150. Valiente-Mullor C, Beamud B, Ansari I, et al. One is not enough: On the effects of reference genome for the mapping and subsequent analyses of short-reads. *PLOS Comput Biol*. 2021;17(1):e1008678. doi:10.1371/journal.pcbi.1008678

151. Schierup MH, Hein J. Consequences of recombination on traditional phylogenetic analysis. *Genetics*. 2000;156(2):879-891. doi:10.1093/genetics/156.2.879

1

PART 1

USING GENOMICS FOR ANTIMICROBIAL RESISTANCE SURVEILLANCE

2

SIMULTANEOUS DETECTION OF *NEISSERIA GONORRHOEAE* AND FLUOROQUINOLONE RESISTANCE MUTATIONS TO ENABLE RAPID PRESCRIPTION OF ORAL ANTIBIOTICS

Jolinda de Korne-Elenbaas, Arno Pol, Jacqueline Vet, Mirjam Dierdorp, Alje P. van Dam, Sylvia M. Bruisten

Sexually Transmitted Diseases 2020; 47(4):238-242

ABSTRACT

BACKGROUND

Absence of rapid antimicrobial resistance testing of *Neisseria gonorrhoeae* hinders personalised antibiotic treatment. To enable rapid ciprofloxacin prescription, a real-time polymerase chain reaction (PCR) for simultaneous detection of *N. gonorrhoeae* and fluoroquinolone resistance-associated gyrA-S91F mutation was evaluated.

METHODS

Analytical NG^{RES} quantitative PCR kit (NYtor BV) performance was assessed on 50 *N. gonorrhoeae* transcription-mediated amplification (TMA)-negative and 100 *N. gonorrhoeae* TMA-positive samples. To assess clinical use, 200 samples were prospectively analysed, in parallel to routine diagnostic tests. Also, 50 urine, 50 anal, 50 pharyngeal and 50 vaginal *N. gonorrhoeae* TMA-positive samples were retrospectively analysed. To assess if patients carried strains with different ciprofloxacin sensitivity at different anatomical locations, 50 urine/anal or vaginal/anal sample pairs collected during a single visit were analysed.

RESULTS

The NG^{RES} qPCR kit showed 97% sensitivity and 100% specificity for *N. gonorrhoeae* detection and 92% sensitivity and 99% specificity for *gyrA*-S91F detection. Relative to TMA results, 85% *N. gonorrhoeae* detection sensitivity and 99% specificity were found. Regarding the 200 prospectively analysed clinical samples, 13 were *N. gonorrhoeae* positive, of which 10 were also tested for antibiotic susceptibility by culture. The kit showed concordance for *GyrA*-S91F detection in 9 of 10 samples. *N. gonorrhoeae* was detected in 96% and 94% of vaginal and urine TMA-positive samples, in 84% of anal samples and only in 22% of pharyngeal samples. Discordant ciprofloxacin sensitivity was found for 2 of 26 characterised urine/anal sample pairs.

CONCLUSION

The NG^{RES} qPCR kit can be implemented in diagnostic testing for vaginal, urine and anal *N. gonorrhoeae* TMA-positive samples to enable rapid prescription of oral ciprofloxacin.

INTRODUCTION

Gonorrhoea, the sexually transmitted infection (STI) caused by *Neisseria gonorrhoeae*, is one of the most common bacterial STIs worldwide[1]. *N. gonorrhoeae* infection is treated with antimicrobial therapy; however, antimicrobial-resistant *N. gonorrhoeae* is a major global public health concern[2]. Treatment recommendations are based on antimicrobial resistance surveillance data and change according to the prevalence of resistant *N. gonorrhoeae* strains. Because of the high genomic plasticity of *N. gonorrhoeae* and the exposure to antibiotics without susceptibility testing before treatment, *N. gonorrhoeae* antimicrobial resistance develops continuously[3]. The World Health Organization regularly updates their guidelines for the treatment of *N. gonorrhoeae* infection to reduce the spread of multi drug-resistant gonococcal strains. Guidelines of 2016 included the recommendation of single-dose intramuscular ceftriaxone and oral azithromycin as first-choice treatment[4]. However, azithromycin- or ceftriaxone-resistant *N. gonorrhoeae* strains have been reported all over the world[2,5-8]. Because alternative antimicrobial treatment options are scarce, the current treatment needs to change towards a more personalised approach[9].

In the past, fluoroquinolones such as ciprofloxacin were effectively used as treatment of *N. gonorrhoeae* infection. Ciprofloxacin treatment had several advantages over other antimicrobials such as oral administration and relatively few adverse effects. However, fluoroquinolones were eliminated as first-choice treatment once the World Health Organization -defined threshold for the prevalence of fluoroquinolone resistant *N. gonorrhoeae* strains was reached[10]. Nevertheless, still large proportions of *N. gonorrhoeae* strains remain susceptible to ciprofloxacin: 70% in the United States, more than 70% in Australia and 53% in Europe[11-13]. In the Netherlands, 65% of all *N. gonorrhoeae* strains collected in 2018 were ciprofloxacin sensitive[14]. Because ciprofloxacin is highly efficacious in susceptible *N. gonorrhoeae* infections of any anatomic site, the British Association for Sexual Health and HIV (BASHH) guidelines still suggest the use of ciprofloxacin when susceptibility is known before treatment[15]. Rapid detection of susceptibility status of clinical *N. gonorrhoeae* isolates could therefore reintroduce treatment with oral ciprofloxacin and provide tailor-made therapy.

The mechanism of ciprofloxacin resistance in *N. gonorrhoeae* is well defined by the presence of mutations in serine codon 91 of the DNA gyrase subunit A (*gyrA)* gene[16,17]. Previous studies have demonstrated the potency of targeting the *gyrA* gene by using a real-time PCR assay to assess ciprofloxacin susceptibility of clinical *N. gonorrhoeae* isolates[18-20]. For even faster analysis, multiplex real-time PCR assays were developed to simultaneously detect *N. gonorrhoeae* and *gyrA* mutations[21,22]. However, most of these assays are not easily available because of in-house development or because extensive validation for use in routine diagnostics has not been performed. In this study, we extensively assessed the performance of the NG^{RES} qPCR kit,

developed by the Dutch company NYtor BV, for simultaneous detection of *N. gonorrhoeae* and the ciprofloxacin resistance-associated *gyrA*-S91F mutation for use in diagnostic routine. The NG^{RES} qPCR kit enables rapid ciprofloxacin susceptibility testing and could allow for the prescription of oral ciprofloxacin based on the results of a single qPCR assay.

METHODS

SAMPLE SELECTION

NG^{RES} qPCR kit (NYtor BV, Nijmegen, the Netherlands) performance was assessed on clinical samples obtained from patients visiting the STI outpatient clinic of Amsterdam and on samples obtained from general practitioners who requested *N. gonorrhoeae* testing. For samples of STI clinic visitors, routine diagnostic tests consisted of a transcription mediated amplification (TMA) test on the Aptima Combo 2 CT/NG assay (Hologic, Marlborough, MA). *N. gonorrhoeae* culture was performed when gram-negative diplococci were present in a gram-stained smear or with material collected at a return visit when the TMA test showed an *N. gonorrhoeae*-positive result. For *N. gonorrhoeae* culture, the clinical sample was inoculated on a plate with BBL™ GC-Lect™ Agar (Becton, Dickinson and Company, Franklin Lakes, NJ). A pure *N. gonorrhoeae* colony was picked for growth on a GC agar plate, enriched with 1% IsoVitaleX (BioTRADING Benelux BV, Mijdrecht, the Netherlands), and ciprofloxacin minimum inhibitory concentration (MIC) values were routinely determined with an e-test according to the manufacturer's instructions (bioMérieux SA, Marcy l'Étoile, France). According to the European Committee on Antimicrobial Susceptibility Testing (EUCAST), an MIC ≤0.03 µg/mL is considered as sensitive and an MIC ≥0.06 µg/mL as resistant to ciprofloxacin. Samples obtained from general practitioners were only tested with the TMA according to routine diagnostics because cultures could not be performed on samples from TMA collection tubes and additional culture material was unavailable.

Analytical performance of the NG^{RES} qPCR kit was assessed with 100 *N. gonorrhoeae* TMA-positive and 50 *N. gonorrhoeae* TMA-negative samples from STI clinic visitors that were randomly selected and retrospectively analysed. Samples were run at the Public Health Laboratory in Amsterdam on the Rotorgene (Qiagen, Holden, Germany) qPCR platform. The same samples were also run on the CFX96 (Bio-Rad, Hercules, CA) by NYtor BV, and results were compared. Successful cultures were performed with additional material from these 100 *N. gonorrhoeae*-positive patients so MIC values were available.

To compare results from the NG^{RES} qPCR kit and routine diagnostic tests, 100 urine and 100 anal samples were selected from male visitors of the STI clinic and prospectively analysed with the kit in parallel to routine diagnostic tests and culture. Results from the Rotorgene qPCR platform were compared with TMA and culture-based results.

To assess NG^{RES} qPCR kit performance on samples from different anatomical locations, 200 *N. gonorrhoeae* TMA-positive samples (50 urinal, 50 anal, 50 vaginal, 50 pharyngeal) were retrospectively analysed with the kit on the Rotorgene qPCR platform. All samples were obtained from general practitioners and collected in TMA collection tubes. Because additional culture material was unavailable, MIC values were unknown.

Lastly, we determined the fraction of patients that carried *N. gonorrhoeae* strains with different ciprofloxacin sensitivity at different anatomical locations. From 41 male patients, urine and anal samples collected during a single visit were selected and paired as well as vaginal and anal samples from 9 female patients. All resulting 50 sample pairs were analysed with the NG^{RES} qPCR kit, and pairwise comparisons were made.

SAMPLE PREPARATION

Previously collected samples had been stored in Aptima Combo 2 collection buffer (Hologic, Marlborough, USA) at -20°C according to routine storage. Retrospectively analysed samples that were taken from storage, thawed and vortexed. DNA was extracted from 200 µL of each sample by isopropanol precipitation after lysis with NucliSENS® easyMAG® Lysis buffer (bioMérieux SA, Marcy l'Étoile, France) enriched with glycogen (40 µg/mL). The pellet was washed with 70% ethanol twice, dissolved in 50µL Tris-HCl at pH 8.0 and stored at -20°C until use in amplification experiments. During each DNA extraction experiment, a negative control was included by adding 200 µL Aptima Combo 2 buffer to the lysis buffer.

NG^{RES} QPCR KIT

The NG^{RES} qPCR kit, developed by NYtor BV, contained a primer/probe mix designed for specific detection of the *N. gonorrhoeae adenylate kinase (adk)* gene and ciprofloxacin resistance-associated *gyrA* single nucleotide polymorphism S91F. During analytical validation, an additional mutation A92P in the *gyrA* probe binding region was detected, which prevented the original probe from binding. Subsequently, NYtor BV provided an adapted NG^{RES} qPCR kit, which effectuated probe binding even in the presence of the *gyrA*-A92P mutation. The adapted kit was then used for further clinical validation. Positive signals were detected in the FAM (*adk*) and VIC (*gyrA*-S91F) channels. The corresponding internal control was detected in the Cy5 channel, and the positive control generated signals in both the FAM and VIC channels.

A qPCR mix was prepared by mixing 10 µL master mix solution and 5 µL primer/probe/internal control solution (provided in NG^{RES} qPCR kit), and 5 µL DNA extract was added. PCR reactions were performed on the CFX96 and/or the Rotorgene with the following program: 1 cycle 3 minutes 95°C and 45 cycles 15 seconds 95°C and 60 seconds 60°C. Raw data were manually checked and analysed with Bio-Rad CFX Manager software version 3.1 or Rotorgene software version 1.7.0.75, respectively. According to diagnostic routine guidelines of the Public Health Laboratory,

detection of *N. gonorrhoeae* and/or *gyrA*-S91F was confirmed in case of well-defined sigmoidal curves and Ct-values <36 for both samples and controls. Samples that showed sigmoidal curves with Ct-values >36 and <40 or badly shaped sigmoidal curves were repeated, and a Ct-value <40 was then determined positive.

RESULTS

CONCORDANCE BETWEEN RESULTS OF CFX96 AND ROTORGENE QPCR PLATFORMS

To assess the analytical performance of the NG^{RES} qPCR kit, 100 *N. gonorrhoeae* TMA-positive and 50 *N. gonorrhoeae* TMA-negative samples were randomly selected for retrospective analysis. One *N. gonorrhoeae*-positive sample could not be traced back and was excluded from analysis. For all 99 *N. gonorrhoeae* TMA-positive samples, ciprofloxacin MICs were determined on routine cultures of additional material from the same patient. All 149 samples were analysed on both the CFX96 and the Rotorgene qPCR platforms. One sample was inhibited on both qPCR platforms. The NG^{RES} qPCR kit detected *N. gonorrhoeae* in 96/98 (98%) and 94/98 (96%) *N. gonorrhoeae* TMA-positive samples on the CFX96 and Rotorgene qPCR platform, respectively. Regarding the *N. gonorrhoeae* TMA-negative samples, 50/50 (100%) were also negative on both platforms (Table 1).

TABLE 1. COMPARISON OF RESULTS FROM DIAGNOSTIC TESTS AND THE NG^{RES} QPCR KIT.

	Results diagnostic tests	***Ng* TMA pos Culture CIP sensitive[a] (n=53[b])**	***Ng* TMA pos Culture CIP resistant[a] (n=45)**	***Ng* TMA neg (n=50)**
CFX96	*Ng* pos / S91F neg	52	4[c,d]	0
	Ng pos / S91F pos	1[e]	39	0
	Ng neg	0	2	50
Rotorgene	*Ng* pos / S91F neg	51	3[c,d]	0
	Ng pos / S91F pos	0	40	0
	Ng neg	2	2	50

[a] *CIP sensitive = MIC ≤0.03 µg/mL, CIP resistant = MIC ≥0.06 µg/mL*
[b] *Of the 100 N. gonorrhoeae TMA positive samples, one sample could not be traced back and one sample was inhibited on both qPCR platforms.*
[c] *Including 2 samples with A92P mutation in probe binding region (see text).*
[d] *Including samples that showed results in accordance to culture-based results after repetition (see text).*
[e] *Resistance was not confirmed in cultured strain (see text).*
Ng indicates N. gonorrhoeae, CIP indicates ciprofloxacin; neg, negative; pos, positive.

Ciprofloxacin sensitivity of the *N. gonorrhoeae* strains in the samples was determined for all 100 *N. gonorrhoeae* TMA-positive samples on routine cultures of additional material from the same patient. Of all clinical samples that contained a ciprofloxacin-sensitive *N. gonorrhoeae*

strain and that were *N. gonorrhoeae* positive with the NG^{RES} qPCR kit (CFX96: n=53, Rotorgene: n=51), 52/53 (98%) and 51/51 (100%) were also *gyrA*-S91F negative, so ciprofloxacin sensitive, on the CFX96 or Rotorgene qPCR platform. One *N. gonorrhoeae* sample was *gyrA*-S91F positive on the CFX96 but negative on the Rotorgene (Table 1). Cultured urethra material of the same patient was subsequently tested as ciprofloxacin sensitive, indicating *gyrA*-S91F absence. Of all 43 clinical samples that contained a resistant *N. gonorrhoeae* strain and were *N. gonorrhoeae* positive with the NG^{RES} qPCR kit, *gyrA*-S91F was detected in 39/43 (91%) and 40/43 (93%) samples on the CFX96 or Rotorgene qPCR platform respectively. The *gyrA*-S91F mutation was thus not detected in 4 samples on the CFX96 and in the same 3 samples on the Rotorgene (Table 1). For the 4 samples that had shown discordant genotype results, DNA was again extracted from fresh cultures and analysed by NYtor BV on the CFX96 platform. The *gyrA*-S91F mutation was detected in 2/4 samples after repetition, which was in accordance with the culture-based results. For the other 2 samples, the NG^{RES} qPCR kit was still unable to detect the *gyrA*-S91F mutation, although these samples contained ciprofloxacin resistant *N. gonorrhoeae* strains with MIC values ≥0.06 µg/mL. The *gyrA* genes from the 4 discordant samples were also sequenced, and it seemed that the 2 samples that remained discordant after repetition carried an additional mutation, as described hereinafter.

FIGURE 1. SEQUENCES OF THE *GYRA* PROBE REGION WITHOUT MUTATION (WT), WITH SINGLE S91F OR WITH BOTH S91F AND A92P MUTATIONS. WT INDICATES WILD TYPE.

GYRA-A92P MUTATION PREVENTED PROBE BINDING

For 2 samples, no *gyrA*-S91F mutation was detected by using the NG^{RES} qPCR kit on both qPCR platforms (Table 1), although the MIC values obtained from concomitant cultures were >32.0, which indicates ciprofloxacin resistance. These samples were collected from the anus and vagina of the same patient. The *gyrA* genes of these samples were partly sequenced, and an additional *gyrA*-A92P mutation was found, which prevented probe binding during the qPCR reaction (Figure 1). NYtor BV provided an adapted NG^{RES} qPCR kit, which effectuated probe binding even in the presence of a *gyrA*-A92P mutation. The adapted kit was then used for further clinical validation.

TABLE 2. COMPARISON OF RESULTS FROM NG^{RES} QPCR KIT AND ROUTINE DIAGNOSTIC TESTS: TRANSCRIPTION-MEDIATED AMPLIFICATION (TMA) TEST AND CULTURE-BASED ANTIBIOTIC SUSCEPTIBILITY TESTING.

		Results TMA/ *Ng* culture			
		***Ng* TMA neg (n= 187)**	***Ng* TMA pos (n=13[a])**	***Ng* TMA pos Culture CIP sensitive[b] (n=5[a])**	***Ng* TMA pos Culture CIP resistant[b] (n=5[a])**
Results NG^{RES} qPCR kit	*Ng* pos	1	11	-	-
	Ng neg	186	2	-	-
	Ng pos/ S91F neg	-	-	4	0
	Ng pos/ S91F pos	-	-	1	5

[a] *From 3 out of 13 N. gonorrhoeae positive sample, cultures were not available so ciprofloxacin sensitivity could not be compared with the detected gyrA genotype.*
[b] *CIP sensitive = MIC≤0.03 µg/mL, CIP resistant = MIC ≥0.06 µg/mL*
Ng indicates N. gonorrhoeae, CIP indicates ciprofloxacin; neg, negative; pos, positive.

COMPARISON OF RESULTS FROM NG^{RES} QPCR KIT AND ROUTINE DIAGNOSTIC TESTS

NG^{RES} qPCR kit performance was compared with routine diagnostic test results with 100 urine and 100 anal samples. These were selected from male visitors of the STI clinic and prospectively analysed with the kit in parallel to routine diagnostic tests, which included culture and MIC assessment. Regarding the total number of samples, 186/187 (99%) were negative and 11/13 (85%) positive with the kit compared with TMA results (Table 2). For 3/13 *N. gonorrhoeae*-positive samples, cultures were not available so the ciprofloxacin MIC values could not be compared with the *gyrA* genotype. For the other 10 samples, ciprofloxacin sensitivity was assessed on routine cultures of additional material from the same patient. The kit detected *gyrA*-S91F in 5/5 (100%) samples that contained a ciprofloxacin-resistant *N. gonorrhoeae* strain, and no mutation was detected in 4/5 (80%) samples that contained a ciprofloxacin-sensitive *N. gonorrhoeae* strain. For one sample, the kit detected the *gyrA*-S91F mutation, indicating ciprofloxacin resistance, although the *N. gonorrhoeae* strain had been assessed as ciprofloxacin sensitive on concomitant culture (Table 2).

HIGH NG^{RES} QPCR KIT SENSITIVITY IN URINE AND VAGINAL SAMPLES

To assess NG^{RES} qPCR kit performance on samples from various anatomical locations, 200 *N. gonorrhoeae* TMA-positive samples were retrospectively analysed with the kit. Samples were collected from urine (n=50), vagina (n=50), anus (n=50) or pharynx (n=50). *N. gonorrhoeae* was detected in 48/50 (96%) urine samples, 47/50 (94%) vaginal samples and 42/50 (84%) anal samples. Only 11/50 (22%) pharyngeal samples were *N. gonorrhoeae* positive with the kit. The *gyrA*-S91F mutation was detected in 50/148 (34%) samples that were *N. gonorrhoeae* positive

(Table 3). Additional material from these patients was not available for culture so *gyrA*-S91F detection could not be validated.

PATIENTS MOSTLY CARRIED SIMILAR *N. GONORRHOEAE* STRAINS AT TWO ANATOMICAL LOCATIONS

To determine the fraction of patients that carried *N. gonorrhoeae* strains with discordant ciprofloxacin sensitivity at different anatomical locations, urine/anal or vaginal/anal samples that were collected during a single clinic visit were selected from 41 men and 9 women. All 50 sample pairs were analysed with the NG^{RES} qPCR kit. For 15/41 men, the kit could not detect *N. gonorrhoeae* in the anal sample so these sample pairs could not be further characterised. For the other 26 men, 14 men (54%) carried gyrA-S91F mutant *N. gonorrhoeae* strains and 10 men (38%) carried gyrA wild-type *N. gonorrhoeae* strains in the urine and anus. In the other 2 men, strains with different gyrA types were found in the urine and anus: 1 man carried a gyrA-S91F mutant strain in the urine and a gyrA wild-type strain in the anus, and 1 man carried a gyrA- S91F mutant strain in the anus and a gyrA wild-type strain in the urine. For 2 of 9 women, the kit could not detect *N. gonorrhoeae* in the anal sample so these sample pairs could not be further characterised. The other 7 women carried *gyrA* wild-type *N. gonorrhoeae* strains in both the vagina and the anus. Regarding ciprofloxacin susceptibility, 10/26 (38%) men carried only ciprofloxacin-sensitive *N. gonorrhoeae* strains in contrast to 7/7 (100%) women.

TABLE 3. NG^{RES} QPCR KIT ANALYSIS RESULTS OF 200 *N. GONORRHOEAE* TMA-POSITIVE SAMPLES ORIGINATING FROM DIFFERENT ANATOMICAL LOCATIONS.

	NG^{RES} qPCR kit results	
	***Ng* positive**	***gyrA* mutant[a]**
Urine (n=50)	48 (96%)	16 (33%)
Vagina (n=50)	47 (94%)	9 (19%)
Rectum (n=50)	42 (84%)	18 (43%)
Pharynx (n=50)	11 (22%)	7 (64%)
Total (n=200)	**148 (74%)**	**50 (34%)**

[a]Ciprofloxacin sensitivity as detected with the NG^{RES} qPCR kit because cultures were not available.
Ng indicates N. gonorrhoeae

DISCUSSION

We assessed the performance of the NG^{RES} qPCR kit for simultaneous detection of *N. gonorrhoeae* and the ciprofloxacin resistance associated *gyrA*-S91F mutation. Use of this kit may allow for the prescription of oral ciprofloxacin based on the results of a single qPCR assay, which takes 4 hours including DNA extraction from the clinical sample. Implementation of the kit in diagnostic screening would increase the turnaround time with about half a day. However, this facilitates more rapid antibiotic sensitivity analysis compared with the current susceptibility testing based on bacterial cultures, which require at least 2 days. A detailed cost-benefit analysis for implementation is hard to perform because both costs and benefits, such as the use of an oral- instead of an injectable antibiotic, are very much depending on the local situation.

The NG^{RES} qPCR kit showed good performance compared with our current daily routine tests, which include the highly sensitive TMA test for *N. gonorrhoeae* detection and culture for antimicrobial sensitivity testing. Only 5 samples showed discordant results between culture-based determination of ciprofloxacin sensitivity and detection of the *gyrA*-S91F mutation with the NG^{RES} qPCR kit (Tables 1,2). These samples either showed concordant results after repeated culture, DNA extraction and analysis, or the discrepancy was caused by an additional *gyrA*-A92P mutation, which prevented probe binding in the initial, non-optimised kit format. Assessing the type of errors in the discrepant results is important to determine its clinical impact. When typing is impossible, treatment recommendation remains ceftriaxone. False-positive *gyrA*-S91F detection would result in not using ciprofloxacin, which is not of clinical concern. However, false-negative *gyrA*-S91F detection would result in using ciprofloxacin to treat resistant infections, which certainly is of clinical concern. In our study, only 2 samples were false-negative for *gyrA*-S91F (Table 1). For these samples, the Ct-value for *N. gonorrhoeae* detection was high (>36), which indicates a low bacterial load. When using the NG^{RES} qPCR kit in clinical practice, samples with such high Ct-values should be repeated or assessed as non-typable. Because only 2 samples gave false-negative results, we are confident that the kit performance is sufficient and will increase if a certain cut-off Ct-value is established by further validation in clinical practice.

SpeeDx recently developed the ResistancePlus® assay for dual detection of *N. gonorrhoeae* and *gyrA*-S91F mutation, which is already approved and commercially available[22]. Ebeyan et al. [22] also reported 2 samples that gave discordant results when comparing culture and genotyping with this ResistancePlus® assay. The results of one sample became concordant after repetition and one sample remained discrepant[22]. A possible explanation for the discrepancies observed by Ebeyan et al. and these in our study could be that patients are infected with multiple *N. gonorrhoeae* strains at the same anatomical location, and that strains with discrepant ciprofloxacin sensitivity were taken for culturing and for genotyping with the NG^{RES} qPCR kit. Such mixed infections have been reported to occur in high-risk populations, albeit at low

prevalence[23,24]. When implementing the kit in clinical practice, it should be further validated to what extent mixed infections influence the genotyping results.

When assessing the performance of the NG^{RES} qPCR kit on samples from different anatomical locations, the NG^{RES} qPCR kit showed excellent results for urine and vaginal samples. *N. gonorrhoeae* detection sensitivity was lower in anal samples, but for pharyngeal samples the sensitivity was clearly insufficient, suggesting a low bacterial load in throat swabs as described previously[25,26]. Compared with the previously published real-time PCR assay developed by Hemarajata et al[18], the NG^{RES} qPCR kit demonstrated higher ability to characterise the genotype. Compared to the assay developed by Ellis et al[20], higher genotyping sensitivity was found for urine and rectal samples but lower sensitivity for pharyngeal samples. The ResistancePlus® assay reported by Ebeyan et al[22] showed better *gyrA*-S91F detection sensitivity and specificity, especially in pharyngeal samples. However, anal samples were missing in the validation of this assay, which seemed difficult to characterise in our study and could have an effect on the proportion of characterised samples. Also, initial *N. gonorrhoeae* screening was not done with the Aptima Combo 2 but with other commercial assays, so direct comparison is not possible. Recently, Allan-Blitz et al[27] showed that the ResistancePlus® assay was able to genotype some of the samples that were non-typable in the method described by Ellis et al. Comparing the NG^{RES} qPCR kit with the ResistancePlus® assay would be valuable. Although the NG^{RES} qPCR kit could not characterise a small proportion of the anal samples, by analysing urine/anal and vaginal/anal sample pairs that were obtained from one patient during a single visit, we showed that only 2 of 33 patients carried *N. gonorrhoeae* strains that differed in ciprofloxacin sensitivity, whereas the other 31 patients had strains with similar ciprofloxacin sensitivity. This indicated that analysis of only the urine or vaginal sample without the anal sample would be suitable for guided ciprofloxacin treatment for most cases.

When assessing the performance of the kit on samples from different anatomical locations, 148 of 200 samples could be characterised. In 66% of these samples, a *gyrA* wild-type *N. gonorrhoeae* strain was detected, which is in accordance with the previously published fraction of 65% of ciprofloxacin-sensitive strains in the Netherlands[14]. The presence of this significant fraction of ciprofloxacin-sensitive *N. gonorrhoeae* strains shows that ciprofloxacin can still be used for treatment of *N. gonorrhoeae* detection if rapid ciprofloxacin-sensitivity detection is part of routine diagnostics. We suggest to implement the NG^{RES} qPCR kit in addition to a highly sensitive assay. In our setting, we used the TMA test which proved to be very suitable for first screening of *N. gonorrhoeae*-suspected samples, followed by analysis of *N. gonorrhoeae* TMA-positive samples with the NG^{RES} qPCR kit. Based on the genotype determined by the NG^{RES} qPCR kit, physicians could be advised if treatment with ciprofloxacin is indicated. Effectiveness of the assay should be carefully validated by comparing results with current diagnostic test results, and ciprofloxacin use could be monitored by a test-of-cure study.

Results of the analysis of sample pairs from 26 men and 7 women showed that 38% of the men and 100% of the women carried *gyrA* wild-type *N. gonorrhoeae* strains. Although the number of men included in the study was much higher than the number of women, this result suggests that men are more often carrying resistant *N. gonorrhoeae* strains. It has been reported that pharyngeal *N. gonorrhoeae* infections are particularly important in transmission networks of men who have sex with men and that *Neisseria* species could exchange resistance determinants at the pharyngeal site[28]. Because 80% of the STI clinic visitors are men who have sex with men, this could explain the high percentage of men carrying ciprofloxacin-resistant strains. However, further research on transmission networks of ciprofloxacin resistant strains among general populations is needed.

In conclusion, our results show that the NG^{RES} qPCR kit is suitable for concomitant detection of *N. gonorrhoeae* and *gyrA*-S91F, especially in urine and vaginal clinical samples, to enable rapid prescription of ciprofloxacin.

REFERENCES

1. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. Bull World Health Organ. 2019;97:548–62.
2. Wi T, Lahra MM, Ndowa F, Bala M, Dillon J-AR, Ramon-Pardo P, et al. Antimicrobial resistance in *Neisseria gonorrhoeae*: Global surveillance and a call for international collaborative action. PLOS Med. 2017;
3. Buono SA, Watson TD, Borenstein LA, Klausner JD, Pandori MW, Godwin HA. Stemming the tide of drug-resistant *Neisseria gonorrhoeae*: the need for an individualized approach to treatment. J Antimicrob Chemother. 2015;70(2):374–81.
4. World Health Organization. WHO guidelines for the treatment of *Neisseria gonorrhoeae*. Available at: http://www.ncbi.nlm.nih.gov/books/NBK379221/ Accessed August 5, 2019.
5. Shimuta K, Unemo M, Nakayama S, Morita-Ishihara T, Dorin M, Kawahata T, et al. Antimicrobial Resistance and Molecular Typing of *Neisseria gonorrhoeae* Isolates in Kyoto and Osaka, Japan, 2010 to 2012: Intensified Surveillance after Identification of the First Strain (H041) with High-Level Ceftriaxone Resistance. Antimicrob Agents Chemother. 2013;57(11):5225–5232.
6. Ko KKK, Chio MT, Goh SS, Tan AL, Koh TH, Abdul Rahman NB. First Case of Ceftriaxone-Resistant Multidrug-Resistant *Neisseria gonorrhoeae* in Singapore. Antimicrob Agents Chemother. 2019;63(5).
7. Unemo M, Golparian D, Nicholas R, Ohnishi M, Gallay A, Sednaoui P. High-level cefixime- and ceftriaxone-resistant *Neisseria gonorrhoeae* in France: novel *penA* mosaic allele in a successful international clone causes treatment failure. Antimicrob Agents Chemother. 2012;56(3):1273–1280.
8. Eyre DW, Town K, Street T, Barker L, Sanderson N, Cole MJ, et al. Detection in the United Kingdom of the *Neisseria gonorrhoeae* FC428 clone, with ceftriaxone resistance and intermediate resistance to azithromycin, October to December 2018. Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull. 2019;24(10).
9. Trembizki E, Guy R, Donovan B, Kaldor JM, Lahra MM, Whiley DM. Further evidence to support the individualised treatment of gonorrhoea with ciprofloxacin. Lancet Infect Dis. 2016;16(9):1005–6.
10. Centers for Disease Control and Prevention (CDC). Update to CDC's sexually transmitted diseases treatment guidelines, 2006: fluoroquinolones no longer recommended for treatment of gonococcal infections. MMWR Morb Mortal Wkly Rep. 2007;56(14):332–6.
11. Harris SR, Cole MJ, Spiteri G, Sánchez-Busó L, Golparian D, Jacobsson S, et al. Public health surveillance of multidrug-resistant clones of *Neisseria gonorrhoeae* in Europe: a genomic survey. Lancet Infect Dis. 2018;18(7):758–768.
12. Lahra MM, Enriquez R, George CRR. Australian Gonococcal Surveillance Programme Annual Report, 2017. Commun Dis Intell 2018. 2019;43.
13. Centers for Disease Control and Prevention (CDC). Sexually Transmitted Disease Surveillance 2018. Atlanta: U.S. Department of Health and Human Services; 2018. Available at: www.cdc.gov/std/stats18/gonorrhea.htm Accessed December 16, 2019.
14. National Institute for Public Health and the Environment. Sexually Transmitted Infections in the Netherlands in 2018. National Institute for Public Health and the Environment; 2019 p. 189. Report No.: 2019-0007. Available at: www.rivm.nl/bibliotheek/rapporten/2019-0007.pdf Accessed August 5, 2019.
15. British Association for Sexual Health and HIV. BASHH Guidelines. Available at: https://bashh.org/guidelines/ Accessed December 16, 2019.
16. Tanaka M, Otsuki M, Nishino T, Kobayashi I, Matsumoto T, Kumazawa J. Mutation in DNA gyrase of norfloxacin-resistant clinical isolates of *Neisseria gonorrhoeae*. Sex Transm Infect. 1996;72(4):295–7.
17. Allan-Blitz L-T, Wang X, Klausner JD. Wild-Type Gyrase A Genotype of *Neisseria gonorrhoeae* Predicts In Vitro Susceptibility to Ciprofloxacin: A Systematic Review of the Literature and Meta-Analysis. Sex Transm Dis. 2017;44(5):261–265.
18. Hemarajata P, Yang S, Soge OO, Humphries RM, Klausner JD. Performance and Verification of a Real-Time PCR Assay Targeting the *gyrA* Gene for Prediction of Ciprofloxacin Resistance in *Neisseria gonorrhoeae*. J Clin Microbiol. 2016;54(3):805–8.
19. Siedner MJ, Pandori M, Castro L, Barry P, Whittington WLH, Liska S, et al. Real-Time PCR Assay for Detection of Quinolone-Resistant *Neisseria gonorrhoeae* in Urine Samples. J Clin Microbiol. 2007;45(4):1250–4.

20. Ellis O, Hemarajata P, Shahkolahi A, Masinde G, Buchs K, Humphries RM, et al. A multisite implementation of a real-time polymerase chain reaction assay to predict ciprofloxacin susceptibility in *Neisseria gonorrhoeae*. Diagn Microbiol Infect Dis. 2019;94(3):213–7.
21. Perera SR, Khan NH, Martin I, Taheri A, Parti RP, Levett PN, et al. Multiplex Real-Time PCR Assay for Simultaneous Identification of *Neisseria gonorrhoeae* and Its Ciprofloxacin Susceptibility Status. J Clin Microbiol. 2017;55(11):3201–3209.
22. Ebeyan S, Windsor M, Bordin A, Mhango L, Erskine S, Trembizki E, et al. Evaluation of the ResistancePlus GC (beta) assay: a commercial diagnostic test for the direct detection of ciprofloxacin susceptibility or resistance in *Neisseria gonorrhoeae*. J Antimicrob Chemother. 2019;74(7):1820–4.
23. Goire N, Kundu R, Trembizki E, Buckley C, Hogan TR, Lewis DA, et al. Mixed gonococcal infections in a high-risk population, Sydney, Australia 2015: implications for antimicrobial resistance surveillance? J Antimicrob Chemother. 2017;72(2):407–9.
24. Martin IMC, Ison CA. Detection of mixed infection of *Neisseria gonorrhoeae*. Sex Transm Infect. 2003;79(1):56–58.
25. Alexander S. The challenges of detecting gonorrhoea and chlamydia in rectal and pharyngeal sites: could we, should we, be doing more? Sex Transm Infect. 2009;85(3):159–60.
26. Hananta IPY, De Vries HJC, van Dam AP, van Rooijen MS, Soebono H, Schim van der Loeff MF. Persistence after treatment of pharyngeal gonococcal infections in patients of the STI clinic, Amsterdam, the Netherlands, 2012-2015: a retrospective cohort study. Sex Transm Infect. 2017;93(7):467–71.
27. Allan-Blitz LT, Ellis OL, Wee R, Truong A, Ebeyan SM, Tan LY et al. Improved determination of *Neisseria gonorrhoeae* gyrase A genotype results in clinical specimens. J Antimicrob Chemother. 2019:74(10):2913-2915.
28. Whittles LK, White PJ, Paul J, Didelot X. Epidemiological Trends of Antibiotic Resistant Gonorrhoea in the United Kingdom. Antibiotics. 2018;7(3).

2


3
1.5
2
2.5
3
3.5
4
mL
0.5
1
1.5
2
2.5
3
3.5
4
mL
0.5
1
1.5
2

EMERGENCE OF A *NEISSERIA GONORRHOEAE* CLONE WITH REDUCED CEPHALOSPORIN SUSCEPTIBILITY BETWEEN 2014 AND 2019 IN AMSTERDAM, THE NETHERLANDS, REVEALED BY GENOMIC POPULATION ANALYSIS

Jolinda de Korne-Elenbaas, Sylvia M. Bruisten,
Henry J. C. de Vries, Alje P. van Dam

Journal of Antimicrobial Chemotherapy 2021; 76(7):1759-1768

ABSTRACT

BACKGROUND

Emerging resistance to cephalosporins in *Neisseria gonorrhoeae* is a major public health threat, since these are considered antibiotics of last resort. Continuous surveillance is needed to monitor the circulation of resistant strains and those with reduced susceptibility.

OBJECTIVES

For the purpose of epidemiological surveillance, genomic population analysis was performed on *N. gonorrhoeae* isolates from Amsterdam with a focus on isolates with reduced susceptibility to ceftriaxone.

METHODS

WGS data were obtained from 318 isolates from Amsterdam, the Netherlands between 2014 and 2019. Isolates were typed according to MLST, *N. gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) and *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) schemes and additional resistance markers were identified. Phylogenetic trees were created to identify genetic clusters and to compare Dutch and non-Dutch MLST7827 isolates.

RESULTS

MLST7363 and MLST1901 were the predominant strains having reduced susceptibility to ceftriaxone during 2014–16; MLST7827 emerged and dominated during 2017–19. NG-STAR38 and NG-MAST2318/10386 were predominant among MLST7827 isolates. MLST7827 reduced susceptibility isolates carried a non-mosaic 13.001 *penA* allele with an A501V mutation and *porB1b* G120K/A121D mutations, which were lacking in susceptible MLST7827 isolates. Phylogenetic analysis of all publicly available MLST7827 isolates showed strong genetic clustering of Dutch and other European MLST7827 isolates.

CONCLUSIONS

MLST7827 isolates with reduced ceftriaxone susceptibility have emerged during recent years in Amsterdam. Co-occurrence of *penA* A501V and *porB1b* G120K/A121D mutations was strongly associated with reduced susceptibility to ceftriaxone. Genetic clustering of Dutch and other European MLST7827 isolates indicates extensive circulation of this strain in Europe. Close monitoring of the spread of this strain having an alarming susceptibility profile is needed.

INTRODUCTION

The emergence of resistance in *Neisseria gonorrhoeae* poses a major public health threat. Current treatment recommendation is the last-resort extended-spectrum cephalosporin ceftriaxone, together with azithromycin as dual therapy[1]. Since the benefit of dual therapy is not evidence-based and emerging high-level azithromycin resistance has been found in many countries[2], Dutch, French and UK treatment guidelines recommend ceftriaxone monotherapy. WHO guidelines endorse monotherapy as well, provided that local resistance data confirming susceptibility to ceftriaxone are available[1]. However, single cases of ceftriaxone-resistant isolates have been reported over recent years, underlining the need for continuous surveillance of circulating strains[3–6].

Although ceftriaxone-resistant isolates have only been reported in a few countries, a drift towards higher ceftriaxone MICs has been observed worldwide, indicating a global reduction in susceptibility[7]. Ceftriaxone resistance can be determined by mutations in several genes, such as *porB*, *ponA*, *mtrR*, *rpoB* and *rpoD*, but especially by A501 mutations or mosaicism in the *penA* gene[8–11]. Additional mutations outside mosaic *penA* genes might add to the resistant phenotype, although their effects remain to be proven[8]. Genotypic characterisation of the ceftriaxone-resistant isolates found in Japan, France and the UK identified mosaic *penA* 37.001, 42.001 or 60.001 alleles in these isolates, which belonged to MLST7363/1901/1903 and *N. gonorrhoeae* Multi-Antigen Sequence Type (NG-MAST) 4220/1407/3435[3,4,6]. A ceftriaxone-resistant isolate from Singapore belonged to different STs but contained the same mosaic *penA* 60.001 allele[5].

Isolates belonging to MLST1901/NG-MAST1407 with mosaic *penA* alleles were highly prevalent among strains with reduced susceptibility to ceftriaxone also, suggesting that tracking this strain is most important for monitoring emerging resistance. However, Osnes et al.[12] recently reported the emergence of a strain with reduced susceptibility between 2016 and 2018 in Norway, belonging to MLST7827 and carrying a non-mosaic *penA* 13.001 allele with A501V mutation. They showed that this strain, with an alarming antimicrobial resistance profile, likely originated from Asia and potentially circulates in Europe. Analysing Dutch isolates with reduced susceptibility from 2009–17, De Laat et al.[13] found a shift from a mosaic *penA* allele towards a non-mosaic *penA* allele with A501 mutation. We have now further examined the genetic change among isolates with reduced susceptibility in Amsterdam. WGS data were used for genomic characterisation isolates with reduced ceftriaxone susceptibility and a representative part of the susceptible gonococcal population isolated from 2014–19 in Amsterdam. We aimed to identify genomic characteristics associated with reduced susceptibility to ceftriaxone in the Amsterdam gonococcal population.

METHODS

ISOLATE DETAILS AND SELECTION

Isolates were collected from *N. gonorrhoeae*-positive visitors to the sexually transmitted infection (STI) outpatient clinic of the Public Health Service of Amsterdam. MICs of azithromycin, ciprofloxacin and ceftriaxone were routinely determined for all isolates using Etests according to the manufacturer's instructions (bioMérieux SA). Ciprofloxacin clinical breakpoints were determined according to EUCAST clinical breakpoints v11.0. For azithromycin, isolates with MIC<0.5 mg/L were assigned as susceptible, MIC=0.5 mg/L as intermediate and MIC≥1.0 mg/L as resistant (epidemiological cut-off; ECOFF). For ceftriaxone, isolates with MIC≤0.016 mg/L were assigned as susceptible, MIC=0.023–0.064 mg/L as intermediate and MIC≥0.094 mg/L as having reduced susceptibility or as resistant in the case of MIC>0.125 mg/L. During the study period of January 2014 to July 2019, ceftriaxone-resistant isolates were not found among the 7323 isolates that were cultured and stored.

For genomic characterisation of strains with reduced ceftriaxone susceptibility circulating in Amsterdam, all 82 isolates with ceftriaxone MIC≥0.094 mg/L obtained during the study period were selected for WGS. To characterise the gonococcal population circulating in Amsterdam, 244 isolates (3.4% of all available isolates) with ceftriaxone MIC<0.094 mg/L obtained during the study period were also selected, resulting in a total selection of 326 isolates for WGS. Isolates were randomly selected after stratification on year of isolation and ceftriaxone MIC; for each reduced susceptibility strain, three isolates with MIC<0.094 mg/L from the same year of isolation were randomly selected. Stratification on ceftriaxone MIC was done to get a distribution of MICs that were <0.094 mg/L in the selection similar to the distribution of MICs that were <0.094 mg/L in the total Amsterdam gonococcal population.

DNA ISOLATION AND WGS

Selected isolates were taken from −80°C storage, grown overnight on chocolate blood agar plates and DNA was extracted from pure cultures. Isolates were sequenced on the Illumina MiSeq or Illumina NovaSeq 6000 platform (the latter was chosen for higher throughput). For Illumina MiSeq sequencing, DNA was extracted using isopropanol precipitation after lysis with NucliSENS easyMAG Lysis Buffer (bioMérieux SA) with glycogen (40 mg/L). The pellet was washed twice in 70% EtOH and dissolved in 50 µL of Tris-HCl at pH 8.0. DNA sequencing libraries were prepared with the KAPA HTP Library Preparation Kit (Roche Life Sciences) and Nextflex Dual-Indexed DNA barcodes (Bioo Scientific) and 300 bp paired-end sequenced. Regarding Illumina NovaSeq 6000 sequencing, DNA was extracted from harvested bacteria in DNA/RNA Shield buffer using the ZymoBIOMICS™ MagBead DNA Kit (ZYMO RESEARCH). DNA sequencing libraries were prepared with the Nextera XT DNA Library Preparation Kit with IDT for Illumina

DNA/RNA UD Indexes (Illumina) and 150 bp paired-end sequenced. All raw reads are available in the European Nucleotide Archive under accession number PRJEB40983.

BIOINFORMATIC ANALYSES

Default settings were used unless noted otherwise. Raw sequence reads were filtered, trimmed and adapters were removed with fastp v0.20.0.[14] Reads were mapped to reference genome FA1090 (NC_002946.2) with BWA-MEM2 v2.2.1 to calculate coverage using the SAMtools package v1.11.[15,16] Isolates were excluded if coverage was <95%. Reads were assembled with Skesa v2.3.0 with a minimum contig length of 500 bp and assembly quality was assessed with QUAST v5.0.2.[17,18] For isolates with a total assembly length of >2.1 Mbp, Kraken2 v2.0.8 was used to check for contamination.[19] Variants were called with Snippy v4.4.0 using reference genome FA1090 and a full core-genome alignment was created with the snippy-core option (https://github.com/tseemann/snippy). Gubbins v2.3.4 was used to identify regions of recombination in this alignment and to create a phylogenetic tree based on a recombination-filtered variant alignment, by using the general time-reversible model with gamma distribution (GTR-GAMMA) in RAxML v8.2.12.[20,21] The phylogenetic tree with metadata was visualised using iTOL and legends were added with PDF Pro.[22] Bayesian Analysis of Population Structure (BAPS) was performed using the rhierBAPS package v1.1.2 in R v3.6.3 [settings: maximum depth=2; maximum number of populations (n.pops)=75].[23]

Isolates were uploaded to the PubMLST database and MLST, NG-MAST and *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) STs were extracted.[24,25] Novel MLSTs and NG-STAR STs were submitted to the PubMLST and NG-STAR databases, respectively. Annotation of resistance genes (*penA*, *porB*, *ponA*, *gyrA*, *parC*, 23S rRNA, *mtrA/R/C/D/E*, *rplD*, *rplV*, *rpmH*, *rpoB*, *rpoD)* was done using either the allele annotation available in the PubMLST database or by manually aligning allele sequences from PubMLST and identifying previously reported resistance mutations or mosaicism.[9,26–29] Raw reads were mapped against all 23S rRNA reference sequences in the PubMLST database to identify heterogeneous A2058/2059G/C2611T mutations in the four different alleles using Ariba v2.14.4.[30] Snakemake v5.6.0 was used for workflow management.[31] The pipeline is freely available at https://github.com/jolindadekorne/Genomic-population-analysis-of-Neisseria-gonorrhoeae.

COMPARISON OF MLST7827 ISOLATES

The genetic relatedness of Dutch and non-Dutch MLST7827 isolates was assessed. PubMLST contains a total of 224 *N. gonorrhoeae* isolates belonging to MLST7827 (August 2020), of which 63 are Dutch and 161 are non-Dutch isolates. For 14 non-Dutch isolates, only contigs were available and downloaded from PubMLST. For the other 147 non-Dutch isolates, raw sequence reads were downloaded. Subsequently, fastp was used for filtering, trimming and adapter removal. One

Dutch MLST7827 isolate was hybrid assembled using Unicycler v0.4.8 with Illumina MiSeq and MinION Nanopore sequence data, yielding a circular chromosome.[32] Variants were called with Snippy v4.6.0 using the Dutch hybrid assembly as reference genome, either using raw reads or contigs with the —ctgs option. A recombination-corrected phylogenetic tree was created and visualised as described above. Median SNP distance per main genetic cluster was calculated using snp-dists v0.7.0 on the filtered variant alignment (https://github.com/tseemann/snp-dists).

STATISTICAL ANALYSES

Associations between patient and/or isolate characteristics were identified with two-tailed chi-squared or Fisher's exact tests using a 95% CI. The Bonferroni correction method was applied in the case of multiple testing. All statistical analyses were performed in R v3.6.3.

ETHICS

According to the Dutch Medical Research Act Involving Human Subjects, no additional ethical approval was required for this study (W20_451 # 20.498).

RESULTS

SEQUENCING DATA

Out of 326 isolates selected for WGS, 4 were excluded due to non-viable cultures, 1 due to >95% read contamination and 3 due to coverage of <95%. For the resulting 318 isolates, 252441 reads were obtained on average per isolate, with an average coverage of 98.7% (Table S1). The phylogenetic tree was created based on a recombination-filtered variant alignment of 18683 sites.

PATIENT CHARACTERISTICS

The 318 isolates were derived from 314 patients: 8 isolates were obtained from 4 patients from two different anatomical locations. Isolates were mainly obtained from MSM (82%) and isolated from the anus (48%). The median patient age was 30 years and the majority of patients were aged between 24 and 34 years (50%) (Table 1).

GENOMIC EPIDEMIOLOGY AND CHARACTERISATION OF RESISTANCE MUTATIONS

A midpoint-rooted phylogenetic tree was created based on the recombination-filtered variant alignment and two separate lineages were identified (Figure 1). The majority of isolates in the main lineage A (*n*=216) were from patients reporting homosexual or bisexual intercourse (96%).

TABLE 1. PATIENT AND ISOLATE CHARACTERISTICS

Patient characteristics	***N*=314**
Sex, *n* (%)	
male	286 (91)
female	28 (9)
Age, years	
median (range)	30 (16-65)
<24, *n* (%)	59 (19)
24–34, *n* (%)	159 (50)
35, *n* (%)	93 (30)
NA[a], *n* (%)	3 (1)
Sexual preference *n* (%)	
MSM	258 (82)
heterosexual	41 (13)
bisexual	12 (4)
NA[a]	3 (1)
Isolates sequenced	***N*=318[b]**
Year of isolation, *n* (%)	
2014–16	128 (40)
2017–19	190 (60)
Anatomical location *n* (%)	
Anus	152 (48)
Urethra	89 (28)
Vagina/cervix	14 (4)
Tonsil	62 (19.5)
Other	1 (0.5)

[a]*NA = not available.*
[b]*From four patients, two isolates were obtained from two anatomical locations.*

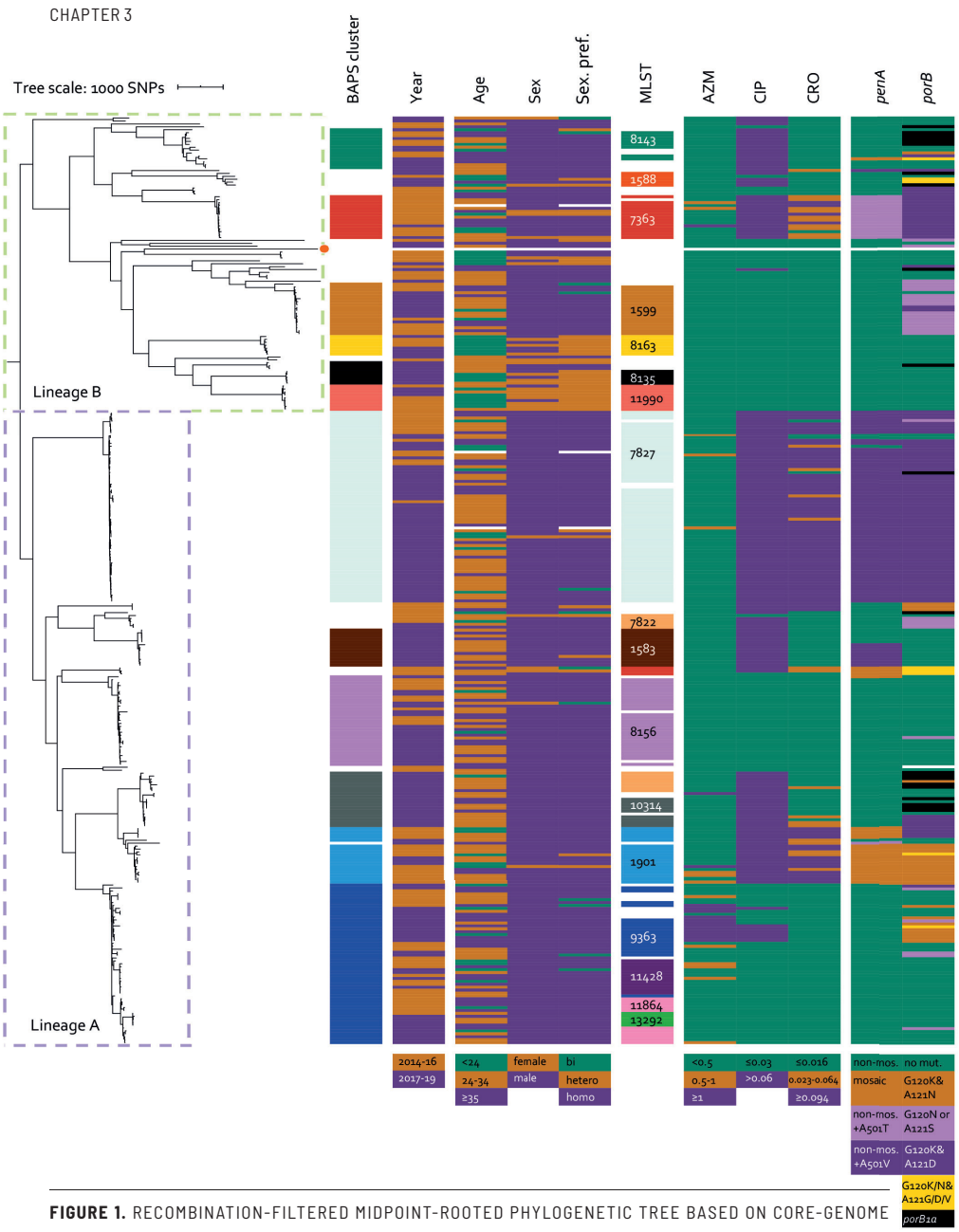


FIGURE 1. RECOMBINATION-FILTERED MIDPOINT-ROOTED PHYLOGENETIC TREE BASED ON CORE-GENOME SNPS INCLUDING 318 *N. GONORRHOEAE* ISOLATES FROM 2014–19 FROM AMSTERDAM, THE NETHERLANDS. THE FA1090 STRAIN WAS USED AS THE REFERENCE STRAIN AND ITS BRANCH IS VISUALISED WITH AN ORANGE DOT. METADATA INCLUDES: MAIN CLUSTERS DETERMINED WITH BAPS AN ALYSIS AT LEVEL 1; YEAR OF ISOLATION; PATIENT CHARACTERISTICS (AGE, SEX AND SEXUAL PREFERENCE); MLST CLUSTERS CONTAINING ≥5 ISOLATES; MICS IN MG/L FOR AZITHROMYCIN (AZM), CIPROFLOXACIN (CIP) AND CEFTRIAXONE (CRO); *PENA* TYPE AND *PORB* TYPE. *PORB1A* IS GIVEN IN BLACK; ALL OTHER COLOURS REPRESENT DIFFERENT *PORB1B* TYPES. PHENOTYPIC DATA ARE VISUALISED AS GREEN FOR SUSCEPTIBLE, ORANGE FOR INTERMEDIATE SUSCEPTIBLE AND PURPLE FOR RESISTANT/REDUCED SUSCEPTIBLE. WHITE BARS INDICATE MISSING DATA. TWO SEPARATE LINEAGES ARE DEFINED WITH DASHED-LINE BOXES.

Isolates from 2017–19 and isolates resistant to azithromycin, ciprofloxacin or ceftriaxone were significantly overrepresented in lineage A. Isolates in lineage B (*n*=102) were significantly associated with being female, aged <24 years and reporting heterosexual intercourse (Table S2).

BAPS clustering resulted in 14 clusters at level 1, of which 12 main clusters are visualised in Figure 1. MLSTs were determined for 317/318 (100%) isolates, yielding 56 different MLSTs, of which 26 were found for a single isolate. One isolate could not be assigned an MLST because of one incomplete locus. The 17 MLST clusters that contained ≥5 isolates were defined as main MLST clusters, with MLST7827 being the largest cluster, containing 63 isolates. MLST8135, MLST8163 and MLST11990 were only found in patients reporting heterosexual intercourse and MLST11990 was significantly associated with female patients (Table S3). NG-MAST and NG-STAR types were obtained for 318 (100%) and 304 (96%) isolates, respectively. Isolates not typable according to the NG-STAR scheme (*n*=14) carried heterogeneous 23S rRNA alleles. The most prevalent NG-MAST and NG-STAR types in each MLST cluster are shown in Table 2.

Isolates belonging to MLST1901 were significantly overrepresented during 2014–16 (15/19) and none of these were isolated after 2018. Also, MLST7363 isolates were significantly overrepresented during 2014–16 (15/17). Remarkably, MLST7827 isolates were significantly overrepresented during 2017–19 (50/63), of which 41/50 (82%) were isolated during 2018–19 (Table S3). These results indicate a recent emergence of MLST7827, which became the dominant strain with reduced ceftriaxone susceptibility in Amsterdam during 2017–19, instead of the previously dominating MLST1901 and MLST7363 strains.

AZITHROMYCIN

Azithromycin resistance was found for 15/318 (5%) isolates and resistance was significantly associated with MLST9363 (Table S3). Remarkably, a 23S rRNA C2611T mutation was only identified once in a susceptible isolate. Mosaicism in the *mtrR* promoter and gene was identified in 13/15 (87%) resistant isolates (Table 3). An additional fully mosaic *mtrC/D/E* operon was found in 12/13 (92%) and these mainly belonged to MLST9363 (9/12; 75%). The mosaic *mtrR* promoter and gene were only found in one susceptible isolate, belonging to MLST7367. The other two resistant isolates carried either a 35A deletion in the *mtrR* promoter or an *mtrR* A39T mutation and a non-mosaic or partly mosaic *mtrC/D/E* operon, but these alleles were also highly prevalent among susceptible isolates (Table 3). Mutations in *rplV, rmpH* and *mtrA* were not found at all (Table S1) and mutations in *rplD* and *mtrC* were not associated with resistance (Table 3).

TABLE 2. MOST PREVALENT NG-MAST AND NG-STAR TYPES FOUND IN EACH MAIN MLST CLUSTER

MLST cluster	Number of isolates	Main NG-MAST type (%)	Main NG-STAR type (%)
1583	13	15589 (92)	1340 (54)
1588	5	NA	NA
1599	17	11461 (65)	520 (76)
1901	19	1407 (58)	90 (32)
7363	17	2400 (35)	158 (53)
7822	12	14994 (33)	1387 (42)
7827	63	10368/2318 (35/33)	38 (89)
8135	5	387 (60)	729 (60)
8143	8	5624 (50)	426 (50)
8156	28	5441 (71)	442 (89)
8163	7	2 (29)	84 (86)
9363	18	12302 (28)	168 (28)
10314	9	NA	1387 (44)
11428	12	2992 (58)	63 (92)
11864	11	18234 (27)	439 (91)
11990	9	14376 (56)	962 (56)
13292	5	9208 (80)	439 (80)

NA, no dominant ST was found in that MLST cluster.

CIPROFLOXACIN

Ciprofloxacin resistance was found for 173/318 (54%) isolates and significant associations were found with MLST1583, MLST1901, MLST7363 and MLST7827. Susceptibility was significantly associated with MLST1599, MLST8156, MLST11428, MLST11864 and MLST11990 (Table S3). All resistant isolates carried the *gyrA* S91F mutation (173/173), which was found in only 2/145 (1%) susceptible isolates, showing its importance in ciprofloxacin resistance. All isolates with an additional D95G mutation were resistant, whereas the two susceptible isolates carried a D95A/N mutation. *ParC* D86N/S87/S88P mutations were highly prevalent among resistant isolates (147/173), although these are not required for resistance, given its absence in 15% of resistant isolates (Table 3).

TABLE 3. PHENOTYPIC CHARACTERISATION VERSUS IDENTIFIED RESISTANCE MUTATIONS.

Azithromycin	**Susceptible (MIC<0.5 mg/L)**	**Intermediate (MIC=0.5 mg/L)**	**Resistant (MIC≥1.0 mg/L)**
	***N*=289**	***N*=14**	***N*=15**
23S rRNA (*n*=318)[a]			
no A2058/A2059/C2611 mutation	288 (99.7)	14 (100)	15 (100)
C2611T in 1/4 alleles[b]	1 (0.3)		
mtrR promoter (–35A) and gene (A39T, G45D mutation) (*n*=318)[a]			
no –35A, A39T, G45D, mosaicism	35 (12.1)	–	–
–35A	66 (23)	5 (36)	1 (6.5)
–35A, A39T	1 (0.3)	–	–
–35A, G45D	68 (23.5)	3 (21)	–
mosaic promoter+gene	1 (0.3)	–	13 (87)
A39T	114 (39.4)	6 (43)	1 (6.5)
G45D	4 (1.4)	–	–
mtrC/D/E operon (*n*=316)[a]			
non-mosaic	248 (86.5)	8 (57)	1 (6.7)
non-mosaic, GC del in *mtrC*	3 (1)	–	–
mosaic	1 (0.3)	–	12 (80)
partly mosaic	35 (12.2)	6 (43)	2 (13.3)
rplD (*n*=318)[a]			
no G68/70 mutations	287 (99.3)	14 (100)	15 (100)
G70D	2 (0.7)	–	–
Ciprofloxacin	**Susceptible (MIC≤0.03 mg/L)**		**Resistant (MIC>0.06 mg/L)**
	***N*=145**		***N*=173**
gyrA (*n*=318)[a]			
no S91/D95 mutations	143 (99)		–
S91F, D95G	–		98 (57)
S91F, D95A	1 (0.5)		73 (42)
S91F, D95N	1 (0.5)		2 (1)
parC (*n*=318)[a]			
no D86, S87, S88 mutations	140 (96.6)		26 (15)
D86N	–		85 (49)
S87I/N/R	4 (2.8)		60 (35)
S87R, S88P	1 (0.6)		2 (1)

TABLE 3. PHENOTYPIC CHARACTERISATION VERSUS IDENTIFIED RESISTANCE MUTATIONS. (CONTINUED)

Ceftriaxone	**Susceptible (MIC≤0.016 mg/L)**	**Intermediate (MIC0.023–0.064 mg/L)**	**Reduced susceptibility (MIC≥0.094 mg/L)**
	***N*=212**	***N*=26**	***N*=80**
penA (*n*=318)[a]			
non-mosaic/no A501 mutation	198 (93)	6 (23)	3 (4)
mosaic 34.001	—	3 (11.5)	11 (14)
mosaic 10.001	1 (0.5)	2 (8)	4 (5)
mosaic 63.001/92.001	2 (1)	—	—
non-mosaic 18.001+A501T	—	1 (4)	—
non-mosaic 44.001+A501T	1 (0.5)	10 (38)	4 (5)
non-mosaic 12.004+A501V	—	1 (4)	—
non-mosaic 13.001+A501V	2 (1)	3 (11.5)	58 (72)
non-mosaic 43.002+A501V	8 (4)	—	—
porB (*n*=317)[a]			
porB1a no G120, A121 mutations	20 (9)	1 (4)	—
porB1b no G120, A121 mutations	137 (65)	2 (8)	—
porB1b G120K, A121N	10 (5)	2 (8)	14 (17.5)
porB1b G120K, A121D	10 (5)	18 (69)	66 (82.5)
porB1b G120N or A121S	29 (14)	—	—
porB1b G120K/N,A121G/D/V	5 (2)	3 (11)	—
mtrR promoter (–35A) and gene (A39T, G45D mutation) (*n*=318)[a]			
no –35A, A39T, G45D, mosaic	33 (16)	2 (8)	—
–35A	37 (17)	13 (50)	22 (27.5)
–35A, A39T	—	1 (4)	—
–35A, G45D	9 (4)	4 (15)	58 (72.5)
mosaic promoter+gene	13 (6)	1 (4)	—
A39T	116 (55)	5 (19)	—
G45D	4 (2)	—	—
ponA (*n*=318)[a]			
no L421P mutation	146 (69)	—	—
L421P	66 (31)	26 (100)	80 (100)

All shown as n (%).

[a]*Gene in which mutations were identified (number of isolates this gene was characterised in).*

[b]*As determined by mapping raw reads against 23S rRNA reference sequences for identification of heterogeneous mutations.*

CEFTRIAXONE

As a result of the selection strategy of this study, 80/318 (25%) isolates had reduced susceptibility to ceftriaxone and 26/318 (8%) had intermediate susceptibility. Intermediate and reduced susceptibility were significantly associated with MLST1901, MLST7827 and MLST7363 and susceptibility with MLST1599, MLST8156 and MLST9363 (Table S3). NG-MAST1407/NG-STAR90 was predominant among MLST1901 isolates, NG-MAST2400/NG-STAR158 among MLST7363 isolates and NG-MAST10386/NG-STAR38 and NG-MAST2318/NG-STAR38 among MLST7827 isolates (Table 2). Seventy-four percent (157/212) of the susceptible isolates carried a *porB1a* (20/157) or *porB1b* (137/157) gene without G120/A121 mutations and a non-mosaic *penA* gene without an A501 mutation (93%). The *porB1a* without G120/A121 mutations was only found in intermediate (1/21) or susceptible (20/21) isolates. *PorB1b* G120K and A121D/N mutations were found in 100% of the isolates with reduced susceptibility, 77% of the isolates with intermediate susceptibility and 10% of the susceptible isolates, indicating their importance in the resistance mechanism (Table 3). Nineteen percent (15/80) of the isolates with reduced susceptibility carried the mosaic *penA* allele 34.001 (11/15) or 10.001 (4/15) and belonged to MLST1901. Notably, these mosaic *penA* alleles were also found in intermediate MLST1901 (3/4) and MLST7363 isolates (2/11). Three susceptible isolates carried a mosaic *penA* allele, of which 63.001 and 92.001 were identified only once. In 42% of the isolates with intermediate susceptibility (11/26), either the non-mosaic *penA* allele 44.001 (10/11) or 18.001 (1/11) with A501T mutation was found, of which 82% (9/11) belonged to MLST7363. The majority of isolates with reduced susceptibility carried non-mosaic *penA* allele 13.001 with A501V mutation (72%), of which 97% (56/58) belonged to MLST7827 (Table 3, Figure 1). Three susceptible MLST7827 isolates were identified, of which one carried *porB1a* and the other two carried *porB1b* without G120/A121 mutations and non-mosaic *penA* without A501 mutation. *PorB1b* G120K/A121D mutations found outside the MLST7827 cluster and co-occurring with a non-mosaic *penA* allele without mutations were not associated with intermediate or reduced ceftriaxone susceptibility, except for three isolates in the MLST10314 cluster (Figure 1). However, intermediate or reduced susceptibility was observed outside the MLST7827 cluster when both *porB1b* mutations and either *penA* mosaicism or A501 mutations co-occurred (e.g. in MLST1901 and MLST7363 clusters). Moreover, isolates carrying a *penA* A501V mutation but lacking the *porB1b* mutations were susceptible (e.g. in the MLST1583 cluster). These findings show the interplay between mutations in *penA* and *porB* and that some of these mutations show a stronger effect on ceftriaxone MIC than others, thus gradually influencing the susceptibility. *PonA* L421P mutation, 35A deletion in the *mtrR* promoter and G45D in *mtrR* were found in the majority of intermediate and reduced susceptibility isolates. However, there was no direct association with reduced susceptibility since these mutations were also prevalent among susceptible isolates (Table 3). Mutations in *rpoB* and *rpoD* genes were not found (Table S1).

CLOSE GENETIC RELATEDNESS BETWEEN DUTCH AND OTHER EUROPEAN MLST7827 STRAINS

The genetic relatedness of the 63 Dutch and 161 non-Dutch MLST7827 isolates publicly available in the PubMLST database was assessed. Available metadata showed that a large proportion of the non-Dutch MLST7827 isolates were from 2011–13 (43%) and from Asia (42%). The Dutch isolates from this study accounted for 28% of the MLST7827 isolates in the database. Regarding available phenotypic data, 100% (192/193) of the isolates were ciprofloxacin resistant but only 3% (3/119) were azithromycin resistant. Regarding ceftriaxone, 47% (91/195) showed reduced susceptibility and 1% (2/195) were resistant (Table 4).

Recombination-filtered variant alignment resulted in 7526 sites, on which the phylogenetic tree was based (Figure 2). The midpoint-rooted phylogenetic tree showed two lineages: main lineage A with three distinct clusters and lineage B with mainly Asian isolates from 2011–13. Clusters 1 and 3 mainly contained Dutch, Norwegian and European+UK isolates from 2014–19 and cluster 2 mainly contained Asian, but also American, Norwegian and European+UK isolates from 2011–13. Dutch isolates from 2017–19 were only found in cluster 1, together with Norwegian and European+UK isolates and one American isolate. The median SNP distances within clusters 1, 2 and 3 were 34, 89 and 33, respectively. This indicated stronger genetic relatedness among Dutch, Norwegian and European+UK isolates in clusters 1 and 3 than among isolates in cluster 2, which were mainly Asian. Cluster 1 contained most of the reduced ceftriaxone susceptibility isolates carrying non-mosaic *penA* 13.001 alleles with A501V mutation and *porB1b* G120K/A121D mutations. In contrast, lineage A mainly contained susceptible isolates carrying non-mosaic *penA* and *porB1a* alleles without mutations. Isolates in clusters 2 and 3 mainly carried non-mosaic *penA* 13.001 with A501V mutations; however, a variety of *porB1b* mutations were found in these clusters. Overall, reduced susceptibility to ceftriaxone was associated with the co-presence of the *penA* 13.001 A501V and *porB1b* G120/A121 mutations among global MLST7827 isolates. Metadata for the MLST7827 isolates are available in Table S4.

FIGURE 2. RECOMBINATION-FILTERED MIDPOINT-ROOTED PHYLOGENETIC TREE BASED ON CORE-GENOME SNPS INCLUDING ALL 224 PUBLICLY AVAILABLE *N. GONORRHOEAE* ISOLATES BELONGING TO MLST7827. A DUTCH MLST7827 ISOLATE WAS USED AS THE REFERENCE STRAIN AND ITS BRANCH IS VISUALISED WITH AN ORANGE DOT. METADATA INCLUDES: COUNTRY/CONTINENT; YEAR OF ISOLATION; CEFTRIAXONE (CRO) MICS IN MG/L ARE VISUALISED AS GREEN FOR SUSCEPTIBLE, ORANGE FOR INTERMEDIATE SUSCEPTIBLE, PURPLE FOR REDUCED SUSCEPTIBLE AND BLACK FOR RESISTANT; *PENA* TYPE; AND *PORB* TYPE. *PORB1A* IS GIVEN IN BLACK; ALL OTHER COLOURS REPRESENT DIFFERENT *PORB1B* TYPES. WHITE BARS INDICATE MISSING DATA. DASHED-LINE BOXES DEFINE SEPARATE LINEAGES AND CLUSTERS.

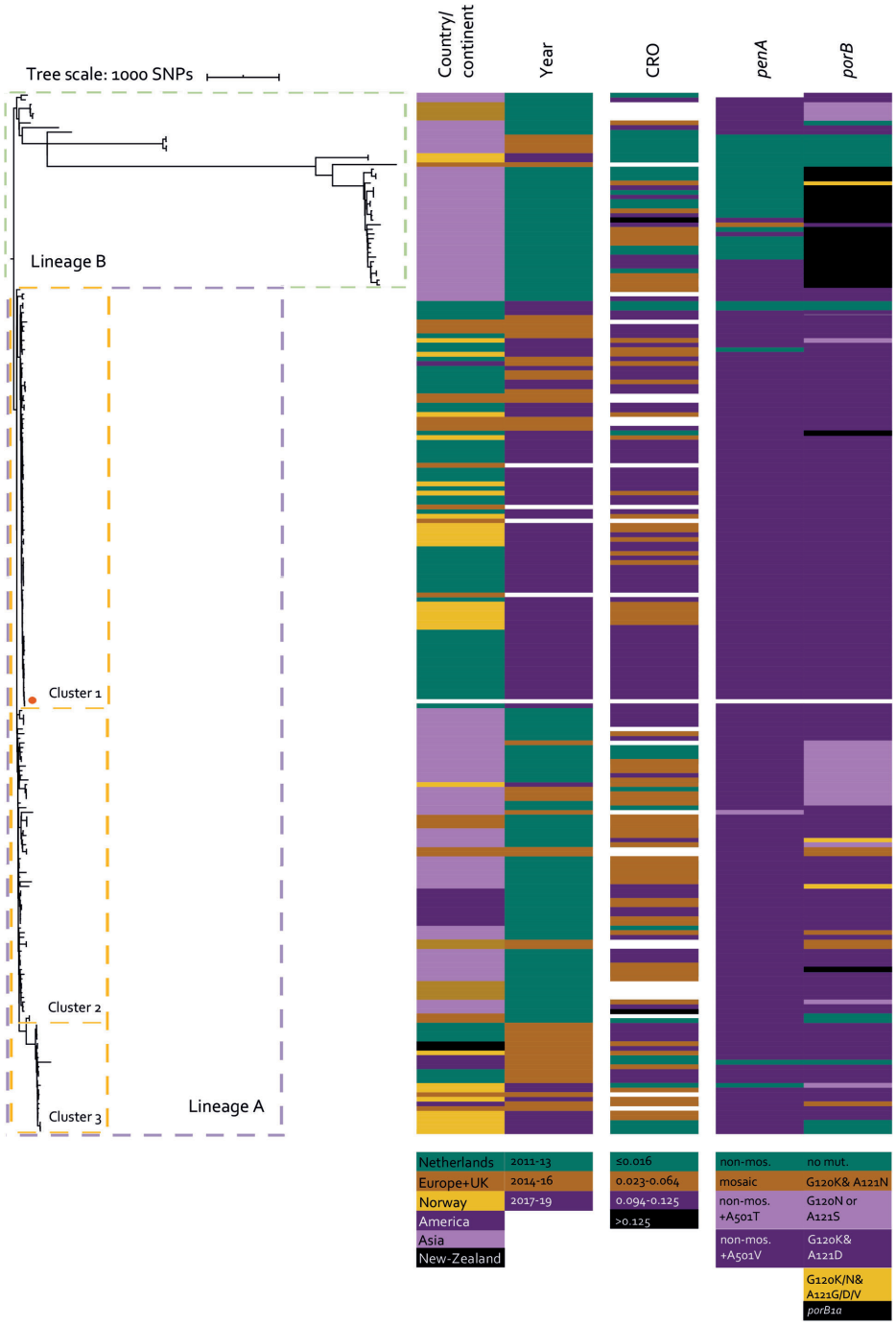

Tree scale: 1000 SNPs
Lineage B
Cluster 1
Cluster 2
Cluster 3
Lineage A
Country/ continent
Year
CRO
penA
porB
Netherlands
Europe+UK
Norway
America
Asia
New-Zealand
2011-13
2014-16
2017-19
≤0.016
0.023-0.064
0.094-0.125
>0.125
non-mos.
mosaic
non-mos. +A501T
non-mos. +A501V
no mut.
G120K& A121N
G120N or A121S
G120K& A121D
G120K/N& A121G/D/V
porB1a

TABLE 4. METADATA AVAILABLE FOR *N. GONORRHOEAE* ISOLATES BELONGING TO MLST7827 OBTAINED FROM THE PUBMLST DATABASE IN AUGUST 2020.

Number of MLST7827 isolates in PubMLST database	**224**	
Country/continent of isolation, *n* (%)		
Netherlands	63	(28)
Norway	30	(13)
Other European countries+UK	22	(10)
America (continent)	13	(6)
Asia	94	(42)
New Zealand	2	(1)
Year of isolation, *n* (%)		
2011–13	96	(43)
2014–16	45	(20)
2017–19	79	(35)
NA[a]	4	(2)
Ciprofloxacin MIC (mg/L), *n*		
≤0.03	1	
>0.06	192	
NA[a]	31	
Azithromycin MIC (mg/L), *n*		
<0.5	101	
0.5	15	
≥1.0	3	
NA[a]	105	
Ceftriaxone MIC (mg/L), *n*		
≤0.016	33	
0.023–0.064	69	
≥0.094–0.125	91	
>0.125	2	
NA[a]	29	

[a]*NA, data not available in PubMLST database.*

DISCUSSION

This genomic population study extended the previous NG-MAST and *penA* typing study, which identified NG-MAST and *penA* shifts among *N. gonorrhoeae* isolates with reduced ceftriaxone susceptibility from Amsterdam, obtained up until 2017.[13] Here we studied this phenomenon in more detail using WGS and including more recent isolates. The results showed that previous observations represented a shift from MLST1901 to MLST7363 and more recently to MLST7827. The emergence of the MLST7827 strain with reduced susceptibility to cephalosporins and resistance to ciprofloxacin in Amsterdam is in line with previously published surveillance articles. Peng *et al.*[33] reported MLST7827 as already being the predominant MLST in China during 2012–13, although at that time this MLST was not particularly associated with reduced susceptibility to cephalosporins. When this strain emerged in Norway during 2016–18 it was associated with reduced susceptibility to cephalosporins and we now confirm that the same emergence has occurred in the Netherlands during the last 3 years.[12] Since recent isolates from other European countries are scarce, these results can only suggest circulation of this strain in other parts of Europe.

Among the Dutch isolates, the co-occurrence of a non-mosaic *penA* 13.001 allele with A501V mutation and *porB1b* G120K/A121D mutations was associated with reduced susceptibility to ceftriaxone. Remarkably, these mutations were already found in isolates with reduced cephalosporin susceptibility from Korea during 2001–07, belonging to different NG-MAST STs.[34] From 2007 onwards, cephalosporin-resistant isolates belonging to MLST1901 or MLST7363 were identified that carried mosaic *penA* 37.001, 42.001 and 60.001 alleles; however, these *penA* alleles were not identified in this study. Other MLST and *penA* types have been found among isolates with reduced ceftriaxone susceptibility as well. In China, isolates belonging to MLST7363 and carrying a mosaic *penA* were responsible for reduced susceptibility during 2012–13.[33] We found both MLST1901 and MLST7363 to be associated with reduced susceptibility from 2014–16 in Amsterdam as well, confirming that these strains were the predominant strains with reduced susceptibility worldwide in previous years. In 2017, Abrams *et al.*[35] reported an isolate with reduced susceptibility that belonged to a different ST and lacked the mosaic *penA* allele. The *rpoB* and *rpoD* mutations found to be the genetic basis for reduced susceptibility in this isolate by Palace *et al.*[9] were not found among the Dutch isolates in the present study. Instead, reduced susceptibility was associated with MLST7827 isolates carrying *penA* and *porB1b* mutations, showing the interplay between these mutations in the resistance mechanism. This multifactorial nature also suggests additional and, as yet, unresolved genetic variations involved in cephalosporin resistance.

Previous research on Dutch isolates from 2008–15 showed a high prevalence of 23S rRNA mutations among azithromycin-resistant isolates with variable genetic backgrounds.[36] More

recently, the influence of *mtr* mosaicism on azithromycin resistance has been described.[28] The results of this study suggest a replacement of mutations in 23S rRNA by mosaic *mtr* genes as the main determinant in azithromycin-resistant strains circulating in Amsterdam, using a limited number of azithromycin-resistant strains. Further research on larger numbers of azithromycin-resistant isolates is needed to confirm this observation.

Phylogenetic analysis of the 318 Dutch isolates revealed two separate lineages. Isolates in lineage A were significantly associated with bisexual or homosexual intercourse and resistance or reduced susceptibility to azithromycin, ciprofloxacin and ceftriaxone was significantly overrepresented. This lineage distinction was also seen in other isolate collections.[37] Previous studies state that MSM are more often infected with MDR isolates, probably because of the higher prevalence of bacterial STI in MSM. This leads to higher antibiotic exposure and increases selection pressure for antimicrobial resistance.[38]

Importantly, because of the MIC-based selection strategy, the percentage of isolates showing reduced ceftriaxone susceptibility in this study is not representative of the *N. gonorrhoeae* population found among all STI clinic visitors in Amsterdam. Routine susceptibility testing showed that only 82 of 7323 (1.1%) isolates routinely obtained at the STI clinic had reduced ceftriaxone susceptibility during the study period. Although ceftriaxone-resistant isolates have not yet been found in the Netherlands, the emergence of MLST7827 isolates raises the question of whether this strain will evolve towards being a resistant strain. The two ceftriaxone-resistant MLST7827 isolates from China show the ability of this strain to become resistant according to the EUCAST clinical threshold, although these isolates did not cause therapy failure. High recombination rates in *N. gonorrhoeae* enable the exchange of resistance mutations, which could cause a further reduction in susceptibility and ultimately lead to clinical resistance. Identification of the MLST7827 strain in multiple European countries over recent years shows its ability to spread quickly, underlining the need for global surveillance to track the prevalence and development of this strain.

ACKNOWLEDGEMENTS

We would like to thank Boas van der Putten for his valuable input and careful reading of the manuscript.

FUNDING

This work was supported by the Public Health Service and the Public Health Laboratory of Amsterdam.

TRANSPARENCY DECLARATIONS

None to declare.

REFERENCES

1. WHO. WHO guidelines for the treatment of *Neisseria gonorrhoeae*. 2016. https://www.ncbi.nlm.nih.gov/books/NBK379221/pdf/Bookshelf_NBK379221.pdf.
2. Derbie A, Mekonnen D, Woldeamanuel Y *et al*. Azithromycin resistant gonococci: a literature review. *Antimicrob Resist Infect Control* 2020; **9**:138.
3. Shimuta K, Unemo M, Nakayama S *et al*. Antimicrobial resistance and molecular typing of *Neisseria gonorrhoeae* isolates in Kyoto and Osaka, Japan, 2010 to 2012: intensified surveillance after identification of the first strain (H041) with high-level ceftriaxone resistance. *Antimicrob Agents Chemother* 2013; **57**:5225–32.
4. Unemo M, Golparian D, Nicholas R *et al*. High-level cefixime- and ceftriaxone-resistant *Neisseria gonorrhoeae* in France: novel *penA* mosaic allele in a successful international clone causes treatment failure. *Antimicrob Agents Chemother* 2012; **56**:1273–80.
5. Ko KKK, Chio MT, Goh SS *et al*. First case of ceftriaxone-resistant multidrug-resistant *Neisseria gonorrhoeae* in Singapore. *Antimicrob Agents Chemother* 2019; **63**: e02624-18.
6. Eyre DW, Town K, Street T *et al*. Detection in the United Kingdom of the *Neisseria gonorrhoeae* FC428 clone, with ceftriaxone resistance and intermediate resistance to azithromycin, October to December 2018. *Euro Surveill* 2019; **24**: 1900147.
7. Kenyon C, Laumen J, Van Den Bossche D *et al*. Where have all the susceptible gonococci gone? A historical review of changes in MIC distribution over the past 75 years. *BMC Infect Dis* 2019; **19**: 1085.
8. Unemo M, Golparian D, Eyre DW. Antimicrobial resistance in *Neisseria gonorrhoeae* and treatment of gonorrhea. *Methods Mol Biol* 2019; **1997**: 37–58.
9. Palace SG, Wang Y, Rubin DHF *et al*. RNA polymerase mutations cause cephalosporin resistance in clinical *Neisseria gonorrhoeae* isolates. *eLife* 2020; **9**: e51407.
10. Zhao S, Duncan M, Tomberg J *et al*. Genetics of chromosomally mediated intermediate resistance to ceftriaxone and cefixime in *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother* 2009; **53**: 3744–51.
11. Tomberg J, Fedarovich A, Vincent LR *et al*. Alanine-501 mutations in penicillin-binding protein 2 from *Neisseria gonorrhoeae*: structure, mechanism, and effects on cephalosporin resistance and biological fitness. *Biochemistry* 2017; **56**: 1140–50.
12. Osnes MN, Didelot X, de Korne-Elenbaas J *et al*. Sudden emergence of a *Neisseria gonorrhoeae* clade with reduced susceptibility to extended-spectrum cephalosporins, Norway. *Microb Genomics* 2020; **6**: e000480.
13. De Laat MM, Wind CM, Bruisten SM *et al*. Ceftriaxone reduced susceptible *Neisseria gonorrhoeae* in the Netherlands, 2009 to 2017: from *penA* mosaicism to A501T/V nonmosaicism. *Sex Transm Dis* 2019; **46**: 594–601.
14. Chen S, Zhou Y, Chen Y *et al*. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018; **34**: i884–90.
15. Li. H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013; arXiv:1303.3997v2.
16. Li H, Handsaker B, Wysoker A *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–9.
17. Souvorov A, Agarwala R, Lipman DJ. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol* 2018; **19**: 153.
18. Gurevich A, Saveliev V, Vyahhi N *et al*. QUAST: quality assessment tool for genome assemblies. *Bioinforma Oxf Engl* 2013; **29**: 1072–5.
19. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol* 2019; **20**: 257.
20. Croucher NJ, Page AJ, Connor TR *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015; **43**: e15.
21. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014; **30**: 1312–3.
22. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 2016; **44**: 242–5.

23. Tonkin-Hill G, Lees JA, Bentley SD *et al*. RhierBAPS: an R implementation of the population clustering algorithm hierBAPS. *Wellcome Open Res* 2018; **3**: 93.
24. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 2018; **3**: 124.
25. Demczuk W, Sidhu S, Unemo M *et al*. *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance, a novel antimicrobial resistance multilocus typing scheme for tracking global dissemination of *N. gonorrhoeae* strains. *J Clin Microbiol* 2017; **55**: 1454–68.
26. Ma KC, Mortimer TD, Duckett MA *et al*. Increased power from conditional bacterial genome-wide association identifies macrolide resistance mutations in *Neisseria gonorrhoeae*. *Nat Commun* 2020; **11**: 5374.
27. Pham CD, Nash E, Liu H *et al*. Atypical mutation in *Neisseria gonorrhoeae* 23S rRNA associated with high-level azithromycin resistance. *Antimicrob Agents Chemother* 2020; **65**: e00885-20.
28. Wadsworth CB, Arnold BJ, Sater MRA *et al*. Azithromycin resistance through interspecific acquisition of an epistasis-dependent efflux pump component and transcriptional regulator in *Neisseria gonorrhoeae*. *mBio* 2018; **9**: e01419-18.
29. Laumen JGE, Manoharan-Basil SS, Verhoeven E *et al*. Molecular pathways to high-level azithromycin resistance in *Neisseria gonorrhoeae*. *bioRxiv* 2020; doi.org/10.1101/2020.12.02.409193.
30. Hunt M, Mather AE, Sánchez-Busó L *et al*. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb Genomics* 2017; **3**: e000131.
31. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* 2012; **28**: 2520–2.
32. Wick RR, Judd LM, Gorrie CL *et al*. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genomics* 2017; **3**: e000132.
33. Peng J-P, Yin Y-P, Chen S-C *et al*. A whole-genome sequencing analysis of *Neisseria gonorrhoeae* isolates in China: an observational study. *E Clinical Medicine* 2019; **7**: 47–54.
34. Lee S-G, Lee H, Jeong SH *et al*. Various *penA* mutations together with *mtrR*, *porB* and *ponA* mutations in *Neisseria gonorrhoeae* isolates with reduced susceptibility to cefixime or ceftriaxone. *J Antimicrob Chemother* 2010; **65**: 669–75.
35. Abrams AJ, Kirkcaldy RD, Pettus K *et al*. A case of decreased susceptibility to ceftriaxone in *Neisseria gonorrhoeae* in the absence of a mosaic penicillin-binding protein 2 (*penA*) allele. *Sex Transm Dis* 2017; **44**: 492–4.
36. Wind CM, Bruisten SM, Schim van der Loeff MF *et al*. a case-control study of molecular epidemiology in relation to azithromycin resistance in *Neisseria gonorrhoeae* isolates collected in Amsterdam, the Netherlands, between 2008 and 2015. *Antimicrob Agents Chemother* 2017; **61**: e02374-16.
37. Sánchez-Busó L, Golparian D, Corander J *et al*. The impact of antimicrobials on gonococcal evolution. *Nat Microbiol* 2019; **4**: 1941–50.
38. Fingerhuth SM, Bonhoeffer S, Low N *et al*. Antibiotic-resistant *Neisseria gonorrhoeae* spread faster with more treatment, not more sexual partners. *PLoS Pathog* 2016; **12**: e1005611.

SUPPLEMENTARY MATERIAL

TABLE S1. METADATA OF ALL ISOLATES INCLUDED IN THIS STUDY: UNIQUE IDENTIFIERS, SEQUENCE DATA QUALITY PARAMETERS, PATIENT DATA, PHENOTYPIC DATA AND TYPING RESULTS.

This table (55 columns and 318 rows) is available with the online version of this article (https://tinyurl.com/458jfddx).

TABLE S2. SIGNIFICANCE OF ASSOCIATIONS BETWEEN LINEAGE A (N=216 ISOLATES) AND LINEAGE B (N=102 ISOLATES) AND PATIENT CHARACTERISTICS, TESTED WITH CHI-SQUARE TESTS.

		Prevalence in lineage A	Prevalence in lineage B	p-value
Age[a]	<24	12%	33%	
	24-34	55%	42%	
	≥35	33%	25%	**2.74E-05**
Sex	Female	3%	22%	
	Male	97%	78%	**1.53E-07**
Sexual preference	Exclusively heterosexual	5%	34%	
	MSM or bisexual	95%	66%	**6.57E-12**
Year of isolation	2014-2016	34%	53%	
	2017-2019	66%	47%	**0.002**
AZM[b]	Susceptible/Intermediate	94%	99%	
	Resistant	6%	1%	**0.04**
CIP	Susceptible/Intermediate	38%	61%	
	Resistant	62%	39%	**0.0003**
CRO	Susceptible/Intermediate	65%	95%	
	Reduced susceptible	35%	5%	**2.38E-08**

[a]*Groups defined for each characteristic as follows: Age: 1) <24, 2) 24-34 and 3) ≥35 years; Sexual preference: 1) exclusively reported heterosexual intercourse and 2) reported bi- or homosexual intercourse; Year of isolation: 1)2014-2016 and 2) 2017-2019; Phenotype: 1) Isolates with MIC assigned as susceptible or intermediate susceptible 2) Isolates with MICs assigned as resistant or reduced susceptible (in case of ceftriaxone).*

[b]*Fisher's exact test performed because one of the groups contained <5 isolates*

TABLE S4. METADATA OF ALL MLST 7827 ISOLATES COMPARED IN THIS STUDY: UNIQUE IDENTIFIERS, NUMBER OF SEQUENCE READS, YEAR OF ISOLATION, PHENOTYPES AND TYPING RESULTS.

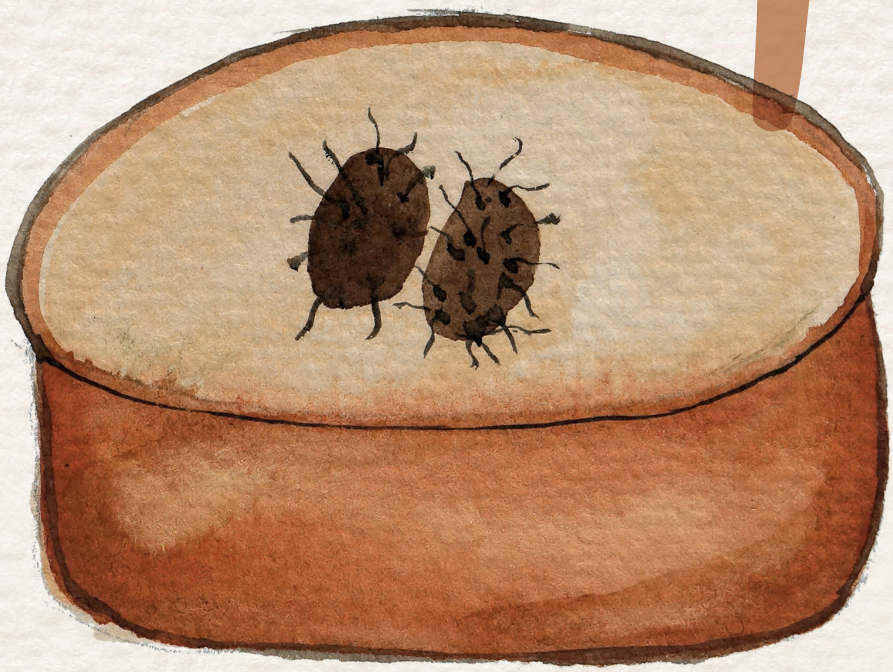
This table (17 columns and 225 rows) is available with the online version of this article (https://tinyurl.com/458jfddx).

TABLE S3. SIGNIFICANCE OF ASSOCIATIONS BETWEEN MAIN MLST CLUSTERS AND PATIENT- OR STRAIN CHARACTERISTICS, TESTED WITH FISHER'S EXACT TESTS WITH BONFERRONI CORRECTION FOR MULTIPLE TESTING.

		Year of isolation[a]		**Sex**		**Age**		**Sexual preference**		**AZM**		**CIP**		**CRO**	
		2017-2019		Male		≤30 years		MSM or bisexual		Resistant		Resistant		Reduced susceptibility	
MLST cluster	**Isolates in cluster**	Prevalence	p-value	Prevalence	p-value	Prevalence	p-value	Prevalence	p-value	Prevalence	p-value	Prevalence	p-value	Prevalence	p-value
1583	13	**100%**	**0.039**	100%	1	46%	1	92%	1	0%	1	**100%**	**0.006**	0%	0.104
1588	5	80%	1	80%	1	80%	1	60%	1	0%	1	80%	1	0%	1
1599	17	59%	1	100%	1	47%	1	100%	1	0%	1	**0%**	**1.56E-05**	**0%**	**0.019**
1901	19	**21%**	**0.009**	95%	1	79%	0.88	89%	1	5%	1	**100%**	**1.87E-04**	**79%**	**5.13E-09**
7363	17	**12%**	**6.79E-04**	71%	0.198	35%	1	71%	1	6%	1	**94%**	**0.011**	**24%**	**4.55E-05**
7822	12	75%	1	92%	1	58%	1	92%	1	0%	1	92%	0.14	0%	1
7827	63	**79%**	**0.009**	98%	0.421	43%	1	94%	0.292	0%	1	**100%**	**9.7E-19**	**89%**	**2.81E-30**
8135	5	100%	1	40%	0.099	100%	1	**0%**	**5.83E-04**	0%	1	0%	0.335	0%	1
8143	8	38%	1	100%	1	38%	1	100%	1	0%	1	100%	0.161	0%	0.1
8156	28	64%	1	96%	1	39%	1	100%	1	0%	1	**0%**	**9.4E-10**	**0%**	**1.59E-04**
8163	7	57%	1	57%	0.308	100%	0.286	**0%**	**7.94E-06**	0%	1	0%	0.066	0%	1
9363	18	61%	1	100%	1	33%	1	100%	1	**50%**	**0.0042**	50%	1	**0%**	**0.011**
10314	9	100%	0.229	100%	1	33%	1	100%	1	0%	1	100%	0.083	0%	1
11428	12	33%	1	100%	1	50%	1	100%	1	0%	1	**0%**	**0.001**	0%	0.182
11864	11	55%	1	100%	1	45%	0.072	100%	1	0%	1	**0%**	**0.002**	0%	0.331
11990	9	33%	1	**22%**	**1.14E-05**	100%	1	**0%**	**9.74E-08**	0%	1	**0%**	**0.013**	0%	0.572
13292	5	80%	1	100%	1	20%	1	100%	1	0%	1	0%	0.335	0%	1
<5 isolates[b]	59	53%	1	92%	1	64%	1	86%	1	93%	1	41%	0.367	**8%**	**0.004**
Total	**317[c,d]**														

[a]Groups defined for each characteristic as follows: Year of isolation: 1)2014-2016 and 2) 2017-2019; Age:1) ≤30 years and 2) >30 years; Sexual preference: 1) exclusively reported heterosexual intercourse and 2) reported bi- or homosexual intercourse; Phenotype: 1) Isolates with MIC assigned as susceptible or intermediate susceptible 2) Isolates with MICs assigned as resistant or reduced susceptible (in case of ceftriaxone). [b]MLST clusters which included <5 isolates were taken together. [c]One strain could not be typed according to the MLST typing scheme since one locus was only partially assembled. [d]Duplicate patients were added only once when tested for associations with patient characteristics (n=314 patients)

4



INCREASED CLONALITY AMONG *NEISSERIA GONORRHOEAE* ISOLATES DURING THE COVID-19 PANDEMIC IN AMSTERDAM, THE NETHERLANDS

Hélène C.A. Zondag*, Jolinda de Korne – Elenbaas*,
Sylvia M. Bruisten, Henry J.C. de Vries, Alje P. van Dam
**equal contribution*

Submitted for publication

ABSTRACT

Distancing measures during the COVID-19 lockdown have led to a temporary decrease of casual sex partners among clients of the Centre for Sexual Health (CSH) of Amsterdam, the Netherlands. We investigated the effect of this change on the genotypic and phenotypic distribution of *Neisseria gonorrhoeae* isolates from CSH patients. From each *N. gonorrhoeae*-positive patient we sequenced one isolate, resulting in 322 isolates which constituted two groups: 181 isolates cultured from January 15th–February 29th 2020 (before the first lockdown) and 141 cultured from May 15th–June 30th 2020 (during the first lockdown). Patient characteristics showed significantly more symptomatic patients and significantly less reported sex partners during the lockdown. Phenotypic data showed an increase in low-level azithromycin resistance and ceftriaxone susceptibility during the lockdown, and this remained after the study period. The isolate groups had similar variety of multi-locus sequence types (MLSTs). However, a shift occurred from MLST 8156 being predominant before lockdown to MLST 9362 during lockdown. SNP distances between isolates with identical MLSTs were significantly lower during lockdown, specifically due to remarkably low SNP distances between MLST 9362 isolates. These findings reflect the change in sexual behaviour of CSH clients during the lockdown, with a potential increased local transmission of the MLST 9362 strain during this period, which led to genotypic and phenotypic changes in the *N. gonorrhoeae* population. This shows that public health measures have far-reaching consequences and should be considered in the surveillance of other infectious diseases.

INTRODUCTION

The COVID-19 pandemic has had an extraordinary global impact on public health[1]. Also, the epidemiology of sexually transmitted infections (STIs) has been impacted by the COVID-19 pandemic[2]. Gonorrhoea, caused by the sexually transmitted pathogen *Neisseria gonorrhoeae,* is one of the most prevalent bacterial STIs worldwide and its prevalence is globally on the rise[3]. This rising trend has also been reported in Europe, including the Netherlands, where before the pandemic the prevalence of gonorrhoea and other bacterial STIs increased, especially among men who have sex with men (MSM)[4-6]. However, interpreting the number of STI cases during the COVID-19 pandemic has been challenging, since social measures and large scale COVID-19 testing temporarily restricted STI testing at Centres for Sexual Health (CSH)[4,7].

In the Netherlands, efforts to reduce and contain the transmission of SARS-CoV-2 started with gradual measures since the first notified infection on February 27th, 2020. The first lockdown with additional measures was put into place on the 16th of March 2020, from then all events were cancelled, gatherings of 3 and more persons were discouraged and all stores (except grocery stores), schools and leisure areas were closed. In general, 1.5 meters of physical distance was enforced both inside and outside, and no more than 3 adult visitors were allowed in a household per day. International travel was restricted to essential travel only and also nationally, people were encouraged to stay at home. From mid-May onwards, schools, stores, museums, bars and restaurants were gradually reopened with restrictions. However, during this time, any form of travel and especially international travel remained highly reduced[8]. The healthcare at the CSH in Amsterdam was also impacted by the restrictions of the lockdown. Between March 23rd and June 1st routine HIV/STI testing was halted for asymptomatic clients unless urgent PrEP prescription was needed[9].

The perceived risk of SARS-CoV-2 infection and public health measures were found to be associated with a change in sexual behaviour[10,11]. A recent study from Bilsen et al. (2021) found that distancing measures during lockdown led to a temporary decrease of casual sex partners and a relatively low STI positivity rate among clients of the CSH of Amsterdam[9]. Besides the influence of this behavioural change on number of STI cases, it could also have had an influence on STI epidemiology due to changes in transmission networks. In this study, we aimed to investigate the influence of the change in sexual behaviour of CSH clients during the first lockdown in Amsterdam on the genotypic and phenotypic distribution of *N. gonorrhoeae* isolates from CSH patients.

METHODS

ISOLATE SELECTION

Two study periods of equal duration were selected before and during the lockdown based on similarity of patient groups regarding sex and sexual orientation. The study periods were also chosen based on the number of *N. gonorrhoeae* isolates obtained during each period, because during the first 8 weeks of the lockdown, access to the CSH was highly restricted which subsequently strongly reduced the number of gonorrhoea cases. *N. gonorrhoeae* isolates from all culture-positive patients who visited the CSH between January 15th and February 29th and between May 15th and June 30th in 2020 were included. The period January 15th - February 29th 2020 will be referred to as "before lockdown" and the period May 15th - June 30th 2020 will be referred to as "during lockdown" throughout this study. Patients who had *N. gonorrhoeae* both before and during lockdown were included in both groups, since these can be considered as two separate infection episodes. From each *N. gonorrhoeae*-positive patient, one isolate was selected with prioritisation of the rectal isolate in case multiple isolates were obtained from a single patient. To prevent SARS-CoV-2 transmission at the CSH, the sampling policy no longer included pharyngeal sampling during the lockdown. Therefore, pharyngeal *N. gonorrhoeae* isolates obtained before the lockdown were not included in this study.

PHENOTYPIC ANALYSIS

At the Public Health Laboratory, minimum inhibitory concentrations (MICs) for azithromycin and ceftriaxone are routinely determined for all *N. gonorrhoeae* isolates obtained from CSH patients, using E-tests according to manufacturer's instruction (bioMerieux SA). The clinical breakpoint used in this study for ceftriaxone resistance was MIC >0.125 mg/L and for azithromycin MIC ≥1 mg/L (ECOFF), according to EUCAST guidelines v12.

Phenotypes from anogenital isolates obtained before and after the study period might indicate whether phenotypic trends observed during lockdown were already seen before the lockdown in 2019 and still hold in the months afterwards. Therefore, azithromycin and ceftriaxone MIC distributions of isolates during the study period were plotted next to phenotypic distributions of anogenital isolates obtained 2.5 months before (15 August – 30 September 2019) and 2.5 months after the study period (15 September – 1 November 2020). The interval of 2.5 months was chosen because the study periods before- and during lockdown also had an interval of 2.5 months. Pharyngeal isolates were excluded from the phenotypic data because pharyngeal isolates were not included during the study period.

ISOLATE PREPARATION AND WGS SEQUENCING

All selected samples were taken from the -80°C freezer and grown overnight on chocolate blood agar plates. DNA was extracted from pure cultures in DNA/RNA shield buffer using the

ZymoBIOMICS TM Magbead DNA kit (ZYMO RESEARCH). The Nextera XT DNA Library Preparation kit with Integrated DNA Technologies for Illumina DNA/RNA Unique Dual Indexes (Illumina) was used for library preparation. WGS was performed by 150 base-pairs (bp) paired-end sequencing using Illumina NovaSeq 6000 at the sequencing facility BaseClear B.V. (L457; NEN-EN-ISO/IEC 17025).

QUALITY CONTROL AND ASSEMBLY

Raw reads were filtered and trimmed using fastp v0.20.1[12]. The mean coverage depth and the percentage of covered reference genome bases were calculated with the SAMtools package v1.15 by first mapping the reads to reference genome FA1090 (NC_002946.1) with BWA-MEM2 v2.1[13]. Genomes were assembled using the 'isolate' option in SPAdes v.3.15.3[14]. QUAST v5.0.2 was used to assess the quality of the assembly and to identify assemblies with aberrant assembly lengths or GC content, which could be indications of read contamination[15]. For these isolates, reads not belonging to *N. gonorrhoeae* (taxid:485) were identified and filtered out with Kraken2 v2.1.1[16]. Filtered reads were assembled again.

TYPING AND DETERMINATION OF SNP DISTANCES

All assemblies were annotated automatically by uploading them to the PubMLST Neisseria database. Multi-locus sequence types (MLST) were extracted and novel MLSTs were assigned novel STs[17]. Snippy v.4.6.0 (https://github.com/tseemann/snippy) was used for variant calling, by first mapping reads to reference genome FA1090 followed by identifying SNPs between the isolate and reference genome. Only SNPs with a minimum base quality of 13, a minimum read coverage of 10x and a read concordance of 90% were reported (default settings). The 'core' option of Snippy was used to create a core genome alignment of all isolates. Snp-dists v.0.7.0 (https://github.com/tseemann/snp-dists) was used to determine SNP distances between all possible combinations of isolates. For maximum resolution, recombination was kept in the SNP distance analyses to determine the genetic diversity between isolates. Recombination filtering was applied using Gubbins v2.4.1 to identify isolate pairs that differed <10 recombination filtered SNPs, a threshold previously defined for identical isolates that are putatively transmitted between individuals[18,19]. The bioinformatic pipeline was managed using Snakemake v7.2.1[20].

STATISTICAL ANALYSES

Sociodemographic and clinical data were extracted from electronic patient files. Chi square or Fisher's exact tests were used to compare participant characteristics from before and during lockdown and to analyse its associations with the most prevalent MLST during the lockdown. A p-value ≤0.05 was deemed significant. SNP distances between groups were compared with the Wilcoxon rank test, for which a p-value ≤0.001 was deemed significant. Data was analysed using RStudio (version 1.2.5033).

ETHICS STATEMENT

All clients of the CSH in Amsterdam were informed of the "opt-out" system regarding research on remnants of patient material. All data were pseudonymised before analysis.

RESULTS

PATIENT CHARACTERISTICS

A total of 322 *N. gonorrhoeae* isolates were included in the study, consisting of 181 isolates before lockdown and 141 during lockdown. Demographic characteristics of the patient groups before (181 patients) and during (141 patients) lockdown were comparable regarding sex and sexual orientation, as expected since study periods were chosen based on similarity of these characteristics. Five patients were *N. gonorrhoeae*-positive both before and during lockdown and therefore included in both groups. The patient groups significantly differed in HIV status (p=0.007), number of sex partners (p=0.009) and symptomatology (p=0.006) (Table 1). A higher proportion of patients had an unknown HIV status in the group during lockdown. During the lockdown, patients reported less sex partners in the past 6 months and a higher proportion of patients was symptomatic, compared to before the lockdown.

DISTRIBUTION OF MLSTS CHANGED DURING LOCKDOWN

All 322 isolates were typed according to the MLST scheme. The overall diversity in MLSTs was similar before (34 MLSTs) to during (30 MLSTs) lockdown (Figure 1). The majority of the types occurred less than 5 times, which was the case in 22/34 (65%) MLSTs before lockdown and 21/30 (70%) during lockdown. Before lockdown, a variety of MLSTs was prevalent among the population, with MLST 8156 predominantly occurring in 22/181 (12%) of the isolates. During lockdown, MLST 8156 prevalence was reduced to 13/141 (9%), whereas MLST 9362 became predominant. MLST 9362 went from being present in 3/181 (2%) of the isolates before the lockdown to 29/141 (21%) during the lockdown. MLST 9362 was significantly associated with MSM (Table S2). A decline was also found for MLST 7827, which went from being present in 10/181 (6%) isolates before the lockdown to 3/141 (2%) isolates during the lockdown. The low prevalent MLSTs 1587, 1599, 7363, 9363, 7827 and 10314 were no longer found >5 times during the lockdown, whereas types 7359, 9362 and 15183 occurred <5 times before the lockdown and their prevalence increased during lockdown.

TABLE 1. CLINICAL AND DEMOGRAPHIC DATA OF PATIENTS BEFORE (JANUARY 15TH – FEBRUARY 29TH) AND DURING (MAY 15TH – JUNE 30TH) THE LOCKDOWN IN 2020.

	Before lockdown N = 181 (%)	**During lockdown N = 141 (%)**	**OR/ P-value**
Sex			0.966
Male	171 (94)	133 (94)	
Female	10 (6)	8 (6)	
Age			0.421
<25	34 (19)	36 (26)	
25-34	79 (44)	62 (44)	
35-44	47 (26)	30 (21)	
≥45	21 (12)	13 (9)	
Anatomical location			0.850[a]
Cervix	3 (2)	3 (2)	
Anus	117 (65)	87 (62)	
Urethra	54 (30)	47 (33)	
Vagina	7 (4)	4 (3)	
Country of origin			0.697[a]
The Netherlands	93 (51)	65 (46)	
Suriname + Dutch Antilles	17 (9) 28 (15)	19 (13) 28 (20)	
Europe + Turkey	20 (11)	11 (8)	
M-S America	14 (8)	10 (7)	
Asia	7 (4)	5 (4)	
Africa	2 (1)	2 (1)	
Other	0 (0)	2 (1)	
Unknown			
Sexual orientation			0.707[a]
MSM	158 (86)	122 (87)	
MSW	11 (6)	11 (8)	
Female	10 (6)	8 (6)	
Transgender	2 (1)	0 (0)	
HIV status			0.007*
Negative	136 (75)	105 (75)	
Positive	39 (22)	20 (14)	
Unknown	6 (3)	16 (11)	
Sex work			
Yes	15 (8)	11 (8)	0.859
No	166 (92)	130 (92)	
Number of sex partners in last 6 months			0.009*
0-1	6 (3)	10 (7)	
2-4	38 (21)	47 (33)	
5-9	46 (25)	37 (26)	
10-19	32 (18)	24 (17)	
20-49	43 (24)	16 (11)	
≥50	15 (8)	7 (5)	
Symptomaticity			0.006*
Symptomatic	48 (27)	58 (41)	
Asymptomatic	133 (73)	84 (59)	

[a]*Fisher's Exact Test | * p ≤0.05*

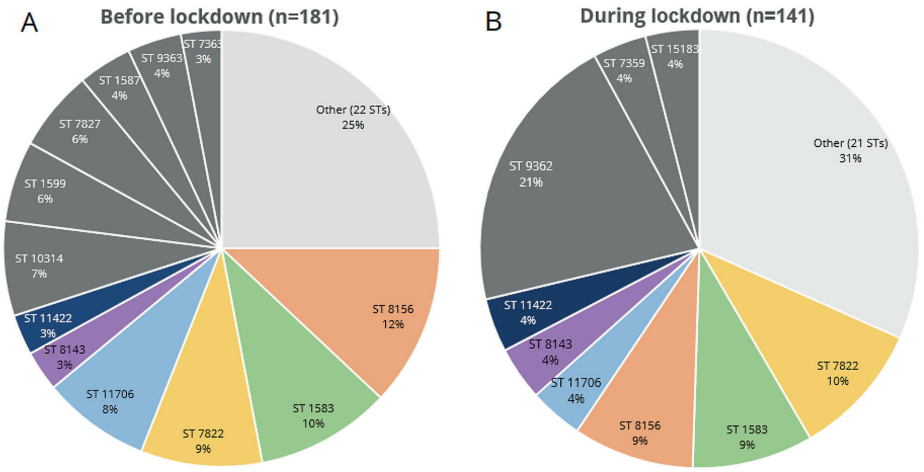


FIGURE 1. MLST DISTRIBUTION BEFORE (A) AND DURING (B) LOCKDOWN. PREVALENCE ARE SHOWN FOR MLSTS THAT OCCURRED >5 TIMES IN THE ISOLATES BEFORE- OR DURING LOCKDOWN. MLSTS THAT OCCURRED <5 TIMES WERE CATEGORISED AS 'OTHER'. COLOURED MLSTS WERE FOUND >5 TIMES BEFORE- AND DURING THE LOCKDOWN, WHEREAS GREY MLSTS ARE ONLY FOUND >5 TIMES IN ONE OF THE PERIODS.

SNP DISTANCES BETWEEN ISOLATES WITH IDENTICAL MLSTS SIGNIFICANTLY DECREASED DURING LOCKDOWN

Pairwise SNP distances were determined between all isolates obtained before- and during lockdown. The SNP distances between isolates with different MLSTs significantly increased during lockdown, with median SNP distances of 5289 before lockdown and 5603 during lockdown (Figure 2A). In contrast, SNP distances significantly decreased between isolates with identical MLSTs during lockdown, with median SNP distances of 503 before lockdown and 98 during lockdown. When further categorizing on MLST, remarkable low SNP distances were found between MLST 9362 isolates obtained during the lockdown (median 17 SNPs), despite the high number of MLST 9362 isolates found in that period (Figure 2B and 2C). Low SNP distances were also found between MLST 15183 isolates (median 29 SNPs), however, this MLST was found only 6 times during the lockdown. Remarkably, diverse SNP distances were found between isolates with MLSTs 7363, 7822, 7827 and 8143 (Figure 2C). Using recombination filtered SNP distances, 1% (160/16,290) of the compared isolate pairs before lockdown contained identical isolates differing with <10 recombination filtered SNPs. This strikingly increased to 3.5% (353/9,870) during lockdown with the vast majority (84%, 295/353) of these isolate pairs belonging to MLST 9362.

ASSOCIATION BETWEEN GENOTYPE AND PHENOTYPE

Isolates belonging to the same MLST mainly have similar MICs for azithromycin and ceftriaxone (Figure 3A, 3B, 3C), showing an association between genotype and phenotype. Azithromycin resistance (MIC ≥1 mg/L) was predominantly found in isolates belonging to the MLSTs 9362,

9363 and 11422. Mosaicism in the *mtrR* gene and/or the mtrCDE operon were found to be the resistance determinants in these three MLSTs (Figure 3D). Azithromycin resistance-associated 23S rRNA mutations were not found. The highest ceftriaxone MICs were found for MLST 7827 isolates, which predominantly carried the *penA* A501V mutation. Isolates belonging to MLSTs 8156 and 9362 were highly susceptible to ceftriaxone. No ceftriaxone resistant isolates were identified in this study.

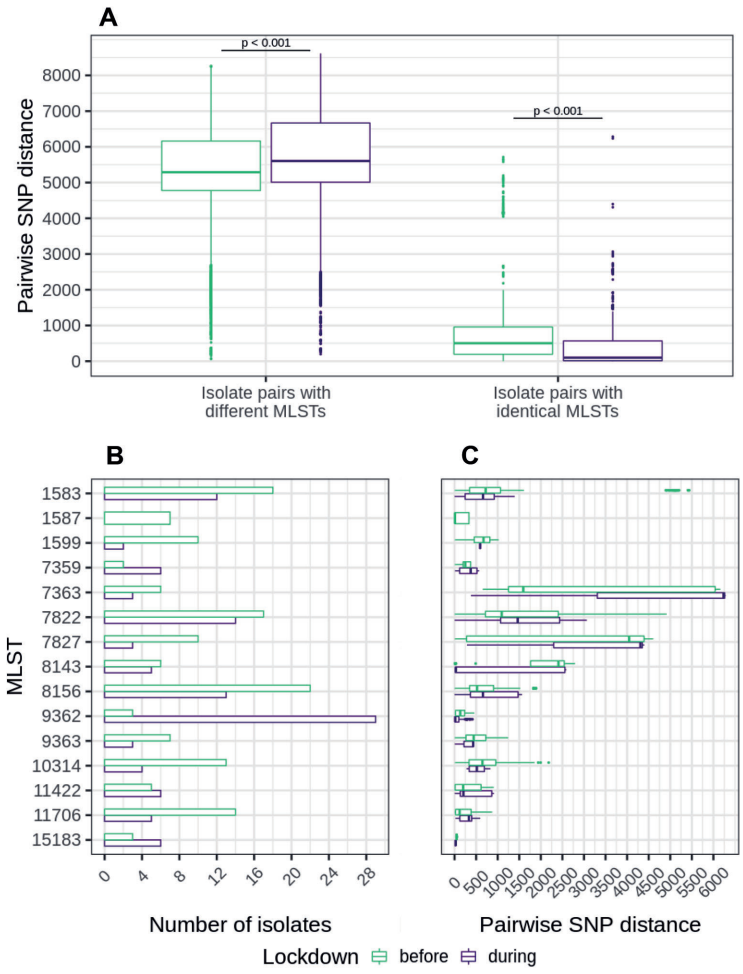


FIGURE 2. PAIRWISE SNP DISTANCES BETWEEN ISOLATES OBTAINED BEFORE- OR DURING LOCKDOWN. A. BOXPLOTS OF PAIRWISE SNP DISTANCES BETWEEN ISOLATES WITH DIFFERENT MLSTS (15,389 PAIRWISE COMPARISONS BEFORE LOCKDOWN, 9,128 DURING LOCKDOWN) OR IDENTICAL MLSTS (901 PAIRWISE COMPARISONS BEFORE LOCKDOWN, 742 DURING LOCKDOWN). **B**. NUMBER OF ISOLATES PER MLST BEFORE- OR DURING THE LOCKDOWN. MLSTS OCCURRING <5 TIMES WERE NOT SHOWN. **C**. BOXPLOTS OF PAIRWISE SNP DISTANCES BETWEEN ISOLATES WITHIN MLST GROUPS BEFORE OR DURING LOCKDOWN.

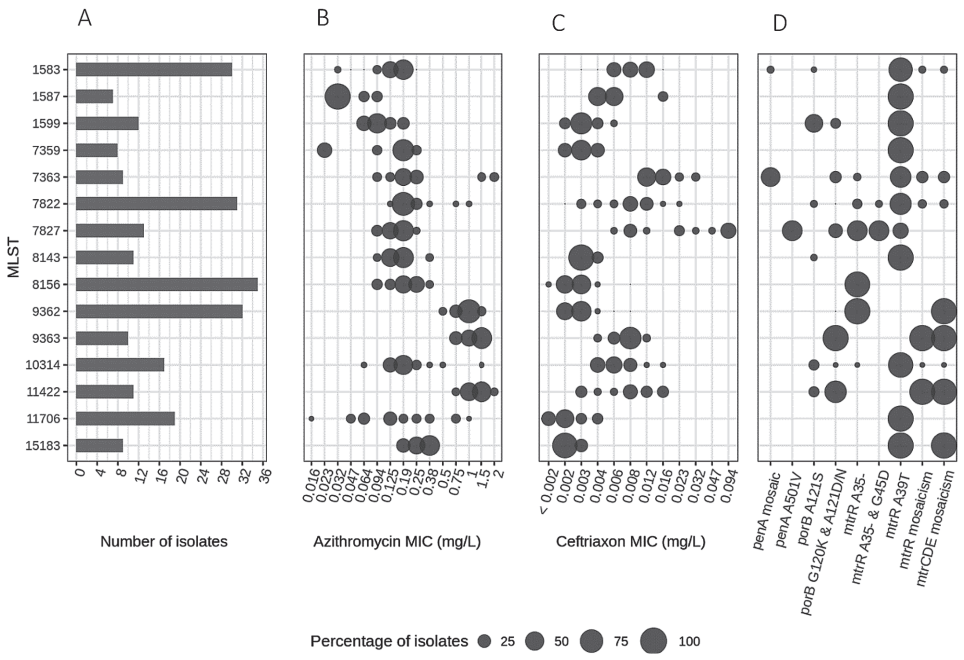


FIGURE 3. DISTRIBUTION OF PHENOTYPES AND RESISTANCE DETERMINANTS AMONG MLSTS. A. TOTAL NUMBER OF ISOLATES PER MLST (NUMBERS BEFORE- AND DURING LOCKDOWN WERE TAKEN TOGETHER). B. PREVALENCE OF AZITHROMYCIN MICS PER MLST. C. PREVALENCE OF CEFTRIAXONE MICS PER MLST. D. PREVALENCE OF AZITHROMYCIN- AND CEFTRIAXONE RESISTANCE DETERMINANTS IN *PENA*, *PORB*, *MTRR* GENES AND THE MTRCDE OPERON PER MLST. ONLY MLSTS WITH >5 ISOLATES ARE SHOWN.

INCREASED AZITHROMYCIN RESISTANCE AND CEFTRIAXONE SUSCEPTIBILITY DURING THE LOCKDOWN

The proportion of isolates with low-level azithromycin resistance and ceftriaxone susceptibility increased during the lockdown (Figure 4A and B). Phenotypes from isolates obtained before and after the study period demonstrate whether trends in phenotype observed during lockdown were already seen before the first lockdown or still hold in the months afterwards, giving an indication about the associated genotypic distribution. For this purpose, we compared MIC distribution of isolates before and during the lockdown with MIC data regarding 163 anogenital isolates cultured 2.5 months before our study period and 167 anogenital isolates cultured 2.5 months after our study period. The azithromycin MIC distribution showed that isolates with high-level azithromycin resistance (MIC >4 mg/L) already disappeared before the study period (Figure 4A). However, the proportion of low-level resistant isolates (MIC 1-4 mg/L) strongly increased during the lockdown and the high prevalence remained 2.5 months after the study period. An opposite trend was observed for ceftriaxone MICs, with an increase in ceftriaxone susceptible isolates and a decrease in intermediate susceptible (MIC 0.012-0.064 mg/L) and reduced susceptible (MIC >0.064 mg/L) isolates over time. Highly susceptible isolates were

predominant during lockdown. Reduced susceptible isolates were no longer found after the lockdown (Figure 4B).

DISCUSSION

In this study we investigated whether changed sexual behaviour of CSH clients during the first lockdown in Amsterdam, the Netherlands, had led to changes in genotypic and phenotypic distribution within the *N. gonorrhoeae* population. Although the MLST diversity did not change during the lockdown, a change was observed in MLST distribution with the largest shift for MLST 9362, which went from a prevalence of 2% before- to 21% during the lockdown and was exclusively present among MSM. When comparing the MLST distribution found in our study to the one found in 2018 in Amsterdam as part of the European Gonococcal Antimicrobial

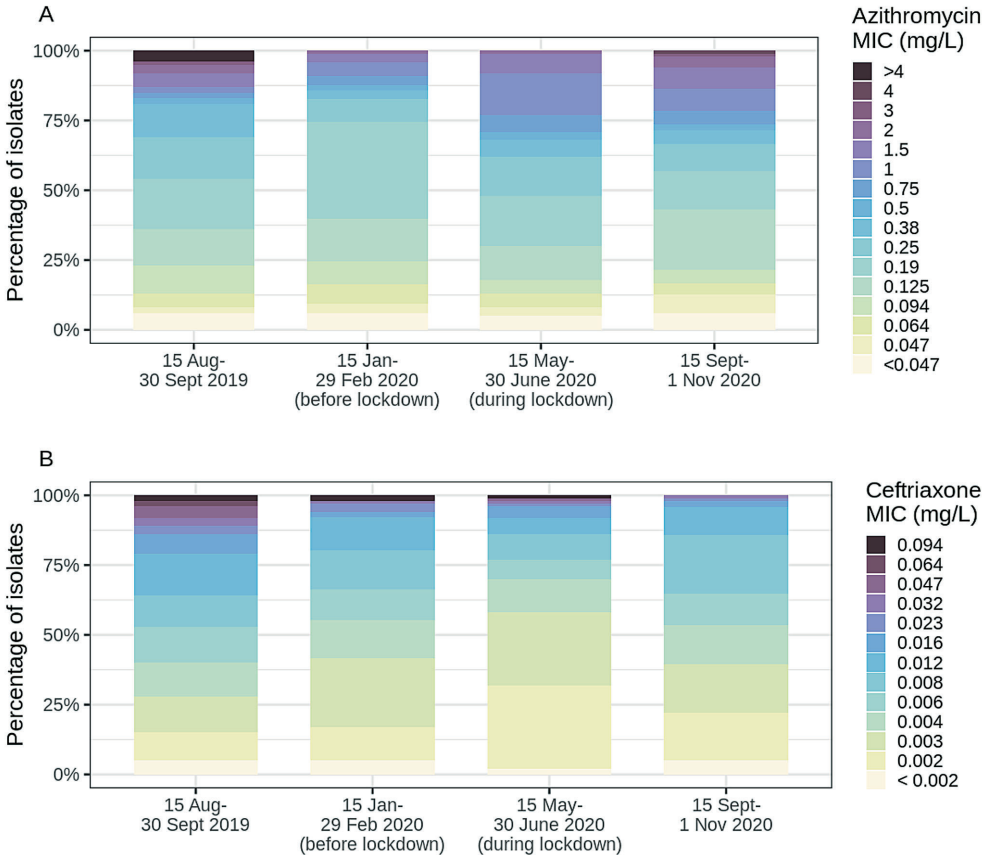


FIGURE 4. DISTRIBUTION OF AZITHROMYCIN (A) AND CEFTRIAXONE (B) MICS BEFORE, DURING AND AFTER THE STUDY PERIODS WITH INTERVALS OF 2.5 MONTHS.

Surveillance Programme (Euro-GASP), the MLST distribution before lockdown was much more similar to the one in 2018 than the MLST distribution found during the lockdown (Figure S1)[21]. Most MLSTs that were prevalent in the period before the lockdown were already prevalent in 2018 and no isolates from the Netherlands in the 2018 collection belonged to MLST 9362. Therefore, the rapid emergence of MLST 9362 does not just reflect a natural trend in MLST patterns but is most likely an impact of the COVID-19 related measures. In addition, the limited genetic diversity found among the MLST 9362 isolates during the lockdown also suggested extensive local transmission of this strain during the lockdown period. This could be explained by the significantly reduced number of casual sex partners during the lockdown, as described by Van Bilsen et al[9], and the restricted national- and international travel, which logically reduced introduction of *N. gonorrhoeae* isolates from sexual networks outside of Amsterdam and the Netherlands. The behavioural changes were only temporary[22], whereas phenotypic data showed that the prevalence of low-level azithromycin resistance stayed high after the study period (Figure 4A), suggesting that MLST 9362 remained prevalent. Additional genotypic data is needed to confirm whether easing of measures has led to a decrease of this MLST over time or whether this MLST is still predominant in Amsterdam.

MLST 9362 was associated with low-level azithromycin resistance and ceftriaxone susceptibility and its emergence therefore drove the observed shift in phenotypic data, with increased low-level resistance to azithromycin and increased susceptibility to ceftriaxone during the lockdown. This is further enhanced by the decline in MLSTs associated with azithromycin susceptibility during the lockdown, such as MLSTs 1587, 1599 and 10314 (Figure 2). The phenotypic shift towards increased low-level azithromycin resistance in *N. gonorrhoeae* that has been observed in the Netherlands over the years 2019 and 2020 has also been observed in other European countries and worldwide[21,23]. European genomic surveillance, which formed part of the Euro-GASP, found that this phenotypic change was caused by the emergence of MLSTs 9363 and 11422, both carrying mosaic *mtrR* and *mtrD* genes which have been associated with azithromycin resistance[24]. MLST 9363 became the predominant MLST in Europe in 2018[21]. Interestingly, in our study we found MLST 9362 being mainly responsible for the increased azithromycin MICs instead, which only carried the mosaic *mtrD* gene and not the *mtrR* mosaic gene. However, we might have missed MLST 9363 isolates because our study did not include pharyngeal swabs, and MLST 9363 was associated with pharyngeal *N. gonorrhoeae* infections in MSM[21]. This could have led to a bias in the MLST distribution that we observed in our study population, however it would only enforce the observed phenotypic shift towards increased azithromycin MICs.

The increased ceftriaxone susceptibility observed in this study was reflected by a decline in MLST 7827, which emerged among the Dutch *N. gonorrhoeae* population between 2017 and 2019, and was associated with ceftriaxone reduced susceptibility[25]. The trend towards increased ceftriaxone susceptibility observed in this study was also supported by European surveillance

data. The Euro-GASP showed a strong decline in MLST 1901 isolates carrying mosaic *penA* genes in 2018 compared to 2013[21]. MLST 1901 isolates with mosaic *penA* genes have globally been associated with ceftriaxone resistance[26-28], although the Netherlands and Norway reported MLST 7827 as a predominant ceftriaxone reduced susceptible strain[25,29]. Remarkably, on a global level the emergence of ceftriaxone reduced susceptible strains continued, with an increasing number of contributing countries reporting reduced susceptible strains from 2014 to 2018[23]. This can be explained by differences in AMR profiles across continents, which indicates that the threat of emerging resistance to cephalosporins still exists and underlines the need for continuous surveillance on global, international and national levels.

During the lockdown, more symptomatic patients were seen at the CSH, which may be have been a temporary effect from the prioritisation of symptomatic patients during the strict lockdown, due to the reduced capacity for STI testing. To our knowledge, no association has been found between having symptoms and *N. gonorrhoeae* genotype, thus we do not expect a major effect of this change in patient population on the results of this study. Although we cannot exclude that the composition of the patient population influenced the *N. gonorrhoeae* epidemiology, we tried to minimise this effect by selecting study periods based on number of isolates that were obtained at the CSH and the similarity of patient populations regarding sex and sexual orientation. Only few isolates were obtained in the first months of the lockdown, due to precautions taken at the CSH, therefore this period was not chosen. For these reasons, we selected May and June as period 'during lockdown' instead of the first months of the lockdown (second half of March and April), in which the number of obtained *N. gonorrhoeae* isolates was strongly reduced and the patient population was much less representative. Since we included isolates from all of the patients who visited the CSH during the study periods, we did not introduce any selection bias in our study.

This study did not include pharyngeal isolates because pharyngeal swabs were no longer taken during lockdown to prevent SARS-CoV-2 transmission. Treatment failures often occurred with pharyngeal *N. gonorrhoeae* infections[30]. Pharyngeal *N. gonorrhoeae* infections are also important because of the possibility for AMR development due to possible exchange of genetic material between *Neisseria* species at the oropharyngeal site[31]. As mentioned above, the Euro-GASP study indeed found an association between pharyngeal isolates from MSM and low-level resistance to azithromycin[21]. The lack of pharyngeal isolates in this study could therefore have led to missed isolates with elevated MICs.

In conclusion, a major change in genotypic and phenotypic distribution was identified during the first COVID-19 lockdown among *N. gonorrhoeae* isolates from CSH patients in Amsterdam. The phenotypic shift towards increased azithromycin resistance and increased ceftriaxone susceptibility is in line with European surveillance data, however, the lockdown might have

enhanced these trends as it has led to expansion of the MLST 9362 strain with low-level azithromycin resistance and high ceftriaxone susceptibility. The low SNP distances between MLST 9362 isolates suggested local transmission of this strain in Amsterdam, reflecting the previously identified change in sexual behaviour and subsequent more local STI transmission networks during the lockdown[9]. This shows that public health measures also have consequences for the epidemiology of other infectious diseases which should be taken in consideration for public health surveillance.

DATA SUMMARY

Raw sequence data are available in the European Nucleotide Archive under project number PRJEB55899 and individual accession numbers can be found in Table S1. Assemblies are available in the PubMLST Neisseria database (https://pubmlst.org/organisms/neisseria-spp). PubMLST identifiers and additional metadata can also be found in Table S1. The bioinformatic pipeline used in this study can be found on Github (https://github.com/jolindadekorne/Diversity-In-Neisseria-Gonorrhoeae-during-Lockdown).

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

FUNDING INFORMATION

This work was funded internally by the Public Health Laboratory of Amsterdam.

REFERENCES

1. Miranda MNS, Pingarilho M, Pimentel V, et al. A Tale of Three Recent Pandemics: Influenza, HIV and SARS-CoV-2. Front Microbiol. 2022;13:889643. doi:10.3389/fmicb.2022.889643
2. Yan J, Li Y, Zhou P. Impact of COVID-19 pandemic on the epidemiology of STDs in China: based on the GM(1,1) model. BMC Infect Dis. 2022;22(1):519.
3. Kirkcaldy RD, Weston E, Segurado AC, Hughes G. Epidemiology of gonorrhoea: a global perspective. Sex Health. 2019;16(5):401-11. doi:10.1186/s12879-022-07496-y
4. Geretti AM, Mardh O, de Vries HJC, et al. Sexual transmission of infections across Europe: appraising the present, scoping the future. Sex Transm Infect. 2022. doi:10.1136/sextrans-2022-055455
5. van Bilsen WPH, Boyd A, van der Loeff MFS, et al. Diverging trends in incidence of HIV versus other sexually transmitted infections in HIV-negative MSM in Amsterdam. AIDS. 2020;34(2):301-9. doi:10.1097/QAD.0000000000002417
6. Staritsky LE, Visser M, van Aar F, et al. Sexually Transmitted Infections in the Netherlands in 2020. Rijksinstituut voor Volksgezondheid en Milieu RIVM; 2021. doi:10.21945/RIVM-2021-0052
7. Crane MA, Popovic A, Stolbach AI, et al. Reporting of sexually transmitted infections during the COVID-19 pandemic. Sex Transm Infect. 2021;97(2):101-2. doi:10.1136/sextrans-2020-054805
8. Rijksoverheid. Coronavirus tijdlijn. Published online 2022; Accessed July 24th, 2022. https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn.
9. van Bilsen WPH, Zimmermann HML, Boyd A, et al. Sexual Behavior and Its Determinants During COVID-19 Restrictions Among Men Who Have Sex With Men in Amsterdam. J Acquir Immune Defic Syndr. 2021;86(3):288-96. doi:10.1097/QAI.0000000000002581
10. Storer D, Prestage G, McManus H, et al. Relationship Between Sexual Behaviors with Non-committed Relationship Partners and COVID-19 Restrictions and Notification Rates: Results from a Longitudinal Study of Gay and Bisexual Men in Australia. Sex Res Social Policy. 2022:1-12. doi:10.1007/s13178-022-00733-8
11. De Vries DC, Zimmermann HML, Druckler S, et al. Barriers and Motives for Complying With "Sexual Distancing" Among Men Who Have Sex With Men During the First COVID-19 Pandemic Lockdown in Amsterdam: A Qualitative Study. Sex Transm Dis. 2022;49(7):497-503. doi:10.1097/OLQ.0000000000001636
12. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i90. doi:10.1093/bioinformatics/bty560
13. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. doi:10.1093/bioinformatics/btp352
14. Prjibelski A, Antipov D, Meleshko D, et al. Using SPAdes De Novo Assembler. Curr Protoc Bioinformatics. 2020;70(1):e102. doi:10.1002/cpbi.102
15. Gurevich A, Saveliev V, Vyahhi N, et al. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5. doi:10.1093/bioinformatics/btt086
16. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20(1):257. doi:10.1186/s13059-019-1891-0
17. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res. 2018;3:124. doi:10.12688/wellcomeopenres.14826.1
18. Croucher NJ, Page AJ, Connor TR, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015;43(3):e15. doi:10.1093/nar/gku1196
19. de Korne-Elenbaas J, Bruisten SM, de Vries HJC, et al. Within-Host Genetic Variation in *Neisseria gonorrhoeae* over the Course of Infection. Microbiol Spectr. 2022;10(3):e0031322. doi:10.1128/spectrum.00313
20. Koster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. Bioinformatics. 2012;28(19):2520-2. doi:10.1093/bioinformatics/bts480

21. Sanchez-Buso L, Cole MJ, Spiteri G, et al. Europe-wide expansion and eradication of multidrug-resistant *Neisseria gonorrhoeae* lineages: a genomic surveillance study. Lancet Microbe. 2022;3(6):e452-e63. doi:10.1016/S2666-5247(22)00044-1

22. Jongen VW, Zimmermann HML, Boyd A, et al. Transient Changes in Preexposure Prophylaxis Use and Daily Sexual Behavior After the Implementation of COVID-19 Restrictions Among Men Who Have Sex With Men. J Acquir Immune Defic Syndr. 2021;87(5):1111-8. doi:10.1097/QAI.0000000000002697

23. Unemo M, Lahra MM, Escher M, et al. WHO global antimicrobial resistance surveillance for *Neisseria gonorrhoeae* 2017-18: a retrospective observational study. Lancet Microbe. 2021;2(11):e627-e36. doi:10.1016/S2666-5247(21)00171-3

24. Wadsworth CB, Arnold BJ, Sater MRA, et al. Azithromycin Resistance through Interspecific Acquisition of an Epistasis-Dependent Efflux Pump Component and Transcriptional Regulator in *Neisseria gonorrhoeae*. mBio. 2018;9(4). doi:10.1128/mBio.01419-18

25. de Korne-Elenbaas J, Bruisten SM, de Vries HJC, et al. Emergence of a *Neisseria gonorrhoeae* clone with reduced cephalosporin susceptibility between 2014 and 2019 in Amsterdam, The Netherlands, revealed by genomic population analysis. J Antimicrob Chemother. 2021;76(7):1759-68. doi:10.1093/jac/dkab082

26. Shimuta K, Unemo M, Nakayama S, et al. Antimicrobial resistance and molecular typing of *Neisseria gonorrhoeae* isolates in Kyoto and Osaka, Japan, 2010 to 2012: intensified surveillance after identification of the first strain (H041) with high-level ceftriaxone resistance. Antimicrob Agents Chemother. 2013;57(11):5225-32. doi:10.1128/AAC.01295-13

27. Eyre DW, Town K, Street T, et al. Detection in the United Kingdom of the *Neisseria gonorrhoeae* FC428 clone, with ceftriaxone resistance and intermediate resistance to azithromycin, October to December 2018. Euro Surveill. 2019;24(10). doi:10.2807/1560-7917.ES.2019.24.10.1900147

28. Unemo M, Golparian D, Nicholas R, et al. High-level cefixime- and ceftriaxone-resistant *Neisseria gonorrhoeae* in France: novel penA mosaic allele in a successful international clone causes treatment failure. Antimicrob Agents Chemother. 2012;56(3):1273-80. doi:10.1128/AAC.05760-11

29. Osnes MN, Didelot X, Korne-Elenbaas J, et al. Sudden emergence of a *Neisseria gonorrhoeae* clade with reduced susceptibility to extended-spectrum cephalosporins, Norway. Microb Genom. 2020;6(12). doi:10.1099/mgen.0.000480

30. Unemo M, Lahra MM, Cole M, et al. World Health Organization Global Gonococcal Antimicrobial Surveillance Program (WHO GASP): review of new data and evidence to inform international collaborative actions and research efforts. Sex Health. 2019;16(5):412-25. doi:10.1071/SH19023

31. Lewis DA. Will targeting oropharyngeal gonorrhoea delay the further emergence of drug-resistant *Neisseria gonorrhoeae* strains? Sex Transm Infect. 2015;91(4):234-7. doi:10.1136/sextrans-2014-051731

SUPPLEMENTARY MATERIAL

TABLE S1. METADATA OF ALL ISOLATES INCLUDED IN THIS STUDY, INCLUDING INDIVIDUAL ACCESSION NUMBERS, PATIENT DATA, PHENOTYPIC DATA AND MLSTS.

This table (17 columns and 323 rows) is available from Figshare (https://tinyurl.com/mtpjhuh7).

TABLE S2. STATISTICAL ASSOCIATION OF MLST 9362 WITH EPIDEMIOLOGICAL CHARACTERISTICS OF PATIENTS FROM THE CSH IN AMSTERDAM BEFORE AND DURING THE COVID-19 LOCKDOWN.

	Isolates with MLSTs other than MLST 9362 (N = 290)	**MLST 9362 isolates (N = 32)**	**P - value**
Sex			0.235
Male	272	32	
Female	18	0	
Age			0.880
<25	63	7	
25-34	127	14	
35-44	68	9	
≥45	32	2	
Country of origin			0.516
The Netherlands	141	17	
Suriname + Dutch Antilles	35	1	
Europe + Turkey	51	5	
M-S America	28	3	
Asia	19	5	
Africa	11	1	
Other	4	0	
Unknown	1	0	
Sexual orientation			0.021*
MSM + Transgender	250	32	
MSW + Female	40	0	
HIV status			0.503
Negative	214	27	
Positive	55	4	
Unknown	21	1	
Sex work			0.492
Yes	25	1	
No	265	31	

TABLE S2. STATISTICAL ASSOCIATION OF MLST 9362 WITH EPIDEMIOLOGICAL CHARACTERISTICS OF PATIENTS FROM THE CSH IN AMSTERDAM BEFORE AND DURING THE COVID-19 LOCKDOWN. (CONTINUED)

	Isolates with MLSTs other than MLST 9362 (N = 290)	**MLST 9362 isolates (N = 32)**	**P - value**
Number of sex partners in last 6 months			0.592
0-1	14	2	
2-4	80	5	
5-9	72	11	
10-19	49	7	
20-49	55	5	
≥50	20	2	
Symptomaticity			0.442
Symptomatic	97	8	
Asymptomatic	193	24	

** $p \leq 0.05$*

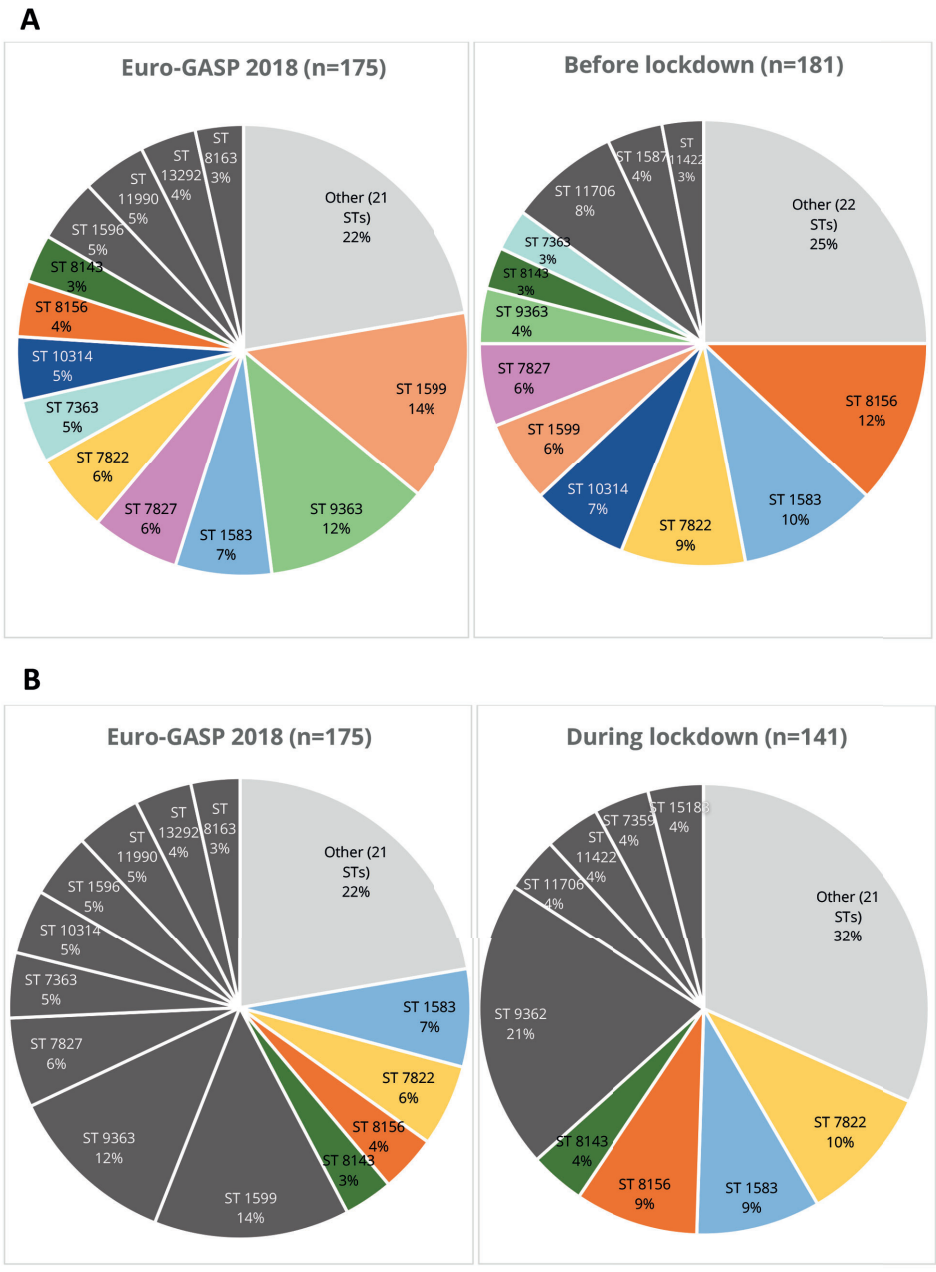


FIGURE S1. MLST DISTRIBUTION OF ISOLATES FROM AMSTERDAM IN THE EURO-GASP 2018 COLLECTION COMPARED TO THE MLST DISTRIBUTION OF ISOLATES OBTAINED BEFORE (A) AND DURING (B) LOCKDOWN. PREVALENCE ARE SHOWN FOR MLSTS THAT OCCURRED >5 TIMES IN THE ISOLATES BEFORE- OR DURING LOCKDOWN. MLSTS THAT OCCURRED <5 TIMES WERE CATEGORIZED AS 'OTHER'. COLOURED MLSTS WERE FOUND >5 TIMES IN BOTH PERIODS THAT WERE COMPARED, WHEREAS GREY MLSTS ARE ONLY FOUND >5 TIMES IN ONE OF THE PERIODS.

5

PUTATIVE TRANSMISSION OF EXTENDED-SPECTRUM β-LACTAMASE-PRODUCING *ESCHERICHIA COLI* AMONG MEN WHO HAVE SEX WITH MEN IN AMSTERDAM, THE NETHERLANDS

Jolinda de Korne-Elenbaas*, Boas C. L. van der Putten*,
Niels D. M. Boek, Amy Matser, Constance Schultsz,
Sylvia M. Bruisten, Alje P. van Dam
*equal contribution

Submitted for publication

ABSTRACT

OBJECTIVES

A previous study showed higher prevalence of *Escherichia coli* expressing extended-spectrum β-lactamases (ESBL-Ec) among men who have sex with men (MSM) in Amsterdam, the Netherlands, compared to the general Dutch population. We here genetically characterised the ESBL-Ec isolates and investigated whether the increased prevalence could be explained by transmission between participants.

METHODS

Whole-genome sequences were obtained from 93 unique ESBL-Ec isolates, cultured from rectal swabs of 79 participants. Isolates were typed according to the Achtman MLST scheme and ESBL- and virulence genes were identified. Pairwise SNP distances were determined between isolates. Isolate pairs with ≤25 SNPs were considered part of a putative transmission event, and events between multiple participants formed putative transmission clusters. To investigate whether putatively transmitted isolates belonged to globally expanded lineages, the level of hierarchical clustering with international isolates was assessed using core genome MLST (cgMLST) implemented on the Enterobase platform.

RESULTS

Most frequently detected *E. coli* types were ST131:$bla_{\text{CTX-M-15}}$ (16/117, 13.5%), ST131:$bla_{\text{CTX-M-27}}$, ST3075:$bla_{\text{CTX-M-15}}$ and ST14:$bla_{\text{SHV-12}}$ (all 6/117, 6.5%). Fourteen putative transmission events were identified, forming 4 putative transmission clusters. The largest putative transmission cluster contained ST131 isolates which clustered with multiple international isolates in SNP and cgMLST analysis. One other transmission cluster (ST14:$bla_{\text{SHV-12}}$) and 2 transmission events (ST14:$bla_{\text{SHV-12}}$ and ST394:$bla_{\text{CTX-M-15}}$) contained strains very rarely found in public data.

CONCLUSIONS

The identification of unique ESBL-Ec strains involved in putative transmission and carried by multiple participants demonstrate a high probability of ESBL-Ec transmission between MSM in Amsterdam. Therefore, ESBL-Ec infection should be considered in case of sexually active MSM having associated symptoms.

INTRODUCTION

Enterobacteriaceae resistant to third-generation cephalosporins have disseminated worldwide and contribute significantly to the antimicrobial resistance crisis. This type of antimicrobial resistance is often conferred by extended-spectrum β-lactamases (ESBL), which are enzymes that break down β-lactam antibiotics. Infections with ESBL-producing Enterobacteriaceae (ESBL-E) are an important cause of urinary tract infections in the community setting, and bloodstream infections associated with these infections occur frequently[1,2]. Resistance to third-generation cephalosporins can lead to delay of correct antibiotic treatment, prolonged morbidity and increased mortality.

The increase in community onset infections with ESBL-E is a major public health concern in multiple continents, including Europe[1,3]. In the Netherlands, the reported prevalence of ESBL-E carriage in the general population is 5.0% and in Amsterdam 8.6%[4,5]. A strikingly higher prevalence of 16.3% was found in a cohort of men who have sex with men (MSM) in Amsterdam[6]. This increased prevalence of ESBL-E has also been found in other MSM populations and among HIV-positive men compared to HIV-negative men[7,8]. Associated risk factors were a high number of sex partners and specific sexual behaviours with casual partners, such as rimming and fisting, suggesting the increase in ESBL-E carriage may in part be due to sexual transmission between MSM.

To further examine whether the increased ESBL-E prevalence in MSM is driven by transmission within this population, we aimed to identify putative transmission events between MSM who were identified as ESBL-E carriers in the previous Dutch study[6]. We identified putative transmission events and subsequently transmission clusters using WGS data. We described the isolates in detail on phenotypic and genotypic levels to identify prevalent ESBL-positive *Escherichia coli* (ESBL-Ec) strains in Amsterdam and to compare these strains with ESBL-Ec lineages that have been expanding globally. Comparing isolates from our dataset with other publicly available isolates enabled us to assess whether the observed clonality in our population is due to putative transmission or due to clonality of globally disseminated lineages.

METHODS

STUDY POPULATION

Since 1984, the Amsterdam Cohort Studies have aimed to study the prevalence, incidence and risk factors for HIV infection and other sexually transmitted infections among sexually active MSM involved in MSM-related activities in Amsterdam. Every 3-6 months, participants fill out a questionnaire on sexual behaviour and are tested for HIV and several other sexually transmitted

diseases. During study visits scheduled between April and December 2018, 583 participants were additionally tested for ESBL-E carriage on self-collected anal swabs[6]. A small proportion of participants were tested again for ESBL-E carriage after 1 year, between September and October 2019.

ESBL-DETECTION

ESBL phenotypes were previously established[6]. In short, Enterobacteriaceae were isolated from self-obtained rectal swabs (Copan Diagnostics Inc.) which were cultured in brain-heart infusion broth supplemented with 16 mg/L amoxicillin (Mediaproducts BV, Groningen, The Netherlands). Species were identified using MALDI Bio-typer (Bruker). Colonies were tested for susceptibility using the Vitek-2XL system (bioMérieux) and three e-test combinations: ceftazidime, cefotaxime and cefepime with and without clavulanic acid on Mueller Hinton II agar (Becton Dickinson). ESBL presence was confirmed if at least one of the three combinations of e-tests had a ratio of MICs with- and without clavulanic acid of ≥8 mg/L, in accordance with EUCAST rules. Antibiotic susceptibility was defined according to EUCAST clinical breakpoints v12.0. Colonies confirmed as ESBL-E were stored at -80°C. When multiple colonies taken from one culture showed differences in morphology and/or phenotype, colonies were separately analysed and stored. Therefore, multiple isolates could have been obtained from one participant.

ISOLATE SELECTION AND WGS

In this study, only ESBL-Ec isolates were used because of the small number of other ESBL-E species previously identified[6]. ESBL-Ec isolates were cultured overnight at 37°C on MacConkey agar. DNA was isolated from pure cultures in DNA/RNA shield buffer using the ZymoBIOMICS TM Magbead DNA kit (ZYMO RESEARCH). DNA sequencing libraries were prepared with the Nextera XT DNA Library Preparation Kit for Illumina DNA/RNA Unique Dual Indexes and sequenced using Illumina NovaSeq 6000 short-read sequencing (BaseClear BV, Leiden, The Netherlands).

QUALITY CONTROL AND TYPING

Raw sequence reads were filtered and trimmed using fastp v0.20.1[9] and assembled with Shovill v1.1.0 (https://github.com/tseemann/shovill), using SPAdes v3.14.1[10]. Quality of the assemblies was assessed using QUAST v5.0.2[11] and BUSCO v5.1.3[12] and contamination was detected using kraken2 v2.1.2[13]. Isolates were typed according to the Achtman seven-gene MLST scheme for E. coli as integrated in Enterobase[14] using the mlst script (https://github.com/tseemann/mlst). ESBL genes were identified in genome assemblies with AMRfinderplus v3.10.5[15], phylogroups were predicted with EzClermont v0.6.3[16] and virulence factors were identified with abricate v1.0.1 (https://github.com/tseemann/abricate) using the virulence factor database (vfdb)[17].

IDENTIFICATION OF PUTATIVE TRANSMISSION EVENTS

Trimmed raw sequence reads of isolates were mapped with Snippy v4.6.0 (https://github.com/tseemann/snippy) to a reference genome with identical MLST, selected with ReferenceSeeker v1.7.3[18]. Core genome alignments were created for each MLST with the snippy-core option in Snippy v4.6.0. Isolates were compared pairwise and SNPs between each set of isolates were identified and calculated with snp-dists v0.7.0 (https://github.com/tseemann/snp-dists). We did not remove or mask recombination events. Putative transmission events were identified as two isolates from different participants having ≤25 SNP differences between them[19]. When multiple isolates were obtained from one participant at the same time point but these isolates had identical MLST and had ≤25 SNP differences between them, only one of the isolates was included in the analyses. Isolates taken from one participant at multiple time points were included separately. The workflow was managed using Snakemake v5.27.4[20] and is available on Github (https://github.com/boasvdp/ESBL-Ec_transmission_ABRACS/).

GENETIC RELATEDNESS OF ISOLATES AND CORRESPONDING METADATA

To assess genetic relatedness between all isolates, all trimmed raw reads were mapped against reference genome EC958 (GenBank: HG941718.1) with Snippy and a core genome alignment was created. ClonalFrameML was used to create a recombination filtered phylogenetic tree using the default settings, which was visualised in iTOL[21]. Metadata on sexual behaviour was obtained during the previous study, which described a latent class analysis (LCA) to assess associations between ESBL-Ec carriage and sexual behaviour[6]. The LCA analysis revealed distinct groups of sexual behaviours with steady partner(s): i) no sex with steady partner(s), ii) fellatio alone, iii) fellatio, rimming and/or anal sex. Groups of sexual behaviours with casual partner(s) were: i) no sex with casual partner(s), ii) fellatio and anal sex, iii) fellatio, anal sex, rimming and/or frottage and iv) multiple behaviours including toy use and fisting. Further details were described previously[6].

HIERARCHICAL CLUSTERING OF PUTATIVELY TRANSMITTED ISOLATES AND OTHER PUBLICLY AVAILABLE ISOLATES

After identification of putative transmission events (≤25 SNP differences between isolates), linkage of events formed putative transmission clusters. We determined whether isolates in putative transmission clusters belonged to predominant and clonal ESBL-Ec lineages that have been expanding globally. In that case, extensive spread of strains belonging to the same clone might explain the genetic relatedness between clustered isolates instead of transmission. First, isolates were uploaded to the Enterobase database (uberstrains ESC_XA1161AA-ESC_XA1275AA, Table S1)[14]. Core genome MLST (cgMLST) was performed using the Enterobase pipeline and cgMLST allelic profiles were downloaded. Differences between cgMLST profiles were calculated using cgmlst-dists (https://github.com/tseemann/cgmlst-dists). Similarity

between cgMLST profiles can be expressed at multiple levels, expressed as "hierarchical clonal complexes" (hierCCs)[22]. For example, if two isolates differ 14 cgMLST loci from each other, these isolates would belong to the same hierCC at the 20 loci level (HC20), because they differ in 20 or fewer cgMLST loci. However, these isolates would be assigned to different clusters at the HC10 levels, as the isolates differ in more than 10 cgMLST loci. The hierCC method allows efficient comparisons in large datasets, such as the >180,000 *E. coli* genomes in Enterobase. Reference-based read mapping is impossible for datasets of this size due to computational constraints.

For all putative transmission clusters that we identified in our dataset, we determined at which HC level the transmission cluster could be captured. Additionally, Enterobase was searched for other publicly available isolates that clustered with isolates from putative transmission clusters in our collection (accessed on March 23rd, 2022). Based on the analysis of strains included in this study (Figure S1), we assumed that if strains differed in more than 10 cgMLST loci, these strains did not fall below the similarity threshold of ≤25 SNP differences (Figure S1). Therefore, we focused on clustering at levels HC2, HC5, HC10 and HC20.

RESULTS

QUALITY OF SEQUENCE DATA AND ISOLATE CHARACTERISTICS

ESBL-E were previously detected in 95 unique participants. In total, 139 unique isolates were obtained, of which 130 (94%) were identified as *E. coli* and used in this study[6]. Pure cultures and raw sequence reads could be obtained for 117 ESBL-Ec isolates. After assembly and quality checks, 2 isolates were excluded due to contamination and a draft genome size of >6 Mbp, whereas the mean genome size of the other 115 isolates was 5.2 Mbp. Sequence reads of these 115 isolates had a mean coverage depth of 238x and were assembled, yielding a mean of 153 contigs (Table S1). Twenty-two isolates were identified as replicate isolates, as these were isolated from the same participants and differed ≤25 SNPs, and therefore excluded from further analyses. This resulted in a total of 93 isolates obtained from 79 unique participants of whom 4 were sampled at multiple time points. Forty-seven percent of the participants were aged >45, and the majority had a Dutch ethnicity (78%) and a negative HIV status (92%) (Table 1).

With regard to antibiotic susceptibility, most isolates were resistant to ceftazidime (88%), cefotaxime (97%), cefuroxime (98%), ciprofloxacin (70%) and trimethoprim (73%). Lower prevalence of resistance was found against gentamicin (24%) tobramycin (27%), fosfomycin (4%) and nitrofurantoin (2%). All isolates were resistant to amoxicillin, which was expected because anal swabs were cultured in broth containing amoxicillin prior to isolation. All isolates were susceptible to carbapenems and colistin. Typing resulted in 40 different MLST types, with most STs found for <5 isolates. ST131 was most frequently found among the isolates (26%),

followed by ST38, ST14 and ST3075 (6.5% of isolates in each group). The majority of isolates belonged to phylogroup B2 (43%), followed by phylogroups A and D (25% and 21%, respectively). The remaining phylogroups B1, E and F constituted only a minor part of our collection (11% all together) (Table 2).

IDENTIFICATION OF PUTATIVE TRANSMISSION EVENTS AND CLUSTERS

We identified 14 isolate pairs with ≤25 SNPs difference between them, resulting in a total of 14 putative transmission events between participants. Twelve of these events were linked through a total of ten shared participants, forming two putative transmission clusters of six and four participants (Figure 1A). Two separate putative transmission events could not be linked to other transmission events and thus comprised only two participants each (Figure 1B). The isolates in the largest putative transmission cluster of six participants were typed as ST131 and included one isolate from 2019. The putative transmission cluster consisting of four participants was typed as ST14 and the two remaining putative transmission events were typed as ST14 and ST394. All isolates in these clusters were from 2018. The cluster belonging to ST131 differed in 6-25 SNPs and 2-8 cgMLST loci. The clusters belonging to ST14 differed in 2-20 SNPs and 0-7 cgMLST loci and the cluster belonging to ST394 differed in 5 SNPs and 2 cgMLST loci.

TABLE 1. CHARACTERISTICS OF PARTICIPANTS FROM WHOM ESBL-EC ISOLATES WERE OBTAINED.

Characteristic	Number of participants (N=79)
Age	
16-35	25 (31.5%)
35-45	17 (21.5%)
45+	37 (47%)
Ethnicity	
Dutch	62 (78%)
Non-Dutch	14 (18%)
Unknown	3 (4%)
HIV-status	
Negative	73 (92%)
Positive	6 (8%)
Number of isolates per participant	
1 isolate from a single time point	69 (87%)
2 isolates from a single time point	6 (8%)
2 isolates from 2 time points	1 (1%)
3 isolates from 2 time points	2 (3%)
4 isolates: 2 from each of the 2 time points	1 (1%)

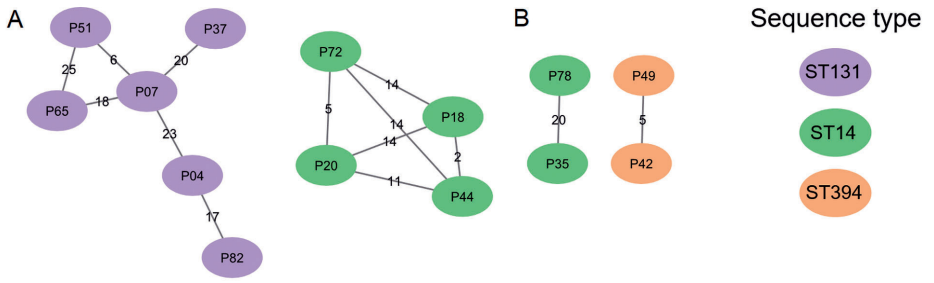


FIGURE 1. IDENTIFIED PUTATIVE TRANSMISSION CLUSTERS (A) AND EVENTS (B) WITHIN 93 ESBL-EC ISOLATES OBTAINED FROM A COHORT OF MSM IN AMSTERDAM. NODES INDICATE UNIQUE PARTICIPANT NUMBERS. NUMBERS ON EDGES CONNECTING NODES ARE THE LOWEST SNP DIFFERENCE OBSERVED BETWEEN ISOLATES FROM PARTICIPANTS. COLOURS REPRESENT THE MLST TO WHICH THE ISOLATES IN THE TRANSMISSION CLUSTERS (A) OR EVENTS (B) BELONG.

GENETIC RELATEDNESS, AMR GENE DISTRIBUTION AND ANTIBIOTIC SUSCEPTIBILITY

Isolates with identical STs strongly clustered in the phylogenetic tree whereas different phylogroups were distributed throughout the tree (Figure 2). AMR genes were identified, showing that the $bla_{\text{CTX-M-15}}$ (harboured by 54% of isolates) and $bla_{\text{CTX-M-27}}$ (harboured by 17% of isolates) were the most prevalent $bla_{\text{CTX-M}}$ genes. $bla_{\text{CTX-M-15}}$ presence was distributed throughout the population although ST131 was most frequently associated with $bla_{\text{CTX-M-15}}$ (16 strains). The $bla_{\text{CTX-M-27}}$ gene was mostly found in isolates belonging to ST1193, ST131 and ST38. With regard to antibiotic susceptibility, isolates carrying the $bla_{\text{CTX-M-15}}$ or $bla_{\text{CTX-M-27}}$ gene were all resistant to ceftazidime and cefotaxime and the majority (97%) to cefuroxime as well. The ST14 strains which formed two distinct putative transmission clusters did not carry $bla_{\text{CTX-M}}$ genes, but did harbour the ESBL gene $bla_{\text{SHV-12}}$. All ST14 strains were resistant to ceftazidime (with MICs of at least 32 mg/L), but displayed varying levels of susceptibility to cefotaxime (range: 1-32 mg/L).

All but a single strain which were non-susceptible to ciprofloxacin (MIC of at least 0.5 mg/L) harboured a resistance determinant for ciprofloxacin. Fifty-three strains had one or more mutations in *gyrA* residues S83, N87 or *parC* residue S80, while eleven strains harboured a *qnr* gene (*qnrS1* or *qnrB19*). Nine isolates were phenotypically susceptible to ciprofloxacin but did harbour a S83L mutation in *GyrA* (Figure 2). Notably, of the 18 strains which harboured *qnrS1*, 12 strains harboured *qnrS1* on the same assembly contig as $bla_{\text{CTX-M-15}}$. In these strains, $bla_{\text{CTX-M-15}}$ and *qnrS1* were associated with transposon Tn2 (Figure S2). These 12 strains belonged to 11 different STs, indicating extensive horizontal transfer of this mobile gene element. NCBI BLAST identified this locus in publicly available genome sequences of *E. coli* and *Shigella* spp. Finally, resistance to trimethoprim was distributed throughout the population, whereas resistance to tobramycin was mainly observed for ST131 isolates (Figure 2).

TABLE 2. CHARACTERISTICS OF UNIQUE ESBL-EC ISOLATES INCLUDED IN THIS STUDY.

Characteristic	Number of isolates (N=93*)
Minimum Inhibitory Concentrations	
Penicillin	
Amoxicillin	
≤8 mg/L	-
>8 mg/L	93 (100%)
Cephalosporins	
Ceftazidime	
≤1 mg/L	11 (12%)
>4 mg/L	82 (88%)
Cefotaxime	
≤1 mg/L	3 (3%)
>1 mg/L	90 (97%)
Cefuroxime	
≤8 mg/L	2 (2%)
>8 mg/L	91 (98%)
Fluoroquinolones	
Ciprofloxacin	
≤0.25 mg/L	28 (30%)
0.5 mg/L	19 (20%)
>0.5 mg/L	36 (49%)
Carbapenems	
Imipenem	
≤2 mg/L	93 (100%)
>4 mg/L	-
Meropenem	
≤2 mg/L	93 (100%)
>8 mg/L	-
Aminoglycosides	
Gentamicin	
≤2 mg/L	71 (76%)
>2 mg/L	22 (24%)
Tobramycin	
≤2 mg/L	68 (73%)
>2 mg/L	25 (27%)
Miscellaneous agents	
Fosfomycin	
≤8 mg/L	89 (96%)
>8 mg/L	4 (4%)
Nitrofurantoin	
≤64 mg/L	91 (98%)
>64 mg/L	2 (2%)
Trimethoprim	
≤4 mg/L	25 (27%)
>4 mg/L	68 (73%)
Colistin	
≤2 mg/L	93 (100%)
>2 mg/L	-

Characteristic	Number of isolates (N=93*)
MLST	
ST131	24 (26%)
ST38	6 (6.5%)
ST14	6 (6.5%)
ST3075	6 (6.5%)
ST1193	5 (5%)
ST10	5 (5%)
Isolates in ST group with <5 isolates	41 (44%)
Phylogroup	
A	24 (26%)
B1	7 (7.5%)
B2	39 (42%)
D	20 (21.5%)
E	1 (1%)
F	2 (2%)

**Duplicate isolates from a single time point were excluded from the analyses (n=22).*

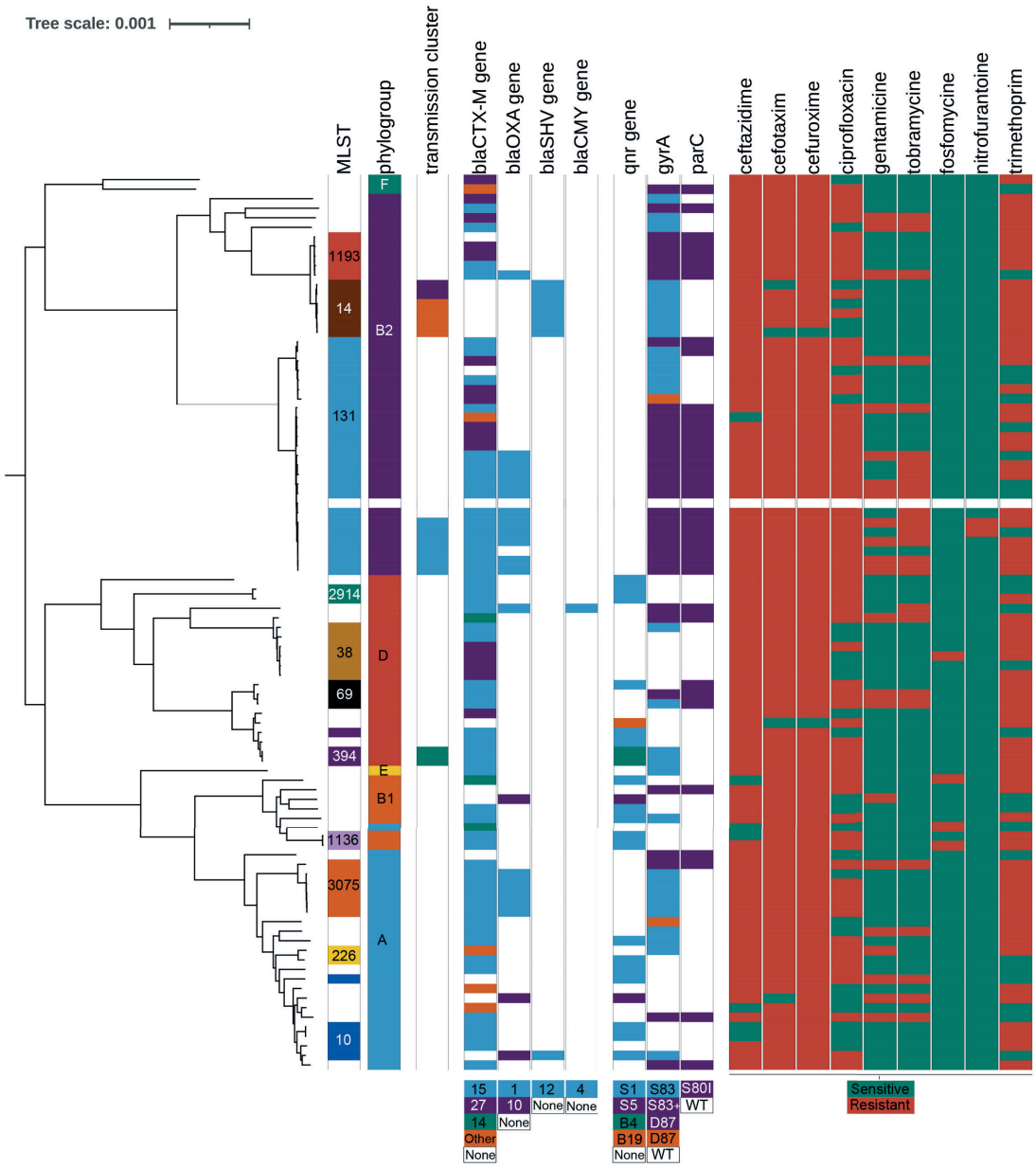


FIGURE 2. RECOMBINATION FILTERED MIDPOINT-ROOTED PHYLOGENETIC TREE BASED ON CORE GENOME SNPS OF 93 ESBL-EC ISOLATES. EC958 WAS USED AS REFERENCE STRAIN. METADATA INCLUDES: MLST CLUSTERS CONTAINING ≥5 ISOLATES, PHYLOGROUPS, IDENTIFIED PUTATIVE TRANSMISSION CLUSTERS OR EVENTS, *BLA* GENES, *QNR* GENES, *GYRA* AND *PARC* MUTATIONS AND PHENOTYPIC DATA CATEGORIZED AS SUSCEPTIBLE OR RESISTANT.

TABLE 3. SEXUAL BEHAVIOURS AND THE NUMBER OF ISOLATES OBTAINED FROM PARTICIPANTS REPORTING THIS BEHAVIOUR.

	Number of isolates N=93*
Steady partner(s)	
No sex with steady partner(s)	40 (43%)
Fellatio alone	4 (4%)
Fellatio, rimming and/or anal sex	37 (40%)
No data available	12 (13%)
Casual partner(s)	
No sex with casual partner(s)	7 (8%)
Fellatio and anal sex	2 (2%)
Fellatio, anal sex, rimming, frottage	44 (47%)
Multiple behaviours including toy use and fisting	25 (27%)
No data available	15 (16%)

**Participants can have multiple isolates with different MLST, with >25 SNPs difference or from different time points.*

SEXUAL BEHAVIOUR OF PARTICIPANTS WHOSE ISOLATES WERE INVOLVED IN PUTATIVE TRANSMISSION EVENTS

The majority of isolates was obtained from participants who had sex with casual partners (71/93 isolates; 76%) and who reported to engage in sexual techniques such as rimming, frottage, toy use and fisting (69/93; 74%) (Table 3). All isolates in putative transmission clusters or events were obtained from participants who reported sex with casual partners and engaged in multiple sexual techniques with casual partners. Five isolates involved in putative transmission were from participants who reported no sex with steady partner. No sexual behaviour was exclusively found for participants whose isolates fell within putative transmission clusters or events.

HIERARCHICAL CLUSTERING OF PUTATIVELY TRANSMITTED DUTCH ISOLATES AND GLOBAL ISOLATES

Within all putative transmission clusters- and events in our collection, isolates belonged to the same HC20 and HC10 clusters (Figure 3A and 3B). Clustering at the HC5 level revealed distinct HC5 clusters within the ST14 putative transmission event (Figure 3C). At HC2 level, some isolates still belonged to the same types (Figure 3D). Isolates from the largest ST131 transmission cluster clustered with many public isolates, highlighting a high degree of circulation of this strain. At HC5 level, the ST131 transmission cluster still clustered with 4 public isolates from the USA and Guadeloupe. Using our SNP analysis approach, we found that the three USA strains differed exactly 25 SNPs with a single Dutch strain, while all other strains were different in more than 25 SNP positions. The raw reads were not available for the Guadeloupe strain. The ST14 transmission cluster of 4 participants was shared with only a single Spanish isolate differing in two cgMLST loci (HC2 level). Isolates in the ST14 transmission event with 2 participants clustered

with 5 Spanish isolates at HC20 level, whereas isolates in the ST394 transmission event did not cluster with any public isolate at that level. All other public isolates differed in >20 cgMLST loci from the Dutch isolates involved in putative transmission.

DISCUSSION

We previously reported an increased carriage of ESBL-positive Enterobacteriaceae (ESBL-E) in a cohort of MSM in Amsterdam, the Netherlands[6]. As ESBL-Ec constituted 94% of identified ESBL-E, we now employed WGS to further examine whether the increased prevalence is potentially driven by transmission within this population. Additionally, we described these ESBL-Ec isolates in great detail on phenotypic and genotypic levels using WGS. This enabled us to identify prevalent strains among MSM in Amsterdam, which also informed us about ESBL-Ec strains circulating in the general population of Amsterdam and the Netherlands.

We identified high genetic relatedness between several isolates in our collection, contained in two clusters and two events. The largest cluster of isolates belonged to ST131, which is a globally disseminated lineage of *E. coli*[23]. In public data, we found four additional isolates from outside of Europe which closely clustered with our ST131 isolates based on cgMLST and SNPs. This suggests that the transmission of this ST131 strain is not limited to this MSM population. Analysing our putatively transmitted isolates together with genomic data of 153 carriage and invasive ST131 isolates from a recent Dutch study[24] revealed that no isolates clustered with ≤25 SNPs or at HC5 level in a cgMLST analysis, thus no strains were shared between studies (data not shown). This might indicate that the ST131 we observed in our study is not simply a prevalent ST131 clone that swept the ESBL-Ec population in the Netherlands.

We also identified a putative transmission cluster and two transmission events containing ST14 or ST394 isolates. These strains were rare in public data, so finding these strains in multiple participants of our study strongly suggests transmission between participants. Interestingly, ST14 was not reported by both recent Dutch studies (5,24). In Enterobase, only three ST14 isolates from the Netherlands are reported: one isolated from an asylum seeker arriving in the Netherlands in 2016 and two from an epidemiological study on colistin-resistant Enterobacterales in the Netherlands from 2019[25,26].

Previous studies which report whole-genome sequencing of ESBL-E among MSM have mainly focused on sexually transmitted enteric infections by *Shigella* spp. or *E. coli*[8,27-32]. However, ESBL-E carriage rather than infection might play an important role in the transmission of antimicrobial resistance. Based on genetic relatedness and the uniqueness of strains, we report at least three putative transmission clusters within our population. This transmission

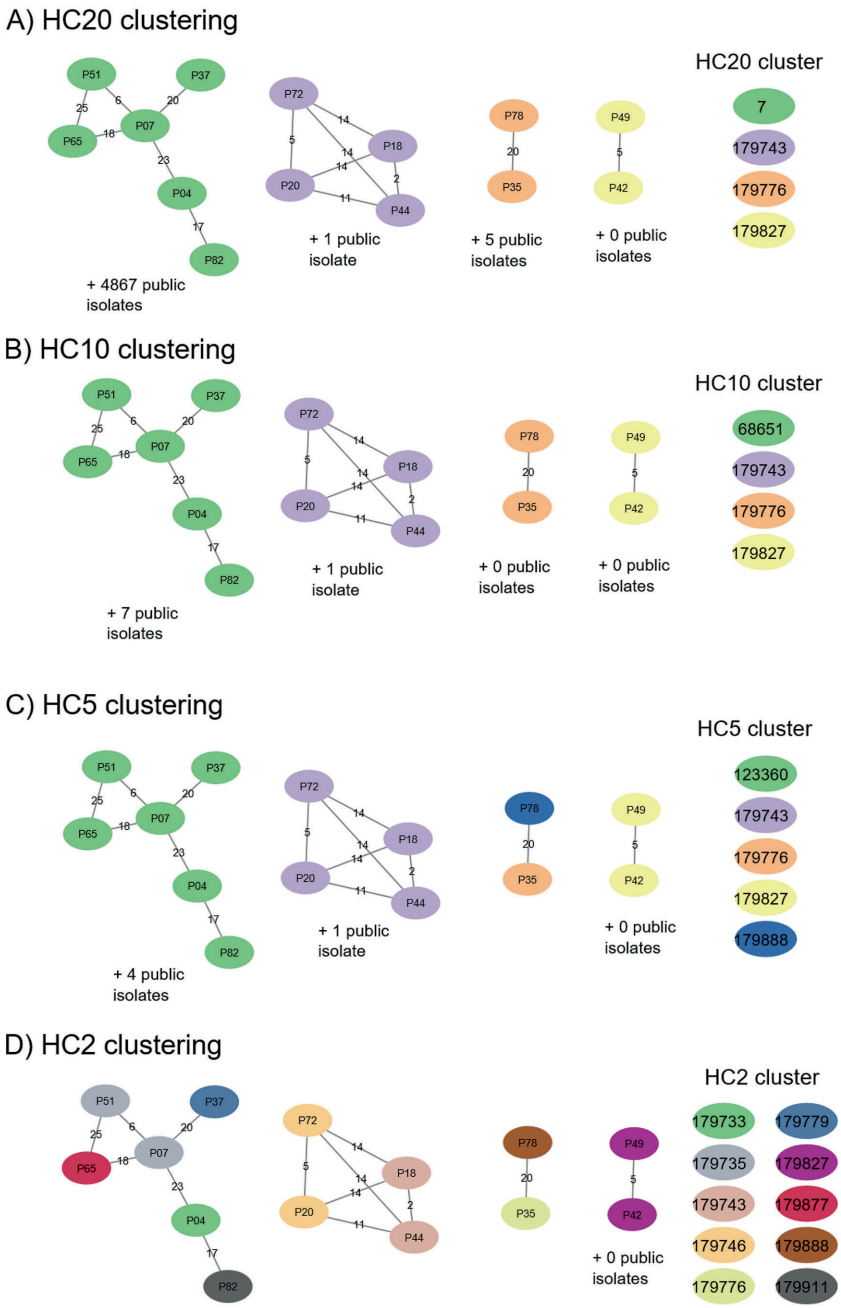


FIGURE 3. HIERARCHICAL CLUSTERING OF ISOLATES IN PUTATIVE TRANSMISSION CLUSTERS ON A) HC20, B) HC10, C) HC5 AND D) HC2 LEVELS. FOR EVERY PUTATIVE TRANSMISSION CLUSTER WHICH CORRESPONDS WITH A SINGLE HC CLUSTER, THE NUMBER OF PUBLIC ENTEROBASE ISOLATES BELONGING TO THE SAME CLUSTER IS INDICATED. NODES INDICATE UNIQUE PARTICIPANT NUMBERS. NUMBERS ON EDGES CONNECTING NODES ARE THE LOWEST SNP DIFFERENCE OBSERVED BETWEEN ISOLATES FROM PARTICIPANTS.

could contribute to the increased carriage rate of ESBL-E among MSM. Since number of sex partners and certain sexual behaviours were associated with ESBL-E carriage, we speculate that sexual transmission might have taken place between MSM in our study[6]. However, other routes of transmission between participants cannot be excluded.

Although we only sampled strains based on resistance to third generation cephalosporins, we observed a high degree of multidrug resistance in our dataset. For example, 70% of the ESBL-Ec strains was also non-susceptible to ciprofloxacin, approximately a quarter was also resistant to aminoglycosides and 73% was resistant to trimethoprim. $bla_{CTX\text{-}M\text{-}15}$ was identified in association with qnrS1 and transposon Tn2 in twelve strains, belonging to eleven distinct STs. Highly similar loci could be identified through NCBI BLAST in *E. coli* and *Shigella* spp. strains of which genetic sequences are available in the public domain. The genetic linkage of genes conferring resistance to different classes of antimicrobials has the potential to increase multidrug resistance. This concept is colloquially known as "genetic hitchhiking" and comprises the maintenance of genetically linked AMR genes through antibiotic pressure from a single antimicrobial[33].

The majority of strains in our dataset harboured $bla_{CTX\text{-}M\text{-}15}$, a pandemic ESBL gene that has increased in frequency in recent years[34]. As has been reported before[35], the success of $bla_{CTX\text{-}M\text{-}15}$ seems to be linked to the success of widespread E. coli clones, such as ST131. Among our study population, we report a 26% prevalence of ST131 harbouring $bla_{CTX\text{-}M\text{-}15}$. This percentage is similar to recent studies on ESBL-Ec carriage in the general Dutch population, 2014-2017[24] and in the general population of Amsterdam in 2011[5]. However, we identified $bla_{CTX\text{-}M\text{-}15}$ in 24 other STs as well. $bla_{CTX\text{-}M\text{-}14}$, another ESBL gene frequently reported globally, was only detected in four strains. Most STs only harboured a single $bla_{CTX\text{-}M}$ variant, although ST38, ST131 and ST1193 all harboured both $bla_{CTX\text{-}M\text{-}15}$ and $bla_{CTX\text{-}M\text{-}27}$. Two of the putative transmission clusters that we defined, consisted of ST14 strains which did not harbour $bla_{CTX\text{-}M}$ genes like most other strains in our dataset, but harboured the ESBL gene $bla_{SHV\text{-}12}$[35]. Our findings show non-CTX-M ESBL genes are also transmitted readily between humans and might represent an understudied population of ESBL-Ec which is important to monitor[36].

Although our SNP analysis indicated very high similarity between strains of different participants, transmission cannot be proven based on genomics alone. This study did not have any contact tracing information available which would aid in confirming transmission clusters. Additionally, since 583 participants from a regularly tested cohort were sampled, there is a high chance we might have missed more cases of transmission of ESBL-Ec in individuals not belonging to the cohort. Future studies should expand our efforts at a larger scale to enable more detailed transmission tracing. Also, the studied MSM cohort has known biases, as it consists mainly of Dutch, highly-educated MSM who were included in a manner prone to convenience sampling

bias. Therefore, the participants with their ESBL-Ec strains might not be fully representative of MSM in Amsterdam or in the Netherlands.

To conclude, the majority of putatively transmitted strains from MSM in Amsterdam belong to predominant global lineages of ESBL-Ec, although we also identified a strain involved in putative transmission being unique to the Netherlands and carried by multiple participants. Therefore, this study demonstrates a high probability of ESBL-Ec transmission between MSM. As certain sexual behaviours were previously identified to increase the ESBL-Ec carriage rate, these behaviours potentially increase ESBL-Ec transmission as well. This implies that sexually active MSM could have increased risk for ESBL-E carriage, which should be taken into account in case of symptoms associated to ESBL-E infection.

FUNDING

This work was funded by the Public Health Laboratory of Amsterdam. BP was funded through an internal AMC grant ("Flexibele OiO beurs").

TRANSPARENCY DECLARATIONS

The authors have nothing to declare.

REFERENCES

1. Pitout JD, Laupland KB. Extended-spectrum β-lactamase-producing Enterobacteriaceae: an emerging public-health concern. Lancet Infect Dis. 2008;8(3):159–66.
2. Trecarichi EM, Cauda R, Tumbarello M. Detecting risk and predicting patient mortality in patients with extended-spectrum β-lactamase-producing Enterobacteriaceae bloodstream infections. Future Microbiol. 2012;7(10):1173–89.
3. Coque TM, Baquero F, Canton R. Increasing prevalence of ESBL-producing Enterobacteriaceae in Europe. Euro Surveill 2008;13(47):19044.
4. van den Bunt G, van Pelt W, Hidalgo L, et al. Prevalence, risk factors and genetic characterisation of extended-spectrum beta-lactamase and carbapenemase-producing Enterobacteriaceae (ESBL-E and CPE): a community-based cross-sectional study, the Netherlands, 2014 to 2016. Euro Surveill. 2019;24(41):1800594.
5. Reuland EA, al Naiemi N, Kaiser AM, et al. Prevalence and risk factors for carriage of ESBL-producing Enterobacteriaceae in Amsterdam. J Antimicrob Chemother. 2016;71(4):1076–82.
6. van Bilsen WPH, van Dulm E, Matser DA, et al. High carriage of ESBL-producing Enterobacteriaceae associated with sexual activity among men who have sex with men. Int J Antimicrob Agents. 2021;106276.
7. Reinheimer C, Keppler OT, Stephan C, et al. Elevated prevalence of multidrug-resistant gram-negative organisms in HIV positive men. BMC Infect Dis. 2017;17(1):206.
8. Mook P, McCormick J, Bains M, et al. ESBL-Producing and Macrolide-Resistant Shigella sonnei Infections among Men Who Have Sex with Men, England, 2015. Emerg Infect Dis 2016;22(11): 1948–1952.
9. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90.
10. Prjibelski A, Antipov D, Meleshko D, et al. Using SPAdes De Novo Assembler. Curr Protoc Bioinforma 2020; 70(1)
11. Gurevich A, Saveliev V, Vyahhi N, et al. QUAST: quality assessment tool for genome assemblies. Bioinforma Oxf Engl. 2013;29(8):1072–5.
12. Manni M, Berkeley MR, Seppey M, et al. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol Biol Evol. 2021;38(10): 4647–4654.
13. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20(1):257.
14. Zhou Z, Alikhan N-F, Mohamed K, et al. The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. Genome Res. 2020;30(1):138–52.
15. Feldgarden M, Brover V, Gonzalez-Escalona N, et al. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci Rep. 2021;11(1):12728.
16. Waters NR, Abram F, Brennan F, et al. Easy phylotyping of Escherichia coli via the EzClermont web app and command-line tool. Access Microbiol. 2020;2(9):acmi000143.
17. Chen L, Yang J, Yu J, et al. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 2005;33:D325-8.
18. Schwengers O, Hain T, Chakraborty T, et al. ReferenceSeeker: rapid determination of appropriate reference genomes. J Open Source Softw. 2020;5(46):1994.
19. Gorrie CL, Silva AGD, Ingle DJ, et al. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. Lancet Microbe 2021;2(11):e575-e583.
20. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. Bioinformatics. 2012;28(19):2520–2.
21. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(1):242–5.
22. Zhou Z, Charlesworth J, Achtman M. HierCC: a multi-level clustering scheme for population assignments based on core genome MLST. Bioinformatics 2021; 37(20):3645-3646.
23. Nicolas-Chanoine MH, Bertrand X, Madec JY. Escherichia coli ST131, an Intriguing Clonal Group. Clin Microb Rev. 2014; 27(3):543-7

24. Verschuuren TD, van Hout D, Arredondo-Alonso S, et al. Comparative genomics of ESBL-producing Escherichia coli (ESBL-Ec) reveals a similar distribution of the 10 most prevalent ESBL-Ec clones and ESBL genes among human community faecal and extra-intestinal infection isolates in the Netherlands (2014–17). J Antimicrob Chemother. 2021;76(4):901–908.
25. Louka C, Ravensbergen SJ, Ott A, et al. Predominance of CTX-M-15-producing ST131 strains among ESBL-producing Escherichia coli isolated from asylum seekers in the Netherlands. J Antimicrob Chemother. 2021;76(1):70–6.
26. Vendrik K, de Haan A, Witteveen S, et al. Genomic Epidemiology of Colistin-resistant Enterobacterales from Dutch Patients: A Prospective Matched Case-control Study. Preprint DOI: 10.21203/rs.3.rs-870731/v1(2021)
27. Hoffmann C, Sahly H, Jessen A, et al. High rates of quinolone-resistant strains of Shigella sonnei in HIV-infected MSM. Infection. 2013;41(5):999–1003.
28. Fischer N, Maex M, Mattheus W, et al. Genomic epidemiology of persistently circulating MDR Shigella sonnei strains associated with men who have sex with men (MSM) in Belgium (2013–19). J Antimicrob Chemother. 2022;77(1):89–97.
29. Ingle DJ, Andersson P, Valcanis M, et al. Prolonged Outbreak of Multidrug-Resistant Shigella sonnei Harboring $bla_{CTX-M-27}$ in Victoria, Australia. Antimicrob Agents Chemother. 2020; 64(12):e01518-20.
30. Locke RK, Greig DR, Jenkins C, et al. Acquisition and loss of CTX-M plasmids in Shigella species associated with MSM transmission in the UK. Microb Genomics. 2021;7(8):000644.
31. Simms I, Gilbart VL, Byrne L, et al. Identification of verocytotoxin-producing Escherichia coli O117:H7 in men who have sex with men, England, November 2013 to August 2014. Euro Surveill. 2014;19(43):20946.
32. Baker KS, Dallman TJ, Thomson NR, et al. An outbreak of a rare Shiga-toxin-producing Escherichia coli serotype (O117:H7) among men who have sex with men. Microb Genomics. 2018;4(7):e000181.
33. Kantele A, Mero S, Kirveskari J, et al. Fluoroquinolone antibiotic users select fluoroquinolone -resistant ESBL-producing Enterobacteriaceae (ESBL-PE) – Data of a prospective traveller study. Travel Med Infect Dis. 2017;16:23–30.
34. Bevan ER, Jones AM, Hawkey PM. Global epidemiology of CTX-M β-lactamases: temporal and geographical shifts in genotype. J Antimicrob Chemother. 2017;72(8):2145–55.
35. Nüesch-Inderbinen MT, Kayser FH, Hächler H. Survey and molecular genetics of SHV beta-lactamases in Enterobacteriaceae in Switzerland: two novel enzymes, SHV-11 and SHV-12. Antimicrob Agents Chemother. 1997;41(5):943.
36. Liakopoulos A, Mevius D, Ceccarelli D. A Review of SHV Extended-Spectrum β-Lactamases: Neglected Yet Ubiquitous. Front Microbiol. 2016;7:1374.

SUPPLEMENTARY MATERIAL

TABLE S1. ISOLATE CHARACTERISTICS.

Sample nr	Participant nr	Uberstrain (Enterobase)	Year of isolation	Assembly length	Nr of contigs	MLST	Phylo-group (clermont typing)	Trans-mission cluster	amoxicillin	cefuroxime	cefotaxime	ceftazidime	ciprofloxacin	colistin	fosfomycin
ESBL001	P01	ESC_XA1161AA	2018	5094934	96	131	B2		R	R	R	R	R	S	S
ESBL002	P02	ESC_XA1162AA	2018	5568824	193	648	F		R	R	R	R	R	S	S
ESBL003	P03	ESC_XA1163AA	2018	5390162	104	131	B2		R	R	R	R	R	S	S
ESBL004	P04	ESC_XA1164AA	2018	5365012	138	131	B2	C1	R	R	R	R	R	S	S
ESBL005	P05	ESC_XA1165AA	2018	5373323	147	38	D		R	R	R	R	S	S	S
ESBL006	P06	ESC_XA1166AA	2018	5447688	226	38	D		R	R	R	R	S	S	S
ESBL008	P07	ESC_XA1167AA	2018	5374097	111	131	B2	C1	R	R	R	R	R	S	S
ESBL011	P08	ESC_XA1171AA	2018	4970137	139	6438	A		R	R	R	R	S	S	S
ESBL012	P09	ESC_XA1172AA	2018	4976156	103	131	B2		R	R	R	R	R	S	S
ESBL013	P10	ESC_XA1174AA	2018	5174150	100	1193	B2		R	R	R	R	R	S	S
ESBL014	P11	ESC_XA1173AA	2018	5231690	82	998	B2		R	R	R	R	R	S	S
ESBL015	P12	ESC_XA1175AA	2018	4622837	93	3519	A		R	R	R	S	S	S	R
ESBL016	P12	ESC_XA1176AA	2018	4766049	44	1081	B1		R	R	R	S	R	S	R
ESBL017	P13	ESC_XA1177AA	2018	4658631	44	4681	B1		R	R	R	R	S	S	S
ESBL019	P14	ESC_XA1179AA	2018	5139048	135	2914	D		R	R	R	R	R	S	S
ESBL020	P15	ESC_XA1178AA	2018	5117676	158	10	A		R	R	R	R	S	S	S
ESBL021	P16	ESC_XA1180AA	2018	5100472	99	131	B2		R	R	R	S	R	S	S
ESBL022	P17	ESC_XA1188AA	2018	5242643	110	131	B2		R	R	R	R	R	S	S
ESBL023	P18	ESC_XA1181AA	2018	5192776	82	14	B2	C2	R	S	S	R	S	S	S
ESBL024	P19	ESC_XA1183AA	2018	5086723	126	131	B2		R	R	R	R	S	S	S
ESBL025	P20	ESC_XA1182AA	2018	5194551	105	2914	D		R	R	R	R	R	S	S
ESBL026	P20	ESC_XA1185AA	2018	5196019	91	14	B2	C2	R	R	R	R	R	S	S
ESBL027	P21	ESC_XA1186AA	2018	5099395	96	2599	B1		R	R	R	R	R	S	S
ESBL028	P22	ESC_XA1187AA	2018	4805900	294	3075	A		R	R	R	R	R	S	S
ESBL030	P23	ESC_XA1191AA	2018	5472291	128	2076	D		R	S	S	R	R	S	S
ESBL031	P24	ESC_XA1189AA	2018	4787688	107	744	A		R	R	R	R	R	S	S
ESBL032	P25	ESC_XA1193AA	2018	4848238	132	93	A		R	R	R	R	S	S	S
ESBL033	P26	ESC_XA1192AA	2018	5028226	108	131	B2		R	R	R	R	S	S	S
ESBL034	P27	ESC_XA1194AA	2018	5006663	178	10	A		R	R	R	R	R	S	S
ESBL035	P28	ESC_XA1195AA	2018	5244989	84	5148	D		R	R	R	R	R	S	S

R = resistant, S = susceptible, WT = Wild-type

gentamicin	imipenem	meropenem	nitrofurantoin	tobramycin	trimethoprim	blaCTX-M genes	blaOXA gene	blaCMY gene	blaSHV gene	gyrA S83& D87	parC S80	qnr gene
S	S	S	S	S	R	blaCTX-M-27	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	S	blaCTX-M-15 blaCTX-M-14	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	R	R	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-27	-	-	-	WT	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	-
S	S	S	R	R	S	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	D87Y	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L	parC_S80I	-
S	S	S	S	S	R	-	-	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-27	-	-	-	S83L	WT	-
S	S	S	S	S	S	blaCTX-M-14	-	-	-	WT	WT	-
S	S	S	S	S	R	blaCTX-M-14	-	-	-	WT	WT	qnrS1
S	S	S	S	S	S	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	-
S	S	S	S	S	R	blaCTX-M-134	-	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	S	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	-	-	-	blaSHV-12	S83L	WT	-
S	S	S	S	S	S	blaCTX-M-27	-	-	-	D87N	WT	-
S	S	S	S	S	S	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	-	-	-	blaSHV-12	S83L	WT	-
S	S	S	S	S	R	-	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-15	blaOXA-1	-	-	S83L	WT	-
S	S	S	S	S	R	-	-	-	-	WT	WT	qnrB19
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	S	-	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	S	-	-	-	-	S83L	WT	-
R	S	S	S	R	S	-	-	-	-	WT	WT	qnrS1
S	S	S	S	S	S	blaCTX-M-15	-	-	-	WT	WT	qnrS1

TABLE S1. ISOLATE CHARACTERISTICS. (CONTINUED)

Sample nr	Participant nr	Uberstrain (Enterobase)	Year of isolation	Assembly length	Nr of contigs	MLST	Phylo-group (clermont typing)	Trans-mission cluster	amoxicillin	cefuroxime	cefotaxime	ceftazidime	ciprofloxacin	colistin	fosfomycin
ESBL036	P29	ESC_XA1196AA	2018	5044858	167	1178	A		R	R	R	R	R	S	S
ESBL038	P30	ESC_XA1198AA	2018	4873094	79	554	A		R	R	R	R	S	S	S
ESBL039	P31	ESC_XA1201AA	2018	5015165	103	131	B2		R	R	R	R	R	S	S
ESBL042	P32	ESC_XA1203AA	2018	5272115	92	131	B2		R	R	R	R	R	S	S
ESBL044	P33	ESC_XA1204AA	2018	5268328	118	131	B2		R	R	R	R	R	S	S
ESBL045	P34	ESC_XA1205AA	2018	5132160	92	1193	B2		R	R	R	R	R	S	S
ESBL047	P35	ESC_XA1208AA	2018	5010534	73	131	B2		R	R	R	R	R	S	S
ESBL048	P35	ESC_XA1207AA	2018	5212412	80	14	B2	C3	R	R	R	R	R	S	S
ESBL050	P37	ESC_XA1209AA	2018	5472636	114	131	B2	C1	R	R	R	R	R	S	S
ESBL051	P38	ESC_XA1210AA	2018	5228433	117	2003	D		R	R	R	R	R	S	S
ESBL052	P38	ESC_XA1211AA	2018	5355323	158	131	B2		R	R	R	R	R	S	S
ESBL053	P39	ESC_XA1212AA	2018	4742666	271	3075	A		R	R	R	R	R	S	S
ESBL054	P40	ESC_XA1214AA	2018	5311962	129	5041	D		R	R	R	R	R	S	S
ESBL055	P41	ESC_XA1213AA	2018	5133495	181	3036	E		R	R	R	R	R	S	S
ESBL057	P42	ESC_XA1216AA	2018	5063521	71	394	D	C4	R	R	R	R	R	S	S
ESBL058	P43	ESC_XA1217AA	2018	4689448	82	10	A		R	R	R	R	R	S	S
ESBL059	P44	ESC_XA1218AA	2018	5195449	83	14	B2	C2	R	R	R	R	S	S	S
ESBL061	P45	ESC_XA1219AA	2018	4742480	280	3075	A		R	R	R	R	S	S	S
ESBL062	P45	ESC_XA1221AA	2018	4861114	265	3075	A		R	R	R	R	R	S	S
ESBL063	P46	ESC_XA1220AA	2018	4949369	83	636	B2		R	R	R	R	S	S	S
ESBL064	P47	ESC_XA1222AA	2018	5246266	136	1490	B1		R	R	R	R	R	S	S
ESBL066	P48	ESC_XA1231AA	2018	5388469	119	38	D		R	R	R	R	S	S	R
ESBL067	P49	ESC_XA1226AA	2018	5058461	70	394	D	C4	R	R	R	R	R	S	S
ESBL068	P50	ESC_XA1228AA	2018	5155466	83	131	B2		R	R	R	R	R	S	S
ESBL069	P51	ESC_XA1223AA	2018	5389489	115	131	B2	C1	R	R	R	R	R	S	S
ESBL070	P52	ESC_XA1224AA	2018	5202072	170	10	A		R	R	R	S	S	S	S
ESBL071	P53	ESC_XA1225AA	2018	5250729	112	69	D		R	R	R	R	R	S	S
ESBL073	P54	ESC_XA1229AA	2018	5178436	105	9791	A		R	R	S	R	S	S	S
ESBL075	P55	ESC_XA1232AA	2018	5023102	71	58	B1		R	R	R	R	S	S	S
ESBL076	P56	ESC_XA1234AA	2018	5321615	228	452	B2		R	R	R	R	R	S	S
ESBL077	P57	ESC_XA1236AA	2018	5485769	143	405	D		R	R	R	R	R	S	S
ESBL078	P58	ESC_XA1239AA	2018	4577600	75	226	A		R	R	R	R	R	S	S

R = resistant, S = susceptible, WT = Wild-type

gentamicin	imipenem	meropenem	nitrofurantoin	tobramycin	trimethoprim	blaCTX-M genes	blaOXA gene	blaCMY gene	blaSHV gene	gyrA S83& D87	parC S80	qnr gene
S	S	S	S	S	S	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-2	-	-	-	WT	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	R	R	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	S	blaCTX-M-27	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-27	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-27	-	-	-	S83L	WT	-
S	S	S	S	S	R	-	-	-	blaSHV-12	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-14	-	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-15	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-15	blaOXA-1	-	-	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L	WT	qnrB4
S	S	S	S	S	S	-	blaOXA-10	-	blaSHV-12	S83L	WT	qnrS1
S	S	S	S	S	R	-	-	-	blaSHV-12	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	blaOXA-1	-	-	S83L	WT	-
R	S	S	S	R	R	blaCTX-M-15	-	-	-	S83L&D87Y	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-15	-	-	-	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83A	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-27	-	-	-	WT	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L	WT	qnrB4
S	S	S	S	R	R	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
R	S	S	R	R	R	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	parC_S80I	qnrS1
R	S	S	S	R	R	-	blaOXA-10	-	-	WT	WT	qnrS5
R	S	S	S	S	S	-	blaOXA-10	-	-	WT	WT	qnrS5
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L&D87Y	parC_S80I	-
S	S	S	S	R	S	blaCTX-M-15	blaOXA-1	blaCMY-4	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	S	blaCTX-M-15	-	-	-	WT	WT	qnrS1

TABLE S1. ISOLATE CHARACTERISTICS. (CONTINUED)

Sample nr	Participant nr	Uberstrain (Enterobase)	Year of isolation	Assembly length	Nr of contigs	MLST	Phylo-group (clermont typing)	Trans-mission cluster	amoxicillin	cefuroxime	cefotaxime	ceftazidime	ciprofloxacin	colistin	fosfomycin
ESBL079	P58	ESC_XA1235AA	2018	5053125	66	1193	B2		R	R	R	R	R	S	S
ESBL081	P59	ESC_XA1238AA	2018	5520729	110	1380	D		R	R	R	R	S	S	S
ESBL082	P60	ESC_XA1240AA	2018	5037156	71	131	B2		R	R	R	R	R	S	S
ESBL083	P60	ESC_XA1242AA	2018	4844790	161	227	A		R	R	R	R	R	S	S
ESBL084	P61	ESC_XA1241AA	2018	5225677	124	131	B2		R	R	R	R	R	S	S
ESBL085	P62	ESC_XA1243AA	2018	5269751	127	394	D		R	R	R	R	S	S	S
ESBL086	P63	ESC_XA1244AA	2018	5014630	140	4060	B2		R	R	R	R	R	S	S
ESBL087	P64	ESC_XA1245AA	2018	4799082	98	34	A		R	R	R	S	S	S	S
ESBL088	P65	ESC_XA1246AA	2018	5395287	129	131	B2	C1	R	R	R	R	R	S	S
ESBL089	P65	ESC_XA1248AA	2018	5510985	115	69	D		R	R	R	R	R	S	S
ESBL090	P66	ESC_XA1249AA	2018	5461444	207	1136	B1		R	R	R	R	R	S	R
ESBL091	P67	ESC_XA1251AA	2018	5493504	126	38	D		R	R	R	R	S	S	S
ESBL093	P68	ESC_XA1275AA	2018	5152830	66	1193	B2		R	R	R	R	R	S	S
ESBL094	P69	ESC_XA1274AA	2018	5274703	168	69	D		R	R	R	R	R	S	S
ESBL095	P70	ESC_XA1273AA	2018	5199665	168	10	A		R	R	R	S	S	S	S
ESBL096	P71	ESC_XA1252AA	2018	4802820	298	3075	A		R	R	R	R	R	S	S
ESBL097	P72	ESC_XA1253AA	2018	5195048	91	14	B2	C2	R	R	R	R	S	S	S
ESBL099	P73	ESC_XA1255AA	2018	4804038	208	757	A		R	R	R	R	R	S	S
ESBL100	P74	ESC_XA1257AA	2018	5353916	113	38	D		R	R	R	R	R	S	S
ESBL102	P74	ESC_XA1256AA	2018	5268425	88	131	B2		R	R	R	R	R	S	S
ESBL103	P75	ESC_XA1258AA	2018	5182559	128	131	B2		R	R	R	R	R	S	S
ESBL104	P76	ESC_XA1259AA	2018	4748097	282	3075	A		R	R	R	R	R	S	S
ESBL105	P77	ESC_XA1261AA	2018	5026718	65	1193	B2		R	R	R	R	R	S	S
ESBL107	P78	ESC_XA1264AA	2018	5213142	129	14	B2	C3	R	R	S	R	S	S	S
ESBL109	P79	ESC_XA1263AA	2018	5512412	190	38	D		R	R	R	R	S	S	S
ESBL110	P20	ESC_XA1267AA	2019	5517651	195	62	F		R	R	R	R	S	S	S
ESBL111	P81	ESC_XA1265AA	2019	4829271	128	226	A		R	R	R	R	R	S	S
ESBL113	P45	ESC_XA1269AA	2019	5299969	208	1136	B1		R	R	R	S	R	S	S
ESBL114	P45	ESC_XA1266AA	2019	5407489	151	131	B2	C1	R	R	R	R	R	S	S
ESBL115	P42	ESC_XA1271AA	2019	5127621	80	131	B2		R	R	R	R	R	S	S
ESBL116	P65	ESC_XA1270AA	2019	5136911	189	746	A		R	R	R	R	S	S	S

R = resistant, S = susceptible, WT = Wild-type

gentamicin	imipenem	meropenem	nitrofurantoin	tobramycin	trimethoprim	blaCTX-M genes	blaOXA gene	blaCMY gene	blaSHV gene	gyrA S83& D87	parC S80	qnr gene
R	S	S	S	R	S	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-27	-	-	-	WT	WT	-
S	S	S	S	S	S	blaCTX-M-15	-	-	-	S83L	WT	-
R	S	S	S	R	S	blaCTX-M-15	-	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	S	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	S	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-27	-	-	-	S83L	WT	-
S	S	S	S	S	S	blaCTX-M-1	-	-	-	WT	WT	-
R	S	S	S	R	R	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-15	-	-	-	S83L	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	S	blaCTX-M-27	-	-	-	WT	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-15	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-15	blaOXA-1	-	-	S83L	WT	-
S	S	S	S	S	R	-	-	-	blaSHV-12	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L	WT	qnrS1
S	S	S	S	S	R	blaCTX-M-27	-	-	-	WT	WT	-
R	S	S	S	R	S	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-27	-	-	-	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	blaOXA-1	-	-	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-27	-	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	-	-	-	blaSHV-12	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-27	-	-	-	WT	WT	-
R	S	S	S	S	R	blaCTX-M-24	-	-	-	S83L	WT	-
S	S	S	S	S	R	blaCTX-M-15	-	-	-	WT	WT	qnrS1
R	S	S	S	R	R	blaCTX-M-15	blaOXA-1	-	-	S83L&D87N	parC_S80I	-
S	S	S	S	S	R	blaCTX-M-27	-	-	-	S83L&D87N	parC_S80I	-
R	S	S	S	R	R	blaCTX-M-15	-	-	-	S83L	WT	-

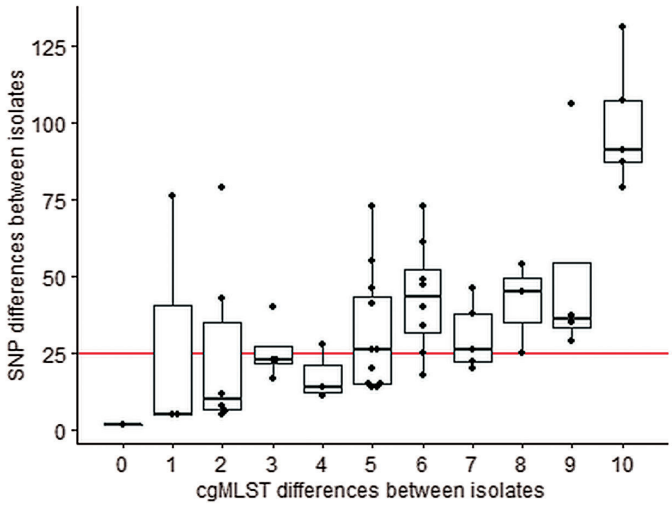


FIGURE S1. THE RELATION BETWEEN CGMLST DIFFERENCES AND SNP DIFFERENCES FOR 53 ISOLATE PAIRS WHICH DIFFERED IN 10 OR FEWER CGMLST LOCI IN OUR COLLECTION. THE RED HORIZONTAL LINE INDICATES THE SNP CUT-OFF FOR PUTATIVE TRANSMISSION USED IN THIS STUDY.

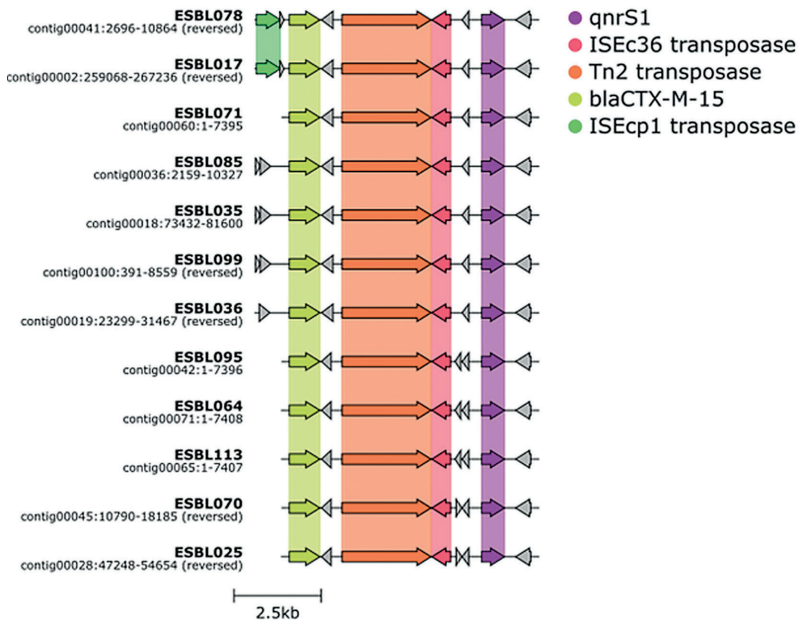


FIGURE S2. ALIGNMENT OF TN2 HARBOURING $BLA_{CTX-M-15}$ AND *QNRS1*. GREY ARROWS INDICATE HYPOTHETICAL PROTEINS. VISUALISED USING CLINKER.

5

PART II

USING GENOMICS TO UNDERSTAND GONOCOCCAL BIOLOGY


6

WITHIN-HOST GENETIC VARIATION IN *NEISSERIA GONORRHOEAE* OVER THE COURSE OF INFECTION

Jolinda de Korne-Elenbaas, Sylvia M. Bruisten, Henry J. C. de Vries, Alje P. van Dam

Microbiology Spectrum 2022; 10(3):e00313-22

ABSTRACT

Knowledge of within-host genetic variation informs studies on transmission dynamics. We studied within-host genetic variation in *Neisseria gonorrhoeae* over the course of infection and across different anatomical locations. Isolates were obtained during a clinical trial, and isolates from consecutive time points reflected persistent infections after treatment failure. We compared sequence types (STs) and recombination unfiltered- and filtered core genome single nucleotide polymorphism (SNP) distances in 65 within-host isolate pairs from the same anatomical location over time—obtained with a median interval of 7 days—and 65 isolate pairs across different anatomical locations at one time point. Isolates with different Multi-Locus Sequence Types (MLST), NG-Sequence Types for Antimicrobial Resistance (NG-STAR) and NG-Multi Antigen Sequence Types (NG-MAST) had a median of 1466 recombination filtered SNPs, whereas a median of 1 SNP was found between isolates with identical STs or a different NG-MAST only. The threshold for differentiating between strains was set at 10 recombination filtered SNPs, showing that isolates from persistent infections could have different NG-MASTs. Antibiotic pressure applied through treatment did not lead to an increase in genetic variation in specific genes or in overall extent of variation, compared to variation across anatomical locations. Instead, within-host genetic variation was proposedly driven by the host immune response, as it was concentrated in genomic regions encoding surface exposed proteins involved in host-microbe interaction. Ultimately, 15/228 (6.5%) between-host pairs contained a single strain, suggesting between-host transmission. However, patient reported data are needed to differentiate within-host persistence from between-host transmission.

INTRODUCTION

Neisseria gonorrhoeae infection is one of the most common bacterial sexually transmitted infections worldwide and the global increasing prevalence causes a high burden to public health[1]. Because of the development of multidrug resistance in *N. gonorrhoeae*, limiting the spread of *N. gonorrhoeae* infections is needed a fortiori. Therefore, prevention strategies as well as prompt treatment of gonorrhoea are essential. However, this is complicated by asymptomatic infections that often stay unnoticed. These asymptomatic infections, mainly rectal or pharyngeal, predominantly occur in men who have sex with men (MSM) and in women, and drive ongoing transmission[2,3].

Research on transmission dynamics informs public health interventions targeting key populations. These interventions are based on epidemiological and behavioural factors associated with transmission within these populations, such as sexual behaviour (e.g., chemsex), sex work, and PrEP use[4,5]. To understand transmission dynamics, ascertainment of a transmission event between two individuals is crucial. To study *N. gonorrhoeae* transmission, the *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) scheme has been created, based on two hypervariable genes[6]. Comparing core genomes, instead of single genes, increases the resolution for assessing genetic relatedness between isolates and for identification of putative transmission events. For *N. gonorrhoeae*, core genome SNP distance thresholds have been suggested for variation between isolates linked by transmission[7]. These thresholds can be defined by comparing genetic variation that occurs within and between hosts, with an example of within-host genetic variation being the variation observed across different anatomical locations that are infected with the same bacterial strain. Definition of a fixed threshold is hampered by high recombination rates in *N. gonorrhoeae*, since recombination events can lead to high SNP counts between two isolates that are identical in the rest of their genomes. Recombination filtering is therefore often applied when calculating SNP distances, although this could inflate the number of closely related isolates[8].

SNP distance thresholds to distinguish between *N. gonorrhoeae* strains can be used for a variety of applications. As mentioned, key populations for *N. gonorrhoeae* infection are defined based on transmission networks, identified by determining SNP distances between isolates from that population[4,5]. Modern applications provide improved partner notification based on SNP distances between isolates, potentially leading to enhanced identification of partner links[9]. SNP-based methods can also be used to assess whether a bacterial strain has persisted over time or is acquired through reinfection, as done for *Mycobacterium tuberculosis* and different *Shigella* species[10,11]. This also provides insight in the course of infection and in variable genomic regions involved in the host-microbe interaction. For the latter purposes and to assess within-host genetic variation, the current study examined a unique collection of *N.*

gonorrhoeae isolates obtained during study visits of the New Antibiotic Treatment Options for Uncomplicated Gonorrhoea clinical trial (Dutch acronym: NABOGO). This trial was performed from 2017 to 2020 at the Centre for Sexual Health of Amsterdam, the Netherlands, and assessed whether gentamicin, ertapenem or fosfomycin were novel treatment options for gonorrhoea[12]. From cases of treatment failure, within-host isolates could be obtained from consecutive time points. To verify that treatment was truly in-effective and participants had persistent infections, reinfections with a different strain were previously ruled out based on NG-MAST typing[12]. In the current study, we performed more detailed and comprehensive genetic analyses using these isolates. The availability of these isolates enabled the examination of within-host genetic changes that occur over the course of a *N. gonorrhoeae* infection. Also, isolates from the same individual from multiple anatomical locations were obtained at one visit during the NABOGO trial. We compared the genomic variation found over time and across multiple anatomical locations to assess whether the human immune response or antibiotic pressure induced genetic variation during the infection period, either across the genome or at specific genomic regions. Ultimately, we compared within- and between-host genetic variation, to determine if cases of within-host persistence and between-host transmission could be differentiated based on core genome SNP distances.

METHODS

STUDY PARTICIPANTS AND ISOLATES

N. gonorrhoeae isolates were obtained during the NABOGO clinical trial, performed from 18 September 2017 to 5 June 2020 at the Centre for Sexual Health of Amsterdam, the Netherlands. Centre visitors who tested *N. gonorrhoeae*-positive with Nucleic Acid Amplification Test (NAAT) were asked to participate and to refrain from sexual intercourse during the study period. At the inclusion visit, participants were treated according to random assignment to one of the treatment arms (ertapenem, gentamicin, fosfomycin or ceftriaxone). Test-of-cure diagnostics were performed 7–14 days after treatment. Participants could come back before or after the test-of-cure visit in case of persisting or worsening symptoms and escape medication with ceftriaxone was then given. Routine diagnostic tests were also performed on swabs obtained during these visits. More details on the study procedure have been described earlier[12].

At each study visit, anal-, pharyngeal and urethral or vaginal swabs were obtained from participants for NAAT and for phenotypic characterisation using culturing. Whereas in the previous report[12] we defined treatment failure as a *N. gonorrhoeae*-positive NAAT at the test-of-cure visit, here we could only include participants from whom *N. gonorrhoeae* isolates from before and after treatment were available. In addition to the isolates from consecutive time points, *N. gonorrhoeae* isolates were included from all participants who had isolates available

from multiple anatomical locations at one time point. As a result, the within-host isolate pairs obtained from a single individual were derived from i) two consecutive time points from the same anatomical location or ii) different anatomical locations at the same time point, referred to as, respectively, i) times-pair or ii) locations-pair throughout the manuscript.

WGS AND QUALITY ASSESSMENT

After phenotypic characterisation, isolates were stored at –80°C. For WGS, isolates were taken from storage and cultured overnight on a chocolate blood agar plate. DNA was extracted from pure cultures in DNA/RNA shield buffer using the ZymoBIOMICS TM Magbead DNA kit (ZYMO RESEARCH). DNA sequencing libraries were prepared using the Nextera XT DNA Library Preparation kit with Integrated DNA Technologies for Illumina DNA/RNA Unique Dual Indexes (Illumina). Short-read sequencing was done using Illumina NovaSeq 6000.

Raw reads were trimmed and filtered using fastp v0.20.1[13]. Reads were mapped to reference genome FA1090 (NC_002946.2) using BWA-MEM2 v2.2.1 to calculate the percentage of bases covered and the mean coverage depth using the SAMtools package v1.9[14,15]. Reads were assembled with Skesa v2.4.0 and assembly quality was assessed with QUAST v5.0.2[16,17]. In case of a final assembly length of >2.1 Mb, Kraken2 v2.1.1 was used to identify contamination and to filter out reads that did not belong to *N. gonorrhoeae* (taxid:485)[18]. Filtered reads were again assembled. Isolates with >95% coverage of reference genome and with a mean coverage depth of >50× were included in the analyses.

TYPING AND SNP DISTANCE DETERMINATION

Assemblies were uploaded to the PubMLST Neisseria database and automatically annotated, after which Multi-Locus Sequence Types (MLST), NG-Sequence Types for Antimicrobial Resistance (NG-STAR), NG-MAST v2.0 STs and core genome MLST (cgMLST) v1.0 alleles were extracted[19]. When alleles were not annotated in PubMLST, sequences were manually extracted and aligned to determine similarity of alleles between paired isolates. CgMLST allele distances were determined using cgmlst-dists v0.4.0 (https://github.com/tseemann/cgmlst-dists). For variant calling, reads were mapped on reference genome FA1090 (NC_002946.2) and SNPs were identified with Snippy v4.6.0 (https://github.com/tseemann/snippy). Default settings were used: SNPs were reported with a minimum read coverage of 10×, a minimum base quality of 13 and a read concordance of 90%. A core genome alignment was created using the Snippy-core option. Recombination was filtered out using Gubbins v2.4.1[20] and masked in the core genome alignment using the maskrc-svg script v0.5 (https://github.com/kwongj/maskrc-svg). Recombination filtered and unfiltered SNP distances between all isolates were determined using the masked- or unmasked core genome alignment with snp-dists v0.7.0 (https://github.com/tseemann/snp-dists). Snakemake v5.31.1 was used for workflow management[21].

COMPARING GENE-BASED TYPING, CORE GENOME-BASED TYPING AND SNP DISTANCES

Gene-based typing results (MLST, NG-STAR, and NG-MAST STs), core genome-based allele distances (cgMLST) and recombination unfiltered and filtered SNP distances were compared between within-host paired isolates to access the discriminatory power of the different methods. Also, comparing within-host paired isolates enabled the definition of a SNP threshold to differentiate between strains. Using this threshold, within-host times-pairs with a single strain were identified as persistent infections due to treatment failure. Times-pairs or locations-pairs with distinct strains were identified as, respectively, reinfection- or coinfection with distinct strains (Table 1). These pairs were excluded from analyses on within-host genetic variation, since these are not representative for within-host genetic variation that occurs during infection.

IDENTIFICATION OF VARIABLE GENOMIC REGIONS BY VISUALIZING GENOMIC LOCATIONS OF UNFILTERED SNPS

The genomic locations of unfiltered SNPs found in within-host pairs with a single strain were visualized to identify hot spots of mutations or recombination in the genome, potentially induced by the human immune response or antibiotic pressure. For this purpose, pairs were selected from the core genome alignment and snp-sites v2.5.1 was used to create vcf files with genomic locations of the SNPs found between isolates on reference genome FA1090[31]. The previously defined SNP threshold was used to identify pairs with distinct strains, and these were excluded from this analysis. The FA1090 reference genome was visualized in Artemis v18.1.0 together with the vcf files, to plot the SNP density across the genome[23]. Also, SNP locations in the vcf files were annotated with bcftools v1.9 using the annotate option and the prevalence of SNPs in each gene of the FA1090 genome was extracted from the annotated vcf files[24].

COMPARING WITHIN- AND BETWEEN-HOST GENETIC VARIATION

We assessed whether SNP distances could distinguish within- and between-host isolates. Isolates from two different participants formed between-host isolate pairs. Pairs were only included in the analysis if at least one of the three STs was identical (MLST, NG-STAR or NG-MAST), since the high SNP distances between isolates that differed in all STs were irrelevant. A participant could have multiple isolate pairs when isolates were available from multiple anatomical locations and from multiple time points. To prevent sampling bias caused by multiple isolates from the same participant, each participant (within-host comparisons) or each combination of participants (between-hosts comparisons) was represented only once in this analysis. Between-host transmission was defined as a single strain in a between-host pair, using the previously defined SNP threshold (Table 1).

TABLE 1. DEFINITIONS OF EVENTS BASED ON EXPECTED PATIENT REPORTED METADATA AND MOLECULAR DATA.

	Expected patient reported data	**Expected molecular data**
Within-host persistence (treatment failure)	Patient reports no sexual contact between the sampling time points.	The same strain* at consecutive time points, obtained from a single individual (before and after treatment).
Within-host coinfections with distinct strains at different anatomical locations	Patient reports sexual contact with one or more sexual partners before sampling.	The same strain* at different anatomical locations of a single individual.
Reinfection with a distinct strain	Patient reports sexual contact between the sampling time points (probably with different partners).	Distinct strain# at consecutive time points, obtained from a single individual.
Reinfection with the same strain	Patient reports sexual contact between sampling time points (probably with the same partner).	The same strain* at consecutive time points, obtained from a single individual.
Between-host transmission	Patient reports sexual contact with one or more partners before sampling.	A single strain* obtained from patient and partner.

**Same strain: <10 SNPs between isolates. #Distinct strains: ≥10 SNPs between isolates.*

RESULTS

SEQUENCE QUALITY AND ISOLATE PAIRS

WGS was performed on 203 isolates from 80 unique participants: 74 MSM, 1 heterosexual man, 3 bisexual men and 2 women. All isolates passed quality control with a mean reference genome coverage of 98.7% (range 98.02–99.2%) and mean coverage depth of 439× (range 62–1079×) (Table S1). In total, the 203 isolates formed 65 within-host locations-pairs and 65 within-host times-pairs (Table S2). For 15 participants both locations-pairs and times-pairs were available. The 65 times-pairs were obtained from 41 participants who had *N. gonorrhoeae* isolates available from either 2 (24/41), 3 (16/41) or 4 (1/41) time points. Days between the consecutive time points ranged from 2–23 days (median 7 days). The 65 locations-pairs were obtained from 54 participants, of whom 49 participants had a single locations-pair, 5 participants had locations-pairs from 2 time points and 3 participants had isolates from 3 anatomical locations at a single time point, which constituted 2 locations-pairs per participant. Both times- and locations-pairs were mainly obtained from men (respectively, 100% and 96%) who were MSM (respectively, 93% and 91%) and no times-pairs were obtained from women. Participants with times-pairs were predominantly allocated to the fosfomycin treatment arm (73%) whereas participants with locations-pairs were evenly distributed across the treatment arms (Table 2).

TABLE 2. CHARACTERISTICS OF PARTICIPANTS FROM THE NABOGO CLINICAL TRIAL FROM WHOM TIMES-PAIRS AND/OR LOCATIONS-PAIRS WERE OBTAINED.

	Times-pairs (N=65)	**Locations-pairs (N=65)**
Number of unique participants	**41**	**54**
Sex		
Male	41 (100%)	52 (96%)
Female	-	2 (4%)
Sex group		
MSM	38 (93%)	49 (91%)
Heterosexual	1 (2%)	2 (4%)
Bisexual	2 (5%)	3 (5%)
Allocated treatment arm		
Ceftriaxone	-	14 (26%)
Ertapenem	1 (3%)	18 (33%)
Gentamicin	10 (24%)	14 (26%)
Fosfomycin	30 (73%)	8 (15%)

THE RELATIONSHIP BETWEEN GENE-BASED TYPING, CORE GENOME-BASED TYPING AND SNP DISTANCES

Typing results and unfiltered- and recombination filtered SNPs were compared between within-host paired isolates. Pairs with identical gene-based STs (MLST, NG-STAR and NG-MAST) and pairs that differed in NG-MAST only had a median cgMLST allele distance of 0 (range 0–15), whereas pairs with different STs had a median cgMLST allele distance of 729 (range 215–949) (Figure S1). High unfiltered SNP distances were found for pairs with different MLST, NG-STAR and NG-MAST STs, with a median of 5101 SNPs (range 1790–6432 SNPs). After recombination filtering, the SNP distances decreased to a median of 1466 SNPs (range 186–2345 SNPs)(Figure 1a). Pairs with different NG-MAST only had much lower SNP distances that were similar to the SNP distances between pairs with identical MLST, NG-STAR and NG-MAST STs, with a median of, respectively, 9 (5–145) and 8 (range 0–82) unfiltered SNPs. After recombination filtering, the median SNP distance was 1 SNP for both groups with a maximum of 7 SNPs, showing that outliers were filtered out (Figure 1a). Based on these results, the SNP threshold for isolates of the same strain was set at <10 recombination filtered SNPs. Isolates with <10 SNPs could have different NG-MAST STs but were still considered the same strain. When using <10 recombination filtered SNPs as reference, both MLST and NG-STAR typing methods differentiated strains in the 130 within-host pairs with 100% sensitivity and specificity, whereas NG-MAST had 93% sensitivity and 100% specificity (Table S3).

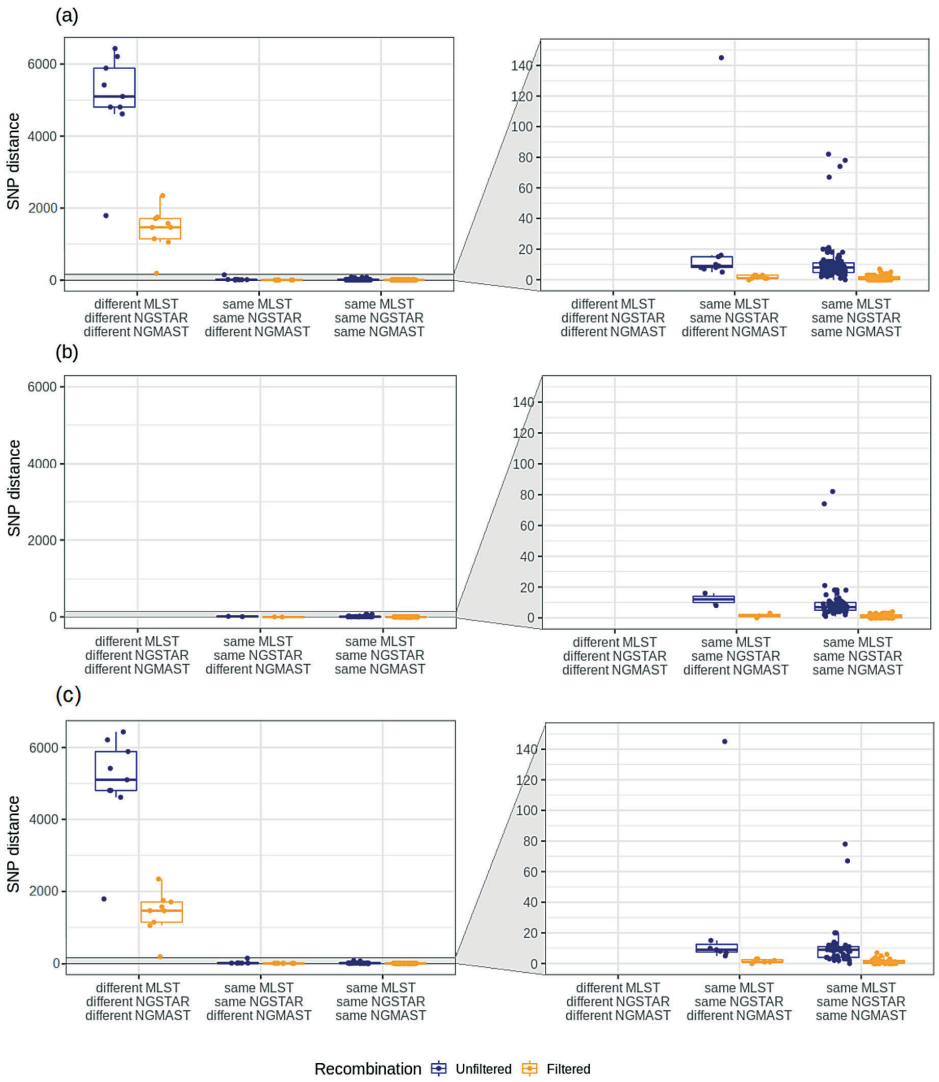


FIGURE 1. GENE-BASED TYPING RESULTS VERSUS RECOMBINATION FILTERED- AND UNFILTERED SNP DISTANCES BETWEEN ISOLATES IN WITHIN-HOST PAIRS. MLST, NG-STAR, NG-MAST STS AND SNP DISTANCES WERE COMPARED BETWEEN WITHIN-HOST PAIRED ISOLATES. RIGHT PANELS MAGNIFY THE LOWER VALUES ON THE Y-AXIS. (A) SNP DISTANCES FOUND IN ALL WITHIN-HOST TIMES-PAIRS. HIGH SNP DISTANCES WERE FOUND BETWEEN ISOLATES WITH DIFFERENT MLST, NG-STAR AND NG-MAST STS WHEREAS ISOLATES THAT DIFFERED IN NG-MAST ONLY OR ISOLATES WITH IDENTICAL MLST, NG-STAR AND NG-MAST STS HAD COMPARABLE SNP DISTANCES. ZOOM-IN ON THE LOWER SNP DISTANCES CONFIRMED THAT SIMILAR SNP DISTANCES WERE FOUND BETWEEN ISOLATES WITH DIFFERENT NG-MAST COMPARED TO ISOLATES WITH IDENTICAL STS. RECOMBINATION FILTERING REDUCED THE SNP DISTANCE TO <10 SNPS FOR COMPARISONS BETWEEN ISOLATES WITH IDENTICAL STS AND ISOLATES THAT DIFFERED IN NG-MAST ONLY. (B) SNP DISTANCES FOUND IN WITHIN-HOST TIMES-PAIRS. NO TIMES-PAIRS WITH DIFFERENT MLST, NG-STAR AND NG-MAST STS WERE FOUND AND ALL TIMES-PAIRS HAD RECOMBINATION FILTERED SNP DISTANCES <10 SNPS. (C) SNP DISTANCES FOUND IN WITHIN-HOST LOCATIONS-PAIRS. ISOLATES WITH DIFFERENT MLST, NG-STAR AND NG-MAST STS AND HIGH SNP DISTANCES WERE EXCLUSIVELY FOUND IN LOCATIONS-PAIRS.

NO REINFECTIONS WITH DISTINCT STRAINS OCCURRED IN PARTICIPANTS OF THE NABOGO TRIAL

Comparing MLST, NG-STAR and NG-MAST STs between within-host paired isolates showed identical STs in 63/65 (97%) times-pairs and 49/65 (75%) locations-pairs (Table 3). Different NG-MAST STs were found in 2/65 (3%) times-pairs and 7/65 (8%) locations-pairs, obtained from, respectively, 2 and 5 participants. The 9 isolate pairs that had identical MLST and NG-STAR STs but different NG-MAST STs all contained a single strain, based on <10 recombination filtered SNPs (Figure 1a). Therefore, no reinfections with distinct strains occurred in participants of the NABOGO trial. However, 9/65 (14%) locations-pairs contained distinct strains, according to different MLST, NG-STAR and NG-MAST STs and high recombination filtered SNP distances (Figure 1c). These were obtained from 8 participants, thus 8/54 (15%) participants were coinfected with distinct strains at different anatomical locations.

TABLE 3. TYPING RESULTS OF ISOLATES IN WITHIN-HOST TIMES-PAIRS AND LOCATIONS-PAIRS.

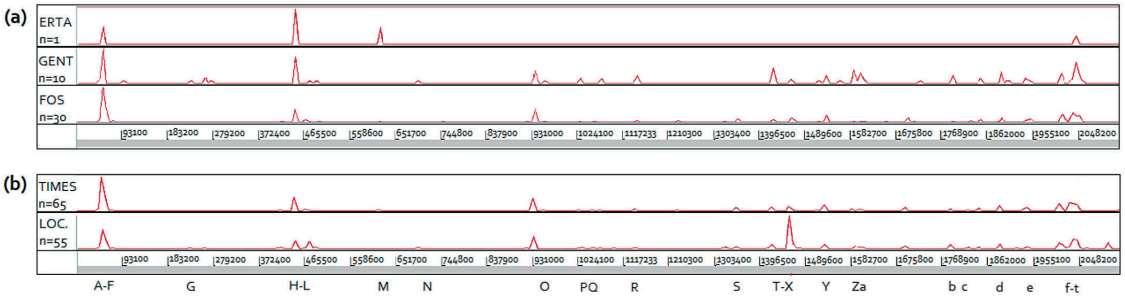
	Times-pairs (n= 65)	**Locations-pairs (n=65)**
Identical MLST, NG-STAR and NG-MAST	63 (97%)	49 (75%)
Different NG-MAST, identical MLST and NG-STAR	2 (3%)	7 (11%)
Different MLST, NG-STAR and NG-MAST	-	9 (14%)

SAME EXTENT OF GENETIC VARIATION FOUND IN TIMES-PAIRS AND LOCATIONS-PAIRS

SNP distances in within-host times-pairs and locations-pairs with a single strain were comparable, with a median of, respectively, 7 (range 1–82 SNPs) and 9 (range 0–145 SNPs) unfiltered SNPs (Figures 1b and c). Both for within-host times-pairs and locations-pairs, the median recombination filtered SNP distance was 1 (ranges, respectively, 0–4 and 0–7 SNPs). These results show that the same extent of genetic variation arose during the infection period of at most 23 days after antibiotic treatment as well as across different anatomical locations at a single time point.

GENETIC VARIATION WAS CONCENTRATED IN GENOMIC REGIONS ENCODING HYPERVARIABLE PROTEINS

When visualizing genomic locations of unfiltered SNPs in within-host pairs with a single strain on reference genome FA1090, times-pairs were categorized into allocated antibiotic treatment arms. Similar genomic regions with high density SNPs were found for the different treatment arms (Figure 2a). Also, similar variable genomic regions were identified in times-pairs and locations-pairs (Figure 2b). Most genetic variation was found in the pilus (assembly) proteins (specifically B: NGO-RS00260 and d: NGO-RS09615), transferrin-binding proteins *tbpA* and *tbpB* (V: NGO-RS07420 and W: NGO-RS07425), bifunctional protein *putA* (Y: NGO-RS07715) and an



(c)

	Genbank locus tag	Gene product	Number of times-pairs with SNP in gene	Number of locations-pairs with SNP in gene		Genbank locus tag	Gene product	Number of times-pairs with SNP in gene	Number of locations-pairs with SNP in gene
A	NGO_RS00235	carbamoyl-phosphate synthase large subunit (carB)	3	1	a	NGO_RS08190	Hypothetical protein	6	5
B	NGO_RS00260	Pilus assembly/ adherence protein (pilC)	22	13	b	NGO_RS09055	PorinB (porB)	6	8
C	NGO_RS00265	Iron-sulfur cluster carrier protein (apbC)	5	3	c	NGO_RS09450	Transposase	9	5
D	NGO_RS00275	Protein disulfide oxidoreductase	3	1	d	NGO_RS09615	Pilus assembly protein	14	11
E	NGO_RS00295	EamA family membrane transporter	3	1	e	NGO_RS09920	Maf protein	9	7
F	NGO_RS00300	Formate-tetrahydrofolate ligase	3	1	f	NGO_RS10310	Pilus protein	3	0
G	NGO_RS01160	Restriction endonuclease	3	2	g	NGO_RS10320	Pilus protein	4	4
H	NGO_RS02325	Replicative DNA helicase	8	6	h	NGO_RS10325	Pilus protein	1	3
I	NGO_RS02330	Tfp pilus assembly protein (fimT/fimU)	7	6	i	NGO_RS10330	Pilus protein	3	0
J	NGO_RS02335	Type IV pilus modification protein (pilV)	5	1	j	NGO_RS10430	Pilus protein	3	2
K	NGO_RS02375	Peroxide stress protein (yaaA)	0	3	k	NGO_RS10440	Pilus protein	2	5
L	NGO_RS02540	Phage portal protein	4	2	m	NGO_RS10455	Pilus protein	3	3
M	NGO_RS03155	Transposase	4	0	n	NGO_RS10465	Pilus protein	2	3
N	NGO_RS03530	Type I RM system subunit M	1	4	o	NGO_RS10480	Pilus protein	3	0
O	NGO_RS04840	Autotransporter outer membrane beta-barrel domain-containing protein	4	0	p	NGO_RS10510	Amino acid permease	15	20
P	NGO_RS05215	Polymorphic toxin MafB class 3	6	2	q	NGO_RS11195	Pilus protein	3	0
Q	NGO_RS05370	Hypothetical protein	3	4	r	NGO_RS11200	Pilus protein	4	1
R	NGO_RS05905	Transposase	9	3	s	NGO_RS11485	Pilus protein	6	1
S	NGO_RS06895	Hypothetical protein	3	0	t	NGO_RS11490	Pilus protein	7	1
T	NGO_RS07235	Transposase	0	3					
U	NGO_RS07270	Hypothetical protein	2	3					
V	NGO_RS07420	Transferrin-binding protein (tbpA)	8	2					
W	NGO_RS07425	Transferrin-binding protein (tbpB)	11	9					
X	NGO_RS07660	Methyltransferase (rsmH)	3	1					
Y	NGO_RS07715	Bifunctional protein (putA)	28	16					
Z	NGO_RS08155	Hypothetical protein	6	1					

FIGURE 2. WITHIN-HOST GENETIC VARIATION WAS MAINLY LOCATED ON REGIONS ENCODING HYPERVARIABLE PROTEINS. THE DENSITY OF SNPS FOUND FOR WITHIN-HOST (A) TIMES-PAIRS, CATEGORIZED ON THE ALLOCATED TREATMENT ARMS, AND (B) TIMES-PAIRS AND LOCATIONS-PAIRS, VISUALIZED WITH THE POSITIONS ON REFERENCE GENOME FA1090. LETTERS BELOW THE GRAPHS CORRESPOND TO GENES LOCATED IN THAT REGION AND IN WHICH MULTIPLE SNPS WERE FOUND, FURTHER SPECIFIED IN (C). (C) LEGEND SHOWING THE GENE ID AND NAMES THAT CORRESPOND TO LETTERS BELOW GRAPH (B), AND THE NUMBER OF TIMES- AND/OR LOCATION PAIRS THAT HAD VARIATION IN THAT GENE. ONLY GENES ARE SHOWN IN WHICH VARIATION WAS IDENTIFIED IN MORE THAN 2 PAIRS.

amino acid permease (p: NGO-RS10510) and this was found in both times-pairs and locations-pairs (Figure 2c). Remarkably, much variation was present in the genes used for NG-MAST typing (*porB* and *tbpB*)(Figure 2c), which in some cases made up the majority of SNPs found between within-host paired isolates. In regard to the 9 within-host pairs that differed in NG-MAST STs only, 7/9 had mutations in *porB*, 1/9 had mutations in *porB* and *tbpB* and none had mutations in only *tbpB*.

BETWEEN-HOST PAIRS WITH A SINGLE STRAIN SUGGEST CASES OF TRANSMISSION BETWEEN PARTICIPANTS

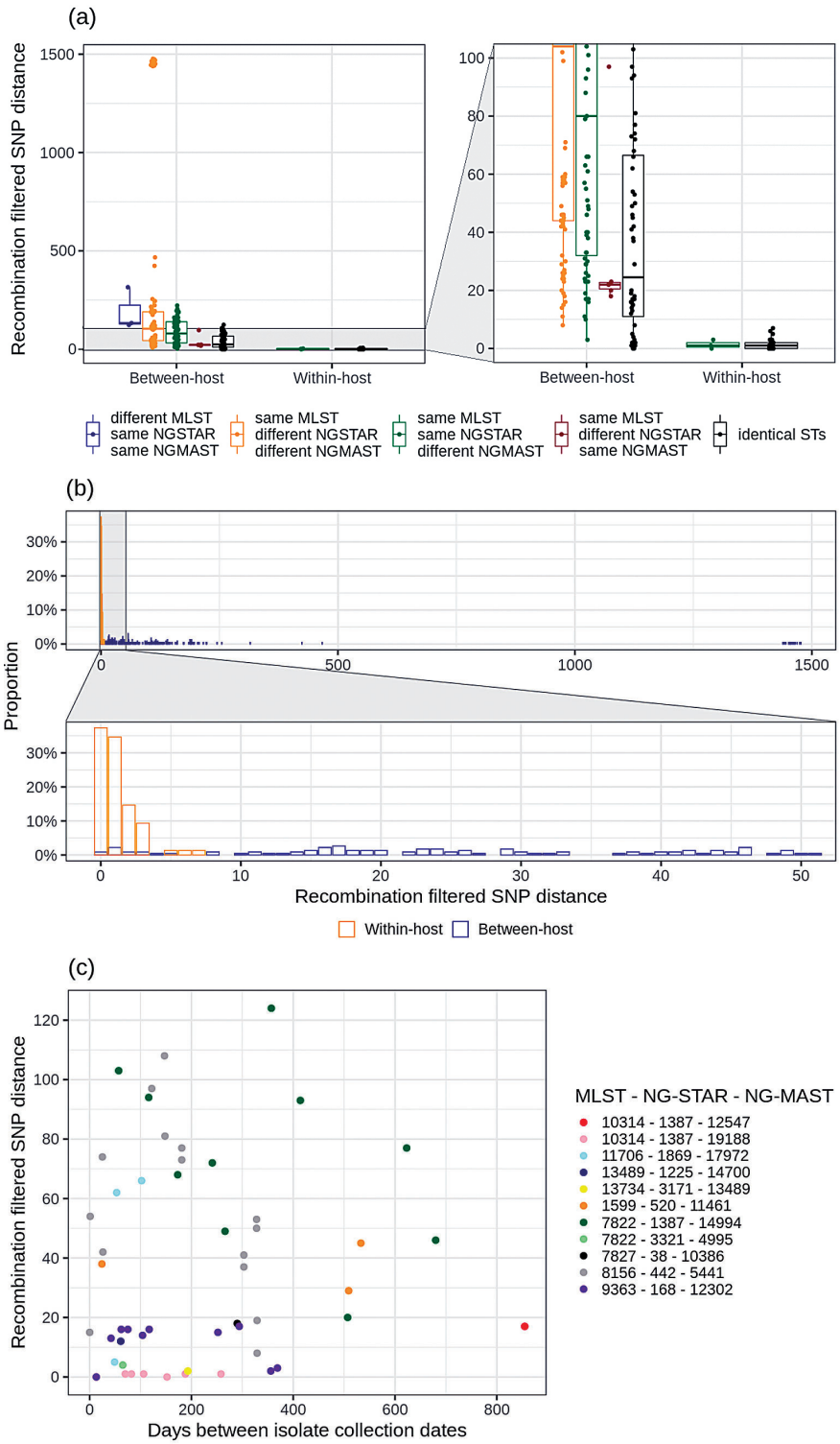
For the comparison of within- and between-host genetic variation, we used paired isolates that had at least one of the MLST, NG-STAR or NG-MAST STs identical. Seventy-five within-host pairs (only one pair per participant) and 228 between-host pairs were included (Table S4). Comparing SNP distances and STs showed broad ranges of SNP distances for between-host pairs with differences in any ST, but also between-host pairs with identical STs had up to 124 SNPs. In contrast to our finding in within-host pairs, all but one between-host pairs with different NG-MAST had ≥10 SNPs (Figure 3a). All 75 within-host pairs had <8 recombination filtered SNPs, with 72% of SNP distances being 0 or 1. Between-host pairs resulted in more diverse recombination filtered SNP distances, with 92% between 0 and 250 SNPs (Figure 3b). Remarkably, 15 between-host pairs (6.5%) had <10 SNPs, meaning that these pairs shared a single strain. Therefore, differentiating between- and within-host isolates is not possible based on SNP distances alone. From the 15 between-host pairs with a single strain, 13 had identical STs and 2 differed in NG-MAST only. Comparable results were found with unfiltered- instead of recombination filtered SNP distances, albeit with higher overall unfiltered SNP distances, indicating that recombination filtering did not inflate the number of closely related isolates (Figure S2). When using <10 recombination filtered SNPs as reference, MLST differentiated between strains with 100% sensitivity but only 1% specificity in the 303 within- and between-host pairs. NG-STAR showed 99% sensitivity and 43% specificity and NG-MAST 94% sensitivity and 79% specificity (Table S3).

The SNP distances <10 SNPs for between-host pairs suggest potential direct or indirect transmission between participants. Interestingly, between-host pairs with a single strain were obtained up to 1 year apart from each other (Figure 3c). Moreover, these pairs mainly contained isolates belonging to the MLST-NG-STAR-NG-MAST profiles 10314-1387-19188 and 9363-168-12302 (Figure 3c), indicating probable transmission between multiple participants within the same sexual network. Between-host paired isolates with identical STs and SNP distances ≥10 SNPs belonged mainly to the MLST-NG-STAR-NG-MAST profiles 8156-442-5441 and 7822-1387-14994, which belonged to the predominant typing profiles in the study population, indicating that these strains circulated more broadly in Amsterdam and were not restricted to a single sexual network.

DISCUSSION

We investigated within-host genetic variation in *N. gonorrhoeae* over the course of infection, using a unique set of within-host isolates from consecutive time points. Isolates were obtained from participants with treatment failure in a randomized clinical trial comparing 4 different antibiotics for the treatment of gonorrhoea[12]. The genetic variation between within-host isolates over time was compared to the genetic variation between within-host isolates from different anatomical locations at a single time point. Paired isolates that differed in all gene-based STs (MLST, NG-STAR, and NG-MAST) had high core genome SNP distances, thus were defined as distinct *N. gonorrhoeae* strains. Fifteen percent of the participants with locations-pairs had distinct strains at different anatomical locations. Similar coinfections have previously been reported in the MSM population of Amsterdam[25,26]. In contrast, a median recombination filtered SNP distance of 1 SNP was found between within-host paired isolates with identical STs or with different NG-MAST only. Since strains that are the same based on core genome SNPs could have different NG-MAST STs, this typing method alone does not identify events such as treatment failure or transmission with complete accuracy. Nevertheless, when comparing the gene-based typing methods, NG-MAST had the highest specificity for differentiating strains in between-host pairs with identical STs for at least one of the three typing methods (Table S3). This showed that NG-MAST is the most appropriate of these methods for transmission studies; however, transmission events might be incorrectly suggested to have taken place or, more rarely, be missed. This underlines the need for high-resolution methods like WGS in these studies.

Genetic variation in times-pairs could potentially be explained by the antibiotic pressure applied through treatment at T0. Due to low numbers of gentamicin- and ertapenem treatment failures, we could not identify whether specific mutations were associated with treatment arm. In addition, genes associated with gentamicin or fosfomycin resistance are not known. Instead, we assessed whether we could identify in times-pairs, compared to locations-pairs, a rise in genetic variation in specific genes or an overall rise in variation due to more general stress responses. These would enable the bacteria to faster pick up resistance genes or create resistance associated mutations (reviewed in[27]). However, in our within-host times-pairs and locations-pairs, similar regions with high SNP densities were found (Figure 2) as well as similar extents of genetic variation (Figures 1b and c). This indicated that the antibiotic pressure applied during infection did not lead to the development of resistance in specific genes or to an increased extent of genetic variation. Importantly, treatment failure was neither caused by a resistant strain at the inclusion visit, since MICs were not associated with treatment failure[12]. Thus, treatment failure was most likely caused by pharmacokinetic and/or dynamic reasons, e.g., suboptimal mucosal antibiotic concentrations.


(a)
Recombination filtered SNP distance
1500
1000
500
0
100
80
60
40
20
0
Between-host
Within-host
Between-host
Within-host
different MLST
same NGSTAR
same NGMAST
same MLST
different NGSTAR
different NGMAST
same MLST
same NGSTAR
different NGMAST
same MLST
different NGSTAR
same NGMAST
identical STs
(b)
Proportion
30%
20%
10%
0%
0
500
1000
1500
0
10
20
30
40
50
Recombination filtered SNP distance
Within-host
Between-host
(c)
Recombination filtered SNP distance
120
100
80
60
40
20
0
0
200
400
600
800
Days between isolate collection dates
MLST - NG-STAR - NG-MAST
10314 - 1387 - 12547
10314 - 1387 - 19188
11706 - 1869 - 17972
13489 - 1225 - 14700
13734 - 3171 - 13489
1599 - 520 - 11461
7822 - 1387 - 14994
7822 - 3321 - 4995
7827 - 38 - 10386
8156 - 442 - 5441
9363 - 168 - 12302

< FIGURE 3. RECOMBINATION FILTERED SNP DISTANCES IN WITHIN-HOST AND BETWEEN-HOST ISOLATE PAIRS WITH IDENTICAL MLST, NG-MAST AND/OR NG-STAR STS. (A) RECOMBINATION FILTERED SNP DISTANCES IN WITHIN-HOST AND BETWEEN-HOST PAIRS, CATEGORIZED ON MLST, NG-STAR AND NG-MAST TYPING RESULTS. ISOLATE PAIRS WITH DIFFERENT MLST, NG-STAR AND NG-MAST STS WERE EXCLUDED. THE RIGHT PANEL MAGNIFIES THE LOWER VALUES ON THE Y-AXIS. (B) PROPORTIONS OF RECOMBINATION FILTERED SNP DISTANCES FOUND AMONG ALL WITHIN-HOST PAIRS (75 PAIRS) AND BETWEEN-HOST PAIRS (228 PAIRS). THE BOTTOM PANEL MAGNIFIES THE LOWER VALUES ON THE X-AXIS. (C) RECOMBINATION FILTERED SNP DISTANCES IN BETWEEN-HOST PAIRS WITH IDENTICAL MLST, NG-STAR AND NG-MAST STS (52 PAIRS) VERSUS THE DAYS BETWEEN ISOLATE COLLECTION DATES, CATEGORISED ON TYPING PROFILE.

Within-host genetic variation could more easily be studied in the closest relative of *N. gonorrhoeae*, *Neisseria meningitidis*, since asymptomatic carriage of *N. meningitidis* is no indication for antibiotic treatment whereas it is for *N. gonorrhoeae*. Several studies reported that during carriage of *N. meningitidis*, the majority of within-host genetic variation was observed in hypervariable genomic regions[28-30]. We also found genetic variation mainly concentrated in hypervariable regions encoding surface exposed proteins, such as type IV pili and transferrin binding proteins. These proteins are known to play important roles in the host-microbe interaction[31,32]. This implies that both in *N. meningitidis* and *N. gonorrhoeae*, a majority of within-host genetic variation leads to antigenic variation, driven by the host immune response.

Since *N. gonorrhoeae* is a pathogen with high recombination rates, it is important to examine the effect of masking recombination when studying within-host variation and between-host transmission dynamics. De Silva et al. used recombination filtered SNP distances and mentioned that recombination filtering was essential for identification of transmission clusters[7]. Kwong et al. used recombination filtered SNP distances to determine genetic relatedness of isolates from men in partnerships and showed that filtering enabled identification of related isolates in partners, whereas unfiltered SNP distances would have distinguished them[33]. Kong et al. showed that recombination filtered SNP distances not only correctly identified partner links, but in addition identified links between individuals that were not identified as partners, most likely through indirect transmission or anonymous contacts. This technique could therefore be useful when identifying transmission networks and when implementing public health outreach interventions[9]. In contrast, Williamson et al. used unfiltered SNP distances to identify potential transmission events, to use a more stringent similarity threshold in a geographically and temporally limited data set. Distinct transmission clusters could be identified using this method[5]. Recombination filtering was discouraged for studies on transmission dynamics of pathogens with high interspecies variability, such as *Escherichia coli*. For these pathogens, calculating SNPs after filtering recombination led to loss of resolution and spurious clustering of isolates[8]. Altogether, these studies show that whether or not to filter out recombination depends on the research question and should therefore be carefully considered and evaluated. The results presented in this study supported the use of recombination filtered SNP distances when

studying within-host genetic diversity, since unfiltered SNP distances include high-density SNP regions between paired isolates. This high SNP density was caused by single recombination events which led to high unfiltered SNP distance between these isolates (Figure 2).

In this study, only SNP distances in the core genome were considered, since all isolates were mapped to reference genome FA1090. This method was chosen to enable the comparison of SNP distance across pairs and ideally, across studies that use the same reference genome. However, when interpreting the results, it should be taken into account that this method does not capture variation in genomic regions that are not in the reference genome. Another limitation was that genomes from multiple colony picks from a single sample were not analysed. As a result, we could not exclude that SNP differences in times-pairs arose during infection or were pre-existing at T0. However, De Silva et al. previously compared multiple colony picks from a single culture and found only 1 SNP, most likely caused by sequencing errors[7]. Thus, analysing multiple colonies would probably not have provided more information. Our threshold of <10 recombination filtered SNPs for within-host pairs was in accordance with results from De Silva et al., who found recombination filtered SNP distances ≤10 SNPs between sequential isolates of 6/113 individuals. They supposed that these isolate pairs resulted from treatment delay or reinfection from the same partner. We confirmed this threshold with times-pairs of 41 individuals.

The between-host isolate pairs with recombination filtered SNP distances <10 SNPs suggested direct- or indirect transmission between participants. Since the within- and between-host SNP distances overlapped, SNP distances alone cannot differentiate between cases of within-host persistence or between-host transmission. This is supported by the studies of Kwong et al. and Kong et al., who found similar relatedness between isolates from men within partnerships, between isolates from multiple anatomical locations of a single individual and between isolates that were probably linked through indirect transmission[9,33]. This shows that patient reported data remains of utmost importance in studies on transmission dynamics and in randomized clinical trials. Since the current study used isolates obtained over time, the possibility of reinfection with the same strain from a steady partner cannot be excluded. However, participants were asked to refrain from sexual intercourse during the study period, which was also recorded. Moreover, the 7 days period between isolate collection was relatively short. For these reasons, patient reported metadata assures us that the within-host paired isolates from consecutive time points were from persistent infections rather than reinfection with the same strain.

CONCLUSIONS

Comparing within-host *N. gonorrhoeae* isolates, our results confirm the previously defined threshold of 10 recombination filtered SNPs to differentiate between strains. Antibiotic pressure applied through treatment did not lead to an increase in genetic variation in specific genes or in extent of variation, compared to variation across anatomical locations. Instead, within-host genetic variation was mainly driven by host immunity as it was concentrated in genomic regions encoding surface exposed proteins involved in host-microbe interaction. Recombination filtered SNP distances <10 SNPs were also found between isolates from different participants, suggesting transmission. To differentiate between within-host persistence and between-host transmission, additional patient reported data remains essential.

DATA AVAILABILITY

Raw sequencing reads are available in the European Nucleotide Archive under project number PRJEB49317. Genome assemblies are available in the PubMLST Neisseria database (https://pubmlst.org/organisms/neisseria-spp). Individual ENA accession numbers, PubMLST IDs and associated metadata can be found in Table S1. The bioinformatic pipeline used in this study is available at Github (https://github.com/jolindadekorne/Within-host-genetic-variation-in-Neisseria-gonorrhoeae).

REFERENCES

1. Kirkcaldy RD, Weston E, Segurado AC, Hughes G. Epidemiology of Gonorrhea: A Global Perspective. *Sex Health* 2019;16:401–11 doi: 10.1071/SH19061
2. Hui B, Fairley CK, Chen M, Grulich A, Hocking J, Prestage G, *et al*. Oral and anal sex are key to sustaining gonorrhoea at endemic levels in MSM populations: a mathematical model. *Sex Transm Infect*. 2015;91:365–9 doi: 10.1136/sextrans-2014-051760
3. Van Liere GAFS, Dukers-Muijrers NHTM, Kuizenga-Wessel S, Wolffs PFG, Hoebe CJPA. Routine universal testing versus selective or incidental testing for oropharyngeal *Neisseria gonorrhoeae* in women in the Netherlands: a retrospective cohort study. *Lancet Infect Dis*. 2021 Jun 1;21:858–67 doi: 10.1016/S1473-3099(20)30594-6
4. Peters J, Cresswell F, Amor L, Cole K, Dean G, Didelot X, *et al*. Whole genome sequencing of *Neisseria gonorrhoeae* reveals transmission clusters involving patients of mixed HIV serostatus. *Sex Transm Infect*. 2018;94:138–43 doi: 10.1136/sextrans-2017-053198
5. Williamson DA, Chow EPF, Gorrie CL, Seemann T, Ingle DJ, Higgins N, *et al*. Bridging of *Neisseria gonorrhoeae* lineages across sexual networks in the HIV pre-exposure prophylaxis era. *Nat Commun*. 2019;10:3988 doi: 10.1038/s41467-019-12053-4
6. Martin IMC, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Rapid sequence-based identification of gonococcal transmission clusters in a large metropolitan area. *J Infect Dis*. 2004;189:1497–505 doi: 10.1086/383047
7. Silva DD, Peters J, Cole K, Cole MJ, Cresswell F, Dean G, *et al*. Whole-genome sequencing to determine *Neisseria gonorrhoeae* transmission : an observational study. *Lancet Infect Dis* 2017;16:1295–303 doi: 10.1016/S1473-3099(16)30157-8
8. Gorrie CL, Silva AGD, Ingle DJ, Higgs C, Seemann T, Stinear TP, *et al*. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. *Lancet Microbe* 2021;E575-E583 doi: 10.1016/S2666-5247(21)00149-X
9. Kong LY, Wilson JD, Moura IB, Fawley W, Kelly L, Walker AS, *et al*. Utility of whole genome sequencing in assessing and enhancing partner notification of *Neisseria gonorrhoeae* infection. *Sex Transm Dis* 2021;48:773-780 10.1097/OLQ.0000000000001419
10. Pérez-Lago L, Monteserin J, Paul R, Maus SR, Yokobori N, Herranz M, *et al*. Recurrences of multidrug-resistant tuberculosis: Strains involved, within-host diversity, and fine-tuned allocation of reinfections. *Transbound Emerg Dis*. 2021 doi: 10.1111/tbed.13982
11. Allen H, Mitchell HD, Simms I, Baker KS, Foster K, Hughes G, *et al*. Evidence for re-infection and persistent carriage of *Shigella* species in adult males reporting domestically acquired infection in England. *Clin Microbiol Infect*. 2021;27:126.e7-126.e13 doi: 10.1016/j.cmi.2020.03.036
12. Vries HJC de, Laat M de, Jongen VW, Heijman T, Wind CM, Boyd A, *et al*. Efficacy of ertapenem, gentamicin, fosfomycin, and ceftriaxone for the treatment of anogenital gonorrhoea (NABOGO): a randomised, non-inferiority trial. *Lancet Infect Dis*. 2022 doi: 10.1016/S1473-3099(21)00625-3
13. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34:i884–90 doi: 10.1093/bioinformatics/bty560
14. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv* 2013; Available from: http://arxiv.org/abs/1303.3997v2
15. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–9 doi: 10.1093/bioinformatics/btp352
16. Souvorov A, Agarwala R, Lipman DJ. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol*. 2018;19:153 doi: 10.1186/s13059-018-1540-z
17. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinforma Oxf Engl*. 2013;29:1072–5 doi: 10.1093/bioinformatics/btt086
18. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20:257 doi: 10.1186/s13059-019-1891-0

19. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res*. 2018;3:124 doi: 10.12688/wellcomeopenres.14826.1

20. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res*. 2015;43:e15 doi: 10.1093/nar/gku1196

21. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics*. 2012;28:2520–2 doi: 10.1093/bioinformatics/bts480

22. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, *et al*. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genomics*. 2016;2:e000056 doi: 10.1099/mgen.0.000056

23. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinforma Oxf Engl*. 2012;28:464–9 doi: 10.1093/bioinformatics/btr703

24. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, *et al*. Twelve years of SAMtools and BCFtools. GigaScience. 2021;10:giab008 doi: 10.1093/gigascience/giab008

25. Kolader M-E, Dukers NHTM, van der Bij AK, Dierdorp M, Fennema JSA, Coutinho RA, *et al*. Molecular Epidemiology of *Neisseria gonorrhoeae* in Amsterdam, The Netherlands, Shows Distinct Heterosexual and Homosexual Networks. *J Clin Microbiol*. 2006 Aug;44:2689–97 doi: 10.1128/JCM.02311-05

26. Heymans R, A Matser A, Bruisten SM, Heijman T, Geskus RB, Speksnijder AGCL, *et al*. Distinct *Neisseria gonorrhoeae* transmission networks among men who have sex with men in Amsterdam, The Netherlands. *J Infect Dis*. 2012 Aug 15;206:596–605 doi: 10.1093/infdis/jis399

27. Blázquez J, Rodríguez-Beltrán J, Matic I. Antibiotic-Induced Genetic Variation: How It Arises and How It Can Be Prevented. *Ann Rev Microbiol* 2018;72:209-230 doi: 10.1146/annurev-micro-090817-062139

28. Green LR, Al-Rubaiawi AA, Al-Maeni MARM, Harrison OB, Blades M, Oldfield NJ, *et al*. Localized Hypermutation is the Major Driver of Meningococcal Genetic Variability during Persistent Asymptomatic Carriage. mBio. 2020 Mar 24;11:e03068-19 doi: 10.1128/mBio.03068-19

29. Bårnes GK, Brynildsrud OB, Børud B, Workalemahu, B, Kristiansen PA, Beyene D, Aseffa A., & Caugant DA. Whole genome sequencing reveals within-host genetic changes in paired meningococcal carriage isolates from Ethiopia. *BMC genomics*. 2017;18:407 doi: 10.1186/s12864-017-3806-3

30. Mustapha MM, Marsh JW, Shutt KA, Schlackman J, Ezeonwuka C, Farley MM, Stephens DS, Wang X, Van Tyne D, Harrison LH. Transmission Dynamics and Microevolution of *Neisseria meningitidis* During Carriage and Invasive Disease in High School Students in Georgia and Maryland, 2006-2007. *J Infect Dis*. 2021;223:2038-2047. doi: 10.1093/infdis/jiaa674

31. Kirchner M, Meyer TF. The *PilC* adhesin of the *Neisseria* type IV pilus-binding specificities and new insights into the nature of the host cell receptor. *Mol Microbiol*. 2005;56:945–57 doi: 10.1111/j.1365-2958.2005.04600.x

32. Cornelissen CN, Kelley M, Hobbs MM, Anderson JE, Cannon JG, Cohen MS, *et al*. The transferrin receptor expressed by gonococcal strain FA1090 is required for the experimental infection of human male volunteers. *Mol Microbiol*. 1998;27:611–6 doi: 10.1046/j.1365-2958.1998.00710.x

33. Kwong JC, Chow EPF, Stevens K, Stinear TP, Seemann T, Fairley CK, *et al*. Whole-genome sequencing reveals transmission of gonococcal antibiotic resistance among men who have sex with men: an observational study. *Sex Transm Infect*. 2018;94:151–7 doi: 10.1136/sextrans-2017-053287

SUPPLEMENTARY MATERIAL

TABLE S1. LIST OF ISOLATES INCLUDED IN THE STUDY WITH CORRESPONDING ISOLATE- AND PARTICIPANT METADATA.

This table (15 columns and 204 rows) is available with the online version of this article or via https://tinyurl.com/3jbtmzwr.

TABLE S2. CHARACTERISTICS OF ALL WITHIN-HOST ISOLATE PAIRS AND THE DIFFERENCES FOUND ON THE LEVELS OF SEQUENCE TYPES, CGMLST ALLELES AND SNPS.

This table (16 columns and 130 rows) is available with the online version of this article or via https://tinyurl.com/3jbtmzwr.

TABLE S3. SENSITIVITY AND SPECIFICITY OF THE GENE-BASED TYPING METHODS MLST, NG-STAR AND NG-MAST.

Within-host pairs (n=130)

	MLST	Same ST	Different ST
<10 recombination filtered SNPs		121	0
≥10 recombination filtered SNPs		0	9
		Sensitivity	100%
		Specificity	100%
	NG-STAR	Same ST	Different ST
<10 recombination filtered SNPs		121	0
≥10 recombination filtered SNPs		0	9
		Sensitivity	100%
		Specificity	100%
	NG-MAST	Same ST	Different ST
<10 recombination filtered SNPs		112	9
≥10 recombination filtered SNPs		0	9
		Sensitivity	93%
		Specificity	100%

Within+ between-host pairs (n=303)

	MLST	Same ST	Different ST
<10 recombination filtered SNPs		90	0
≥10 recombination filtered SNPs		210	3
		Sensitivity	100%
		Specificity	1%
	NG-STAR	Same ST	Different ST
<10 recombination filtered SNPs		89	1
≥10 recombination filtered SNPs		109	104
		Sensitivity	99%
		Specificity	49%
	NG-MAST	Same ST	Different ST
<10 recombination filtered SNPs		85	5
≥10 recombination filtered SNPs		45	168
		Sensitivity	94%
		Specificity	79%

TABLE S4. CHARACTERISTICS OF INCLUDED WITHIN- AND BETWEEN-HOST ISOLATE PAIRS AND THE FOUND ON THE LEVELS OF SEQUENCE TYPES, CGMLST ALLELES AND SNPS.

This table (17 columns and 303 rows) is available with the online version of this article or via https://tinyurl.com/3jbtmzwr.

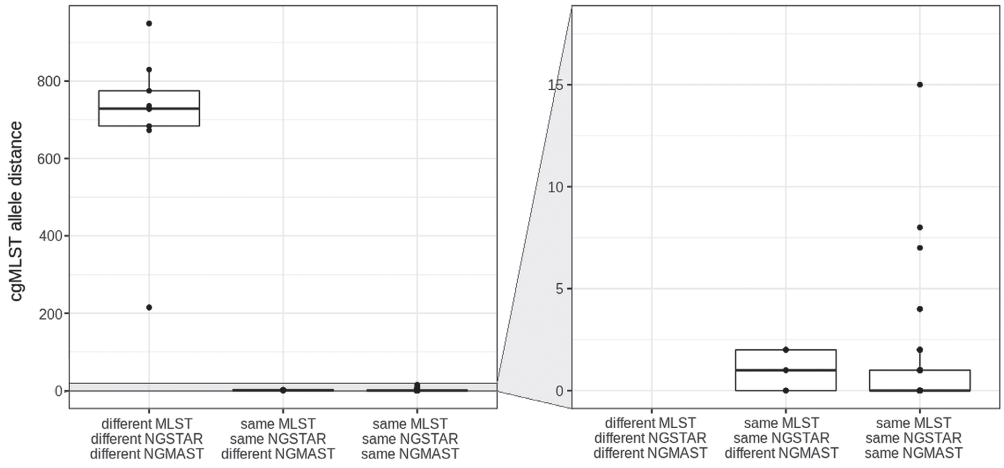


FIGURE S1. GENE-BASED TYPING RESULTS VERSUS CORE GENOME MLST (CGMLST) ALLELE DISTANCES IN WITHIN-HOST ISOLATE PAIRS. ISOLATES IN WITHIN-HOST TIMES-PAIRS AND LOCATIONS-PAIRS WERE COMPARED BY MLST, NG-STAR, NG-MAST STS AND CGMLST ALLELE DISTANCE. HIGH CGMLST ALLELE DISTANCES WERE FOUND BETWEEN ISOLATES WITH DIFFERENT MLST, NG-STAR AND NG-MAST STS WHEREAS ISOLATES THAT DIFFERED IN NG-MAST ONLY OR ISOLATES WITH IDENTICAL MLST, NG-STAR AND NG-MAST STS HAD COMPARABLE CGMLST ALLELE DISTANCES. THE RIGHT PANEL MAGNIFIES THE LOWER VALUES ON THE Y-AXIS, WHICH CONFIRMS THAT COMPARABLE ALLELE DISTANCES WERE FOUND BETWEEN ISOLATES WITH DIFFERENT NG-MAST COMPARED TO ISOLATES WITH IDENTICAL STS.

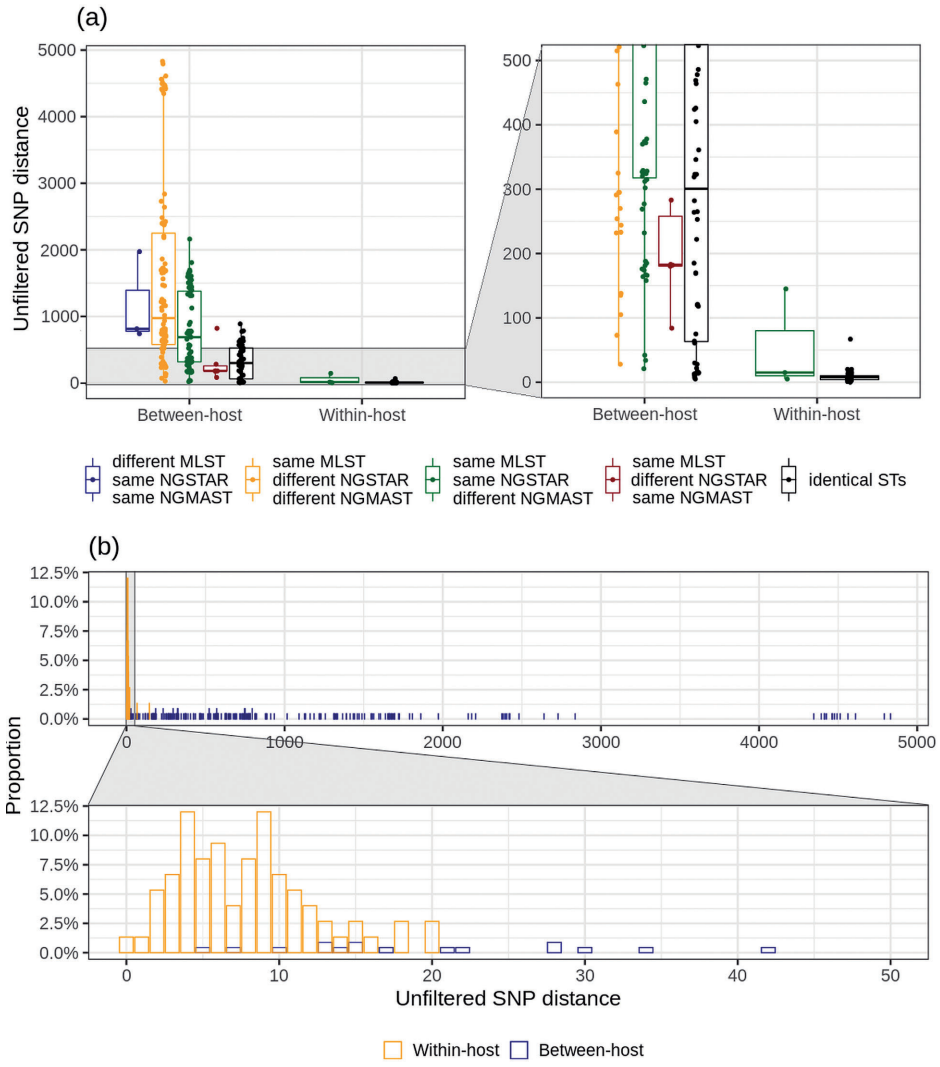
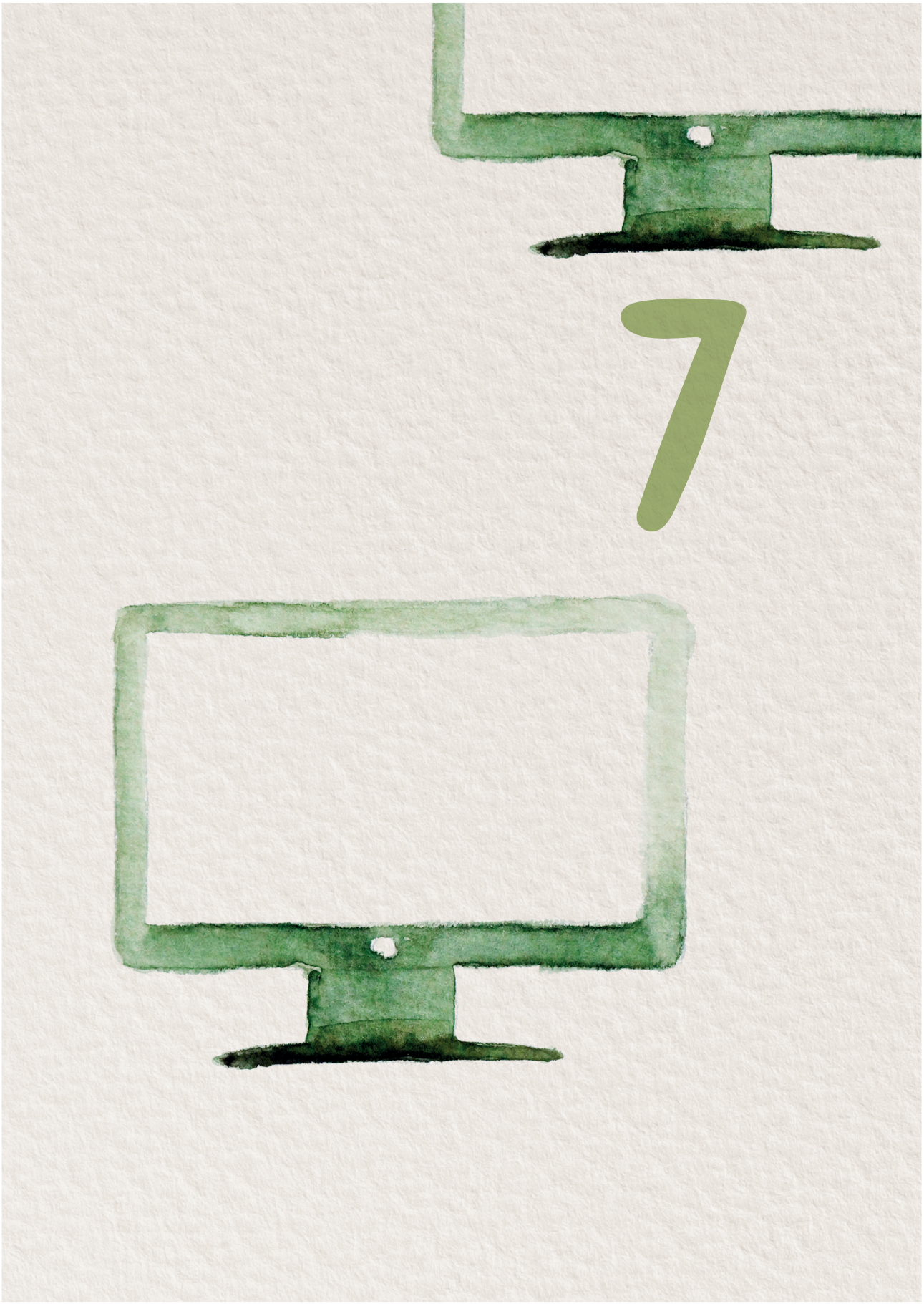


FIGURE S2. SNP DISTANCES IN WITHIN-HOST AND BETWEEN-HOST ISOLATE PAIRS WITHOUT RECOMBINATION FILTERING. (A) UNFILTERED SNP DISTANCES IN WITHIN-HOST AND BETWEEN-HOST ISOLATE PAIRS, CATEGORIZED ON MLST, NG-STAR AND NG-MAST TYPING RESULTS. ISOLATE PAIRS WITH DIFFERENT MLST, NG-STAR AND NG-MAST STS WERE EXCLUDED. (B) PROPORTIONS OF SNP DISTANCES FOUND IN ALL WITHIN-HOST PAIRS (75 PAIRS) AND BETWEEN-HOST PAIRS (228 PAIRS). ISOLATE PAIRS WITH DIFFERENT MLST, NG-STAR AND NG-MAST STS WERE EXCLUDED.


7

THE *NEISSERIA GONORRHOEAE* ACCESSORY GENOME AND ITS ASSOCIATION WITH THE CORE GENOME AND ANTIMICROBIAL RESISTANCE

Jolinda de Korne-Elenbaas, Sylvia M. Bruisten,
Alje P. van Dam, Martin C. J. Maiden, Odile B. Harrison

Microbiology Spectrum 2022; 10(3):e02654-21

ABSTRACT

The bacterial accessory genome provides the genetic flexibility needed to facilitate environment and host adaptation. In *Neisseria gonorrhoeae*, known accessory elements include plasmids which can transfer and mediate antimicrobial resistance (AMR); however, chromosomal accessory genes could also play a role in AMR. Here, the gonococcal accessory genome was characterised using gene-by-gene approaches and its association with the core genome and AMR were assessed. The gonococcal accessory gene pool consisted of 247 genes, which were mainly genes located on large mobile genetic elements, phage associated genes, or genes encoding putative secretion systems. Accessory elements showed similar synteny across genomes, indicating either a predisposition for particular genomic locations or ancestral inheritance that are conserved during strain expansion. Significant associations were found between the prevalence of accessory elements and core genome multi-locus sequence types (cgMLST), consistent with a structured gonococcal population despite frequent horizontal gene transfer (HGT). Increased prevalence of putative DNA exchange regulators was significantly associated with AMR, which included a putative secretion system, methyltransferases and a toxin-antitoxin system. Although frequent HGT results in high genetic diversity in the gonococcus, we found that this is mediated by a small gene pool. In fact, a highly organised genome composition was identified with a strong association between the accessory and core genome. Increased prevalence of DNA exchange regulators in antimicrobial resistant isolates suggests that genetic material exchange plays a role in the development or maintenance of AMR. These findings enhance our understanding of gonococcal genome architecture and have important implications for gonococcal population biology.

INTRODUCTION

Neisseria gonorrhoeae, the gonococcus, is a Gram negative, oxidase positive diplococcus belonging to the Neisseriaceae family. It is arguably the only pathogen in the genus *Neisseria* colonizing the *tractus urogenitalis* as an obligate pathogen, with its close relative *Neisseria meningitidis* more frequently described as an opportunistic or accidental pathogen[1]. Gonococcal urogenital colonization in men invariably results in disease pathology, with untreated gonococcal infection potentially leading to epididymitis. Urogenital infections in women often stay asymptomatic and therefore untreated, potentially causing pelvic inflammatory disease[2,3]. Effective and prompt treatment of infection is therefore essential; however, this is threatened by the emergence of multidrug resistance in the gonococcus which is further exacerbated by an increasing prevalence of *N. gonorrhoeae* infections globally[3]. This is particularly concerning given the emergence of reduced susceptibility against currently used third generation cephalosporins because these are the antibiotics of last resort. In addition, asymptomatic infections of the pharynx, rectum, or female urogenital tract facilitate ongoing transmission of *N. gonorrhoeae* strains with reduced susceptibility. Understanding antimicrobial resistance (AMR) mechanisms in the gonococcus and tracking the spread of resistant strains is therefore important to limit infections globally.

In *N. gonorrhoeae*, AMR is both plasmid and chromosomally mediated, depending on the class of antibiotics. Genetic determinants for fluoroquinolone and macrolide resistance have high predictive value with regard to susceptibility phenotype[4]. In contrast, determinants mediating resistance against third generation cephalosporins are less predictive and still contain unknown "factor X" genetic determinants, indicating that mechanisms of resistance have not been entirely resolved[4]. As a result, to fully expose the mechanisms of AMR, a more conclusive understanding of genome content across the gonococcal population is needed to comprehensively evaluate the evolution and mechanisms of AMR in this species.

For most bacterial species, genome content is not fixed but rather consists of a flexible gene pool, known as the pangenome[5]. The pangenome can be further divided into: (i) the core genome, comprised of genes that are present in all isolates of a species and which are essential for survival, and (ii) the accessory genome, consisting of genes present in only a subset of the population and providing the genetic flexibility to facilitate niche adaptation and phenotypic variation[6,7]. The accessory genome can be shaped through evolutionary and environmental selective pressure resulting in gene acquisition or loss[8]. Antimicrobial selection, for example, plays an important role in shaping bacterial genome content and in structuring bacterial populations, with bacteria acquiring resistance determinants possessing a selective advantage over susceptible ones when exposed to antibiotics. This development has been observed in several bacterial species, including the gonococcus[9].

Studies assessing the gonococcal core genome have identified the presence of a residual clonal structure that retains gene synteny despite frequent horizontal gene transfer (HGT), with some lineages found to be more resistant to antimicrobials than others[10]. In addition, the distribution of mobile genetic elements including plasmids, which form part of the accessory genome, was found to be associated with the gonococcal core genome[11]. Although plasmids are important mobile genetic elements that transfer and mediate AMR, chromosomal accessory genes could also play a role. However, the accessory genome content as well as its implications for phenotypic characteristics in *N. gonorrhoeae* are largely unknown. Here, we characterised the gonococcal accessory genome and examined its association with the core genome and AMR. Our findings reveal that the gonococcal gene pool was smaller than expected with respect to the high genetic diversity known to exist in this species. In addition, genome composition was highly organised with conserved gene synteny and a close association between accessory genome content and the core genome. These findings enhance our understanding of the gonococcal genome architecture and have important implications for gonococcal population biology.

METHODS

ISOLATE COLLECTION AND WHOLE GENOME SEQUENCE DATA

The PubMLST.org/neisseria database contained 8,013 publicly available *N. gonorrhoeae* isolates with whole genome sequence data available (accessed April 2021). A subset of 765 *N. gonorrhoeae* isolates was chosen that was representative of the total *N. gonorrhoeae* genomes found in PubMLST. This data set of 765 isolates included (i) 380 isolates from a Dutch surveillance study containing many different MLSTs[12] and (ii) 385 isolates chosen to be representative of all MLSTs, core genome groups, and geographical origins available in PubMLST (Table S1). The extent to which this data set was representative of the diversity found in all 8,013 publicly available isolates was assessed through the use of a minimum spanning tree generated using the PubMLST plugin GrapeTree[13]. Using this tool, all 8,013 isolates were compared with the *N. gonorrhoeae* cgMLST v1.0 core genome scheme[10], and the distribution of the subset of 765 isolates in the tree was assessed. Tree nodes were positioned through dynamic rendering, meaning that branch lengths are not scaled. Genome assembly metrics were obtained for all isolates to validate genome quality. In addition, species identity was confirmed using the ribosomal MLST species identification tool which also controls for species contamination (pubmlst.org/rmlst).

IDENTIFICATION OF ACCESSORY LOCI

The PubMLST Neisseria database allows genomes across the Neisseria genus to be annotated and as a result contains gene definitions belonging to multiple *Neisseria* species. A total of 2,859 loci have been defined in PubMLST that are representative of the *Neisseria* gene pool. Using the

Genome Comparator tool available in PubMLST, these loci were compared in the 765 isolates[14]. Alleles defined as "incomplete" or "undefined" were considered present because this might be due to hypervariability of the gene or to assembly errors. Core genes were defined as loci found in 95% to 100% of genomes, whereas accessory genes were defined as loci present in 0.6% to 95% of genomes (loci present in <5 genomes were excluded, equalling 0.6%). Loci considered to be core were excluded from further analyses.

A sequence-based approach was also used to identify loci that may not have been defined in PubMLST. Identification and classification of orthologous genes within the 765 isolates was done with PIRATE v1.0.4 using genomes previously annotated with Prokka v1.13-v1.14.5[15,16]. Default settings were used. Pangenome plots were produced with Roary v3.13.0 using default settings except paralogue splitting was disabled[17]. These plots were used to assess the relationship between genes identified and the number of genomes added. Similar criteria as used for Genome Comparator were employed to define core and accessory genes identified by PIRATE. Representative sequences of the identified accessory gene families were queried in PubMLST to find corresponding NEIS loci. Sequences that matched only partially to a NEIS locus were queried against a custom database including all 1,547,524 alleles (accessed June 2021) of all 2,859 NEIS loci using BLASTN with a minimum sequence identity of 80%. Sequences that did not match any NEIS allele with >80% identity were queried against the *N. gonorrhoeae* (taxid: 485) BLASTX database to get (predicted) functional annotations[18]. Novel sequences were then defined in PubMLST using the NEIS nomenclature. BLAST searches were performed to screen isolates for these novel loci using the Gene Presence plugin with the following thresholds: minimum sequence identity 95% and minimum sequence alignment 95%. Finally, *N. gonorrhoeae* typing scheme was defined consisting of the *N. gonorrhoeae*: (i) accessory genome (agMLST); (ii) core genome (cgMLST); and (iii) pangenome (pgMLST) (Workflow visualised in Figure 1).

FILTERING OF IDENTIFIED ACCESSORY LOCI

NEIS loci assigned as accessory by Genome Comparator and PIRATE were filtered as follows: (i) exclusion of loci that were present in <5 isolates; (ii) exclusion of loci that were part of the PubMLST *N. gonorrhoeae* cgMLST v1.0 scheme of PubMLST; these are most likely erroneously in the accessory genome due to incompleteness of the genes in draft genome data; and (iii) exclusion of paralogous, incomplete and hypervariable loci, including those belonging to the multiple adhesin family (*maf*) island, pilus and glycosyltransferase schemes of PubMLST. The paralogous nature and hypervariability of these genes lead to uncertainty when estimating prevalence in draft genome data. After filtering, lists from Genome Comparator and PIRATE were combined to produce a consensus list of accessory NEIS loci (Figure 1).

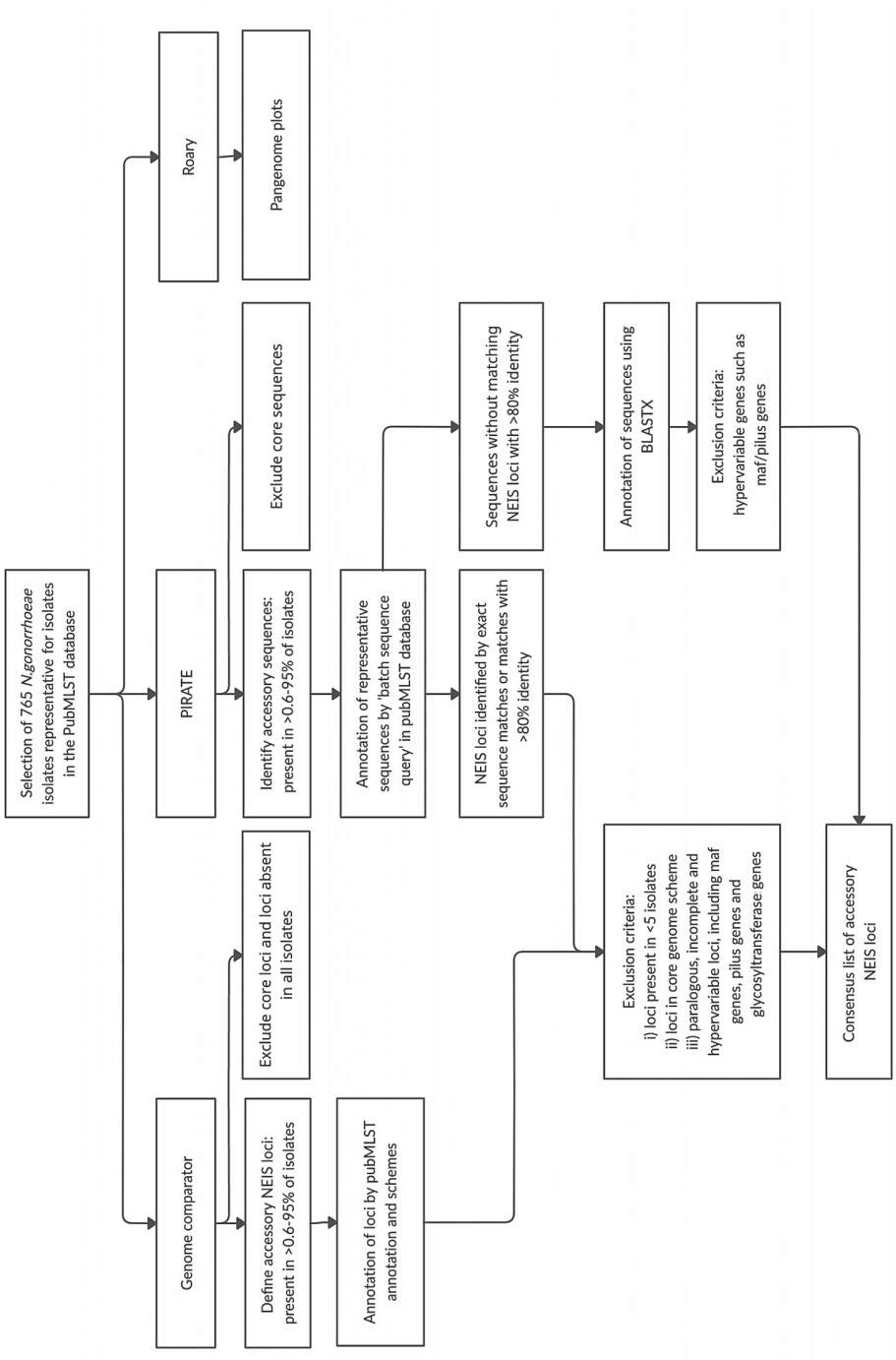


FIGURE 1. WORKFLOW FOR THE CHARACTERISATION OF THE ACCESSORY GENOME IN *N. GONORRHOEAE*. GENOME COMPARATOR, PIRATE, AND ROARY SOFTWARE WERE USED ON A DATA SET OF 765 ISOLATES.

FUNCTIONAL CHARACTERISATION AND DISTRIBUTION OF ACCESSORY LOCI

Identified accessory loci were classified into functional groups using annotations defined in PubMLST. Locations of all accessory loci were visualised using the reference genomes: WHO-F (PubMLST ID: 62968; GenBank accession number: LT591897.1), WHO-Y (PubMLST ID: 88866; GenBank accession number: LT592161.1), and FA1090 (PubMLST ID: 2855; GenBank accession number: AE004969.1) in CGviewer[19]. Loci annotated as "hypothetical" but located on *maf*-islands were excluded from the list of accessory genes as these are paralogous. Duplicate sequences of the same loci were identified by performing sequence queries in PubMLST. The distribution of accessory loci across isolates was assessed using the Gene Presence plugin in PubMLST with the following thresholds: minimum sequence identity 95% and minimum sequence alignment 95%. To identify genomic islands or operons, co-location of loci was assessed based on gene distribution patterns. A genomic island or operon was considered present when >75% of its genes were identified and these were in adjacent locations. Phage associated genes NEIS0027 to 0031 and NEIS2451 were assigned to the Nf1 phage island as defined by Al Suwayyid et al.[20].

ASSOCIATION BETWEEN CORE- AND ACCESSORY GENOME

An agMLST scheme was created in PubMLST, including all identified accessory NEIS loci. The association between core and accessory genome was assessed including all 8,013 isolates, through the use of a minimum spanning tree generated with the GrapeTree tool. The tree was built based on the agMLST scheme with the "rescan undesignated loci" option enabled. Nodes were coloured by cgMLST group using the 300 or fewer locus threshold (Ng_cgc_300) and positioned through dynamic rendering, meaning that branch lengths are not scaled.

ACCESSORY GENOME DISTRIBUTION IN 8,013 PUBLICLY AVAILABLE ISOLATES

Using the agMLST scheme, prevalence and distribution of the accessory NEIS loci were assessed in all 8,013 isolates in PubMLST, using the Gene Presence plugin as described above. Prevalence of accessory elements were determined for cgMLST groups that contained >35 isolates and using the 300 or fewer locus threshold (Ng_cgc_300).

STATISTICS

Fisher's exact tests were performed to identify significant associations between the presence of accessory elements and cgMLST group as well as the association with resistance against penicillin, tetracycline, ciprofloxacin, ceftriaxone, cefixime, and azithromycin. When testing for associations with resistance, only isolates that had phenotypic data for that specific antibiotic were included in analyses. Phenotypic data was categorised as "susceptible" or "resistant" (or "intermediate" for penicillin) according to EUCAST breakpoints v11.0[21]. The Bonferroni correction method was used to correct for multiple testing and P-values were adjusted accordingly. Analyses were performed in R v3.6.3.

RESULTS

QUALITY OF SEQUENCE DATA AND ISOLATE CHARACTERISTICS

For accessory genome characterisation, genome sequence data from 765 publicly available isolates were chosen, dating from 1979 to 2019 and originating from 42 countries (Table S1). Draft genomes were available for 764 of 765 isolates with a median number of 123 contigs, and a complete genome was available for one isolate. Median GC content was 52.5%, the total median assembly length was 2,125,586 bp and a median of 1,632 alleles (range 1,460 to 1,652) was designated. Isolates belonged to 289 different multilocus sequence types (MLST) of which 204 were represented by a single isolate and 24 by ≥5 isolates. Sixty-five isolates belonged to MLST ST-7827, being the predominant MLST in the data set. There were 199 different core genome MLST (cgMLST) groups assigned, when using the threshold of 300 or fewer locus differences to differentiate between cgMLST groups (Ng_cgc_300), of which 148 were represented by a single isolate and 23 by ≥5 isolates. The cgMLST_300 group 3 predominated in the data set with 88 isolates belonging to it. These isolates dated from 2004 to 2019 and mainly originated from the Netherlands (64/88, 79%) (Table S1). The largest proportion of isolates in cgMLST_300 group 3 belonged to MLST ST-1901 (43/88, 49%), similar to the predominant group of isolates in PubMLST which also belong to cgMLST_300 group 3 and MLST ST-1901 (1,140/8,013, 14%, April 2021).

To determine to what extent the subset of 765 isolates was representative for the whole database of 8,013 isolates publicly available in PubMLST (accessed April 2021), a minimum spanning tree was generated including these isolates. Locations of the 765 isolates were visualised, revealing a distribution throughout the tree consistent with the selection of a heterogeneous data set that was representative of all publicly available *N. gonorrhoeae* isolates (Figure S1). The workflow of characterisation of the accessory genome using these 765 isolates is visualised in Figure 1.

THE GONOCOCCAL GENE POOL HAS A LIMITED SIZE

Roary pangenome plots were examined to assess the relationships between numbers of genes identified and numbers of genomes used. These plots showed that saturation in the number of new (Figure 2A) and conserved genes (Figure 2B) identified was reached after including fewer than 20 genomes, consistent with a restricted gonococcal pangenome size. As a feature, Roary splits variants of the same gene into multiple unique genes where these have less than 95% similarity. Therefore, the number of unique (Figure 2A) and total genes (Figure 2B) identified correlated with the number of gene variants defined, which in turn increased as more genomes were examined. This indicates that, in the gonococcus, diversity within genes increases proportionally as more genomes are examined, whereas diversity in genome content saturizes with less than 20 genomes. This also showed that the subset of 765 isolates was sufficient to capture the full extent of the gonococcal pangenome.

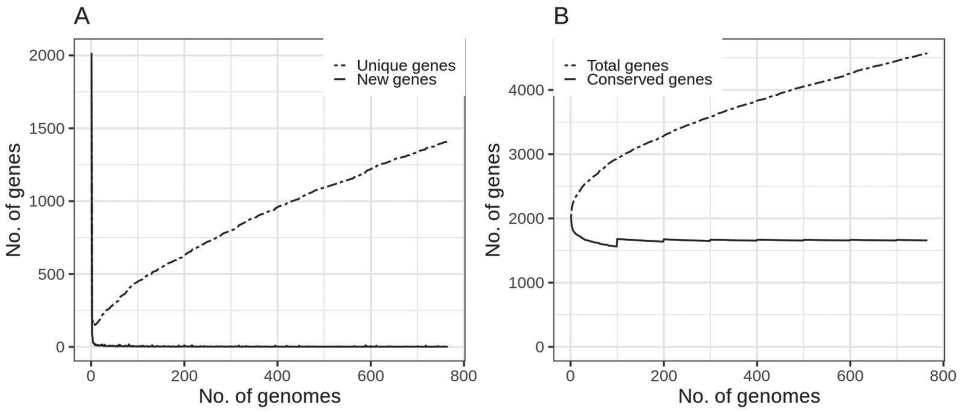


FIGURE 2. ROARY PANGENOME PLOTS DEMONSTRATING RELATIONSHIPS BETWEEN NUMBERS OF GENES IDENTIFIED AND NUMBERS OF GENOMES USED. THE NUMBERS OF NEW (A) AND CONSERVED (B) GENES SATURISED WITH LESS THAN 20 GENOMES. THE NUMBERS OF UNIQUE (A) AND TOTAL (B) GENES WERE POSITIVELY CORRELATED WITH THE NUMBER OF GENOMES INCLUDED IN THE ANALYSIS, AS SHOWN BY A CONSTANT INCREASE OF IDENTIFIED GENES WHEN THE NUMBER OF GENOMES INCREASED.

THE GONOCOCCAL ACCESSORY GENOME CONTAINED 247 LOCI

All 2,859 NEIS loci defined in PubMLST.org/neisseria were compared in the 765 isolates using Genome Comparator. This showed that 573 loci were absent in all 765 isolates, 24 loci were present in fewer than five isolates (prevalence: <0.6%) and 292 were accessory (prevalence: 0.6% to 95%). Filtering resulted in a final list of 196 accessory loci based on Genome Comparator analyses. PIRATE identified 3,189 distinct gene sequences of which 795 were present in fewer than five isolates and 623 were accessory. Querying of these 623 gene sequences in PubMLST showed that there were duplicate sequences of single NEIS loci, falsely identified as separate genes by PIRATE. As a result, the 623 gene sequences matched to 209 unique NEIS loci with exact matches, of which 173 had also been identified as accessory using Genome Comparator. Eight NEIS loci were identified as accessory by PIRATE only; however, these loci were determined as core by Genome Comparator and were therefore excluded from further analyses. Genome Comparator identified 23 accessory loci not identified by PIRATE. For 93 gene sequences identified as accessory by PIRATE, no matches to NEIS loci with >80% identity were found. Sequence annotation and filtering resulted in a list of 82 unique accessory genes identified by PIRATE only, of which 58 were added as NEIS loci to PubMLST (NEIS3177 to 3235, except NEIS3201) and of which 51 were confirmed to be accessory after screening isolates for these loci with the Gene Presence tool available in PubMLST. Twenty-four sequences were not added because these were either predicted to be non-functional or encoded transposases. Taken together, the consensus list contained 247 unique accessory NEIS loci which are provided in Table S2, together with prevalences and functions. A detailed workflow showing all steps taken to obtain the final list of 247 accessory NEIS loci is provided in Figure S2.

THE GONOCOCCAL ACCESSORY GENOME MAINLY CONSISTS OF LARGE MOBILE GENETIC ELEMENTS

Functional annotation of identified accessory genes showed that most of genes were located on large mobile genetic elements, namely: 67 genes on the gonococcal genomic island (GGI) (27%), 51 on the conjugative plasmid (21%), and eight on the beta-lactamase plasmid (3%). Aside from the mobile genetic elements, the major constituent of the accessory genome consisted of 31 hypothetical genes (12.5%), 30 genes encoding the VirB type IV secretion system (VirB T4SS) (12%), and 34 phage associated genes (13.5%) (Table 1). Phage-associated genes were found both as part of prophages (16/34 genes on Nf1, Nf4 and phage island X) or individual genes scattered over the chromosome. One group of seven genes (NEIS0080 to 0089), annotated as hypothetical or genes encoding protein export proteins in PubMLST, formed a genomic island with a constant synteny across genomes. Detailed BLAST annotation showed that this island encodes a putative secretion system. Other accessory genes encode methyltransferases, toxin-antitoxins or proteins involved in DNA transcription and replication (Table 1).

TABLE 1. FUNCTIONAL ANNOTATIONS OF 247 IDENTIFIED ACCESSORY GENES IN *N. GONORRHOEAE*.

Functional category	Subcategory	Number of accessory genes	%
Mobile genetic elements	Gonococcal genomic island (GGI)	67	(27%)
	Conjugative plasmid	51	(21%)
	Beta-lactamase plasmid	8	(3%)
VirB Type IV secretion system		30	(12%)
Putative secretion system		7	(3%)
Methyltransferases (outside genomic islands)		6	(2.5%)
Phage associated genes	Nf1 phage island	6	(2.5%)
	Nf4 phage island	3	(1%)
	Phage island X associated	7	(3%)
	(Putative) phage associated	18	(7%)
Genetic information processing	DNA transcription	3	(1%)
	DNA replication	1	(0.5%)
Toxin-antitoxin system		2	(1%)
TonB- dependent receptors		3	(1%)
Other functions	Alcohol dehydrogenase	1	(0.5%)
	Putative peptidase	1	(0.5%)
	Membrane proteins	1	(0.5%)
	TspB virulence factor	1	(0.5%)
Hypothetical genes		31	(12.5%)
Total number of accessory genes		247	100%

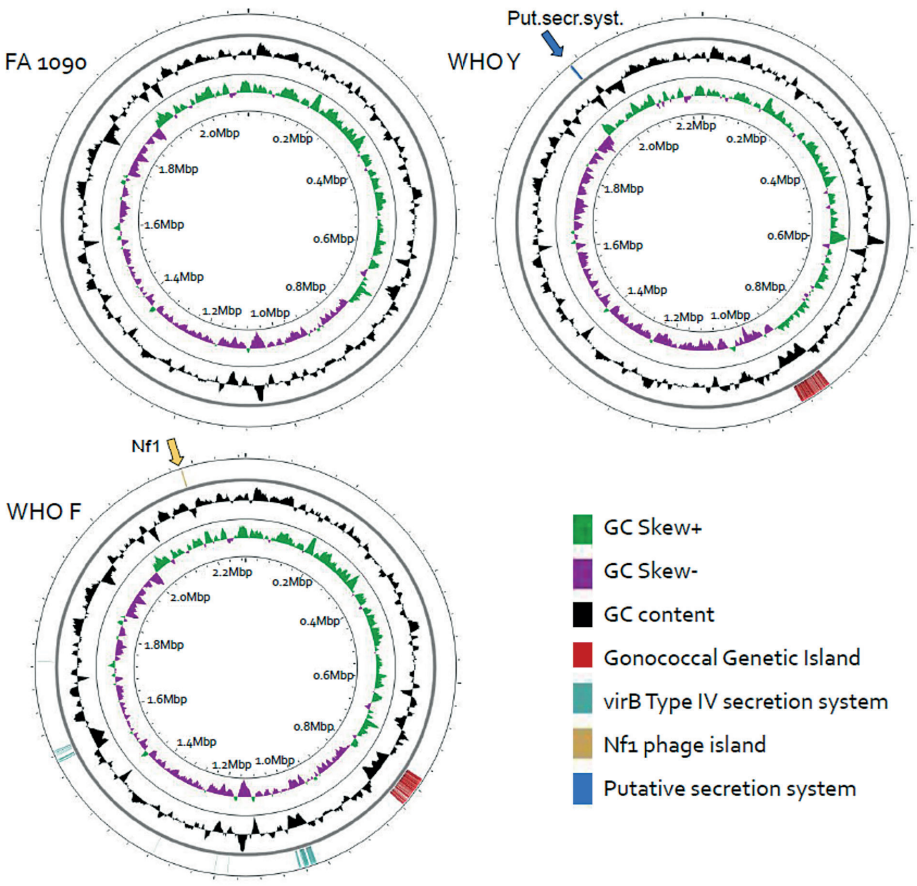


FIGURE 3. GENOMIC LOCATIONS OF CHROMOSOMAL ACCESSORY ELEMENTS BELONGING TO THE GONOCOCCAL ACCESSORY GENOME. LOCATIONS WERE VISUALISED ON CIRCULAR GENOMES OF WELL-KNOWN REFERENCE STRAINS. REFERENCE STRAIN FA1090 LACKS ALL ACCESSORY ELEMENTS AND ITS CHROMOSOMAL ACCESSORY GENOME IS THE SMALLEST OF THREE. ACCESSORY ELEMENTS WERE IDENTIFIED IN WHO-Y AND WHO-F, OF WHICH WHO-F HAS THE LARGEST ACCESSORY GENOME. FROM INNER TO OUTER CIRCLE: GC SKEW, GC CONTENT, READING FRAME WITH ANNOTATED GENES.

CHROMOSOMAL ACCESSORY ELEMENTS SHOW SIMILAR SYNTENY ACROSS GENOMES

The main chromosomal accessory elements were visualised using reference genomes FA1090, WHO-Y, and WHO-F (Figure 3). WHO-F has the largest chromosomal accessory genome, including the GGI, VirB T4SS, and Nf1 phage island. WHO-Y has the GGI and the putative secretion system while FA1090 has the smallest chromosomal accessory genome and lacks all major elements. Interestingly, the accessory elements displayed similar synteny which indicates either predisposition for these elements in these locations or ancestral inheritance, conserved during strain expansion.

THE ACCESSORY GENOME IS HIGHLY STRUCTURED AND ASSOCIATED WITH THE CORE GENOME

After characterisation of the accessory genome in the subset of 765 isolates, associations between the accessory and core genome and/or AMR were assessed using all 8,013 genomes available in PubMLST (accessed April 2021). The accessory genome typing scheme (agMLST v1.0) was created in PubMLST including all identified accessory loci. Based on this scheme, a minimum-spanning tree was generated including all 8,013 isolates. Nodes were coloured by cgMLST group and clustering of colours demonstrated that there was a strong relationship between accessory and core genomes (Figure 4).

Among the 8,013 isolates, the GGI was present in 69% (5,524/8,013), the conjugative plasmid in 27% (2,174/8,013), and the beta-lactamase plasmid in 11% (887/8,013). Associations between cgMLST and accessory elements were examined and results of the main cgMLST groups with >35 isolates are shown in Table 2. When comparing the prevalence of accessory elements within a cgMLST group to the prevalence among all isolates, the GGI, conjugative plasmid, beta-lactamase plasmid, and the genes encoding the putative secretion system were either significantly more prevalent or significantly less prevalent in most of the cgMLST groups (P-value range: <0.05 - <0.0001; Table 2 and Table S3). The Nf1 phage island was significantly more prevalent in isolates belonging to cgMLSTs 18, 415, and 304 (P-values <0.0001), compared with the distribution among all isolates. Isolates belonging to cgMLSTs 3, 16, 18, and 196 showed little variation in accessory genome content, which is striking given the large numbers of isolates in these groups. Isolates belonging to cgMLSTs 21, 8, and 133 had a more variable accessory genome content with almost half of the isolates possessing the GGI or beta-lactamase plasmid (Table 2). The strong association found between accessory and core genome is consistent with a structured gonococcal population. This also indicated ancestral inheritance of accessory elements in certain gonococcal lineages which are conserved during lineage expansion with limited spread to other lineages.

ACCESSORY GENES INVOLVED IN DNA UPTAKE AND EXCRETION ARE ASSOCIATED WITH AMR

Associations between the accessory genome and resistance against penicillin, tetracycline, ciprofloxacin, ceftriaxone, cefixime, and azithromycin were assessed. As expected, the beta-lactamase plasmid was significantly more prevalent among penicillin- (P <0.0001) or tetracycline- (P <0.0001) resistant isolates, compared with susceptible isolates. However, this was also found for ciprofloxacin-resistant isolates (P <0.0001). Interestingly, the genes encoding the putative secretion system (NEIS0080 to 0089), all methyltransferases (NEIS1193 to 1194, NEIS1311, NEIS2691 to 2692) except one (NEIS3182), a DNA transcription protein (NEIS3184) and a toxin-antitoxin system (NEIS3188 to NEIS3232) were significantly more prevalent among isolates resistant to penicillin, tetracycline, ciprofloxacin, cefixime, or azithromycin (P-value

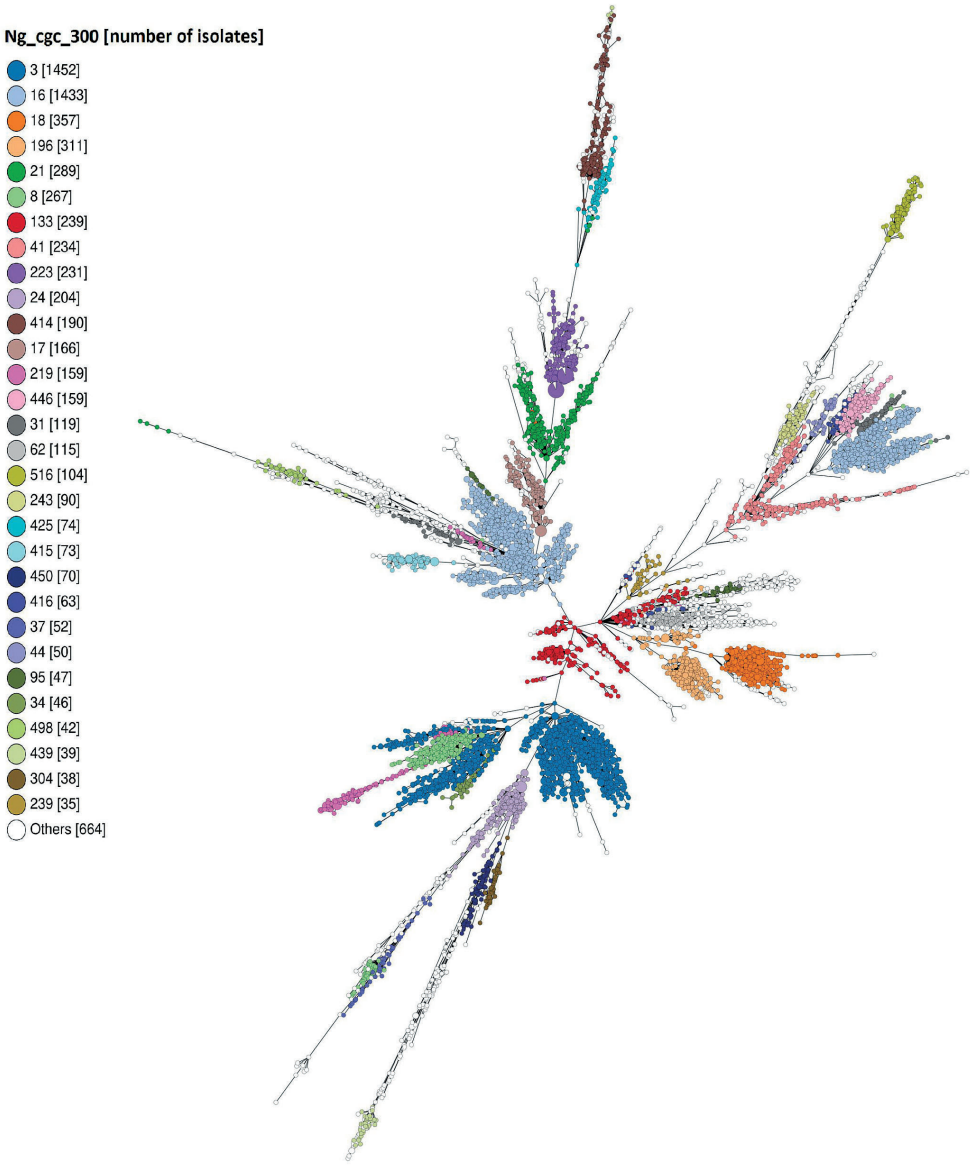


FIGURE 4. MINIMUM-SPANNING TREE BASED ON THE *N. GONORRHOEAE* ACCESSORY GENOME MLST TYPING SCHEME CREATED IN PUBMLST. ALL 8,013 PUBLICLY AVAILABLE *N. GONORRHOEAE* ISOLATES WERE INCLUDED IN THE TREE. COLOURS REPRESENT CORE GENOME MLST GROUPS USING THE 300 OR FEWER LOCUS THRESHOLD (NG_CGC_300). CLUSTERING OF COLOURS INDICATED THAT THE GONOCOCCAL ACCESSORY- AND CORE GENOMES ARE STRONGLY ASSOCIATED, CONFIRMING THAT THE GONOCOCCAL GENOME IS HIGHLY STRUCTURED. TWO SEPARATE CLUSTERS COULD BE IDENTIFIED FOR THE LARGEST CORE GENOME MLST GROUPS 3 AND 16. THE PRESENCE OF THESE TWO CLUSTERS INDICATES DIVERGENCE IN BOTH CORE AND ACCESSORY GENOME CONTENT AND A REQUIREMENT FOR MORE STRINGENT LOCUS THRESHOLDS TO BE USED TO DELINEATE THESE CLUSTERS. THE TREE WAS CREATED WITH GRAPETREE AND NODES WERE POSITIONED THROUGH DYNAMIC RENDERING, MEANING THAT BRANCH LENGTHS ARE NOT SCALED. LEGEND SHOWS THE CORE GENOME MLST GROUP (300 OR FEWER LOCUS THRESHOLD).

TABLE 2. PREVALENCE OF ACCESSORY ELEMENTS AND SIGNIFICANCE OF ASSOCIATIONS WITH CGMLST GROUPS (NG_CGC_300)

		Gonococcal Genetic Island		Conjugative plasmid		Beta-lactamase plasmid		VirB T4SS		Nf1 phage island		Putative secretion system	
CgMLST group	**N. isolates**	**Prevalence**	**sign.[a]**	**Prevalence**	**sign.[a]**	**Prevalence**	**sign.[a]**	**Prevalence**	**sign.[a]**	**Prevalence**	**sign.[a]**	**Prevalence**	**sign.[a]**
3	**1452**	99%	****[c]	0%	****[b]	1%	****[b]	0%	*[b]	0%	****[b]	97%	****[c]
16	**1433**	20%	****[b]	0%	****[b]	0%	****[b]	0%	*[b]	0%	****[b]	92%	****[c]
18	**357**	90%	****[c]	1%	****[b]	2%	****[b]	0%		15%	****[c]	0%	****[b]
196	**311**	100%	****[c]	4%	****[b]	0%	****[b]	0%		0%	**[b]	0%	****[b]
21	**289**	41%	****[b]	93%	****[c]	77%	****[c]	0%		0%	**[b]	0%	****[b]
8	**267**	52%	****[b]	0%	****[b]	0%	****[b]	0%		0%	*[b]	95%	****[c]
133	**239**	59%		100%	****[c]	31%	****[c]	0%		0%		93%	****[c]
41	**234**	99%	****[c]	100%	****[c]	48%	****[c]	0%		0%		0%	****[b]
223	**231**	99%	****[c]	1%	****[b]	0%	****[b]	0%		0%		0%	****[b]
24	**204**	100%	****[c]	2%	****[b]	0%	****[b]	0%		0%		0%	****[b]
414	**190**	100%	****[c]	7%	****[b]	1%	****[b]	0%		0%		0%	****[b]
17	**166**	81%		1%	****[b]	1%	***[b]	0%		0%		95%	****[c]
219	**159**	9%	****[b]	1%	****[b]	4%		0%		0%		0%	****[b]
446	**159**	100%	****[c]	100%	****[c]	1%	***[b]	0%		0%		1%	****[b]
31	**119**	54%		12%	*[b]	1%	**[b]	0%		0%		74%	****[c]
62	**115**	12%	****[b]	100%	****[c]	47%	****[c]	0%		0%		1%	****[b]
516	**104**	100%	****[c]	0%	****[b]	34%	****[c]	0%		0%		0%	****[b]
243	**90**	100%	****[c]	76%	****[c]	2%		0%		0%		0%	****[b]
425	**74**	100%	****[c]	49%	*[c]	3%		0%		1%		1%	****[b]
415	**73**	100%	****[c]	1%	****[b]	0%		0%		100%	****[c]	0%	****[b]
450	**70**	97%	****[c]	100%	****[c]	1%		1%		0%		0%	****[b]

TABLE 2. PREVALENCE OF ACCESSORY ELEMENTS AND SIGNIFICANCE OF ASSOCIATIONS WITH CGMLST GROUPS (NG_CGC_300). (CONTINUED)

		Gonococcal Genetic Island		Conjugative plasmid		Beta-lactamase plasmid		VirB T4SS		Nf1 phage island		Putative secretion system	
416	**63**	8%	****[b]	100%	****[c]	33%	***[c]	0%		0%		0%	****[b]
37	**52**	100%	****[c]	88%	****[c]	79%	****[c]	0%		0%		2%	****[b]
44	**50**	100%	****[c]	14%		0%		0%		0%		0%	****[b]
95	**47**	32%	****[b]	66%	****[c]	4%		0%		0%		98%	****[c]
34	**46**	98%	***[c]	0%	***[b]	0%		0%		0%		89%	****[c]
498	**42**	100%	****[c]	100%	****[c]	17%		0%		0%		0%	****[b]
439	**39**	5%	****[b]	90%	****[c]	0%		0%		0%		0%	****[b]
304	**38**	97%	**[c]	100%	****[c]	3%		5%		89%	****[c]	0%	****[b]
239	**35**	9%	****[b]	100%	****[c]	83%	****[c]	0%		0%		91%	****[c]
Other	**675**												
No cgMLST assigned	**601**												
Overall prevalence	**8013**	**69%**		**27%**		**11%**		**0,7%**		**3,6%**		**48%**	

[a]Adjusted p-values were calculated by multiplying the original p-value by the total number of core genome groups tested (n=250).

*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001

[b]significantly less prevalent in cgMLST group compared to distribution among all isolates

[c]significantly more prevalent in cgMLST group compared to distribution among all isolates

range: <0.05 to <0.0001; Table 3 and Table S4 and S5). The low number of ceftriaxone resistant isolates (n=22) hampered identification of significant associations, although genes encoding toxin-antitoxin systems were significantly more prevalent in ceftriaxone resistant isolates (P <0.05) and trends of increasing prevalence were seen for the other genes. The two hypothetical genes NEIS0364 to 0365, which always occur as a pair, were also significantly more prevalent in isolates resistant to all antibiotics except ceftriaxone (P-values <0.0001). BLAST annotation showed that NEIS0365 encodes a predicted competence protein which interacts with incoming DNA. Altogether, these findings indicated increased prevalence of genes encoding DNA uptake and excretion proteins in resistant isolates, independent of the class of antibiotic.

ACCESSORY GENES INVOLVED IN DNA UPTAKE AND EXCRETION ARE ASSOCIATED WITH AMR

Associations between the accessory genome and resistance against penicillin, tetracycline, ciprofloxacin, ceftriaxone, cefixime, and azithromycin were assessed. As expected, the beta-lactamase plasmid was significantly more prevalent among penicillin- (P <0.0001) or tetracycline- (P <0.0001) resistant isolates, compared with susceptible isolates. However, this was also found for ciprofloxacin-resistant isolates (P <0.0001). Interestingly, the genes encoding the putative secretion system (NEIS0080 to 0089), all methyltransferases (NEIS1193 to 1194, NEIS1311, NEIS2691 to 2692) except one (NEIS3182), a DNA transcription protein (NEIS3184) and a toxin-antitoxin system (NEIS3188 to NEIS3232) were significantly more prevalent among isolates resistant to penicillin, tetracycline, ciprofloxacin, cefixime, or azithromycin (P-value range: <0.05 to <0.0001; Table 3 and Table S4 and S5). The low number of ceftriaxone resistant isolates (n=22) hampered identification of significant associations, although genes encoding toxin-antitoxin systems were significantly more prevalent in ceftriaxone resistant isolates (P <0.05) and trends of increasing prevalence were seen for the other genes. The two hypothetical genes NEIS0364 to 0365, which always occur as a pair, were also significantly more prevalent in isolates resistant to all antibiotics except ceftriaxone (P-values <0.0001). BLAST annotation showed that NEIS0365 encodes a predicted competence protein which interacts with incoming DNA. Altogether, these findings indicated increased prevalence of genes encoding DNA uptake and excretion proteins in resistant isolates, independent of the class of antibiotic.

DISCUSSION

This study used a robust and detailed gene-by-gene approach to characterise the gonococcal accessory genome. Since the introduction of the pangenome concept in bacterial research, multiple computational tools have been developed to examine pangenome content[22]. Tools such as Roary and PIRATE are based on clustering algorithms which use gene similarity thresholds to identify gene families. These thresholds can be difficult to define as a consequence of major

TABLE 3. PREVALENCE OF ACCESSORY ELEMENTS IN ISOLATES WITH AVAILABLE PHENOTYPIC DATA AND THE SIGNIFICANCE OF ASSOCIATIONS.

			Gonococcal genetic island		Conjugative plasmid		Beta-lactamase plasmid		virB T4SS		Nf1 phage island		Putative secretion system	
Susceptibility[a]	**MIC**	**Number of isolates**	**prevalence**	**sign.[b]**	**prevalence**	**sign.[b]**	**prevalence**	**sign.[b]**	**prevalence**	**sign.[b]**	**prevalence**	**sign.[b]**	**prevalence**	**sign.[b]**
PEN S	**≤0.06**	**117**	89%	****[c]	32%	****	0%	****[d]	7%	****[c]	7%		16%	****[d]
PEN I	**>0.06-1.0**	**1896**	64%		23%		2%		1%		3%		52%	
PEN R	**>1.0**	**1074**	74%		33%		33%		0%		3%		53%	
TET S	**≤0.5**	**774**	74%	***[c]	23%	*[d]	8%	****[d]	2%	***[c]	4%		37%	****[d]
TET R	**>0.5**	**2143**	66%		28%		14%		1%		3%		57%	
CIP S	**≤0.03**	**2098**	65%	**[d]	26%		6%	****[d]	2%	****[c]	3%	**[c]	42%	****[d]
CIP R	**>0.06**	**2316**	70%		29%		16%		0%		2%		56%	
CRO S	**≤0.125**	**4868**	67%		26%		11%		1%		2%		51%	
CRO R	**>0.125**	**22**	91%		14%		0%		0%		0%		77%	
CFX S	**≤0.125**	**4004**	66%	****[d]	29%	****[c]	12%	****[c]	1%		3%	****[c]	47%	****[d]
CFX R	**>0.125**	**490**	85%		2%		1%		0%		0%		93%	
AZI S	**<1.0**	**3547**	71%	****[c]	29%	****[c]	13%	****[c]	1%		3%		45%	****[d]
AZI R	**≥1.0**	**898**	60%		8%		2%		0%		1%		75%	

[a] *PEN=penicillin; TET=tetracycline; CIP=ciprofloxacin; CRO=ceftriaxone; CFX=cefixime; AZI=azithromycin; S=susceptible; I= intermediate; R=resistant*

[b] *Adjusted p-values were calculated by multiplying the original p-value by the number of accessory elements tested (n=6). *p<0.05; **p0.01; ***p<0.001; ****p<0.0001.*

[c] *significantly more prevalent among susceptible isolates*

[d] *significantly more prevalent among resistant isolates*

differences in diversity among genes. This often results in over-splitting of divergent alleles of the same gene into multiple clusters or over-clustering of related gene families. PIRATE tries to improve this by measuring sequence diversity within the data set and correcting gene clusters based on this[16]. However, the output obtained here still contained far more gene clusters than genes identified in PubMLST, indicating over-splitting. This was evidenced by querying the identified PIRATE sequences in PubMLST, with multiple sequences matching the same NEIS locus. This therefore shows the necessity for careful examination of raw outputs resulting from pangenome tools for undertaking research on a gene level. Furthermore, gene-by-gene analysis combined with in-depth knowledge of the bacterial species is important to translate pangenome outputs from theoretical to functional genes.

The bacterial accessory genome provides genetic flexibility, facilitating host and environmental adaptation. In the gonococcus, major accessory elements include plasmids, mediating resistance against beta-lactams and tetracycline. The content of the chromosomal accessory genome and its implications in population biology and AMR are largely unknown. Due to frequent intraspecies HGT, high genetic diversity is found in gonococci exemplified by the challenges encountered in vaccine development and rapidity with which AMR develops. Such genetic diversity was also expected in the accessory genome, however, our results showed that a small gene pool constituted the accessory genome with the GGI and plasmids being the main accessory elements. Characterisation of the chromosomal gonococcal accessory genome revealed that it was highly organised consistent with retention of gene synteny despite frequent HGT and that it was associated with the core genome. Therefore, the main driver of allelic variation in the gonococcal population is the high rate of genetic exchange, whereas gene acquisition occurs to a much lesser extent[23]. This is evidenced by the rapid saturation in new genes identified once a small number of genomes were examined, and the increase in gene diversity seen as more genomes were included (Figure 2). Recombination being predominant to gene gain and loss has also been observed for *Chlamydia trachomatis*, suggesting a typical genome constitution for sexually transmitted pathogens[24]. The highly structured genome organization and small gene pool size are indicative of a founder effect resulting from an ancestral gonococcus first transitioning to the urogenital tract, and resulting in a niche switch and a change in genome content that has been selected for and preserved. This could explain why the *Neisseria* genus consists mostly of commensal and symbiotic species residing in the human nasopharynx, with the gonococcus being unusual in that it is principally found in the male and female urogenital tracts[1].

The limited gonococcal gene pool is possibly a consequence of the gonococcus being for the most part an obligate pathogen. Compared with opportunistic pathogens that are highly adaptable to diverse environments through their extensive accessory genomes, obligate pathogens often have limited pangenomes, a consequence of reductive evolution due to

host restriction[5,25-27]. Gene sharing is also more prevalent among species that share the same habitat, and the gene diversity in the habitat determines the size and diversity of the species' gene pool[6,7]. Species existing in diverse communities often have an open and highly variable pangenome, while "niche specialists" that exist in stable, less variable, and host-restricted environments have much smaller and more conserved pangenomes[8]. The gonococcus is such a host-restricted pathogen, predominantly inhabiting the human urogenital tract and despite the male and female urogenital tract being different environments, the gonococcus can be considered a niche specialist. Pharyngeal gonococcal infections nonetheless occur, which will facilitate interspecies genetic exchange between *N. gonorrhoeae*, *N. meningitidis*, and commensal *Neisseria* inhabiting the nasopharynx[28]. These pharyngeal infections are considered to be the main driver for uptake of antibiotic resistance mutations from other *Neisseria* species, facilitated by the high prevalence of these infections in key populations[29]. Genetic exchange has also been shown to occur from gonococci to urogenital associated meningococci, of which increasing cases have been reported over the last decade[30,31]. This shows the potential for genetic exchange between these two *Neisseria* species[32]. Further studies assessing the prevalence of gonococci and meningococci co-existing in the same environment may determine the likelihood of increased HGT between them.

The gonococcal accessory and core genomes were associated, indicating ancestral inheritance of accessory elements in certain gonococcal lineages which are conserved during lineage expansion with limited spread to other lineages. This is consistent with a previous finding showing that plasmid distribution was lineage-associated in the gonococcus[11]. Similar structuring of bacterial populations, with strong associations between lineage and accessory genome content, has been found in other bacterial species, including *Escherichia coli*[33]. This indicates that bacterial lineages are structured by both core and accessory genome content. Genomic elements, such as transposons or insertion sequences, play key roles in intra- and interchromosomal rearrangements[34]. For example, GGI uptake is regulated by the GGI insertion sequence and recombination in this region leads to GGI loss[35]. Further examination of these elements might inform whether lineage specific prevalence of accessory elements is associated with particular insertion sequences. Examination of these sequences, which often contain repetitive regions, is challenging using current short-read sequencing techniques. Such studies are foreseeable in the near future with the advent of cheaper and more accessible long-read sequencing techniques.

Significant associations were found between increased prevalence of DNA uptake and exchange regulators and AMR, suggesting a role for genetic exchange in the development or maintenance of AMR. This also supports the hypothesis of certain genetic backgrounds being predisposed for resistance, as in the gonococcus resistance is highly associated with the core genome and specific lineages are more resistant than others[10]. Domenech et al. suggested reducing the

spread of AMR by inhibiting bacterial competence[36]. Important to note is that the prevalence indicated in this study is only at the gene level, whereas upregulation of DNA exchange and uptake should be confirmed by functional examination at the gene expression level.

CONCLUSIONS

Despite frequent HGT being responsible for high genetic diversity in the gonococcus, the gonococcal gene pool was found to be small with the accessory genome consisting mainly of the GGI and plasmids. In addition, the gonococcal accessory genome was associated with the core genome, consistent with the ordered population structure previously identified. These data indicate that recombination predominantly drives allelic variation and to a much lesser extent gene acquisition or loss. Increased prevalence of genes encoding DNA exchange regulators in resistant gonococcal isolates suggests that genetic material exchange may play a role in the development or maintenance of AMR. Understanding the role of such mechanisms will be important to limit the spread of AMR in the gonococcus. The findings presented here enhance our understanding of gonococcal genome architecture and have important implications for gonococcal population biology.

DATA AVAILABILITY

The data sets supporting the conclusions of this article are available in the PubMLST Neisseria database repository (https://pubmlst.org/organisms/neisseria-spp). PubMLST IDs can be found in Table S1.

FUNDING

The project was funded by the Public Health Laboratory of Amsterdam. M.C.J.M. and O.B.H. are funded by the Wellcome Trust (Grants 218205/Z/19/Z and 214374/Z/18/Z).

REFERENCES

1. McSheffrey GG, Gray-Owen SD. *Neisseria gonorrhoeae*. *Molecular Medical Microbiology*. 2nd ed. Elsevier; 2015
2. Unemo M, Shafer WM. Antimicrobial Resistance in *Neisseria gonorrhoeae* in the 21st Century: Past, Evolution, and Future. *Clin Microbiol Rev*. 2014 Jul 1;27:587–613.
3. Kirkcaldy RD, Weston E, Segurado AC, Hughes G. Epidemiology of Gonorrhea: A Global Perspective. *Sex Health*. 2019;16:401–11.
4. Unemo M, Golparian D, Eyre DW. Antimicrobial Resistance in *Neisseria gonorrhoeae* and Treatment of Gonorrhea. *Methods Mol Biol*. 2019;1997:37–58.
5. Dobrindt U, Hacker J. Whole genome plasticity in pathogenic bacteria. *Curr Opin Microbiol*. 2001;4:550–7.
6. Hooper SD, Mavromatis K, Kyrpides NC. Microbial co-habitation and lateral gene transfer: what transposases can tell us. *Genome Biol*. 2009;10:R45.
7. Popa O, Dagan T. Trends and barriers to lateral gene transfer in prokaryotes. *Curr Opin Microbiol*. 2011;14:615–23.
8. Brockhurst MA, Harrison E, Hall JPJ, Richards T, McNally A, MacLean C. The Ecology and Evolution of Pangenomes. *Curr Biol*. 2019;29:R1094–103.
9. Golparian D, Harris SR, Sánchez-Busó L, Hoffmann S, Shafer WM, Bentley SD, et al. Genomic evolution of *Neisseria gonorrhoeae* since the preantibiotic era (1928–2013): antimicrobial use/misuse selects for resistance and drives evolution. *BMC Genomics*. 2020;21:116.
10. Harrison OB, Cehovin A, Skett J, Jolley KA, Massari P, Genco CA, et al. *Neisseria gonorrhoeae* Population Genomics: Use of the Gonococcal Core Genome to Improve Surveillance of Antimicrobial Resistance. *J Infect Dis*. 2020;222:1816–25.
11. Cehovin A, Jolley KA, Maiden MCJ, Harrison OB, Tang CM. Association of *Neisseria gonorrhoeae* Plasmids With Distinct Lineages and The Economic Status of Their Country of Origin. *J Infect Dis*. 2020;222:1826–36.
12. de Korne-Elenbaas J, Bruisten SM, de Vries HJC, Van Dam AP. Emergence of a *Neisseria gonorrhoeae* clone with reduced cephalosporin susceptibility between 2014 and 2019 in Amsterdam, The Netherlands, revealed by genomic population analysis. *J Antimicrob Chemother*. 2021;76:1759-1768.
13. Zhou Z, Alikhan N-F, Sergeant MJ, Luhmann N, Vaz C, Francisco AP, et al. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Res*. 2018;28:1395–404.
14. Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*. 2010;11:595.
15. Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.
16. Bayliss SC, Thorpe HA, Coyle NM, Sheppard SK, Feil EJ. PIRATE: A fast and scalable pangenomics toolbox for clustering diverged orthologues in bacteria. *GigaScience*. 2019;8:giz119.
17. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31:3691–3.
18. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
19. Grant JR, Stothard P. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res*. 2008;36:W181–4.
20. Al Suwayyid BA, Rankine-Wilson L, Speers DJ, Wise MJ, Coombs GW, Kahler CM. Meningococcal Disease-Associated Prophage-Like Elements Are Present in *Neisseria gonorrhoeae* and Some Commensal *Neisseria* Species. *Genome Biol Evol*. 2020;12:3938–50.
21. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 11.0, 2021. http://www.eucast.org.
22. Vernikos GS. A Review of Pangenome Tools and Recent Studies. In: Tettelin H, Medini D, editors. *The Pangenome: Diversity, Dynamics and Evolution of Genomes*. Cham: Springer International Publishing; 2020. p. 89–112.

23. Hamilton HL, Dillard JP. Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol Microbiol.* 2006;59(2):376–85.
24. Joseph SJ, Didelot X, Gandhi K, Dean D, Read TD. Interplay of recombination and selection in the genomes of *Chlamydia trachomatis*. *Biol Direct.* 2011;6:28.
25. Dobrindt U, Chowdary MG, Krumbholz G, Hacker J. Genome dynamics and its impact on evolution of *Escherichia coli*. *Med Microbiol Immunol (Berl).* 2010;199(3):145–54.
26. Martin RM, Bachman MA. Colonization, Infection, and the Accessory Genome of *Klebsiella pneumoniae*. *Front Cell Infect Microbiol.* 2018;8:4.
27. Moran NA. Microbial Minimalism: Genome Reduction in Bacterial Pathogens. *Cell.* 2002;108(5):583–6.
28. Marri PR, Paniscus M, Weyand NJ, Rendón MA, Calton CM, Hernández DR, et al. Genome Sequencing Reveals Widespread Virulence Gene Exchange among Human *Neisseria* Species. *PLoS ONE*. 2010;5:e11835.
29. Hook EW, Bernstein K. Kissing, saliva exchange, and transmission of *Neisseria gonorrhoeae*. *Lancet Infect Dis*. 2019;19:e367–9.
30. Kanemitsu N, Hayashi I, Satoh N, Hayakawa T, Mitsuya H, Hayase Y, et al. Acute urethritis caused by *Neisseria meningitidis*. *Int J Urol Off J Jpn Urol Assoc*. 2003;10:346–7.
31. Hayakawa K, Itoda I, Shimuta K, Takahashi H, Ohnishi M. Urethritis Caused by Novel *Neisseria meningitidis* Serogroup W in Man Who Has Sex with Men, Japan. *Emerg Infect Dis*. 2014;20:1585–7.
32. Harrison OB, Cole K, Peters J, Cresswell F, Dean G, Eyre DW, et al. Genomic analysis of urogenital and rectal *Neisseria meningitidis* isolates reveals encapsulated hyperinvasive meningococci and coincident multidrug-resistant gonococci. *Sex Transm Infect*. 2017;93:445–51.
33. McNally A, Oren Y, Kelly D, Pascoe B, Dunn S, Sreecharan T, et al. Combined Analysis of Variation in Core, Accessory and Regulatory Genome Regions Provides a Super-Resolution View into the Evolution of Bacterial Populations. *PLoS Genet*. 2016;12:e1006280.
34. Cehovin A, Lewis SB. Mobile genetic elements in *Neisseria gonorrhoeae*: Movement for change. *Pathog Dis*. 2017;75:1–12.
35. Domínguez NM, Hackett KT, Dillard JP. XerCD-Mediated Site-Specific Recombination Leads to Loss of the 57-Kilobase Gonococcal Genetic Island. *J Bacteriol.* 2011;193:377–88.
36. Domenech A, Brochado AR, Sender V, Hentrich K, Henriques-Normark B, Typas A, et al. Proton Motive Force Disruptors Block Bacterial Competence and Horizontal Gene Transfer. *Cell Host Microbe.* 2020;27:544-555.e3.

SUPPLEMENTARY MATERIAL

TABLE S1. ISOLATE CHARACTERISTICS.

This table (7 columns and 766 rows) is available with the online version of this article (https://tinyurl.com/5c7aww9m).

TABLE S2. LIST OF ACCESSORY NEIS LOCI WITH ITS PREVALENCE AND FUNCTIONAL ANNOTATION.

Phage associated	Hypothetical proteins	Mobile Genetic Elements	Toxin-antitoxin system	Methyltransferases	Putative secretion system	TonB dependent receptor	DNA transcription	Other functions

NEIS locus	Duplicate NEIS loci	Prevalence in subset of 765 isolates	Prevalence in all 8013 isolates	(PubMLST) gene annotation	PubMLST functional scheme (NG schemes only)	Final annotation	Identified by
NEIS0027	NEIS1865	3,7%	3,6%	hypothetical integral membrane protein	-	Nf1 phage island	GC + PIRATE
NEIS0028	NEIS1864	3,7%	3,6%	hypothetical integral membrane protein	-	Nf1 phage island	GC + PIRATE
NEIS0029	NEIS1863/NEIS2453/ NEIS2459	3,7%	3,6%	hypothetical protein	-	Nf1 phage island	GC + PIRATE
NEIS0030	NEIS1862/NEIS2452/ NEIS2460	3,7%	3,6%	conserved hypothetical protein	-	Nf1 phage island	GC + PIRATE
NEIS0031	NEIS1861/NEIS0277	3,7%	3,6%	putative phage replication initiation factor	-	Nf1 phage island	GC + PIRATE
NEIS0080		38,4%	48,7%	hypothetical protein	Other loci	Putative secretion system	GC + PIRATE
NEIS0081		38,4%	48,7%	hypothetical protein	Other loci	Putative secretion system	GC
NEIS0083		37,8%	48,0%	hypothetical protein	Other loci	Putative secretion system	GC + PIRATE
NEIS0084		37,9%	48,4%	hypothetical protein	Other loci	Putative secretion system	GC
NEIS0085		38,4%	48,6%	protein export protein	Other loci	Putative secretion system	GC + PIRATE
NEIS0086		38,2%	48,4%	protein export protein	Other loci	Putative secretion system	GC + PIRATE
NEIS0089		38,4%	48,7%	hypothetical protein	Other loci	Putative secretion system	GC + PIRATE
NEIS0364		78,3%	79,6%	hypothetical protein	Other loci	Hypothetical protein	GC + PIRATE
NEIS0365		78,3%	79,3%	hypothetical protein	Other loci	Hypothetical protein	GC + PIRATE
NEIS0486		30,1%	26,0%	alcohol dehydrogenase	-	Alcohol dehydrogenase	GC + PIRATE

TABLE S2. LIST OF ACCESSORY NEIS LOCI WITH ITS PREVALENCE AND FUNCTIONAL ANNOTATION. (CONTINUED)

Phage associated	Hypothetical proteins	Mobile Genetic Elements	Toxin-antitoxin system	Methyltransferases	Putative secretion system	TonB dependent receptor	DNA transcription	Other functions

NEIS locus	Duplicate NEIS loci	Prevalence in subset of 765 isolates	Prevalence in all 8013 isolates	(PubMLST) gene annotation	PubMLST functional scheme (NG schemes only)	Final annotation	Identified by
NEIS0524		52,3%	74,4%	putative peptidase	-	Putative peptidase	GC
NEIS0594		38,2%	35,0%	hypothetical protein	Other loci	Hypothetical protein	GC + PIRATE
NEIS0840		1,2%	0,8%	phage related protein	-	(Putative) phage associated protein	GC
NEIS0953	NEIS1664	27,6%	17,3%	hypothetical protein	-	(Putative) phage associated protein	GC + PIRATE
NEIS0955		80,4%	92,0%	hypothetical protein	Other loci	Nf4 phage island	GC + PIRATE
NEIS1089		28,5%	26,7%	hypothetical protein	-	Hypothetical protein	GC
NEIS1193		91,1%	92,1%	type III restriction-modification system endonuclease protein	Methyltransferases	Methyltransferase	GC + PIRATE
NEIS1194		90,8%	90,7%	Type III restriction/modification system methyltransferase	Methyltransferases	Methyltransferase	GC + PIRATE
NEIS1311		85,0%	84,0%	putative type III restriction/ modification system enzyme	Methyltransferases	Methyltransferase	GC
NEIS1667		22,9%	12,4%	hypothetical protein	-	Nf4 phage island	GC + PIRATE
NEIS2202		36,9%	26,5%	DNA primase	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2203		37,5%	27,1%	transcription elongation factor	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2204		37,6%	27,2%	yegA / hypothetical protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2205		37,6%	27,2%	hypothetical protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2206		37,3%	27,1%	hypothetical protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2207		37,5%	27,2%	hypothetical protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2208		37,0%	26,9%	putative DNA modification methylase	Conjugative Plasmid/ Methyltransferases	Conjugative Plasmid	GC + PIRATE
NEIS2209		37,4%	27,1%	hypothetical protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2210		22,9%	14,3%	tetM / tetracycline resistance protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2211		22,6%	13,8%	hypothetical protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE

NEIS2212	22,4%	13,8%	epsilon_2 antitoxin	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2213	37,5%	27,0%	zeta_2 toxin	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2214	37,8%	27,3%	-	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2215	37,3%	26,9%	zeta_1 toxin	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2216	37,6%	27,2%	marR / transcriptional regulator	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2217	37,5%	27,1%	res / putative resolvase	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2218	37,5%	27,2%	vapD / toxin	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2219	37,5%	27,0%	trbN / lytic transglycosylase	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2220	37,8%	27,0%	trbM / involved in conjugative transfer	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2221	37,3%	26,9%	trbL / virB6-like inner membrane protein	Conjugative Plasmid	Conjugative Plasmid	GC
NEIS2222	37,5%	26,6%	trbK / putative entry exclusion protein	Conjugative Plasmid	Conjugative Plasmid	GC
NEIS2223	37,4%	27,1%	trbJ	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2224	37,3%	26,8%	trbI / virbB10-like core complex component	Conjugative Plasmid	Conjugative Plasmid	GC
NEIS2225	37,6%	27,2%	trbH / virB7-like core complex component	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2226	37,3%	26,9%	trbG / virB9-like core complex component	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2227	37,5%	27,1%	trbF / DNA transfer protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2228	36,9%	26,8%	trbE / virB4-like conjugal transfer ATPase	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2229	37,8%	27,2%	trbD / similar to VirB3	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2230	37,8%	27,2%	trbC / virB2 prepilin	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2231	37,3%	26,8%	trbB / virB11-like conjugal transfer ATPase	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2232	37,9%	27,2%	trbA / transcriptional repressor	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE

TABLE S2. LIST OF ACCESSORY NEIS LOCI WITH ITS PREVALENCE AND FUNCTIONAL ANNOTATION. (CONTINUED)

Phage associated	Hypothetical proteins	Mobile Genetic Elements	Toxin-antitoxin system	Methyltransferases	Putative secretion system	TonB dependent receptor	DNA transcription	Other functions

NEIS locus	Duplicate NEIS loci	Prevalence in subset of 765 isolates	Prevalence in all 8013 isolates	(PubMLST) gene annotation	PubMLST functional scheme (NG schemes only)	Final annotation	Identified by
NEIS2233		37,6%	27,2%	ssb / putative single stranded DNA binding protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2234		37,4%	27,1%	trfA / oriV activator protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2235		37,8%	27,2%	korC / regulator of plasmid partitioning	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2236		37,6%	27,2%	kleE / involved in plasmid partitioning	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2237		37,5%	27,1%	incC2 / ATP binding protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2238		37,4%	27,1%	parB / DNA binding protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2239		35,2%	26,5%	kfrB / putative kfrB/traO	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2240		37,6%	27,2%	kfrC / putative kfrC/traN protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2241		37,4%	27,0%	traM / involved in conjugal transfer	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2242		37,4%	27,1%	traL / involved in conjugal transfer	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2243		37,5%	27,2%	traK / involved in conjugal transfer	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2244		37,9%	27,2%	traJ / oriT	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2245		37,1%	26,6%	traI / DNA relaxase	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2246		36,7%	26,4%	traG / conjugal coupling protein	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2247		37,5%	27,2%	traF / conjugal prepilin peptidase	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2248		37,5%	26,9%	traD / DNA topoisomerase III	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2249		37,5%	27,1%	traD / involved in conjugal transfer	Conjugative Plasmid	Conjugative Plasmid	GC + PIRATE
NEIS2250		63,9%	69,0%	traD / putative docking protein	GGI	GGI	GC + PIRATE
NEIS2251		63,8%	68,9%	traI / putative nicking enzyme	GGI	GGI	GC + PIRATE
NEIS2252		63,9%	68,9%	yaf / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2253		63,9%	68,9%	ltgX / peptidoglycan hydrolase	GGI	GGI	GC + PIRATE
NEIS2254		63,9%	69,0%	yag / outer membrane protein	GGI	GGI	GC + PIRATE

NEIS2255	64,2%	69,0%	traA / putative transfer protein	GGI	GGI	GC + PIRATE
NEIS2256	64,1%	69,0%	traL / pilus assembly	GGI	GGI	GC + PIRATE
NEIS2257	63,9%	69,0%	traE / pilus biogenesis	GGI	GGI	GC + PIRATE
NEIS2258	63,8%	68,9%	traK / pilus assembly	GGI	GGI	GC + PIRATE
NEIS2259	63,8%	68,8%	traB / conjugal transfer	GGI	GGI	GC + PIRATE
NEIS2260	63,9%	68,9%	dsbC / protein disulphide isomerase	GGI	GGI	GC + PIRATE
NEIS2261	63,9%	68,9%	traV / putative transfer protein	GGI	GGI	GC + PIRATE
NEIS2262	63,7%	68,7%	traC / pilus assembly	GGI	GGI	GC + PIRATE
NEIS2263	63,9%	69,0%	ybe / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2264	64,1%	69,0%	trbI / conjugal transfer	GGI	GGI	GC + PIRATE
NEIS2265	63,9%	68,9%	traW / pilus biogenesis	GGI	GGI	GC + PIRATE
NEIS2266	63,1%	68,8%	traU / pilus biogenesis	GGI	GGI	GC + PIRATE
NEIS2267	63,7%	68,7%	trbC / conjugative transfer	GGI	GGI	GC + PIRATE
NEIS2268	63,8%	68,8%	ybi / mating-pair stabilisation	GGI	GGI	GC + PIRATE
NEIS2269	63,5%	68,7%	traN / mating-pair stabilisation	GGI	GGI	GC + PIRATE
NEIS2270	63,8%	68,8%	ycb / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2271	63,5%	68,8%	traF / pilus assembly	GGI	GGI	GC + PIRATE
NEIS2272	63,8%	68,8%	traH / pilus assembly	GGI	GGI	GC + PIRATE
NEIS2273	63,4%	68,8%	traG / pilus assembly; mating-pair stabilisation	GGI	GGI	GC + PIRATE
NEIS2274	58,0%	61,5%	atlA / peptidoglycan transglycosylase	GGI	GGI	GC + PIRATE
NEIS2275	58,2%	61,6%	ych / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2276	57,9%	61,1%	exp1 / exported protein	GGI	GGI	GC
NEIS2277	59,5%	62,1%	cspA / RNA/ssDNA binding protein	GGI	GGI	GC + PIRATE
NEIS2278	57,3%	60,1%	exp2 / hypothetical protein	GGI	GGI	GC + PIRATE

TABLE S2. LIST OF ACCESSORY NEIS LOCI WITH ITS PREVALENCE AND FUNCTIONAL ANNOTATION. (CONTINUED)

Phage associated	Hypothetical proteins	Mobile Genetic Elements	Toxin-antitoxin system	Methyltransferases	Putative secretion system	TonB dependent receptor	DNA transcription	Other functions

NEIS locus	Duplicate NEIS loci	Prevalence in subset of 765 isolates	Prevalence in all 8013 isolates	(PubMLST) gene annotation	PubMLST functional scheme (NG schemes only)	Final annotation	Identified by
NEIS2279		63,5%	68,3%	yda / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2280		63,3%	68,9%	ydbA / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2281		63,7%	68,9%	ydbB / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2282		63,9%	69,1%	ydcA / putative protease	GGI	GGI	GC + PIRATE
NEIS2283		63,8%	69,1%	ydcB / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2284		63,9%	69,1%	ydd / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2285		63,9%	69,2%	ydeA / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2286		63,8%	69,0%	ydeB / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2287		63,7%	69,0%	ydf / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2288		63,9%	69,0%	ydg / putative DNA methylase	GGI/ Methyltransferases	GGI	GC + PIRATE
NEIS2289		63,1%	68,9%	ydhA / putative DNA methylase	GGI/ Methyltransferases	GGI	GC + PIRATE
NEIS2290		63,8%	69,1%	ydhB / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2291		63,7%	69,0%	ydi / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2292		62,9%	68,7%	yea / putative helicase	GGI	GGI	GC + PIRATE
NEIS2293		61,6%	63,1%	yeb / putative N-acetyltransferase	GGI	GGI	GC + PIRATE
NEIS2294		63,7%	69,0%	yecA / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2295		62,0%	67,2%	yecB / repeat containing protein	GGI	GGI	GC + PIRATE
NEIS2296		63,8%	69,1%	yedA / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2297		63,9%	69,0%	yedB / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2298		64,1%	69,1%	yee / hypothetical protein	GGI	GGI	GC
NEIS2299		63,8%	69,1%	yegA / hypothetical protein	GGI	GGI	GC + PIRATE

NEIS2300	63,9%	69,2%	yegB / hypothetical protein	GGI	GGI	GC
NEIS2301	61,2%	67,8%	yeh / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2302	63,7%	68,9%	topB / DNA topoisomerase	GGI	GGI	GC + PIRATE
NEIS2303	64,2%	69,1%	ssbB / single-stranded DNA binding protein	GGI	GGI	GC + PIRATE
NEIS2304	63,8%	68,6%	yfa / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2305	63,9%	68,9%	yfb / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2306	63,9%	68,9%	yfd / putative tonB-like transporter	GGI	GGI	GC + PIRATE
NEIS2307	63,3%	69,0%	yfeA / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2308	64,1%	68,9%	yfeB / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2309	64,1%	69,0%	parB / chromosome partitioning	GGI	GGI	GC + PIRATE
NEIS2310	63,5%	68,9%	parA / chromosome partitioning	GGI	GGI	GC + PIRATE
NEIS2311	5,6%	7,5%	eppA / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2312	5,6%	7,5%	ych1 / hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2313	64,1%	69,1%	hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2314	63,9%	69,1%	hypothetical protein	GGI	GGI	GC + PIRATE
NEIS2315	64,1%	69,1%	conjugal transfer protein	GGI	GGI	GC + PIRATE
NEIS2323	1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2324	1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2325	1,2%	0,6%	-	VirB T4SS	VirB T4SS	GC
NEIS2326	1,2%	0,6%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2327	1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2328	1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2329	1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2330	1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE

TABLE S2. LIST OF ACCESSORY NEIS LOCI WITH ITS PREVALENCE AND FUNCTIONAL ANNOTATION. (CONTINUED)

Phage associated	Hypothetical proteins	Mobile Genetic Elements	Toxin-antitoxin system	Methyltransferases	Putative secretion system	TonB dependent receptor	DNA transcription	Other functions

NEIS locus	Duplicate NEIS loci	Prevalence in subset of 765 isolates	Prevalence in all 8013 isolates	(PubMLST) gene annotation	PubMLST functional scheme (NG schemes only)	Final annotation	Identified by
NEIS2331		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2332		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2333		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2334		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2335		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2336		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2337		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2338		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2339		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2342		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2343		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2345		88,4%	93,5%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2348		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2350		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2351		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2352		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2353		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC
NEIS2354		1,2%	0,7%	-	VirB T4SS	VirB T4SS	GC + PIRATE
NEIS2356		15,4%	13,5%	-	Conjugative Plasmid	Conjugative Plasmid	GC
NEIS2357		16,3%	10,6%	beta lactamase	Beta-lactamase plasmid	Beta-lactamase plasmid	GC + PIRATE
NEIS2358		19,1%	11,1%	replicase A	Beta-lactamase plasmid	Beta-lactamase plasmid	GC + PIRATE
NEIS2359		19,2%	11,3%	resolvase	Beta-lactamase plasmid	Beta-lactamase plasmid	GC + PIRATE

NEIS2360		7,2%	3,4%	-	Beta-lactamase plasmid	Beta-lactamase plasmid	GC + PIRATE
NEIS2451	NEIS2461	3,7%	3,6%	hypothetical protein	Other loci	Nf1 phage island	GC
NEIS2486		89,2%	90,8%	conserved hypothetical protein	Other loci	Hypothetical protein	GC + PIRATE
NEIS2582		20,4%	19,2%	hypothetical protein	Other loci	Hypothetical protein	GC + PIRATE
NEIS2615		49,0%	60,1%	hypothetical protein, putative phage associated protein	Other loci	(Putative) phage associated protein	GC + PIRATE
NEIS2616		65,8%	80,5%	replicative DNA helicase	Other loci	(Putative) phage associated protein	GC + PIRATE
NEIS2618		39,3%	64,6%	hypothetical protein, putative phage associated protein	Other loci	(Putative) phage associated protein	GC + PIRATE
NEIS2619		50,5%	58,5%	hypothetical protein, putative phage associated protein	Other loci	(Putative) phage associated protein	GC + PIRATE
NEIS2621		4,7%	5,9%	hypothetical protein, putative phage associated protein	Other loci	(Putative) phage associated protein	GC + PIRATE
NEIS2640		26,0%	22,6%	hypothetical protein, putative phage associated protein	Other loci	(Putative) phage associated protein	GC
NEIS2643		49,8%	73,4%	hypothetical protein, putative phage associated protein	Other loci	(Putative) phage associated protein	GC + PIRATE
NEIS2646		4,7%	2,7%	putative TonB-dependent receptor	Other loci	TonB dependent receptor	GC
NEIS2647		5,4%	3,8%	hypothetical protein	Other loci	TonB dependent receptor	GC
NEIS2666		82,6%	86,7%	hypothetical protein	Other loci	Hypothetical protein	GC + PIRATE
NEIS2691		87,7%	88,5%	restriction endonuclease R.NgoMIII	Methyltransferases	Methyltransferase	GC + PIRATE
NEIS2692		87,5%	88,5%	DNA cytosine methyltransferase M.NgoMIII	Methyltransferases	Methyltransferase	GC + PIRATE
NEIS2693		82,5%	82,7%	conserved hypothetical protein, putative phage associated protein	Other loci	Nf4 phage island	GC + PIRATE
NEIS2705		56,9%	50,0%	hypothetical protein	Other loci	Hypothetical protein	GC
NEIS2713		28,4%	22,3%	phage associated protein	Other loci	(Putative) phage associated protein	GC + PIRATE
NEIS2714		55,2%	59,6%	phage associated protein	Other loci	(Putative) phage associated protein	GC + PIRATE

TABLE S2. LIST OF ACCESSORY NEIS LOCI WITH ITS PREVALENCE AND FUNCTIONAL ANNOTATION. (CONTINUED)

Phage associated	Hypothetical proteins	Mobile Genetic Elements	Toxin-antitoxin system	Methyltransferases	Putative secretion system	TonB dependent receptor	DNA transcription	Other functions

NEIS locus	Duplicate NEIS loci	Prevalence in subset of 765 isolates	Prevalence in all 8013 isolates	(PubMLST) gene annotation	PubMLST functional scheme (NG schemes only)	Final annotation	Identified by
NEIS2718		47,7%	68,3%	phage associated protein	Other loci	(Putative) phage associated protein	GC
NEIS2744		94,2%	94,2%	TonB dependent receptor (SLAM)	Other loci	TonB dependent receptor	GC
NEIS2859		8,8%	24,3%	hypothetical protein	Other loci	Hypothetical protein	GC + PIRATE
NEIS2960		19,5%	12,1%	hypothetical protein	Beta-lactamase plasmid	Beta-lactamase plasmid	GC
NEIS2961		14,0%	8,3%	DNA strand transferase	Beta-lactamase plasmid	Beta-lactamase plasmid	GC + PIRATE
NEIS2962		13,3%	8,8%	hypothetical protein	Beta-lactamase plasmid	Beta-lactamase plasmid	GC + PIRATE
NEIS2964		19,0%	11,3%	hypothetical protein	Beta-lactamase plasmid	Beta-lactamase plasmid	GC + PIRATE
NEIS3177		0,7%	1,5%	(Putative) phage associated protein		Phage island X	PIRATE
NEIS3178		26,7%	38,6%	Hypothetical protein		Hypothetical protein	PIRATE
NEIS3179		74,6%	93,6%	Hypothetical protein		Hypothetical protein	PIRATE
NEIS3180		85,9%	79,5%	Hypothetical protein		Hypothetical protein	PIRATE
NEIS3181		93,1%	89,5%	(Putative) phage associated protein		(Putative) phage associated protein	PIRATE
NEIS3182		40,0%	57,6%	Restriction modification enzyme		Methyltransferase	PIRATE
NEIS3183		78,3%	93,5%	Hypothetical protein		Hypothetical protein	PIRATE
NEIS3184		74,2%	70,8%	Histone deacetylase		DNA transcription	PIRATE
NEIS3186		0,5%	1,2%	(Putative) phage associated protein		Phage island X	PIRATE
NEIS3188		27,6%	32,9%	Type IV toxin-antitoxin system, putative AbiEii toxin		Toxin-antitoxin system	PIRATE
NEIS3189		87,1%	93,9%	TspB virulence factor		TspB virulence factor	PIRATE
NEIS3190		4,7%	2,3%	Hypothetical protein		Conjugative Plasmid	PIRATE
NEIS3192		17,8%	16,2%	Hypothetical protein		GGI	PIRATE

NEIS3193	70,8%	80,8%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3195	0,7%	1,5%	Hypothetical protein	Phage island X	PIRATE
NEIS3196	89,8%	90,6%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3197	94,5%	94,4%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3198	59,7%	81,7%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3200	37,6%	30,1%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3202	1,6%	1,7%	DNA replication protein	DNA replication	PIRATE
NEIS3203	1,4%	2,0%	(Putative) phage associated protein	(Putative) phage associated protein	PIRATE
NEIS3204	0,7%	1,5%	(Putative) phage associated protein	Phage island X	PIRATE
NEIS3205	1,6%	2,1%	(Putative) phage associated protein	(Putative) phage associated protein	PIRATE
NEIS3206	44,3%	27,4%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3207	0,7%	1,6%	Hypothetical protein	Phage island X	PIRATE
NEIS3208	67,1%	71,0%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3209	61,0%	61,8%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3210	23,0%	23,8%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3211	26,3%	30,9%	Helix-turn-helix transcriptional regulator	DNA transcription	PIRATE
NEIS3212	1,2%	0,7%	Hypothetical protein	VirB T4SS	PIRATE
NEIS3213	2,4%	1,4%	DUF domain-containing protein	Hypothetical protein	PIRATE
NEIS3214	24,1%	20,3%	(Putative) phage associated protein	(Putative) phage associated protein	PIRATE
NEIS3215	84,8%	88,0%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3216	12,5%	13,8%	Hypothetical protein	Hypothetical protein	PIRATE
NEIS3217	1,8%	2,0%	DUF domain-containing protein	Hypothetical protein	PIRATE
NEIS3218	94,4%	94,4%	(Putative) phage associated protein	(Putative) phage associated protein	PIRATE
NEIS3219	1,4%	2,1%	Hypothetical protein	Hypothetical protein	PIRATE

TABLE S2. LIST OF ACCESSORY NEIS LOCI WITH ITS PREVALENCE AND FUNCTIONAL ANNOTATION (CONTINUED).

Phage associated	Hypothetical proteins	Mobile Genetic Elements	Toxin-antitoxin system	Methyltransferases	Putative secretion system	TonB dependent receptor	DNA transcription	Other functions

NEIS locus	Duplicate NEIS loci	Prevalence in subset of 765 isolates	Prevalence in all 8013 isolates	(PubMLST) gene annotation	PubMLST functional scheme (NG schemes only)	Final annotation	Identified by
NEIS3220		94,0%	89,7%	Helix-turn-helix transcriptional regulator		DNA transcription	PIRATE
NEIS3221		1,6%	2,2%	(Putative) phage associated protein		(Putative) phage associated protein	PIRATE
NEIS3222		0,7%	1,5%	Hypothetical protein		Phage island X	PIRATE
NEIS3223		24,1%	20,3%	Hypothetical protein		Hypothetical protein	PIRATE
NEIS3224		0,7%	1,6%	Hypothetical protein		Phage island X	PIRATE
NEIS3225		1,2%	0,7%	ssDNA binding protein		VirB T4SS	PIRATE
NEIS3226		0,8%	0,8%	Membrane protein		Membrane protein	PIRATE
NEIS3227		9,4%	7,3%	Helix-turn-helix transcriptional regulator		Conjugative Plasmid	PIRATE
NEIS3228		86,8%	94,0%	Hypothetical protein		Hypothetical protein	PIRATE
NEIS3230		1,2%	0,7%	Hypothetical protein		VirB T4SS	PIRATE
NEIS3231		1,2%	0,7%	Type II toxin-antitoxin system - Plasmid stabilization system protein - parE toxin		VirB T4SS	PIRATE
NEIS3232		27,7%	32,9%	Type II toxin-antitoxin system - vapC family protein		Toxin-antitoxin system	PIRATE
NEIS3233		94,6%	94,5%	Hypothetical protein		Hypothetical protein	PIRATE
NEIS3234		59,1%	93,8%	Hypothetical protein		Hypothetical protein	PIRATE

TABLE S3. SIGNIFICANCE OF ASSOCIATIONS BETWEEN CGMLST AND ACCESSORY ELEMENTS.

		Gonococcal Genetic Island			Conjugative plasmid			Beta-lactamase plasmid			VirB T4SS			Nf1 phage island			Putative secretion system		
CgMLST group	**Number of isolates**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**
3	**1452**	99%	1.89E-236	****	0%	1.89E-212	****	1%	9.56E-54	****	0%	0,0103	*	0%	7.33E-24	****	97%	0	****
16	**1433**	20%	0	****	0%	1.43E-216	****	0%	1.54E-74	****	0%	0,0105	*	0%	2.14E-23	****	92%	0	****
18	**357**	90%	6.93E-20	****	1%	9.91E-40	****	2%	1.42E-09	****	0%	1		15%	1.93E-18	****	0%	3.79E-104	****
196	**311**	100%	2.04E-47	****	4%	1.54E-24	****	0%	1.01E-12	****	0%	1		0%	0,0039	**	0%	5.57E-90	****
21	**289**	41%	1.12E-20	****	93%	8.68E-130	****	77%	3.11E-158	****	0%	1		0%	0,0087	**	0%	1.87E-83	****
8	**267**	52%	5.57E-07	****	0%	1.74E-33	****	0%	7.66E-12	****	0%	1		0%	0,0197	*	95%	2.79E-62	****
133	**239**	59%	0,4340		100%	4.49E-138	****	31%	5.35E-15	****	0%	1		0%	0,0693		93%	8.03E-49	****
41	**234**	99%	8.89E-33	****	100%	3.29E-132	****	48%	1.28E-44	****	0%	1		0%	0,0665		0%	1.71E-64	****
223	**231**	99%	2.47E-32	****	1%	6.78E-27	****	0%	1.73E-08	****	0%	1		0%	0,1080		0%	7.73E-66	****
24	**204**	100%	1.85E-31	****	2%	3.71E-19	****	0%	1.07E-08	****	0%	1		0%	0,2260		0%	6.1E-58	****
414	**190**	100%	3.49E-29	****	7%	2.79E-10	****	1%	1.61E-06	****	0%	1		0%	0		0%	9.19E-54	****
17	**166**	81%	0,1680		1%	6.35E-18	****	1%	0,0003	***	0%	1		0%	1		95%	1.05E-36	****
219	**159**	9%	9.39E-56	****	1%	1.64E-18	****	4%	1		0%	1		0%	1		0%	1.79E-44	****
446	**159**	100%	3.54E-24	****	100%	2.71E-90	****	1%	0	***	0%	1		0%	1		1%	2.32E-42	****
31	**119**	54%	0,1550		12%	0,0162	*	1%	5,65E-03	**	0%	1		0%	1		74%	3.71E-06	****
62	**115**	12%	7.73E-35	****	100%	1.87E-64	****	47%	4.59E-20	****	0%	1		0%	1		1%	1.32E-29	****

*Adjusted p-values were calculated by multiplying the original p-value by the number of core genome groups tested (n=250). *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.*

significantly less prevalent in cgMLST group compared to distribution among all isolates

significantly more prevalent in cgMLST group compared to distribution among all isolates

TABLE S3. SIGNIFICANCE OF ASSOCIATIONS BETWEEN CGMLST AND ACCESSORY ELEMENTS. (CONTINUED)

		Gonococcal Genetic Island			Conjugative plasmid			Beta-lactamase plasmid			VirB T4SS			Nf1 phage island			Putative secretion system		
CgMLST group	**Number of isolates**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**	**Pre-valence**	**adj. p-value**	**sign.**
516	**104**	100%	4.37E-15	****	0%	1.35E-12	****	34%	1.44E-07	****	0%	1		0%	1		0%	2.1E-28	****
243	**90**	100%	7.91E-13	****	76%	2.35E-19	****	2%	1		0%	1		0%	1		0%	2.89E-24	****
425	**74**	100%	4.14E-10	****	49%	0,0251	*	3%	1		0%	1		1%	1		1%	1.01E-17	****
415	**73**	100%	7.61E-10	****	1%	1.22E-06	****	0%	0		0%	1		100%	1.02E-107	****	0%	3.36E-19	****
450	**70**	97%	1.14E-06	****	100%	2.43E-38	****	1%	1		1%	1		0%	1		0%	2.69E-18	****
416	**63**	8%	6.6E-22	****	100%	2.64E-34	****	33%	0,0005	***	0%	1		0%	1		0%	1.68E-16	****
37	**52**	100%	1.67E-06	****	88%	5.1E-18	****	79%	1.27E-27	****	0%	1		0%	1		2%	1.33E-11	****
44	**50**	100%	2.69E-06	****	14%	1		0%	1		0%	1		0%	1		0%	1.01E-12	****
95	**47**	32%	5.22E-05	****	66%	7.2E-06	****	4%	1		0%	1		0%	1		98%	2.38E-11	****
34	**46**	98%	0,0003	***	0%	2,01E-04	***	0%	1		0%	1		0%	1		89%	1.7E-06	****
498	**42**	100%	0,0001	****	100%	3.01E-22	****	17%	1		0%	1		0%	1		0%	2.13E-10	****
439	**39**	5%	1.24E-14	****	90%	7.48E-14	****	0%	1		0%	1		0%	1		0%	1.59E-09	****
304	**38**	97%	0,0059	**	100%	5.87E-20	****	3%	1		5%	1		89%	2.3E-43	****	0%	3.11E-09	****
239	**35**	9%	2.79E-11	****	100%	3.04E-18	****	83%	2.66E-20	****	0%	1		0%	1		91%	3.06E-05	****
Other	**675**																		
No cgMLST assigned	**601**																		
Overall prevalence	**8013**	**69%**			**27%**			**11%**			**0,7%**			**3,6%**			**49%**		

*Adjusted p-values were calculated by multiplying the original p-value by the number of core genome groups tested (n=250). *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.*

significantly less prevalent in cgMLST group compared to distribution among all isolates

significantly more prevalent in cgMLST group compared to distribution among all isolates

TABLE S4. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY ELEMENTS AND ANTIMICROBIAL RESISTANCE AGAINST PENICILLIN, TETRACYCLINE, CIPROFLOXACIN, CEFTRIAXONE, CEFIXIME AND AZITHROMYCIN.

			Gonococcal genetic island			Conjugative plasmid			Beta-lactamase plasmid			virB T4SS			Nf1 phage island			Putative secretion system		
	MIC	No isolates	prevalence	adj. p-value	sign.	prevalence	adj. p-value	sign.	prevalence	adj. p-value	sign.	prevalence	adj. p-value	sign.	prevalence	adj. p-value	sign.	prevalence	adj. p-value	sign.
PEN S	**≤0.06**	**117**	88.89%	6,66E-13	****	31.63%	4,69E-07	****	0.02%	5,87E-132	****	6.84%	2,00E-06	****	6.83%	0,4890		16.24%	8,28E-14	****
PEN I	**>0.06-1.0**	**1896**	63.98%			23.42%			1.85%			1.21%			3.38%			51.85%		
PEN R	**>1.0**	**1074**	74.02%			32.96%			32.68%			0.19%			2.79%			53.26%		
TET S	**≤0.5**	**774**	74.29%	0,0002	***	23.00%	0,0474	*	8.14%	5,26E-05	****	2.45%	0,0003	***	3.62%	1		36.56%	1,12E-21	****
TET R	**>0.5**	**2143**	66.17%			27.95%			14.19%			0.56%			3.13%			56.98%		
CIP S	**≤0.03**	**2098**	65.3%	0,0034	**	26.02%	0,2568		5.82%	8,28E-26	****	1.76%	5,66E-12	****	3.43%	0,0030	**	41.56%	1,07E-21	****
CIP R	**>0.06**	**2316**	70.16%			28.76%			15.67%			0.00%			1.77%			56.26%		
CRO S	**≤0.125**	**4868**	67.46%	0,1224		25.84%	1		10.56%	0,9600		0.64%	1		2.28%	1		50.92%	0,1032	
CRO R	**>0.125**	**22**	90.9%			13.64%			0.00%			0.00%			0.00%			77.27%		
CFX S	**≤0.125**	**4004**	65.86%	2,08E-19	****	29.00%	5,50E-52	****	12.01%	7,86E-21	****	0.85%	0,2760		2.7%	3,37E-05	****	46.55%	6,84E-98	****
CFX R	**>0.125**	**490**	85.31%			2.04%			0.61%			0.00%			0.00%			93.06%		
AZI S	**<1.0**	**3547**	70.99%	1,78E-09	****	29.01%	2,89E-47	****	12.52%	2,19E-23	****	0.93%	0,1950		2.68%	0,2292		44.8%	7,08E-61	****
AZI R	**≥1.0**	**898**	59.91%			7.57%			2.34%			0.22%			1.45%			75.17%		

*Adjusted p-values were calculated by multiplying the original p-value by the number of accessory elements tested (n=6). PEN=penicillin; TET=tetracycline; CIP=ciprofloxacin; CRO=ceftriaxone; CFX=cefixime; AZI=azithromycin; S=susceptible; I= intermediate; R=resistant. *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR.

Locus	NEIS0364			NEIS0365			NEIS0486			NEIS0524			NEIS0594			NEIS0840		
Functional annotation	Hypothetical protein			Hypothetical protein			Alcohol dehydrogenase			Putative peptidase			Hypothetical protein			(Putative) phage associated		
	prevalence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	prevalencwe	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	31.62%	3.42E-45	****	30.77%	3.74E-47	****	47.01%	8.42E-09	****	47.86%	1.76E-08	****	8.55%	6.83E-72	****	0.85%	0.9594	
PEN I	79.27%			79.01%			21.52%			77.11%			43.25%			1.64%		
PEN R	90.50%			90.69%			28.40%			74.49%			13.59%			0.47%		
TET S	59.82%	2.15E-69	****	59.95%	1.72E-68	****	39.15%	6.20E-23	****	68.09%	1.26E-07	****	40.31%	2.56E-06	****	1.55%	1	
TET R	90.15%			90.11%			19.88%			79.14%			29.35%			0.98%		
CIP S	63.11%	1.18E-173	****	62.77%	5.23E-165	****	29.74%	2.64E-11	****	73.69%	1		44.57%	2.57E-49	****	2.19%	1.44E-08	****
CIP R	95.60%			94.91%			20.25%			75.60%			23.23%			0.22%		
CRO S	81.72%	1		81.12%	1		22.62%	1		74.53%	1		35.09%	1		1.13%	1	
CRO R	95.45%			90.91%			22.73%			86.36%			22.73%			0.00%		
CFX S	79.57%	2.43E-28	****	79.17%	4.63E-21	****	26.30%	7.13E-41	****	76.72%	4.87E-16	****	34.37%	0.0037	**	1.17%	1	
CFX R	97.55%			95.51%			2.65%			92.24%			25.31%			0.20%		
AZI S	79.14%	1.17E-37	****	78.74%	3.99E-39	****	26.87%	7.76E-44	****	74.15%	4.06E-16	****	30.65%	9.28E-19	****	1.04%	1	
AZI R	95.77%			95.77%			6.68%			87.19%			47.44%			1.00%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS0953			NEIS0955			NEIS1089			NEIS1193			NEIS1194			NEIS1311		
Functional annotation	(Putative) phage associate			Nf4 phage island			Hypothetical protein			Methyltransferase			Methyltransferase			Methyltransferase		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	15.38%	1.05E-05	****	82.91%	0.00081	***	22.22%	0.44538	**	83.76%	2.85E-06	****	83.76%	5.02E-07	****	54.70%	8.74E-52	****
PEN I	19.67%			94.30%			22.05%			90.77%			88.92%			79.80%		
PEN R	11.82%			95.44%			27.28%			95.53%			94.79%			96.28%		
TET S	21.71%	1.05E-03	**	90.70%	7.75E-05	****	25.06%	1		91.73%	1		90.96%	1		56.20%	2.00E-132	****
TET R	14.75%			95.71%			23.15%			93.14%			92.49%			95.47%		
CIP S	21.78%	3.85E-06	****	94.47%	4.05E-05	****	20.07%	1.93E-08	****	86.56%	4.93E-58	****	85.22%	1.88E-38	****	68.16%	2.65E-175	****
CIP R	15.37%			90.46%			28.24%			98.58%			96.42%			97.80%		
CRO S	17.50%	1		91.02%	1		23.09%	1		93.26%	1		91.31%	1		84.57%	1	
CRO R	4.55%			95.45%			4.55%			100.00%			100.00%			100.00%		
CFX S	16.38%	3.51E-06	****	93.46%	0.19578		26.40%	2.86E-52	****	93.03%	2.64E-10	****	91.41%	1.26E-10	****	82.94%	1.40E-29	****
CFX R	7.55%			89.59%			0.82%			99.59%			98.98%			99.18%		
AZI S	17.23%	1		90.30%	2.91E-16	****	27.29%	2.64E-43	****	92.59%	7.37E-03	**	90.61%	0.024492	*	83.28%	7.88E-22	****
AZI R	13.81%			98.11%			7.02%			96.10%			94.32%			95.21%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS1667			NEIS2486			NEIS2582			NEIS2615			NEIS2616			NEIS2618		
Functional annotation	Nf4 phage island			Hypothetical protein			Hypothetical protein			(Putative) phage associate			(Putative) phage associate			(Putative) phage associate		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	26.50%	2.11E-05	****	94.02%	0.01521	***	67.52%	1.25E-44	****	41.03%	4.06E-07	****	78.63%	0.0003	***	72.65%	2.74E-09	****
PEN I	10.13%			92.67%			19.51%			66.35%			81.65%			61.60%		
PEN R	8.29%			88.27%			9.03%			70.11%			88.18%			50.09%		
TET S	15.37%	3.68E-06	****	93.15%	1		38.50%	1.34E-64	****	53.75%	8.27E-16	****	72.09%	6.97E-21	****	61.11%	1	
TET R	8.21%			91.46%			9.61%			70.98%			87.96%			57.16%		
CIP S	6.72%	9.59E-29	****	94.76%	8.11E-14	****	34.60%	2.91E-173	****	65.68%	0.12402		85.41%	5.71E-12	****	63.78%	3.14E-08	****
CIP R	18.01%			88.00%			3.41%			61.05%			76.55%			54.49%		
CRO S	14.56%	1		91.78%	1		16.52%	1		63.21%	1		79.66%	1		59.14%	1	
CRO R	9.09%			90.91%			4.55%			63.64%			81.82%			40.91%		
CFX S	12.69%	1		91.36%	1.02E-09	****	18.91%	6.87E-34	****	65.48%	0.8268		81.54%	0.47		63.76%	1	
CFX R	16.53%			98.78%			0.82%			59.59%			86.53%			62.04%		
AZI S	15.96%	5.14E-29	****	90.36%	2.96E-07	****	19.42%	1.01E-35	****	57.88%	1.64E-33	****	76.06%	3.71E-44	****	58.11%	1	
AZI R	3.01%			96.10%			3.79%			79.62%			95.10%			59.91%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS2619			NEIS2621			NEIS2640			NEIS2643			NEIS2646			NEIS2647		
Functional annotation	(Putative) phage associate			(Putative) phage associate			(Putative) phage associate			(Putative) phage associate			TonB dependent receptor			TonB dependent receptor		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	64.10%	8.27E-09	****	13.68%	1.63E-10	****	32.48%	0.029	*	75.21%	0.11388		7.69%	0.00874	**	8.55%	7.02E-04	***
PEN I	61.87%			3.06%			23.26%			67.30%			1.95%			2.27%		
PEN R	49.35%			0.65%			18.72%			73.18%			3.91%			4.93%		
TET S	63.82%	9.75E-04	****	4.91%	5.62E-04	****	31.52%	2.07E-10	****	65.37%	0.07605		5.81%	1.19E-05	****	6.20%	1.73E-04	***
TET R	54.74%			1.73%			18.95%			71.77%			1.87%			2.38%		
CIP S	61.30%	0.01326	*	4.91%	2.54E-10	****	22.64%	1		67.21%	1.39E-06	****	4.77%	7.62E-08	****	5.43%	3.09E-08	****
CIP R	55.70%			1.34%			22.71%			74.91%			1.60%			1.94%		
CRO S	58.01%	1		2.73%	1		22.27%	1		71.12%	0.0975		2.92%	1		3.20%	1	
CRO R	36.36%			9.09%			13.64%			100.00%			4.55%			13.64%		
CFX S	62.51%	1.47E-32	****	3.05%	0.16		22.15%	0.58		74.10%	0.19		3.25%	0.09		3.62%	0.24882	
CFX R	33.47%			0.82%			16.94%			80.41%			0.82%			1.22%		
AZI S	57.68%	4.27E-05	****	3.02%	1		26.11%	1.24E-17	****	70.45%	0.43602		3.47%	3.50E-08	****	3.95%	5.79E-10	****
AZI R	48.33%			2.12%			12.47%			75.17%			0.22%			0.22%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS2666			NEIS2691			NEIS2692			NEIS2693			NEIS2705			NEIS2713		
Functional annotation	Hypothetical protein			Methyltransferase			Methyltransferase			Nf4 phage island			Hypothetical protein			(Putative) phage associate		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	56.41%	8.89E-19	****	43.59%	1.90E-45	****	43.59%	1.16E-44	****	88.89%	2.64E-19	****	70.09%	9.91E-59	****	86.32%	2.00E-79	****
PEN I	84.92%			91.61%			91.56%			89.19%			60.34%			25.63%		
PEN R	91.34%			96.28%			96.09%			75.61%			30.17%			8.66%		
TET S	79.33%	8.35E-11	****	81.78%	9.98E-26	****	81.78%	2.60E-25	****	94.06%	9.75E-20	****	58.01%	1.79E-06	****	50.52%	1.62E-106	****
TET R	89.78%			95.29%			95.19%			80.45%			46.24%			10.64%		
CIP S	75.93%	3.77E-97	****	80.89%	9.98E-104	****	80.79%	1.93E-102	****	91.37%	5.47E-42	****	72.40%	6.60E-202	****	42.14%	4.14E-207	****
CIP R	96.63%			98.92%			98.83%			76.08%			27.29%			4.92%		
CRO S	88.17%	1		91.04%	1		90.96%	1		84.33%	1		46.41%	0.025974	*	20.83%	1	
CRO R	95.45%			100.00%			100.00%			77.27%			9.09%			4.55%		
CFX S	87.04%	2.93E-26	****	89.91%	2.02E-14	****	89.84%	1.27E-14	****	86.11%	7.80E-24	****	50.20%	9.67E-64	****	22.85%	3.99E-38	****
CFX R	100.00%			99.18%			99.18%			65.92%			11.84%			1.63%		
AZI S	85.93%	6.19E-16	****	89.26%	5.12E-29	****	89.23%	4.59E-28	****	83.03%	1		44.88%	8.58E-05	****	24.08%	1.30E-53	****
AZI R	95.55%			99.33%			99.22%			81.29%			54.01%			3.56%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS2714			NEIS2718			NEIS2744			NEIS2859			NEIS3177			NEIS3178		
Functional annotation	(Putative) phage associate			(Putative) phage associate			TonB dependent receptor			Hypothetical protein			Phage island X			Hypothetical protein		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	prevalence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	64.96%	4.08E-04	***	71.79%	1.15E-05	****	88.89%	1		23.08%	1		6.84%	5.82E-07	****	21.37%	1.29E-21	****
PEN I	56.86%			59.28%			94.78%			16.03%			0.53%			43.25%		
PEN R	65.83%			68.90%			93.11%			14.71%			0.00%			25.79%		
TET S	61.24%	1		55.94%	3.27E-05	****	90.57%	3.66E-04	***	14.73%	1		1.55%	1.77E-03	**	30.88%	5.82E-03	**
TET R	59.26%			66.26%			95.33%			16.43%			0.14%			38.92%		
CIP S	58.77%	1		60.92%	4.24E-11	****	92.61%	1.61E-05	****	19.11%	1		1.43%	1.43E-08	****	40.42%	2.70E-07	****
CIP R	59.33%			71.20%			96.20%			19.17%			0.00%			31.87%		
CRO S	58.34%	1		65.69%	0.9282		94.99%	1		18.69%	1		0.66%	1		37.96%	1	
CRO R	68.18%			90.91%			86.36%			9.09%			0.00%			22.73%		
CFX S	56.64%	1.71E-12	****	68.91%	1		94.23%	0.2262		24.08%	9.75E-12	****	0.72%	1		40.83%	3.01E-03	**
CFX R	74.29%			73.27%			97.35%			10.20%			0.00%			31.22%		
AZI S	59.37%	0.21528		66.20%	1		94.02%	0.0194	*	20.98%	3.89E-06	****	0.76%	0.25194		32.08%	4.00E-34	****
AZI R	53.79%			70.27%			96.99%			13.14%			0.00%			55.01%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS3179			NEIS3180			NEIS3181			NEIS3182			NEIS3183			NEIS3184		
Functional annotation	**Hypothetical protein**			**Hypothetical protein**			**(Putative) phage associate**			**Methyltransferase**			**Hypothetical protein**			**DNA transcription**		
	pre-valence	**adj. p-value**	**sign.**	**pre-valence**	**adj. p-value**	**sign.**	**pre-valence**	**adj. p-value**	**sign.**	**pre-valence**	**adj. p-value**	**sign.**	**pre-valence**	**adj. p-value**	**sign.**	**pre-valence**	**adj. p-value**	**sign.**
PEN S	98.29%	2.29E-03	****	92.31%	6.04E-81	****	86.32%	0.01092	*	73.50%	3.12E-04	***	93.16%	1		37.61%	1.65E-16	****
PEN I	93.57%			68.62%			83.23%			57.44%			96.47%			67.88%		
PEN R	89.76%			95.81%			88.83%			51.68%			96.65%			76.82%		
TET S	94.32%	1		77.65%	1		74.03%	1.61E-19	****	53.10%	1		96.51%	1		57.11%	3.13E-21	****
TET R	92.11%			78.81%			88.94%			56.00%			96.78%			76.34%		
CIP S	93.23%	1		63.54%	4.30E-171	****	86.08%	0.00858	**	65.54%	7.39E-33	****	93.85%	1		61.49%	6.48E-20	****
CIP R	91.54%			95.64%			89.90%			47.19%			95.21%			74.91%		
CRO S	91.68%	1		80.85%	1		88.68%	1		58.05%	1		93.22%	1		70.11%	1	
CRO R	100.00%			90.91%			90.91%			81.82%			100.00%			68.18%		
CFX S	95.05%	1		78.30%	5.39E-39	****	86.84%	1.77E-06	****	58.92%	1.14E-07	****	96.08%	1		68.73%	3.63E-14	****
CFX R	92.65%			98.98%			94.90%			72.86%			94.29%			85.51%		
AZI S	90.98%	1		84.47%	4.42E-54	****	86.44%	5.58E-05	****	50.21%	1.76E-50	****	92.08%	1.32E-10	****	68.14%	5.35E-15	****
AZI R	92.65%			59.13%			92.32%			77.73%			98.00%			81.85%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS3186			NEIS3188			NEIS3189			NEIS3193			NEIS3195			NEIS3196		
Functional annotation	Phage island X			Toxin-antitoxin system			TspB virulence factor			Hypothetical protein			Phage island X			Hypothetical protein		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	5.98%	3.88E-06	****	18.80%	9.20E-119	****	98.29%	0.012168	*	86.32%	2.92E-43	****	6.84%	5.82E-07	****	86.32%	6.06E-06	****
PEN I	0.42%			24.74%			96.94%			89.40%			0.53%			89.87%		
PEN R	0.00%			67.50%			93.85%			67.97%			0.00%			95.25%		
TET S	1.29%	4.95E-03	**	16.93%	3.88E-54	****	98.84%	1.47E-05	****	86.56%	4.30E-03	**	1.55%	1.77E-03	**	88.89%	0.019734	*
TET R	0.09%			48.02%			94.91%			80.12%			0.14%			93.19%		
CIP S	1.29%	1.36E-07	****	9.72%	1.22E-285	****	96.04%	3.36E-05	****	92.04%	6.31E-75	****	1.43%	1.43E-08	****	87.89%	3.28E-18	****
CIP R	0.00%			59.93%			92.49%			70.68%			0.00%			95.42%		
CRO S	0.62%	1		35.62%	0.011076	*	94.04%	1		82.21%	1		0.66%	1		92.67%	1	
CRO R	0.00%			77.27%			95.45%			95.45%			0.00%			86.36%		
CFX S	0.65%	1		32.69%	1.44E-59	****	96.08%	7.43E-09	****	80.67%	1.54E-33	****	0.72%	1		92.66%	1	
CFX R	0.00%			71.63%			88.57%			98.98%			0.00%			91.22%		
AZI S	0.68%	0.65442		36.20%	4.20E-04	****	93.71%	1		78.29%	7.53E-36	****	0.76%	0.25194		91.65%	2.64E-08	****
AZI R	0.00%			44.54%			94.21%			94.99%			0.00%			97.22%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS3197			NEIS3198			NEIS3200			NEIS3202			NEIS3203			NEIS3204		
Functional annotation	Hypothetical protein			Hypothetical protein			Hypothetical protein			DNA replication			(Putative) phage associate			Phage island X		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	85.47%	1.45E-18	****	81.20%	1		55.56%	9.59E-07	****	3.42%	4.56E-03	**	2.56%	2.66E-03	**	7.69%	4.36E-08	****
PEN I	96.99%			86.39%			29.85%			5.49%			6.43%			0.53%		
PEN R	88.73%			88.45%			27.47%			2.23%			2.89%			0.00%		
TET S	97.03%	2.04E-04	***	88.37%	1		45.48%	4.21E-27	****	9.30%	3.61E-10	****	10.34%	1.44E-09	****	1.68%	5.69E-04	***
TET R	92.44%			88.61%			23.57%			2.80%			3.55%			0.14%		
CIP S	97.62%	9.59E-21	****	75.50%	8.66E-13	****	37.23%	5.80E-20	****	3.72%	0.60762		3.96%	1		1.48%	6.72E-09	****
CIP R	90.89%			84.76%			23.92%			2.33%			3.15%			0.00%		
CRO S	93.98%	1		78.37%	1		31.37%	1		2.71%	1		3.20%	1		0.68%	1	
CRO R	100.00%			81.82%			22.73%			0.00%			0.00%			0.00%		
CFX S	92.48%	3.03E-14	****	83.02%	1		32.17%	3.93E-10	****	3.27%	4.70E-04	***	3.87%	4.24E-05	****	0.75%	1	
CFX R	100.00%			82.45%			17.55%			0.20%			0.20%			0.00%		
AZI S	93.40%	1.16E-09	****	80.32%	2.11E-03	**	32.03%	5.59E-13	****	3.27%	1	*	3.95%	0.08346		0.79%	0.25818	
AZI R	98.55%			86.30%			19.15%			1.78%			1.78%			0.00%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS3205			NEIS3206			NEIS3207			NEIS3208			NEIS3209			NEIS3210		
Functional annotation	(Putative) phage associate			Hypothetical protein			Phage island X			Hypothetical protein			Hypothetical protein			Hypothetical protein		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	3.42%	1.37E-05	****	7.69%	1.02E-06	****	8.55%	2.54E-09	****	64.96%	2.94E-03	**	18.80%	6.66E-18	****	57.26%	3.07E-13	****
PEN I	7.01%			31.28%			0.53%			68.88%			62.39%			28.27%		
PEN R	2.51%			28.68%			0.00%			60.71%			60.24%			21.51%		
TET S	11.37%	3.43E-12	****	36.82%	3.81E-04	***	1.81%	1.80E-04	***	66.28%	1		43.67%	1.21E-25	****	40.44%	1.28E-19	****
TET R	3.55%			27.90%			0.14%			66.17%			66.17%			22.21%		
CIP S	4.67%	0.14118		26.26%	4.34E-05	****	1.53%	3.17E-09	****	70.97%	0.09126		54.81%	1.83E-12	****	33.51%	8.74E-25	****
CIP R	2.85%			33.16%			0.00%			66.41%			66.06%			19.34%		
CRO S	3.37%	1		29.05%	0.053664		0.70%	1		67.73%	0.8502		60.17%	1		23.60%	1	
CRO R	0.00%			0.00%			0.00%			40.91%			36.36%			9.09%		
CFX S	4.07%	1.12E-05	****	30.14%	1.20E-52	****	0.77%	1		72.85%	1.88E-94	****	63.66%	1.88E-61	****	26.17%	2.11E-38	****
CFX R	0.20%			2.24%			0.00%			24.49%			24.08%			3.06%		
AZI S	4.14%	0.076752		34.76%	2.09E-44	****	0.82%	0.15834		66.48%	4.60E-04	***	58.61%	5.94E-16	****	26.73%	1.29E-08	****
AZI R	1.89%			11.80%			0.00%			74.28%			73.94%			16.70%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS3211			NEIS3213			NEIS3214			NEIS3215			NEIS3216			NEIS3217		
Functional annotation	DNA transcription			Hypothetical protein			(Putative) phage associated			Hypothetical protein			Hypothetical protein			Hypothetical protein		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	13.68%	2.25E-31	****	4.27%	6.85E-05	****	86.32%	5.06E-70	****	84.62%	1.79E-08	****	18.80%	1.63E-20	****	8.55%	2.53E-05	****
PEN I	41.19%			2.06%			21.04%			86.76%			8.54%			1.00%		
PEN R	21.32%			0.19%			9.50%			94.04%			21.51%			0.56%		
TET S	25.32%	1.97E-06	****	4.26%	4.17E-09	****	39.66%	6.97E-61	****	86.30%	6.68E-04	***	8.91%	3.35E-04	***	1.68%	1	
TET R	36.21%			0.56%			10.92%			91.97%			15.40%			0.89%		
CIP S	39.85%	1.73E-24	****	3.57%	2.29E-23	****	36.18%	8.50E-154	****	84.13%	4.98E-21	****	5.20%	2.11E-51	****	2.24%	1.41E-05	****
CIP R	24.91%			0.00%			5.31%			93.39%			20.29%			0.47%		
CRO S	30.98%	1		1.40%	1		18.32%	1		89.93%	1		13.37%	1		1.15%	1	
CRO R	13.64%			0.00%			0.00%			81.82%			18.18%			0.00%		
CFX S	33.02%	5.23E-24	****	1.62%	0.0702		20.60%	6.65E-38	****	90.83%	1		15.11%	1.58E-04	***	1.35%	0.2652	
CFX R	11.43%			0.00%			0.82%			89.80%			7.55%			0.00%		
AZI S	25.88%	3.34E-56	****	1.86%	0.31902		20.92%	3.19E-45	****	88.86%	9.36E-11	****	14.77%	1.69E-03	**	1.30%	4.01E-03	**
AZI R	54.68%			0.56%			3.12%			96.10%			9.47%			0.00%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS3218			NEIS3219			NEIS3220			NEIS3221			NEIS3222			NEIS3223		
Functional annotation	(Putative) phage associate			Hypothetical protein			DNA transcription			(Putative) phage associate			Phage island X			Hypothetical protein		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	85.47%	4.79E-18	****	2.56%	1.95E-04	***	87.18%	1		3.42%	4.13E-05	****	5.98%	7.23E-06	****	86.32%	1.29E-69	****
PEN I	96.99%			6.96%			93.09%			7.49%			0.53%			21.04%		
PEN R	88.92%			2.89%			91.62%			2.98%			0.00%			9.59%		
TET S	97.03%	2.82E-04	***	11.11%	1.01E-10	****	87.98%	2.03E-04	***	12.14%	1.17E-12	****	1.42%	5.42E-03	**	39.66%	1.14E-60	****
TET R	92.53%			3.73%			93.56%			3.92%			0.14%			10.97%		
CIP S	97.62%	4.11E-20	****	4.43%	1		93.57%	1		4.91%	0.44382		1.38%	3.02E-08	****	36.18%	3.92E-153	****
CIP R	91.02%			3.15%			93.18%			3.24%			0.00%			5.35%		
CRO S	93.88%	1		3.41%	1		94.39%	1		3.66%	1		0.64%	1		18.34%	1	
CRO R	100.00%			0.00%			100.00%			0.00%			0.00%			0.00%		
CFX S	92.46%	1.97E-11	****	4.12%	1.15E-05	****	93.28%	4.40E-07	****	4.42%	3.44E-06	****	0.70%	1		20.63%	6.40E-38	****
CFX R	99.59%			0.20%			98.98%			0.20%			0.00%			0.82%		
AZI S	93.35%	2.02E-10	****	4.17%	0.11934		92.44%	3.14E-05	****	4.45%	0.1248		0.73%	0.41106		20.92%	1.81E-44	****
AZI R	98.66%			2.00%			96.88%			2.23%			0.00%			3.23%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

TABLE S5. SIGNIFICANCE OF ASSOCIATIONS BETWEEN ACCESSORY GENES AND AMR. (CONTINUED)

Locus	NEIS3224			NEIS3226			NEIS3228			NEIS3232			NEIS3233			NEIS3234		
Functional annotation	Phage island X			Membrane protein			Hypothetical protein			Toxin-antitoxin system			Hypothetical protein			Hypothetical protein		
	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.	pre-valence	adj. p-value	sign.
PEN S	8.55%	2.85E-09	****	0.85%	1		95.73%	0.1794		18.80%	1.78E-118	****	85.47%	4.79E-18	****	98.29%	0.9048	
PEN I	0.58%			0.74%			95.62%			24.79%			96.99%			99.89%		
PEN R	0.00%			0.37%			97.95%			67.50%			88.92%			99.72%		
TET S	1.81%	1.80E-04	***	1.42%	0.037674	*	96.51%	0.8268		16.93%	3.44E-54	****	97.03%	2.80E-04	***	99.74%	1	
TET R	0.14%			0.23%			98.18%			48.06%			92.58%			99.77%		
CIP S	1.57%	1.50E-09	****	1.33%	6.40E-08	****	92.56%	1		9.77%	6.22E-271	****	97.66%	1.02E-20	****	93.66%	1	
CIP R	0.00%			0.00%			92.79%			59.93%			90.98%			94.08%		
CRO S	0.72%	1		0.60%	1		92.15%	1		35.62%	0.011076	*	94.04%	1		93.28%	1	
CRO R	0.00%			0.00%			86.36%			77.27%			100.00%			100.00%		
CFX S	0.77%	1		0.67%	1		94.63%	1		32.72%	1.61E-59	****	92.58%	4.76E-14	****	99.83%	1	
CFX R	0.00%			0.00%			92.04%			71.63%			100.00%			99.80%		
AZI S	0.82%	0.15834		0.65%	0.63492		92.73%	2.29E-07	****	36.23%	4.24E-04	***	93.46%	4.57E-10	****	91.20%	3.26E-20	****
AZI R	0.00%			0.00%			97.66%			44.54%			98.66%			99.11%		

Adjusted p-values were calculated by multiplying the original p-value by the number of genes tested (n=78)

**p<0.05; **p<0.01; ***p<0.001; ****p<0.0001*

significantly more prevalent among susceptible isolates

significantly more prevalent among resistant isolates

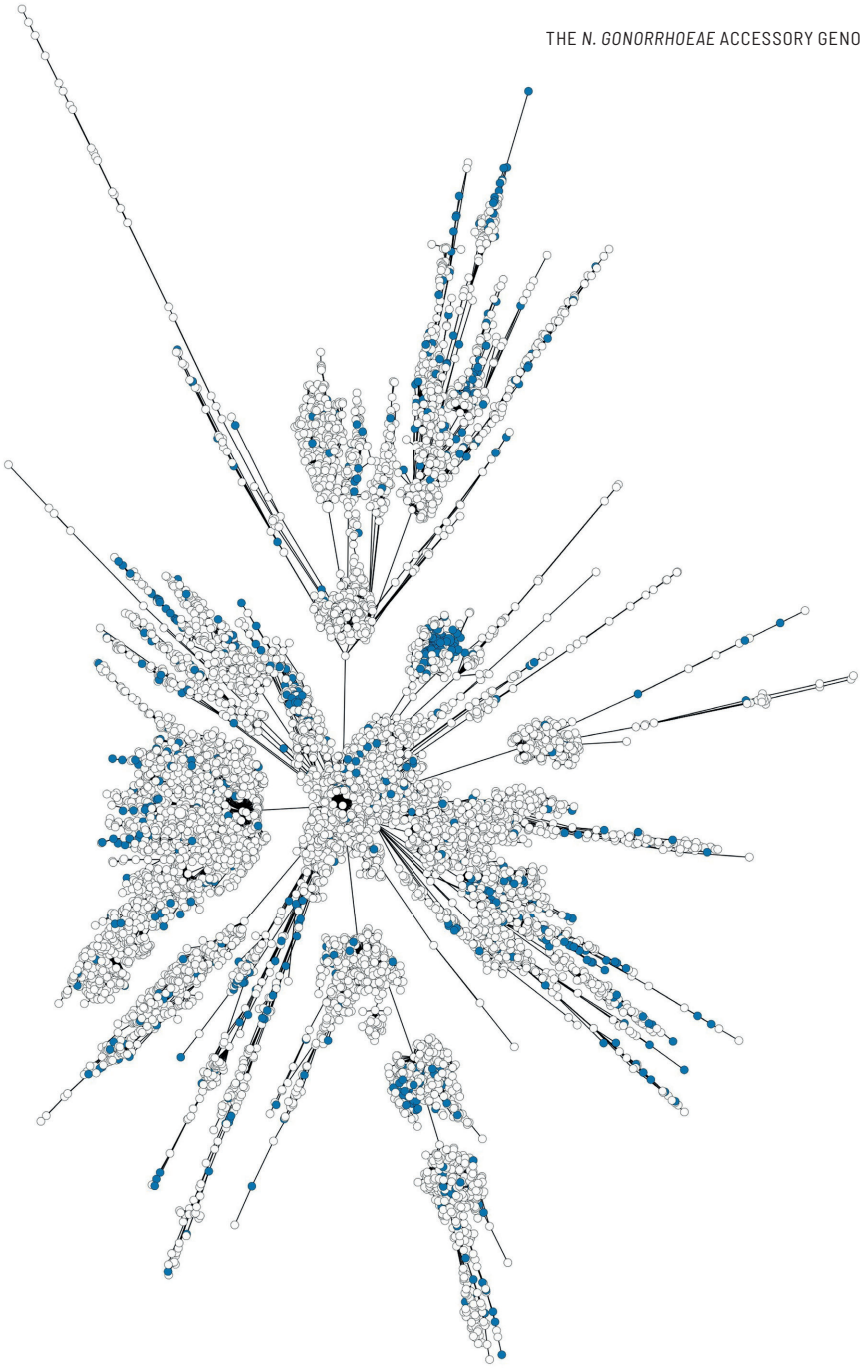


FIGURE S1. MINIMUM-SPANNING TREE BASED ON THE *N. GONORRHOEAE* CGMLST V1.0 TYPING SCHEME INCLUDING ALL 8,013 *N. GONORRHOEAE* ISOLATES AVAILABLE IN PUBMLST. THE 765 ISOLATES INCLUDED IN THE REPRESENTATIVE SUBSET OF ISOLATES USED FOR ACCESSORY GENOME CHARACTERISATION ARE COLOURED IN BLUE. THE BLUE NODES ARE DISTRIBUTED THROUGHOUT THE TREE CONSISTENT WITH HETEROGENEOUS, NON-BIASED SAMPLING. THE TREE WAS CREATED WITH GRAPETREE AND NODES WERE POSITIONED THROUGH DYNAMIC RENDERING, MEANING THAT BRANCH LENGTHS ARE NOT SCALED.

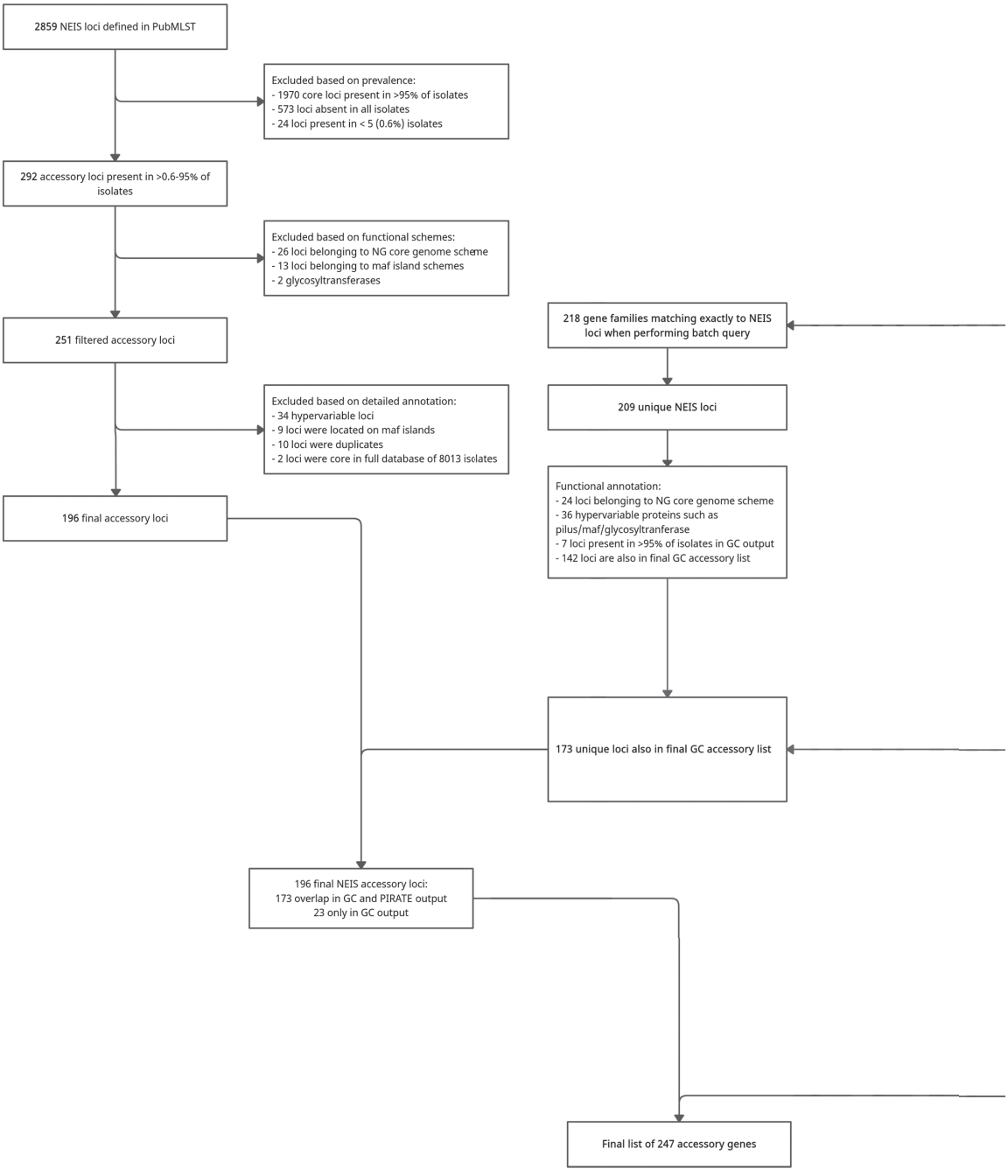
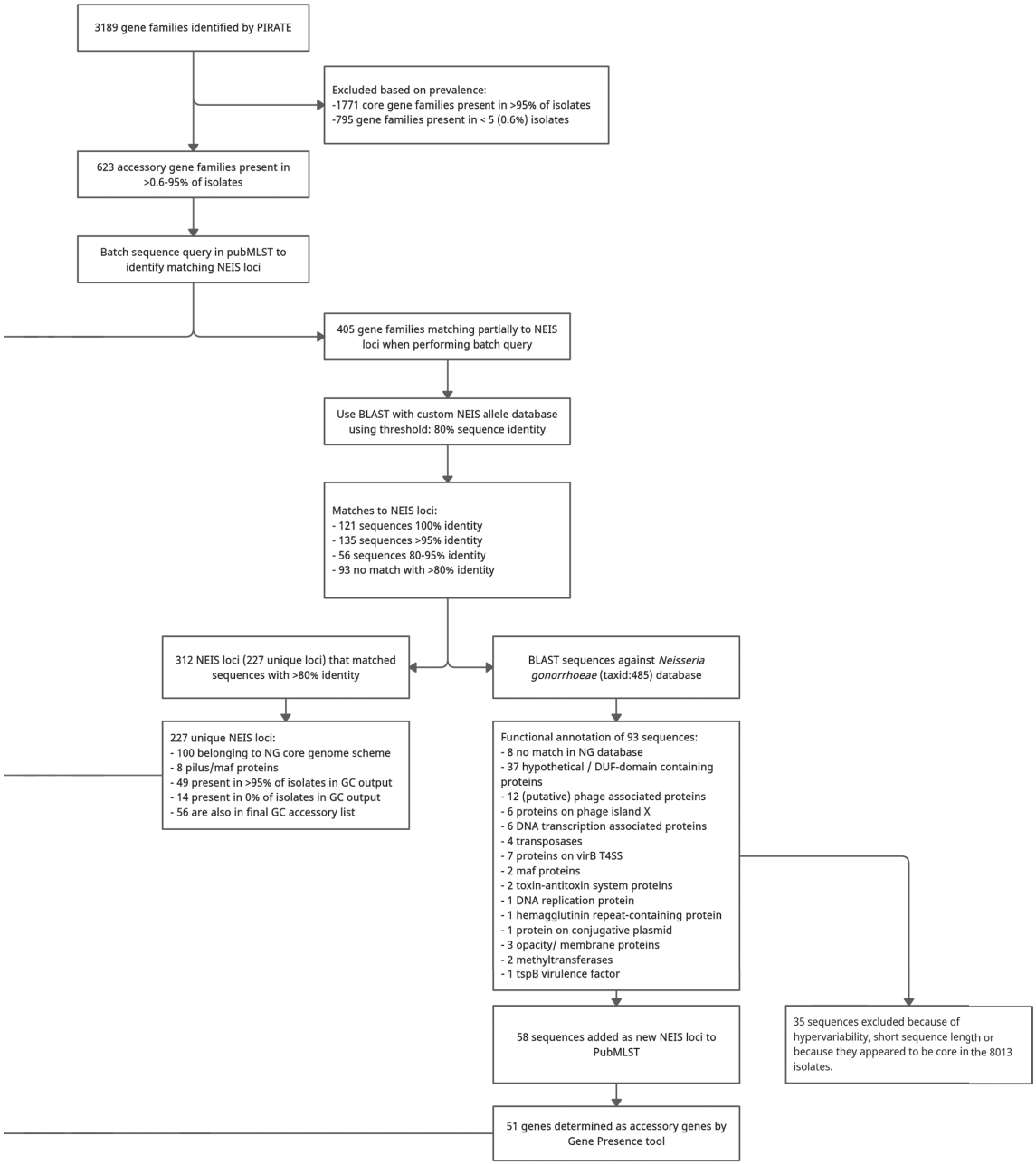
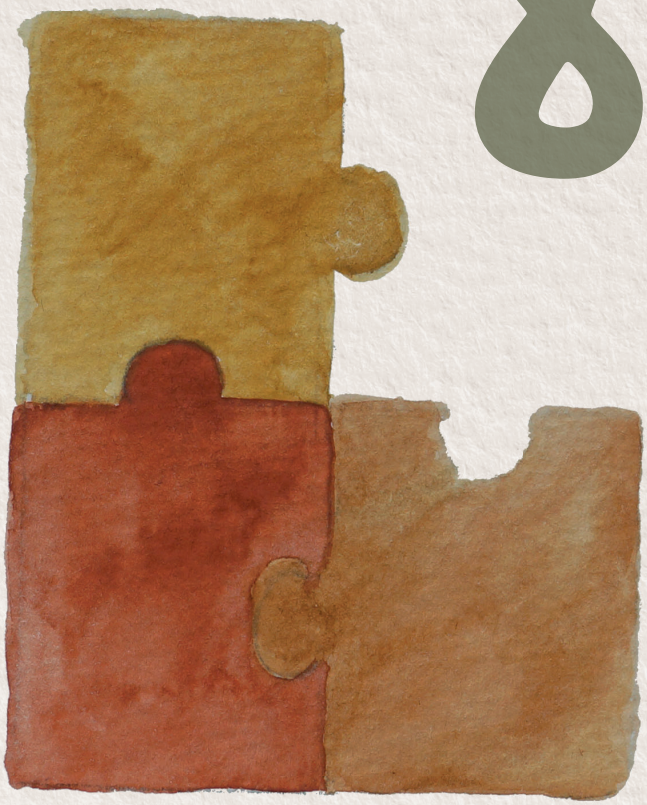


FIGURE S2. DETAILED WORKFLOW AND RESULTS OF EACH FILTERING STEP DURING ACCESSORY GENOME IDENTIFICATION.


3189 gene families identified by PIRATE
Excluded based on prevalence:
-1771 core gene families present in >95% of isolates
-795 gene families present in < 5 (0.6%) isolates
623 accessory gene families present in >0.6-95% of isolates
Batch sequence query in pubMLST to identify matching NEIS loci
405 gene families matching partially to NEIS loci when performing batch query
Use BLAST with custom NEIS allele database using threshold: 80% sequence identity
Matches to NEIS loci:
- 121 sequences 100% identity
- 135 sequences >95% identity
- 56 sequences 80-95% identity
- 93 no match with >80% identity
312 NEIS loci (227 unique loci) that matched sequences with >80% identity
BLAST sequences against Neisseria gonorrhoeae (taxid:485) database
227 unique NEIS loci:
- 100 belonging to NG core genome scheme
- 8 pilus/maf proteins
- 49 present in >95% of isolates in GC output
- 14 present in 0% of isolates in GC output
- 56 are also in final GC accessory list
Functional annotation of 93 sequences:
- 8 no match in NG database
- 37 hypothetical / DUF-domain containing proteins
- 12 (putative) phage associated proteins
- 6 proteins on phage island X
- 6 DNA transcription associated proteins
- 4 transposases
- 7 proteins on virB T4SS
- 2 maf proteins
- 2 toxin-antitoxin system proteins
- 1 DNA replication protein
- 1 hemagglutinin repeat-containing protein
- 1 protein on conjugative plasmid
- 3 opacity/ membrane proteins
- 2 methyltransferases
- 1 tspB virulence factor
58 sequences added as new NEIS loci to PubMLST
35 sequences excluded because of hypervariability, short sequence length or because they appeared to be core in the 8013 isolates.
51 genes determined as accessory genes by Gene Presence tool

8



GENERAL DISCUSSION

In this thesis, I used genomics to perform antimicrobial resistance (AMR) surveillance (**part I**) and to enhance the understanding of *Neisseria gonorrhoeae* biology (**part II**). In this chapter, I first discuss avenues for future treatment of gonorrhoea in the light of the obtained Dutch AMR surveillance data. Second, I place the findings about gonococcal biology in the context of AMR development and spread. Last, I discuss important aspects of the genomic and bioinformatic analyses widely used throughout this thesis, and I propose directions to improve genomic data management.

GONORRHOEA TREATMENT IN AN ERA OF AMR

Worryingly, *N. gonorrhoeae* has developed resistance to all antibiotics used for gonorrhoea treatment in the past, and isolates resistant to the currently used antibiotic ceftriaxone have been reported sporadically. Since cephalosporins are antibiotics of last resort, resistance development forms a major threat to the treatment and control of gonorrhoea. If we are to sustain effective gonorrhoea treatment in the future, avenues to pursue include the improvement of antimicrobial stewardship, the implementation of resistance-guided therapy and the development of novel options for gonorrhoea treatment and prevention.

IMPROVING ANTIMICROBIAL STEWARDSHIP BY REDUCING MACROLIDE CONSUMPTION

Antimicrobial stewardship aims to improve the prescription and subsequent use of antibiotics. The World Health Organisation (WHO) and the European International Union on Sexually Transmitted Infections (STIs) currently recommend empirical dual therapy with ceftriaxone and azithromycin for gonorrhoea[1,2]. Dual therapy ensures that ceftriaxone-resistant strains are still eliminated with azithromycin and therefore combat the spread of ceftriaxone resistance. Dual therapy also aims to overcome the limited efficacy of ceftriaxone monotherapy for pharyngeal gonorrhoea cases (reviewed in[3]), which are of special interest because of the higher chance of acquiring AMR determinants through exchange of genetic material with commensal *Neisseria* spp. at the oropharyngeal site[4]. Resistance to either ceftriaxone or azithromycin has been found as well as sporadic cases of resistance to both antibiotics[5–18]. Despite the sustained transmission of ceftriaxone resistant strains in China, these resistant strains have not yet spread extensively and ceftriaxone resistance is still only rarely found on a global level[19]. Ceftriaxone susceptibility even increased during recent years in the Netherlands (**chapter 4**) and in Europe[20,21].

An opposite pattern was observed for azithromycin during recent years, with an increasing trend in resistance in Europe and world-wide[21,22]. Remarkably, azithromycin resistance also increased in countries where gonorrhoea treatment does not include azithromycin, such as the Netherlands (**chapter 4**). The increase in resistance is presumably due to the widespread use of macrolides to treat *Chlamydia trachomatis* and *Mycoplasma genitalium* infections, non-gonococcal urethritis

and respiratory tract infections[23,24]. Macrolide consumption is especially high in Europe and even higher among populations with high STI rates, such as men who have sex with men (MSM) using pre-exposure prophylaxis (PrEP)[25,26]. High macrolide consumption and subsequent high macrolide exposure levels generally increase the azithromycin resistance selection pressure, especially because azithromycin has a long half-life and may remain in subinhibitory concentrations in the body[27]. This is demonstrated by the high macrolide resistance rates in other sexually transmittable pathogens that are not generally treated with macrolides, such as *Mycoplasma genitalium* and *Treponema pallidum*[28–31]. Dual therapy for gonorrhoea strongly increases macrolide consumption and subsequent exposure levels, especially in populations with high STI rates, thereby contributing to the development of resistance[32]. Several countries currently recommend ceftriaxone monotherapy for gonorrhoea, which has recently been shown to be as effective as dual therapy for extragenital gonorrhoea[33–37]. To reduce macrolide consumption and subsequently slow down the development and spread of macrolide resistance, ceftriaxone monotherapy should be considered as recommended gonorrhoea treatment in more countries, provided that ceftriaxone resistance is closely monitored.

TOWARDS RESISTANCE-GUIDED THERAPY

The use of currently available antibiotics can be improved by determining antimicrobial susceptibility prior to treatment. The WHO rejects an antibiotic as first-line empirical treatment for gonorrhoea when the resistance prevalence exceeds the threshold of 5%[38]. However, at that point the vast majority of isolates are still susceptible and infections with these isolates could still be successfully treated with that antibiotic, although this requires susceptibility testing prior to treatment. Current antimicrobial susceptibility testing strongly relies on the procedure of exposing pure *N. gonorrhoeae* cultures to antibiotics *in vitro,* requiring at least 2 days. Although the time-to-result already hampers susceptibility-based therapy, this is even further complicated by the difficulty to cultivate *N. gonorrhoeae* due to logistic distances between the sample collection site and the laboratory. A faster and easier method for susceptibility testing is the detection of genetic resistance determinants by nucleic acid amplification tests (NAAT), referred to as resistance-guided therapy. This concept is already applied for the treatment of multiple bacterial species, including *Mycobacterium tuberculosis* and *Staphylococcus aureus*[39,40]. In **chapter 2**, an assay for resistance-guided ciprofloxacin therapy for *N. gonorrhoeae* is presented, through detection of the ciprofloxacin-resistance associated *gyrA*-S91F mutation. Ciprofloxacin is no longer empirically used for gonorrhoea because of the high resistance prevalence[21,22]. However, ciprofloxacin is highly effective for gonorrhoea and has advantages such as oral administration and relatively few side effects. Several gonorrhoea treatment guidelines recommend the use of ciprofloxacin when susceptibility is known prior to treatment[3,41]. Using a similar assay to detect *gyrA*-S91 mutations in a prospective clinical trial, resistance-guided ciprofloxacin therapy was shown to be 100% effective for gonorrhoea treatment[42]. This demonstrated the high potential for using such assays in routine diagnostics

to reintroduce previously used antibiotics and subsequently save alternative antibiotics for the future.

Resistance-guided therapy with ceftriaxone is more challenging due to the multifactorial and heterogeneous nature of the cephalosporin resistance mechanisms[43]. The resistance mechanisms are also dynamic, with AMR determinants changing over time as lineages emerge and disappear, even further complicated by differences in dynamics among countries[44]. These AMR dynamics are described for the Netherlands in **chapter 3**, where mutations associated with ceftriaxone reduced susceptibility shifted from a mosaic *penA* to a non-mosaic *penA* with A501V mutation. A Chinese study showed the potential of a molecular assay to detect *penA* mutations associated with ceftriaxone reduced susceptibility in *N. gonorrhoeae*, using samples from a single province in China[45]. A Canadian study developed a molecular assay to determine ceftriaxone susceptibility by screening for mutations in 4 different genes (*ponA, mtrR, penA, porB*)[46]. The assay performed well on Canadian isolates, however its specificity strongly reduced when including isolates from different geographic locations. This showed that generalisation of these assays is difficult due to diverse geographical distributions of gonococcal strains[47]. Algorithms were developed to identify the combination of genetic markers with the highest predictive value for ceftriaxone reduced susceptibility, which could be used for assay development[48]. However, the predictability of markers varies among isolate sets from different periods or different geographical locations, due to the diversity of the mechanisms that mediate reduced susceptibility (**chapter 3**). The mechanism for ceftriaxone resistance, instead of reduced susceptibility, mostly included the mosaic *penA* gene[6–14,16–18]. This could be due to strong clonality of the ceftriaxone-resistant isolates identified so far. However, it could also suggest a more homogeneous mechanism for resistance and a higher potential for assays that discriminate between ceftriaxone susceptible- and resistant isolates. This can currently not be validated because ceftriaxone resistant isolates are rare, however such assays may enable resistance-guided ceftriaxone therapy in the future. These assays should then always be informed by continuous AMR surveillance to ensure that shifts in AMR determinants are not missed.

NOVEL STRATEGIES FOR GONORRHOEA TREATMENT AND PREVENTION

Besides improving the use of currently available antibiotics, novel treatment options are needed to ensure effective gonorrhoea treatment in case ceftriaxone resistance development is inevitable. Multiple clinical trials have provided evidence that alternative antibiotics or newly developed antibiotics may be suitable for gonorrhoea treatment. The New Antibiotic Treatment Options for Uncomplicated Gonorrhoea trial (acronym NABOGO), performed in the Netherlands, and the Gentamicin for the Treatment of Gonorrhoea (G-ToG) trial, performed in the UK, assessed the efficacy of the aminoglycoside gentamicin, either as monotherapy or with azithromycin as dual therapy[49,50]. Gentamicin was non-inferior to ceftriaxone and therefore unsuitable as first-line gonorrhoea treatment, but it still has the potential to effectively treat isolated urogenital

gonorrhoea in case of a cephalosporin allergy. The carbapenem ertapenem was shown to be as effective as ceftriaxone for gonorrhoea treatment. However, ceftriaxone and ertapenem both belong to the β-lactam antibiotics, thus resistance to ceftriaxone might also influence ertapenem susceptibility[49,51]. Ertapenem caused strikingly more side-effects than ceftriaxone and the costs are significantly higher in most settings, both unfavourable aspects for ertapenem to become the first-line gonorrhoea treatment. A potential novel antibiotic is the DNA gyrase B inhibitor zoliflodacin (also referred to as ETX0914 or AZD0914), being the first in a new class of antibiotics and specifically developed for gonorrhoea treatment[52–56]. Zoliflodacin has progressed from *in vitro* studies to clinical trials, which showed successful treatment of urogenital and rectal gonorrhoea and mild side effects[57,58]. The estimated completion time of the phase 3 clinical trial is 2023 (clinicaltrials.gov identifier: NCT03959527), of which the results will show whether zoliflodacin would be a suitable first-line treatment for gonorrhoea. *In vitro* zoliflodacin resistance is caused by mutations in the *gyrB* gene, encoding the main target DNA gyrase B, however these mutations are hitherto barely found in clinical isolates[44,59–62]. The topoisomerase type II inhibitor gepotidacin is another first-in-class antibiotic, developed for the treatment of urogenital tract infections and urogenital gonorrhoea. Despite its promising *in vitro*-activity and its high efficacy against urogenital gonorrhoea, resistance-causing mutations in *gyrA*, encoding the target DNA gyrase A, were already identified during the phase 2 clinical trial[63,64]. *ParC* mutations involved in ciprofloxacin resistance act as stepping-stone to gepotidacin resistance, mediated by an additional *gyrA* mutation which differs from the ciprofloxacin resistance associated *gyrA* mutation[65]. The estimated completion time of the phase 3 clinical trial is 2023 (clinicaltrials.gov identifier: NCT04010539), however AMR is already a concern for this antibiotic.

The new antibiotics are not being developed for the treatment of pharyngeal gonorrhoea. Pharyngeal *N. gonorrhoeae* infection may be an important reservoir of ongoing *N. gonorrhoeae* transmission, hence effective treatment of these infections is essential in gonorrhoea control[4,66]. However, pharyngeal gonorrhoea is the most difficult to treat and had the lowest cure rate in previous clinical trials[49,50,67]. These low cure rates could be influenced by the time set between treatment and test-of-cure with RNA- or DNA based NAATs, as it may take longer to clear *N. gonorrhoeae* from the pharyngeal site than from urogenital or rectal sites[68]. A test-of-cure after the recommended 7-14 days might therefore lead to spurious treatment failures and lower cure rates. Treatment failures might be further validated by genomic analyses, which provide means to separate cases of reinfection with another strain from cases of treatment failure based on genomic profiles of the clinical isolates (isolates used in **chapter 6** were obtained during a clinical trial and reflected treatment failures[49]). Optimisation of methods to identify treatment failure might improve clinical trial outcomes for pharyngeal gonorrhoea, however, a truly higher therapy failure rate is expected for pharyngeal gonorrhoea compared to urogenital gonorrhoea because of suboptimal antibiotic concentrations in the pharyngeal

mucosa[4]. Higher antibiotic doses may therefore be needed to treat pharyngeal gonorrhoea, which requires additional testing in clinical trials.

Prevention is better than cure. Vaccines may become part of strategies to prevent *N. gonorrhoeae* transmission, since vaccines that protect against *Neisseria meningitidis* serogroup B showed potential cross-protection to *N. gonorrhoeae*[69,70]. Effectiveness of the 4CMenB vaccine (also referred to as Bexsero®) against *N. gonorrhoeae* was shown to be 40% in a retrospective observational study[71]. A phase 3 clinical trial is currently prospectively assessing the efficacy of the 4CMenB vaccine against *N. gonorrhoeae* infection, and will be completed in 2024 (clinicaltrials.gov identifier: NCT04415424). Besides, efforts are made towards the development of novel gonococcal vaccins, based on new insights into the immunological response to a gonococcal infection[72,73]. Estimates showed that even with an efficacy of 50%, vaccines could reduce gonorrhoea prevalence with 62% within 2 years, if 30% of the MSM would be fully vaccinated[74]. The vaccine would also be cost-effective in an English public health setting, as shown by a dynamic-transmission model[75]. However, further research is needed on the roll-out of such vaccines in different settings and on the vaccine willingness in key populations.

TACKLING THE DEVELOPMENT OF AMR IN *N. GONORRHOEAE* AND THE SPREAD OF RESISTANT STRAINS

Bacterial genomes are shaped through selective evolutionary pressure, which drives the acquisition and loss of genetic material required for environmental adaptation[76]. Emergence and disappearance of bacterial strains is driven by selective pressure that gives strains with certain genetic characteristics fitness benefit over others. Antimicrobial exposure is such a pressure, with strains that acquired resistance genes having benefit over susceptible strains when exposed to antibiotics. In *N. gonorrhoeae,* antibiotic exposure has driven the evolution towards a more clonal species with distinct strains[77]. Isolates of these strains have limited genetic diversity and mostly similar accessory genomes, as described in **chapter 7**.

THE DEVELOPMENT OF AMR IN *N. GONORRHOEAE*

N. gonorrhoeae has the extraordinary ability to develop AMR, either by taking up resistance determinants through horizontal gene transfer or by *de novo* missense mutations. Although the rate of *de novo* mutations may be low in *N. gonorrhoeae*, selection for resistance mutations can occur after prolonged exposure to subinhibitory concentrations of an antibacterial agent, as shown for the antiseptic chlorhexidine[78]. The need for prolonged exposure is confirmed in **chapter 6,** in which is shown that unsuccessful treatment with fosfomycin, and presumably the subsequent short-term exposure to suboptimal fosfomycin concentrations, did not result in AMR. However, once a beneficial *de novo* AMR-mediating mutation has occurred, the mutation

may easily be transferred among gonococcal strains due to high rates of horizontal gene transfer in *N. gonorrhoeae*. Despite intraspecies horizontal gene transfer being most frequent in *N. gonorrhoeae*, interspecies recombination with closely related *Neisseria* spp. also occurs and has been responsible for the acquisition of mosaic *penA* and *mtr* genes that mediate resistance to respectively cephalosporins and macrolides[79,80].

THE ACQUISITION AND CONSERVATION OF AMR IN GONOCOCCAL STRAINS

AMR determinants are strongly associated with certain gonococcal strains (**chapter 3 and 4**), suggesting its ancestral inheritance or predisposition in certain genetic backgrounds[81]. Genomes predisposed for AMR might contain genes that drive the acquisition of AMR determinants. Indeed, the presence of a type IV secretion system on the gonococcal genetic island (GGI), that enhances horizontal gene transfer, has been associated with AMR to multiple antibiotics[82]. The role of genetic exchange in AMR is confirmed in **chapter 7**, in which associations were identified between AMR and the presence of multiple other accessory genes, besides the GGI, that encode proteins involved in genetic exchange. Resistant strains can subsequently be conserved and expanded once they have a selective advantage over other strains. Some AMR determinants have been shown to improve gonococcal fitness and are therefore likely conserved[83,84]. Other AMR determinants reduce fitness but can still be conserved if the fitness cost is compensated with mutations in other genes[83,85,86]. When the fitness cost is not completely compensated, strains with the AMR determinant may be outcompeted by susceptible strains once the antibiotic pressure is reduced[87].

TACKLING AMR IN *N. GONORRHOEAE*

The spread of gonococcal strains with reduced ceftriaxone susceptibility has been feared during recent years. Indeed, the ceftriaxone reduced susceptible MLST 7827 strain emerged in Amsterdam between 2017 and 2019, as described in **chapter 3**. However, from 2019 onwards an opposite trend towards higher ceftriaxone susceptibility was observed in the Netherlands (**chapter 4**) and in Europe[44]. This raises the questions what factors contribute to this increase in ceftriaxone susceptibility despite ceftriaxone still being used for treatment, and what can be learned from this regarding the spread of AMR.

The increase in susceptibility has been attributed to the implementation of dual therapy with azithromycin for gonorrhoea, which should prevent the spread of ceftriaxone resistant strains[1]. However, this does not explain the trend in the Netherlands, where ceftriaxone monotherapy is recommended[37]. *In vitro* and *in vivo* studies have shown that cephalosporin resistance comes with a fitness cost, presumably leading to resistant strains being outcompeted by susceptible strains once the antibiotic pressure reduces[85]. Indeed, the antibiotic selective pressure might have been changed due to the increase in recommended ceftriaxone dose from 0.25 gram to 0.5 or 1 gram, which appeared to be more effective and prevented the exposure to subinhibitory

concentrations[1,36,37,88,89]. As described above, a major driver of AMR development in *N. gonorrhoeae* is the exposure to subinhibitory antibiotic concentrations with interspecies recombination being mainly responsible for AMR development and intraspecies recombination being the facilitator of AMR transfer among strains. The oropharyngeal site is of particular importance in this regard, as limited antibiotic penetration towards the pharyngeal mucosa may lead to subinhibitory concentrations at this site. Moreover, *N. gonorrhoeae* mostly encounters other commensal *Neisseria* spp. at the oropharyngeal site, facilitating interspecies recombination[4]. The increase in ceftriaxone dose and subsequent prevention of exposing alive *N. gonorrhoeae* bacteria to subinhibitory concentrations could therefore have played an important role in the observed increase in ceftriaxone susceptibility, as it may have blocked the spread of ceftriaxone reduced susceptible strains as well as the development of novel AMR mechanisms. This demonstrates how dose adjustments based on pharmacokinetic and pharmacodynamic parameters could aid in tackling AMR[90]. Optimised doses also contribute to timely gonorrhoea treatment, which in turn contributes to blocking AMR development by limiting the time for intra- and interspecies recombination in *N. gonorrhoeae* in the human body. Shortening the duration of infection should also be supported by public health screening programmes aiming to timely detect and treat asymptomatic *N. gonorrhoeae* infections.

CHALLENGES IN GENOMICS AND BIOINFORMATICS

Genomics and bioinformatics analyses are widely used throughout this thesis. As shown, genomics methodologies can be used for a variety of research objectives, however each application requires specific analyses that come with certain considerations. Several aspects regarding data management and reuse should also be considered when performing bacterial genomics research.

THE DIFFERENT LEVELS OF RESOLUTION TO DETERMINE GENETIC RELATEDNESS

Genomics analyses enable the determination of genetic relationships between bacterial isolates, which can inform outbreak analyses, bacterial population studies, or transmission interventions. Genomics methodologies can also be applied to study the relationship between geno- and phenotype and to unravel genetic mechanisms of AMR. Whatever purpose genomics methodologies may serve, the analyses mainly entail grouping of isolates that are related in a certain way. The level on which relatedness can be studied varies among different methodologies. Typing of multiple genes or gene fragments provides portable and relatively cheap means to study genetic diversity in a bacterial population. However, the selection of typing genes strongly influences the composition of groups of a certain 'sequence type'. For example, in **chapter 6** is demonstrated that isolates with different *N. gonorrhoeae* Multi-Antigen Sequence Types (NG-MAST) had nearly identical core genomes, and that NG-MAST genes of a single isolate

can vary over the course of infection, potentially due to host immune pressure. This typing scheme should therefore not solely be used in longitudinal studies, but rather for transmission studies, which was the intended application of this scheme[91]. The *N. gonorrhoeae* Multi-Locus Sequence Typing (MLST) scheme is suitable to study population dynamics[92], as shown in **chapter 3 and 4**. However, in **chapter 6** is shown that MLST is less suitable for transmission analyses because the low discriminatory value would yield spurious transmission links. Since typing schemes are developed to serve a certain purpose, its resolution for other purposes should always be validated.

Much higher resolution is obtained with the core genome MLST (cgMLST) scheme (used in **chapter 7**), including core genes that are present in nearly all *N. gonorrhoeae* isolates[81]. However, most isolates would be separated when using the allelic profile of all 1668 genes included in the scheme, because isolates with 100% genome identity are very rare due to sequencing errors already causing slight variation[93]. To still be able to group isolates, core genome groups may be formed by clustering isolates using a certain threshold of allelic differences between them[81]. Using a threshold of <400 allelic differences between isolates in a cluster resulted in stable core genome groups. This core genome grouping has a dynamic nature because using another threshold will yield different groups. It is therefore important to clearly state the used thresholds in research reports for correct data interpretation and to enable comparison between studies.

Whole-genome sequencing (WGS) and subsequent Single-Nucleotide Polymorphism (SNP) analyses provide even higher resolution by comparing the whole genome, including the intergenic regions. Its high resolution is promising for transmission analyses, however the definition of a SNP threshold for transmitted isolates is a major challenge, as SNP distances depend on the methods used for SNP distance determination, the bacterial species under study and the time between sampling of the compared isolates[94,95]. The comparative analyses of extended-spectrum β-lactamase-producing *Escherichia coli* (ESBL-Ec) isolates in **chapter 5** used a threshold of 25 SNPs to identify isolates that were potentially transmitted between study participants. This threshold was previously defined when studying long-term *E.coli* carriage in international travellers, with persistent strains differing <25 SNPs[96]. Remarkably, this study also identified isolates with <25 SNPs in unrelated travellers between whom transmission was very unlikely, and these isolates all belonged to a highly clonal *E.coli* subtype. This illustrated that SNP thresholds can even differ between subtypes of a bacterial species, and that SNP distances should always be interpreted in context of the expected genetic diversity in the study population and the epidemiological metadata.

There is ongoing debate on whether or not to mask recombination before comparing genomes. For bacterial species with high recombination rates, such as *N. gonorrhoeae,* recombination events can lead to high SNP counts between otherwise closely related bacterial isolates,

whereas masking recombination could lead to spurious clustering of isolates[95]. Recombination was masked when calculating the SNP distances between *N. gonorrhoeae* isolates in **chapter 6,** and a SNP threshold of <10 recombination filtered SNPs was defined for indistinguishable isolates, either representing a persistent *N. gonorrhoeae* infection over time or an infection with a single *N. gonorrhoeae* strain on multiple anatomical locations. Isolates from different individuals with <10 recombination filtered SNPs suggested *N. gonorrhoeae* transmission between individuals. In **chapter 4**, *N. gonorrhoeae* isolates were compared that were obtained during two periods in time at the Center for Sexual Health of Amsterdam, which showed that the majority of isolates belonging to the MLST 9362 strain, which emerged during the study period, differed <10 recombination filtered SNPs. As it is impossible that this high number of isolates have all been directly transmitted between individuals, isolates with such high similarity might also be part of a transmission network. Therefore, additional epidemiological data are needed to confirm that a direct transmission event has taken place. When epidemiological data are unavailable, SNP analyses might preferably be used to rule out transmission events instead of to identify them.

DATA AVAILABILITY

The wide use of genomics in bacterial research strongly increased the amount of available genomic sequence data. This asked for a good data infrastructure that enabled reusability of data within the research community. FAIR principles were introduced to improve data management and stewardship, by stating that data should be Findable, Accessible, Interoperable and Reusable[97]. Several platforms for bacterial genomic data have been developed that adhere to these principles, thereby enabling the summary and comparison of data on a large scale. Examples are the PubMLST, Enterobase and Pathogenwatch platforms, which contain searchable public databases of sequences with corresponding metadata that can be viewed and analysed in a user-friendly way[98-100]. Public *N. gonorrhoeae* isolates from the PubMLST database were used for comparative analyses in **chapter 3 and 7**, as well as public *E.coli* isolates from Enterobase in **chapter 5**. Maintenance of such databases is very important to keep high data quality, however this is quite labour-intensive and therefore requires time-investment. Since time is money, more dedicated funding should become available to support research institutes in such practices and subsequently support the research community to improve data use.

Besides, databases are only useful when there is sufficient data in it, which requires effort from all researchers who are producing data. With publication of research articles, good quality raw sequence data with corresponding metadata should always be made public to enable reuse. Scientific journals and research funding bodies increasingly demand for publication of raw sequence data, however the way of publication is not always regulated. For instance, corresponding metadata might not be published and raw sequence data without metadata is often useless and do not align with FAIR principles. Since the role of sequence data will be increasingly important in the future, the FAIR principles need to be further implemented in

bacterial genomics research. Despite researchers having the primary responsibility for FAIR data use, scientific journals and research funding bodies can aid in its regulation.

ISOLATE SELECTION

Correct sample selection is essential in good research practice. The selection of bacterial isolates for WGS has a major impact on the study outcomes and should therefore be carefully considered. The selection of isolates for **chapter 3** was based on ceftriaxone susceptibility, in order to identify genetic AMR determinants for reduced ceftriaxone susceptibility. This collection thus contains a disproportionately high number of ceftriaxone reduced susceptible that do not reflect the prevalence of ceftriaxone reduced susceptible isolates in the MSM population of Amsterdam. Since much research is focused on AMR, prioritisation of AMR isolates for sequencing often occurs in research settings. Publishing of such isolate sets leads to a general overrepresentation of AMR isolates in public databases, which is important to take into account when using public data. A completely different selection method based on date of isolation was used in **chapter 4**, to investigate time-dependent changes in the gonococcal population during the COVID-19 pandemic. The isolate selection solely represented the gonococcal population during the study periods, and additional sequence data would be needed to put these isolates in a wider context. Thus, conditions for initial selection of isolates should be considered when reusing these isolate sets.

A lack of genomic data from specific subpopulations can also influence isolate selections and subsequently, influence study outcomes. In **chapter 5,** the *E.coli* isolates of Dutch MSM were compared to publicly available isolates to validate the identified putative transmission events. Isolates involved in putative transmission belonged to a certain type that was very rarely found in the Netherlands before, according to the public database. However, due to population sampling not being done on a regular basis, data on circulating *E.coli* strains in the Netherlands might be incomplete in public databases. For surveillance studies in general, this means that conclusions about the prevalence of certain types are valid with the available data at that moment, although they could change when additional data becomes available.

Also, genomic data from low- and middle-income countries is mostly lacking in public databases due to the limited resources for genomic sequences in these countries[101]. This should be taken into account when studying AMR, as low- and middle-income countries are disproportionately burdened by AMR but genomic AMR data are mostly unavailable[102,103]. Comparative analyses of publicly available *N. gonorrhoeae* isolates from the PubMLST database were performed in **chapter 7,** to identify the gonococcal accessory genome, which consisted of genes present in <95% of isolates in the database. The PubMLST database mainly contains *N. gonorrhoeae* isolates from high-income countries, whereas certain AMR determinants or genomic elements might occur more often in strains circulating in low- and middle income countries, due to differences in antibiotic or environmental pressures. For example, the tetracycline resistance-mediating

tetM plasmid was significantly more prevalent in *N. gonorrhoeae* strains from low- and middle income countries[104]. As a result, the accessory genome identified in **chapter 7** might not be the definitive accessory genome, since additional data from unstudied gonococcal populations could potentially add hitherto undefined accessory genes to it. Additionally, genetic elements can be incorporated into the genome in the future to adapt to environmental pressures that have not been applied before[76]. The composition of the *N. gonorrhoeae* pangenome might therefore differ among different isolate sets, thus instead of defining a strictly separated core- and accessory genome, the pangenome should rather be considered somewhat dynamic[105].

REPRODUCIBILITY OF BIOINFORMATIC ANALYSES

The need for computational algorithms for genomic data analyses grew in parallel with the increase in sequence data. A multitude of bioinformatics tools are available now to address scientific questions, of which the majority has been developed by the research community. Since many microbial bioinformaticians are microbiologists by training and had no extensive training in software engineering, good practices for software engineering are not yet widely adopted within the community. For example, continuously testing the functionality of software prevents computational errors and strongly improves software reproducibility and sustainability, however this practice is not widely applied[106]. As a result, a major concern is the lack of reproducibility of computational analyses with the same software[107]. Reproducibility is also hampered by scientists not publishing the code of their analyses pipelines, thereby not enabling others to repeat their analyses with the exact same parameters. In an attempt to overcome this problem, many open-source pipelines have been published that in theory enable instant data analyses and can easily be shared between scientists (examples are Gen2Epi, Bactopia, rMAP and CABgen[108-111]). However, using these pipelines is often more challenging in practice, for example due to difficult software installation that requires many dependencies or deprecation of included software. These pipelines are also developed for a specific purpose or bacterial species and might still need some adaptation, which in turn requires bioinformatics skills. Since the role of bioinformatics will only increase in the future, dedicated training should become more available for scientists and should form a basic part of educational programmes[112].

CONCLUDING REMARKS

Despite the expectation that AMR will keep evolving in *N. gonorrhoeae* in the future, this development may be slowed down by reducing the consumption of current first-line antibiotics or by reintroducing previously used antibiotics, such as ciprofloxacin. For the latter purpose, molecular methods to rapidly test for susceptibility prior to treatment need to be developed, informed by continuous genomic AMR surveillance to ensure that AMR dynamics are not missed. Clinical trials with the novel antibiotic zoliflodacin for urogenital and rectal gonorrhoea treatment

yield hopeful results, however the need for novel treatment options for pharyngeal gonorrhoea remains. Management of pharyngeal gonorrhoea is essential since pharyngeal gonorrhoea is a potential reservoir of ongoing *N. gonorrhoeae* transmission as well as the main driver of AMR development. Further research is needed on pharmacokinetic and pharmacodynamic parameters of antibiotics in the oropharynx and how antibiotic doses should be optimised for effective treatment of pharyngeal gonorrhoea. In addition, public health interventions should focus on timely treatment by excessively screening populations at high risk for pharyngeal gonorrhoea.

Genomics analyses may support future research on AMR in *N. gonorrhoeae* in multiple ways. As shown in this thesis, genomics analyses enable the identification of AMR mechanisms and provide means to perform genomic AMR surveillance, which informs about gonococcal population- and AMR dynamics. Genomics analyses also aid in improving the understanding of processes involved in AMR development and spread. The obtained knowledge could inform public health interventions targeting key populations for AMR or to develop molecular tests that enable rapid AMR detection in routine diagnostics and that subsequently facilitate individualised treatment. The choice of genomics methodology largely depends on the study objective and on study parameters, such as the species and population under study. However, in general, genomic data only act as supplement to studies and should always be interpreted together with corresponding epidemiological metadata.

REFERENCES

1. M Unemo, JDC Ross, AB Serwin, et al. 2020 European guideline for the diagnosis and treatment of gonorrhoea in adults. *Int J STD AIDS*. Published online 2020. doi:10.1177/0956462420949126
2. World Health Organization, Reproductive Health and Research. *WHO Guidelines for the Treatment of Neisseria Gonorrhoeae.*; 2016. Accessed August 5, 2019. http://www.ncbi.nlm.nih.gov/books/NBK379221/
3. Unemo M, Ross J, Serwin A, et al. Background review for the '2020 European guideline for the diagnosis and treatment of gonorrhoea in adults.' *Int J STD AIDS*. 2021;32(2):108-126. doi:10.1177/0956462420948739
4. Lewis DA. Will targeting oropharyngeal gonorrhoea delay the further emergence of drug-resistant *Neisseria gonorrhoeae* strains? *Sex Transm Infect*. 2015;91(4):234-237. doi:10.1136/sextrans-2014-051731
5. Hook EW, Kirkcaldy RD. A Brief History of Evolving Diagnostics and Therapy for Gonorrhea: Lessons Learned. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2018;67(8):1294-1299. doi:10.1093/cid/ciy271
6. Ohnishi M, Golparian D, Shimuta K, et al. Is *Neisseria gonorrhoeae* initiating a future era of untreatable gonorrhea?: detailed characterization of the first strain with high-level resistance to ceftriaxone. *Antimicrob Agents Chemother*. 2011;55(7):3538-3545. doi:10.1128/AAC.00325-11
7. Unemo M, Golparian D, Nicholas R, et al. High-level cefixime- and ceftriaxone-resistant *Neisseria gonorrhoeae* in France: novel *penA* mosaic allele in a successful international clone causes treatment failure. *Antimicrob Agents Chemother*. 2012;56(3):1273-1280. doi:10.1128/AAC.05760-11
8. Unemo M, Golparian D, Potočnik M, et al. Treatment failure of pharyngeal gonorrhoea with internationally recommended first-line ceftriaxone verified in Slovenia, September 2011. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*. 2012;17(25):20200.
9. Camara J, Serra J, Ayats J, et al. Molecular characterization of two high-level ceftriaxone-resistant *Neisseria gonorrhoeae* isolates detected in Catalonia, Spain. *J Antimicrob Chemother*. 2012;67(8):1858-1860. doi:10.1093/jac/dks162
10. Ko KKK, Chio MT, Goh SS, et al. First Case of Ceftriaxone-Resistant Multidrug-Resistant *Neisseria gonorrhoeae* in Singapore. *Antimicrob Agents Chemother*. 2019;63(5). doi:10.1128/aac.02624-18
11. Eyre DW, Town K, Street T, et al. Detection in the United Kingdom of the *Neisseria gonorrhoeae* FC428 clone, with ceftriaxone resistance and intermediate resistance to azithromycin, October to December 2018. *Euro Surveill*. 2019;24(10):1900147. doi:10.2807/1560-7917.ES.2019.24.10.1900147
12. Xiu L, Yuan Q, Li Y, et al. Emergence of ceftriaxone-resistant *Neisseria gonorrhoeae* strains harbouring a novel mosaic *penA* gene in China. *J Antimicrob Chemother*. 2020;75(4):907-910. doi:10.1093/jac/dkz530
13. Egli K, Roditscheff A, Flückiger U, et al. Molecular characterization of a ceftriaxone-resistant *Neisseria gonorrhoeae* strain found in Switzerland: a case report. *Ann Clin Microbiol Antimicrob*. 2021;20:52. doi:10.1186/s12941-021-00456-5
14. Pleininger S, Indra A, Golparian D, et al. Extensively drug-resistant (XDR) *Neisseria gonorrhoeae* causing possible gonorrhoea treatment failure with ceftriaxone plus azithromycin in Austria, April 2022. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*. 2022;27(24). doi:10.2807/1560-7917.ES.2022.27.24.2200455
15. Fifer H, Cole M, Hughes G, et al. Sustained transmission of high-level azithromycin-resistant *Neisseria gonorrhoeae* in England: an observational study. *Lancet Infect Dis*. 2018;18(5):573-581. doi:10.1016/S1473-3099(18)30122-1
16. Eyre DW, Sanderson ND, Lord E, et al. Gonorrhoea treatment failure caused by a *Neisseria gonorrhoeae* strain with combined ceftriaxone and high-level azithromycin resistance, England, February 2018. *Eurosurveillance*. 2018;23(27):1800323. doi:10.2807/1560-7917.ES.2018.23.27.1800323
17. Zhang L, Zhang C, Zeng Y, et al. Emergence and Characterization of a Ceftriaxone-Resistant *Neisseria gonorrhoeae* FC428 Clone Evolving Moderate-Level Resistance to Azithromycin in Shenzhen, China. *Infect Drug Resist*. 2021;14:4271-4276. doi:10.2147/IDR.S336212

18. Yuan Q, Li Y, Xiu L, et al. Identification of multidrug-resistant *Neisseria gonorrhoeae* isolates with combined resistance to both ceftriaxone and azithromycin, China, 2017-2018. *Emerg Microbes Infect*. 2019;8(1):1546-1549. doi:10.1080/22221751.2019.1681242

19. Lin X, Chen W, Xie Q, et al. Dissemination and genome analysis of high-level ceftriaxone-resistant *penA* 60.001 *Neisseria gonorrhoeae* strains from the Guangdong Gonococcal antibiotics susceptibility Programme (GD-GASP), 2016-2019. *Emerg Microbes Infect*. 2022;11(1):344-350. doi:10.1080/22221751.2021.2011618

20. Harris SR, Cole MJ, Spiteri G, et al. Public health surveillance of multidrug-resistant clones of *Neisseria gonorrhoeae* in Europe: a genomic survey. *Lancet Infect Dis*. 2018;18(7):758-768. doi:10.1016/S1473-3099(18)30225-1

21. Day MJ, Jacobsson S, Spiteri G, et al. Significant increase in azithromycin "resistance" and susceptibility to ceftriaxone and cefixime in *Neisseria gonorrhoeae* isolates in 26 European countries, 2019. *BMC Infect Dis*. 2022;22(1):524. doi:10.1186/s12879-022-07509-w

22. Unemo M, Lahra MM, Escher M, et al. WHO global antimicrobial resistance surveillance for *Neisseria gonorrhoeae* 2017-18: a retrospective observational study. *Lancet Microbe*. 2021;2(11):e627-e636. doi:10.1016/S2666-5247(21)00171-3

23. Kong FYS, Horner P, Unemo M, et al. Pharmacokinetic considerations regarding the treatment of bacterial sexually transmitted infections with azithromycin: a review. *J Antimicrob Chemother*. 2019;74(5):1157-1166. doi:10.1093/jac/dky548

24. Tien V, Punjabi C, Holubar MK. Antimicrobial resistance in sexually transmitted infections. *J Travel Med*. 2020;27(1):taz101. doi:10.1093/jtm/taz101

25. Kenyon C, Buyze J, Wi T. Antimicrobial Consumption and Susceptibility of *Neisseria gonorrhoeae*: A Global Ecological Analysis. *Front Med*. 2018;5. doi:10.3389/fmed.2018.00329

26. Kenyon C, Baetselier ID, Wouters K. Screening for STIs in PrEP cohorts results in high levels of antimicrobial consumption. *Int J STD AIDS*. 2020;31(12):1215-1218. doi:10.1177/0956462420957519

27. Kong FYS, Rupasinghe TW, Simpson JA, et al. Pharmacokinetics of a single 1g dose of azithromycin in rectal tissue in men. *PloS One*. 2017;12(3):e0174372. doi:10.1371/journal.pone.0174372

28. Hetem DJ, Kuizenga Wessel S, Bruisten SM, et al. High prevalence and resistance rates of *Mycoplasma genitalium* among patients visiting two sexually transmitted infection clinics in the Netherlands. *Int J STD AIDS*. 2021;32(9):837-844. doi:10.1177/0956462421999287

29. De Baetselier I, Vuylsteke B, Reyniers T, et al. Worryingly high prevalence of resistance-associated mutations to macrolides and fluoroquinolones in *Mycoplasma genitalium* among men who have sex with men with recurrent sexually transmitted infections. *Int J STD AIDS*. 2022;33(4):385-390. doi:10.1177/09564624211070704

30. Marra CM, Colina AP, Godornes C, et al. Antibiotic selection may contribute to increases in macrolide-resistant *Treponema pallidum*. *J Infect Dis*. 2006;194(12):1771-1773. doi:10.1086/509512

31. Beale MA, Marks M, Sahi SK, et al. Genomic epidemiology of syphilis reveals independent emergence of macrolide resistance across multiple circulating lineages. *Nat Commun*. 2019;10(1):3255. doi:10.1038/s41467-019-11216-7

32. Kenyon C. Dual Azithromycin/Ceftriaxone Therapy for Gonorrhea in PrEP Cohorts Results in Levels of Macrolide Consumption That Exceed Resistance Thresholds by up to 7-Fold. *J Infect Dis*. 2021;224(9):1623-1624. doi:10.1093/infdis/jiab178

33. Aoki T, Mizushima D, Takano M, et al. Efficacy of 1 g Ceftriaxone Monotherapy Compared to Dual Therapy With Azithromycin or Doxycycline for Treating Extragenital Gonorrhea Among Men Who Have Sex With Men. *Clin Infect Dis*. 2021;73(8):1452-1458. doi:10.1093/cid/ciab455

34. Fifer H, Saunders J, Soni S, et al. 2018 UK national guideline for the management of infection with *Neisseria gonorrhoeae*. *Int J STD AIDS*. 2020;31(1):4-15. doi:10.1177/0956462419886775

35. Fouéré S, Cazanave C, Hélary M, et al. Update on French recommendations for the treatment of uncomplicated *Neisseria gonorrhoeae* infections. *Int J STD AIDS*. 2021;32(11):1081-1083. doi:10.1177/09564624211023025

36. Cyr, SS, Barbee L, Workowski KA, et al. Update to CDC's Treatment Guidelines for Gonococcal Infection, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69. doi:10.15585/mmwr.mm6950a6

37. Rijksinstituut voor Volksgezondheid en Milieu. Gonorroe | LCI richtlijnen. Published 2020. Accessed July 13, 2022. http://lci.rivm.nl/richtlijnen/gonorroe
38. World Health Organization. Global action plan to control the spread and impact of antimicrobial resistance in *Neisseria gonorrhoeae*. Published 2012. Accessed August 2, 2022. https://apps.who.int/iris/handle/10665/44863
39. Malhotra-Kumar S, Van Heirstraeten L, Lee A, et al. Evaluation of Molecular Assays for Rapid Detection of Methicillin-Resistant *Staphylococcus aureus*. *J Clin Microbiol*. 2010;48(12):4598-4601. doi:10.1128/JCM.00004-10
40. Nguyen TNA, Anton-Le Berre V, Bañuls AL, et al. Molecular Diagnosis of Drug-Resistant Tuberculosis; A Literature Review. *Front Microbiol*. 2019;10:794. doi:10.3389/fmicb.2019.00794
41. British Association for Sexual Health and HIV. BASHH Guidelines. Accessed December 16, 2019. https://bashh.org/guidelines/
42. Klausner JD, Bristow CC, Soge OO, et al. Resistance-Guided Treatment of Gonorrhea: A Prospective Clinical Study. *Clin Infect Dis*. 2021;73(2):298-303. doi:10.1093/cid/ciaa596
43. Unemo M, del Rio C, Shafer WM. Antimicrobial resistance expressed by *Neisseria gonorrhoeae*: a major global public health problem in the 21st century. *Microbiol Spectr*. 2016;4(3). doi:10.1128/microbiolspec.EI10-0009-2015
44. Sánchez-Busó L, Cole MJ, Spiteri G, et al. Europe-wide expansion and eradication of multidrug-resistant *Neisseria gonorrhoeae* lineages: a genomic surveillance study. *Lancet Microbe*. 2022;3(6):e452-e463. doi:10.1016/S2666-5247(22)00044-1
45. Zhao L, Liu A, Li R, et al. Multiplex TaqMan real-time PCR platform for detection of *Neisseria gonorrhoeae* with decreased susceptibility to ceftriaxone. *Diagn Microbiol Infect Dis*. 2019;93(4):299-304. doi:10.1016/j.diagmicrobio.2018.10.013
46. Peterson SW, Martin I, Demczuk W, et al. Molecular Assay for Detection of Genetic Markers Associated with Decreased Susceptibility to Cephalosporins in *Neisseria gonorrhoeae*. *J Clin Microbiol*. 2015;53(7):2042-2048. doi:10.1128/JCM.00493-15
47. Lin EY, Adamson PC, Klausner JD. Evaluating the generalizability of a multiplex real-time PCR assay for predicting decreased susceptibility to ceftriaxone in a global set of *Neisseria gonorrhoeae* sequences. *J Antimicrob Chemother*. 2021;76(4):1104-1106. doi:10.1093/jac/dkaa525
48. Lin EY, Adamson PC, Deng X, et al. Establishing Novel Molecular Algorithms to Predict Decreased Susceptibility to Ceftriaxone in *Neisseria gonorrhoeae* Strains. *J Infect Dis*. 2020;223(7):1232-1240. doi:10.1093/infdis/jiaa495
49. Vries HJC de, Laat M de, Jongen VW, et al. Efficacy of ertapenem, gentamicin, fosfomycin, and ceftriaxone for the treatment of anogenital gonorrhoea (NABOGO): a randomised, non-inferiority trial. *Lancet Infect Dis*. 2022;0(0). doi:10.1016/S1473-3099(21)00625-3
50. Ross JDC, Brittain C, Cole M, et al. Gentamicin compared with ceftriaxone for the treatment of gonorrhoea (G-ToG): a randomised non-inferiority trial. *Lancet Lond Engl*. 2019;393(10190):2511-2520. doi:10.1016/S0140-6736(18)32817-4
51. Unemo M, Golparian D, Limnios A, et al. In Vitro Activity of Ertapenem versus Ceftriaxone against *Neisseria gonorrhoeae* Isolates with Highly Diverse Ceftriaxone MIC Values and Effects of Ceftriaxone Resistance Determinants: Ertapenem for Treatment of Gonorrhea? *Antimicrob Agents Chemother*. 2012;56(7):3603-3609. doi:10.1128/AAC.00326-12
52. Jacobsson S, Golparian D, Alm RA, et al. High in vitro activity of the novel spiropyrimidinetrione AZD0914, a DNA gyrase inhibitor, against multidrug-resistant *Neisseria gonorrhoeae* isolates suggests a new effective option for oral treatment of gonorrhea. *Antimicrob Agents Chemother*. 2014;58(9):5585-5588. doi:10.1128/AAC.03090-14
53. Kern G, Palmer T, Ehmann DE, et al. Inhibition of *Neisseria gonorrhoeae* Type II Topoisomerases by the Novel Spiropyrimidinetrione AZD0914. *J Biol Chem*. 2015;290(34):20984-20994. doi:10.1074/jbc.M115.663534
54. Papp JR, Lawrence K, Sharpe S, et al. In vitro growth of multidrug-resistant *Neisseria gonorrhoeae* isolates is inhibited by ETX0914, a novel spiropyrimidinetrione. *Int J Antimicrob Agents*. 2016;48(3):328-330. doi:10.1016/j.ijantimicag.2016.05.018

55. Unemo M, Ringlander J, Wiggins C, et al. High in vitro susceptibility to the novel spiropyrimidinetrione ETX0914 (AZD0914) among 873 contemporary clinical *Neisseria gonorrhoeae* isolates from 21 European countries from 2012 to 2014. *Antimicrob Agents Chemother*. 2015;59(9):5220-5225. doi:10.1128/AAC.00786-15
56. Su XH, Wang BX, Le WJ, et al. Multidrug-Resistant *Neisseria gonorrhoeae* Isolates from Nanjing, China, Are Sensitive to Killing by a Novel DNA Gyrase Inhibitor, ETX0914 (AZD0914). *Antimicrob Agents Chemother*. 2016;60(1):621-623. doi:10.1128/AAC.01211-15
57. Basarab GS, Kern GH, McNulty J, et al. Responding to the challenge of untreatable gonorrhea: ETX0914, a first-in-class agent with a distinct mechanism-of-action against bacterial Type II topoisomerases. *Sci Rep*. 2015;5:11827. doi:10.1038/srep11827
58. Taylor SN, Marrazzo J, Batteiger BE, et al. Single-Dose Zoliflodacin (ETX0914) for Treatment of Urogenital Gonorrhea. *N Engl J Med*. 2018;379(19):1835-1845. doi:10.1056/NEJMoa1706988
59. Unemo M, Ahlstrand J, Sánchez-Busó L, et al. High susceptibility to zoliflodacin and conserved target (*GyrB*) for zoliflodacin among 1209 consecutive clinical *Neisseria gonorrhoeae* isolates from 25 European countries, 2018. *J Antimicrob Chemother*. 2021;76(5):1221-1228. doi:10.1093/jac/dkab024
60. Le W, Su X, Lou X, et al. Susceptibility Trends of Zoliflodacin against Multidrug-Resistant *Neisseria gonorrhoeae* Clinical Isolates in Nanjing, China, 2014 to 2018. *Antimicrob Agents Chemother*. 2021;65(3):e00863-20. doi:10.1128/AAC.00863-20
61. Adamson PC, Lin EY, Ha SM, et al. Using a public database of *Neisseria gonorrhoeae* genomes to detect mutations associated with zoliflodacin resistance. *J Antimicrob Chemother*. 2021;76(11):2847-2849. doi:10.1093/jac/dkab262
62. Jacobsson S, Golparian D, Oxelbark J, et al. Pharmacodynamic Evaluation of Zoliflodacin Treatment of *Neisseria gonorrhoeae* Strains With Amino Acid Substitutions in the Zoliflodacin Target *GyrB* Using a Dynamic Hollow Fiber Infection Model. *Front Pharmacol*. 2022;13:874176. doi:10.3389/fphar.2022.874176
63. Farrell DJ, Sader HS, Rhomberg PR, et al. In Vitro Activity of Gepotidacin (GSK2140944) against *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother*. 2017;61(3):e02047-16. doi:10.1128/AAC.02047-16
64. Taylor SN, Morris DH, Avery AK, et al. Gepotidacin for the Treatment of Uncomplicated Urogenital Gonorrhea: A Phase 2, Randomized, Dose-Ranging, Single-Oral Dose Evaluation. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2018;67(4):504-512. doi:10.1093/cid/ciy145
65. Scangarella-Oman NE, Hossain M, Dixon PB, et al. Microbiological Analysis from a Phase 2 Randomized Study in Adults Evaluating Single Oral Doses of Gepotidacin in the Treatment of Uncomplicated Urogenital Gonorrhea Caused by *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother*. 2018;62(12):e01221-18. doi:10.1128/AAC.01221-18
66. Hook EW, Bernstein K. Kissing, saliva exchange, and transmission of *Neisseria gonorrhoeae*. *Lancet Infect Dis*. 2019;19(10):e367-e369. doi:10.1016/S1473-3099(19)30306-8
67. Moran JS. Treating uncomplicated *Neisseria gonorrhoeae* infections: is the anatomic site of infection important? *Sex Transm Dis*. 1995;22(1):39-47. doi:10.1097/00007435-199501000-00007
68. Hananta IPY, De Vries HJC, van Dam AP, et al. Persistence after treatment of pharyngeal gonococcal infections in patients of the STI clinic, Amsterdam, the Netherlands, 2012-2015: a retrospective cohort study. *Sex Transm Infect*. 2017;93(7):467-471. doi:10.1136/sextrans-2017-053147
69. Petousis-Harris H, Paynter J, Morgan J, et al. Effectiveness of a group B outer membrane vesicle meningococcal vaccine against gonorrhoea in New Zealand: a retrospective case-control study. *The Lancet*. 2017;390(10102):1603-1610. doi:10.1016/S0140-6736(17)31449-6
70. Ruiz García Y, Sohn WY, Seib KL, et al. Looking beyond meningococcal B with the 4CMenB vaccine: the *Neisseria* effect. *NPJ Vaccines*. 2021;6(1):130. doi:10.1038/s41541-021-00388-3
71. Abara WE, Bernstein KT, Lewis FMT, et al. Effectiveness of a serogroup B outer membrane vesicle meningococcal vaccine against gonorrhoea: a retrospective observational study. *Lancet Infect Dis*. 2022;22(7):1021-1029. doi:10.1016/S1473-3099(21)00812-4
72. Gala RP, Zaman RU, D'Souza MJ, et al. Novel Whole-Cell Inactivated *Neisseria gonorrhoeae* Microparticles as Vaccine Formulation in Microneedle-Based Transdermal Immunization. *Vaccines*. 2018;6(3):60. doi:10.3390/vaccines6030060

73. Bagwe P, Bajaj L, Gala RP, et al. Assessment of In Vitro Immunostimulatory Activity of an Adjuvanted Whole-Cell Inactivated *Neisseria gonorrhoeae* Microparticle Vaccine Formulation. *Vaccines*. 2022;10(7):983. doi:10.3390/vaccines10070983
74. Hui BB, Padeniya TN, Rebuli N, et al. A Gonococcal Vaccine Has the Potential to Rapidly Reduce the Incidence of *Neisseria gonorrhoeae* Infection Among Urban Men Who Have Sex With Men. *J Infect Dis*. 2022;225(6):983-993. doi:10.1093/infdis/jiab581
75. Whittles LK, Didelot X, White PJ. Public health impact and cost-effectiveness of gonorrhoea vaccination: an integrated transmission-dynamic health-economic modelling analysis. *Lancet Infect Dis*. 2022;22(7):1030-1041. doi:10.1016/S1473-3099(21)00744-1
76. Dobrindt U, Hacker J. Whole genome plasticity in pathogenic bacteria. *Curr Opin Microbiol*. 2001;4(5):550-557. doi:10.1016/S1369-5274(00)00250-2
77. Golparian D, Harris SR, Sánchez-Busó L, et al. Genomic evolution of *Neisseria gonorrhoeae* since the preantibiotic era (1928–2013): antimicrobial use/misuse selects for resistance and drives evolution. *BMC Genomics*. 2020;21(1):116. doi:10.1186/s12864-020-6511-6
78. Laumen JGE, Van Dijck C, Manoharan-Basil SS, et al. Sub-Inhibitory Concentrations of Chlorhexidine Induce Resistance to Chlorhexidine and Decrease Antibiotic Susceptibility in *Neisseria gonorrhoeae*. *Front Microbiol*. 2021;12:776909. doi:10.3389/fmicb.2021.776909
79. Ameyama S, Onodera S, Takahata M, et al. Mosaic-Like Structure of Penicillin-Binding Protein 2 Gene (*penA*) in Clinical Isolates of *Neisseria gonorrhoeae* with Reduced Susceptibility to Cefixime. *Antimicrob Agents Chemother*. 2002;46(12):3744-3749. doi:10.1128/AAC.46.12.3744-3749.2002
80. Wadsworth CB, Arnold BJ, Sater MRA, et al. Azithromycin Resistance through Interspecific Acquisition of an Epistasis-Dependent Efflux Pump Component and Transcriptional Regulator in *Neisseria gonorrhoeae*. *mBio*. 2018;9(4). doi:10.1128/mBio.01419-18
81. Harrison OB, Cehovin A, Skett J, et al. *Neisseria gonorrhoeae* Population Genomics: Use of the Gonococcal Core Genome to Improve Surveillance of Antimicrobial Resistance. *J Infect Dis*. 2020;222(11):1816-1825. doi:10.1093/infdis/jiaa002
82. Harrison OB, Clemence M, Dillard JP, et al. Genomic analyses of *Neisseria gonorrhoeae* reveal an association of the gonococcal genetic island with antimicrobial resistance. *J Infect*. 2016;73(6):578-587. doi:10.1016/j.jinf.2016.08.010
83. Kunz AN, Begum AA, Wu H, et al. Impact of fluoroquinolone resistance mutations on gonococcal fitness and in vivo selection for compensatory mutations. *J Infect Dis*. 2012;205(12):1821-1829. doi:10.1093/infdis/jis277
84. Warner DM, Shafer WM, Jerse AE. Clinically relevant mutations that cause derepression of the *Neisseria gonorrhoeae* MtrC-MtrD-MtrE Efflux pump system confer different levels of antimicrobial resistance and *in vivo* fitness. *Mol Microbiol*. 2008;70(2):462-478. doi:10.1111/j.1365-2958.2008.06424.x
85. Vincent LR, Kerr SR, Tan Y, et al. In Vivo-Selected Compensatory Mutations Restore the Fitness Cost of Mosaic *penA* Alleles That Confer Ceftriaxone Resistance in *Neisseria gonorrhoeae*. *mBio*. 2018;9(2):e01905-17. doi:10.1128/mBio.01905-17
86. Warner DM, Folster JP, Shafer WM, et al. Regulation of the MtrC-MtrD-MtrE efflux-pump system modulates the in vivo fitness of *Neisseria gonorrhoeae*. *J Infect Dis*. 2007;196(12):1804-1812. doi:10.1086/522964
87. Andersson DI. The biological cost of mutational antibiotic resistance: any practical conclusions? *Curr Opin Microbiol*. 2006;9(5):461-465. doi:10.1016/j.mib.2006.07.002
88. Chisholm SA, Mouton JW, Lewis DA, et al. Cephalosporin MIC creep among gonococci: time for a pharmacodynamic rethink? *J Antimicrob Chemother*. 2010;65(10):2141-2148. doi:10.1093/jac/dkq289
89. Connolly KL, Eakin AE, Gomez C, et al. Pharmacokinetic Data Are Predictive of In Vivo Efficacy for Cefixime and Ceftriaxone against Susceptible and Resistant *Neisseria gonorrhoeae* Strains in the Gonorrhea Mouse Model. *Antimicrob Agents Chemother*. 2019;63(3):e01644-18. doi:10.1128/AAC.01644-18
90. Blondeau JM, Hansen G, Metzler K, et al. The Role of PK/PD Parameters to Avoid Selection and Increase of Resistance: Mutant Prevention Concentration. *J Chemother*. 2004;16(sup3):1-19. doi:10.1080/1120009X.2004.11782371
91. Martin IMC, Ison CA, Aanensen DM, et al. Rapid sequence-based identification of gonococcal transmission clusters in a large metropolitan area. *J Infect Dis*. 2004;189(8):1497-1505. doi:10.1086/383047

92. Bennett JS, Jolley KA, Sparling PF, et al. Species status of *Neisseria gonorrhoeae*: evolutionary and epidemiological inferences from multilocus sequence typing. *BMC Biol*. 2007;5:35. doi:10.1186/1741-7007-5-35

93. Stoler N, Nekrutenko A. Sequencing error profiles of Illumina sequencing instruments. *NAR Genomics Bioinforma*. 2021;3(1):lqab019. doi:10.1093/nargab/lqab019

94. Silva DD, Peters J, Cole K, et al. Whole-genome sequencing to determine *Neisseria gonorrhoeae* transmission : an observational study. *Lancet Infect Dis*. 2017;16(11):1295-1303. doi:10.1016/S1473-3099(16)30157-8

95. Gorrie CL, Silva AGD, Ingle DJ, et al. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. *Lancet Microbe*. 2021;2(11):e575-e583. doi:10.1016/S2666-5247(21)00149-X

96. Putten BCL van der, Hattem JM van, Penders J, et al. Extraintestinal pathogenic *Escherichia coli* (ExPEC) are associated with prolonged carriage of extended-spectrum β-lactamase-producing *E. coli* acquired during travel. bioRxiv. Published online October 26, 2021. doi:10.1101/2020.09.23.309856

97. Wilkinson MD, Dumontier M, Aalbersberg IjJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3(1):160018. doi:10.1038/sdata.2016.18

98. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res*. 2018;3:124. doi:10.12688/wellcomeopenres.14826.1

99. Zhou Z, Alikhan NF, Mohamed K, et al. The EnteroBase user's guide, with case studies on Salmonella transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res*. 2020;30(1):138-152. doi:10.1101/gr.251678.119

100. Sánchez-Busó L, Yeats CA, Taylor B, et al. A community-driven resource for genomic epidemiology and antimicrobial resistance prediction of *Neisseria gonorrhoeae* at Pathogenwatch. *Genome Med*. 2021;13(1):61. doi:10.1186/s13073-021-00858-2

101. Helmy M, Awad M, Mosa KA. Limited resources of genome sequencing in developing countries: Challenges and solutions. *Appl Transl Genomics*. 2016;9:15-19. doi:10.1016/j.atg.2016.03.003

102. Murray CJ, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*. 2022;399(10325):629-655. doi:10.1016/S0140-6736(21)02724-0

103. World Health Organization. *Antimicrobial Resistance: Global Report on Surveillance*. World Health Organization; 2014. Accessed August 10, 2022. https://apps.who.int/iris/handle/10665/112642

104. Cehovin A, Jolley KA, Maiden MCJ, et al. Association of *Neisseria gonorrhoeae* Plasmids With Distinct Lineages and The Economic Status of Their Country of Origin. *J Infect Dis*. 2020;222(11):1826-1836. doi:10.1093/infdis/jiaa003

105. Tettelin H, Medini D, eds. *The Pangenome: Diversity, Dynamics and Evolution of Genomes*. Springer International Publishing; 2020. doi:10.1007/978-3-030-38281-0

106. van der Putten BCL, Mendes CI, Talbot BM, et al. Software testing in microbial bioinformatics: a call to action. *Microb Genomics*. 2022;8(3):000790. doi:10.1099/mgen.0.000790

107. Stodden V, Seiler J, Ma Z. An empirical analysis of journal policy effectiveness for computational reproducibility. *Proc Natl Acad Sci U S A*. 2018;115(11):2584-2589. doi:10.1073/pnas.1708290115

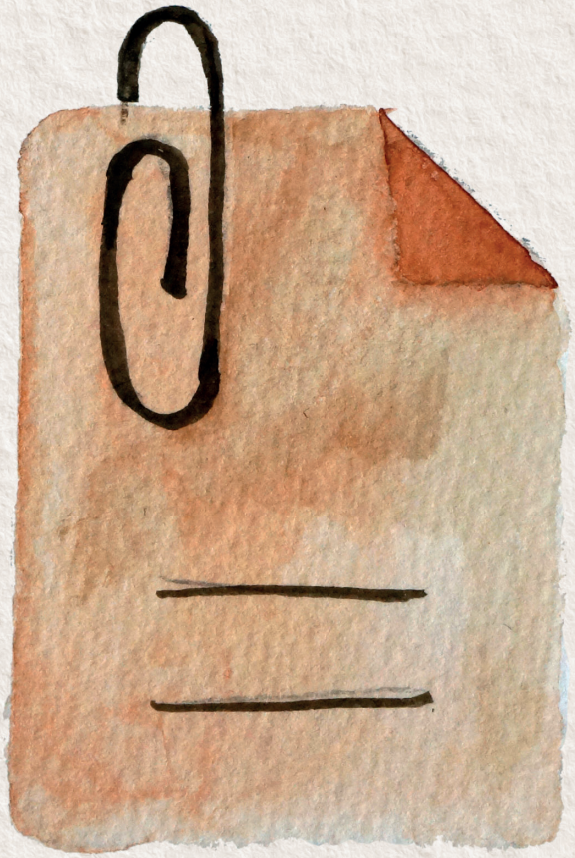
108. Singh R, Dillon JAR, Demczuk W, et al. Gen2Epi: an automated whole-genome sequencing pipeline for linking full genomes to antimicrobial susceptibility and molecular epidemiological data in *Neisseria gonorrhoeae*. *BMC Genomics*. 2019;20(165). doi:10.1186/s12864-019-5542-3

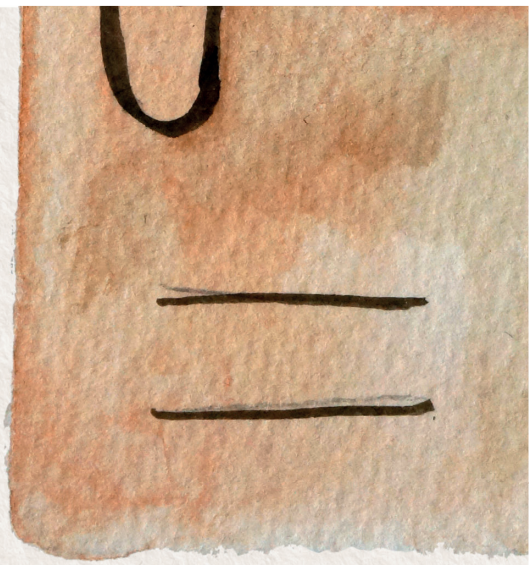
109. Petit RA, Read TD. Bactopia: a Flexible Pipeline for Complete Analysis of Bacterial Genomes. *mSystems*. 5(4):e00190-20. doi:10.1128/mSystems.00190-20

110. Sserwadda I, Mboowa G. rMAP: the Rapid Microbial Analysis Pipeline for ESKAPE bacterial group whole-genome sequence data. *Microb Genomics*. 2021;7(6). doi:10.1099/mgen.0.000583

111. Duré FM, Silveira MC, Rocha-de-Souza CM, et al. CABGen: A Web Application for the Bioinformatic Analysis of Bacterial Genomes. *Front Microbiol*. 2022;13:893474. doi:10.3389/fmicb.2022.893474

112. Attwood TK, Blackford S, Brazas MD, et al. A global perspective on evolving bioinformatics and data science training needs. *Brief Bioinform*. 2019;20(2):398-404. doi:10.1093/bib/bbx100





APPENDICES

SUMMARY

GENOMICS APPROACHES FOR CHARACTERISING AND TRACKING ANTIMICROBIAL RESISTANCE IN *NEISSERIA GONORRHOEAE*

The bacterium *Neisseria gonorrhoeae*, the gonococcus, causes the sexually transmittable infection (STI) gonorrhoea, one of the most prevalent STIs world-wide. Treatment and control of gonorrhoea is currently threatened by antimicrobial resistance (AMR), since *N. gonorrhoeae* has developed resistance to all antibiotics formerly used for gonorrhoea treatment. Resistance to the currently used antibiotic ceftriaxone has already been detected, albeit very rarely. To prevent the spread of antimicrobial-resistant *N. gonorrhoeae* strains, AMR surveillance is indispensable. AMR surveillance can be supplemented by genomics to identify and track genetic AMR determinants. **Part I** of this thesis contains genomics approaches to perform AMR surveillance on a national level in the Netherlands. Similar genomics approaches are used in **part II,** for the purpose of gaining novel insights in gonococcal biology.

In **chapter 1**, several aspects of gonorrhoea are introduced, and the history of antimicrobial treatment of gonorrhoea and the associated genetic mechanisms of resistance are summarised. The genomics methodologies and bioinformatic analyses that are used throughout this thesis are also introduced in this chapter. Molecular typing is described as a tool to study genetic diversity among bacterial populations based on sequences of genes or gene fragments. Whole-genome sequencing (WGS) has a considerably higher resolution and can be used for comparative analyses on a gene-by-gene level or a single nucleotide polymorphism (SNP) level.

PART I: USING GENOMICS FOR ANTIMICROBIAL RESISTANCE SURVEILLANCE

A MOLECULAR ASSAY TO DETERMINE *N. GONORRHOEAE* FLUOROQUINOLONE SUSCEPTIBILITY

Recommendations for the treatment of bacterial infections change according to resistance prevalence. In the past, gonorrhoea was effectively treated with the fluoroquinolone ciprofloxacin, which has advantages over other antibiotics such as oral administration and relatively few side effects. Ciprofloxacin was eliminated as recommended gonorrhoea treatment when fluoroquinolone-resistant *N. gonorrhoeae* strains emerged. However, still a considerable proportion of *N. gonorrhoeae* strains is susceptible to ciprofloxacin and infections with susceptible strains could still be treated with this antibiotic. This would require susceptibility testing prior to treatment, which is impossible with current culture-based susceptibility tests that require at least 2 days. In **chapter 2**, a qPCR assay for simultaneous detection of *N. gonorrhoeae* and the ciprofloxacin resistance-mediating mutation *gyrA*-S91F is evaluated.

We suggest implementation of this qPCR assay in routine diagnostics for fluoroquinolone-susceptibility testing after *N. gonorrhoeae* detection with the currently used TMA assay, as this has a higher detection sensitivity. Besides enabling resistance-guided prescription of ciprofloxacin, the qPCR assay also provides fluoroquinolone susceptibility testing for samples that cannot be cultivated and thereby aids in AMR surveillance.

CEFTRIAXONE REDUCED SUSCEPTIBLE *N. GONORRHOEAE* STRAINS IN AMSTERDAM

The currently used antibiotic ceftriaxone is the last remaining treatment option for gonorrhoea, thus ceftriaxone resistance surveillance is of great importance. Routine ceftriaxone susceptibility testing had shown an increase in ceftriaxone reduced susceptible isolates between 2014 and 2019 in Amsterdam. Typing of the gene encoding the primary ceftriaxone target (*penA* gene) had led to the identification of the mosaic *penA* gene as the initial cause of the reduced ceftriaxone susceptibility among the gonococcal population in Amsterdam, which later changed to a non-mosaic *penA* gene with A501 mutation. In chapter 3, previous data is supplemented with WGS data to further elucidate AMR dynamics in the Amsterdam gonococcal population. All ceftriaxone reduced susceptible isolates (minimum inhibitory concentration (MIC) >0.064 mg/L) obtained between 2014 and 2019 were genetically characterised, as well as a random selection of ceftriaxone susceptible isolates from the same time period. Previous observations regarding the change in *penA* represented a shift in ceftriaxone reduced susceptible *N. gonorrhoeae* strains over time from MLST 1901 with mosaic *penA* to MLST 7363 and more recently to MLST 7827, both with a non-mosaic *penA* with A501 mutation. Additional *porB1b* mutations were frequently found among ceftriaxone reduced susceptible isolates. The ceftriaxone reduced susceptible MLST 7827 strain with *penA* A501V and *porB1b* G120K/A121D mutations emerged in Amsterdam between 2017 and 2019. The MLST 7827 strain had also emerged in Norway, and Dutch and Norwegian MLST 7827 isolates strongly clustered on a WGS level, suggesting extensive spread of the strain in Europe. This demonstrated the need for continuous AMR surveillance to detect changes in AMR mechanisms over time.

GONOCOCCAL POPULATION DYNAMICS DURING THE FIRST COVID-19 LOCKDOWN IN AMSTERDAM

An update on the genomic surveillance of *N. gonorrhoeae* is given in **chapter 4,** in which the *N. gonorrhoeae* population dynamics were investigated during the first COVID-19 lockdown in Amsterdam in 2020. A previous study had shown a temporary reduction in the number of casual sex partners reported by clients of the Centre for Sexual Health (CSH) in Amsterdam during the lockdown. This behavioural change might also have influenced STI transmission networks and might subsequently have changed the composition of sexually transmittable bacteria populations. In this chapter, the genotypic and phenotypic characteristics were compared of the *N. gonorrhoeae* population in Amsterdam before (between January 15th and February 29th 2020) and during (between May 15th and June 30th 2020) the first lockdown.

The composition of the gonococcal population changed, with a shift from the MLST 8156 strain being slightly predominant before lockdown to the vastly predominant MLST 9362 strain during lockdown. The MLST 9362 carried mosaic *mtr* genes associated with azithromycin resistance and had no mutations associated with ceftriaxone reduced susceptibility, which explains the observed increase in azithromycin resistance and the higher ceftriaxone susceptibility during the lockdown. The high genetic similarity of MLST 9362 isolates showed that this strain clonally expanded in Amsterdam, suggesting local transmission of this strain during the lockdown. This chapter demonstrated the consequences of public health measures for the population dynamics of infectious diseases, such as gonorrhoea, which should be taken into account when performing surveillance.

GENOMIC SURVEILLANCE OF EXTENDED-SPECTRUM CEPHALOSPORIN-RESISTANT *ESCHERICHIA COLI*

Similar genomic methodologies were applied to the surveillance of extended-spectrum β-lactamases (ESBL)-producing *Escherichia coli* (ESBL-Ec), resistant to extended-spectrum cephalosporins. *E.coli* is an opportunistic human pathogen that can cause urinary tract infections or gastrointestinal symptoms, but is mainly asymptomatically carried in the intestines. Since ESBL genes can be horizontally transferred between different bacterial species (not towards Neisseriaceae), circulation of ESBL genes are of great relevance to AMR surveillance. A previous Dutch study had identified a strikingly higher carriage of ESBL-Ec among men who have sex with men (MSM) participating in the Amsterdam Cohort Studies (16.3%), compared to the general Dutch population (8.6%). In **chapter 5,** ESBL-Ec isolates obtained from MSM were genetically characterised and it was investigated whether the increased ESBL-Ec prevalence could be explained by transmission between study participants. Genomic characterisation showed that the majority of isolates belonged to the globally expanded MLST 131 lineage. Putative transmission clusters were identified based on genomic similarity between isolates. Two putative transmission clusters contained MLST 14 and MLST 394 isolates, which are strains very rarely found in public data. The identification of these strains in multiple participants, with very low genetic distance between them, supported potential ESBL-Ec transmission between MSM in Amsterdam. This implies that ESBL-Ec infection should be considered when sexually active MSM present with symptoms associated with *E.coli* infection.

PART II: USING GENOMICS TO UNDERSTAND GONOCOCCAL BIOLOGY

ENHANCING THE UNDERSTANDING OF *N. GONORRHOEAE* INFECTION DYNAMICS

Genetic variation that occurs in *N. gonorrhoeae* during an infection can inform about the processes involved in infection. In **chapter 6**, within-host genetic variation was assessed over the course of infection by comparing Multi-Locus Sequence Types (MLST), NG-Multi Antigen

Sequence Types (NG-MAST) and NG-Sequence Types for Antimicrobial Resistance (NG-STAR) and recombination filtered- and unfiltered SNP distances between paired isolates from a single individual from consecutive time points or from multiple anatomical locations at a single time point. Comparisons showed that isolates with different typing profiles also had high SNP distances, whereas low SNP distances were found between isolates with similar profiles or with a different NG-MAST only. Isolates from a single patient were highly similar, with only up to 10 recombination filtered SNPs difference. The genetic variation observed during infection was concentrated in genomic regions that encode hypervariable proteins under immune pressure, demonstrating that the human immune response is the main driver of genetic variation in *N. gonorrhoeae* during infection. *N. gonorrhoeae* isolates from different patients could also differ in less than 10 recombination filtered SNPs, indicating cases of transmission between patients. This showed that genetic variation alone cannot always distinguish between a persistent and a transmitted infection, thus additional patient reported metadata are needed. Knowledge on the extent of genetic variation that could occur within a patient or between patients may be applied in the public health setting, for instance when using genomics data to identify transmission networks.

ENHANCING UNDERSTANDING OF GONOCOCCAL POPULATION BIOLOGY

The bacterial pangenome consist of a core genome, containing genes essential for survival and present in all isolates of a bacterial species, and an accessory genome, consisting of genes that are present in only a subset of isolates as these provide flexibility to the bacterium for environment- or host adaptation. The gonococcal accessory genome contains plasmids associated with AMR, however, the chromosomally encoded gonococcal accessory genome and its role in AMR was hitherto undefined. In **chapter 7**, the chromosomally encoded gonococcal accessory genome was characterised by performing gene-by-gene comparative analyses on *N. gonorrhoeae* genomes publicly available in the PubMLST database. A small accessory genome was identified containing 247 genes that were mainly located on known large mobile genetic elements, such as plasmids. Prevalence of accessory genes was associated with certain core genome types, indicating that the gonococcal population structure is ordered with either predisposition for accessory elements in certain lineages or ancestral inheritance of accessory elements and its conservation during strain expansion. Associations were found between AMR and accessory genes encoding DNA exchange proteins, indicating that the exchange of genetic material plays an important role in the development or maintenance of AMR.

In **chapter 8**, an overview of the insights obtained in this thesis is given and recommendations for future research are provided. Based on the results in this thesis and on recent literature, it is expected that AMR in *N. gonorrhoeae* will continue to develop in the future. To slow AMR development down, the consumption of currently used antibiotics could be reduced and previously used antibiotics could be reintroduced for infections with susceptible strains. Future

research should focus on the development of molecular tests that enable susceptibility testing prior to treatment, and these tests should be continuously informed by AMR surveillance data. Research on gonorrhoea vaccines and on novel antibiotics for gonorrhoea treatment is still necessary, especially for pharyngeal gonorrhoea. Current antibiotic dose regimen might be sub-optimal for pharyngeal gonorrhoea treatment and subsequently facilitate AMR development at the oropharyngeal site. Therefore, future research should focus on optimisation of antibiotic dose regimen for pharyngeal gonorrhoea based on pharmacokinetic and pharmacodynamic parameters. Genomics methodologies can supplement AMR research by providing means for genomic AMR surveillance as well as to improve understanding of processes involved in AMR development and spread. As the choice of genomics methodology largely depends on the study objectives and on study parameters, transparency about the used software and thresholds is important in genomics research. Scientists should therefore always publish raw sequence data and bioinformatic pipelines in public databases, for which maintenance should be supported by dedicated funding.

SAMENVATTING

GENOOMANALYSES OM ANTIBIOTICARESISTENTIE IN *NEISSERIA GONORRHOEAE* TE KARAKTERISEREN EN TE VOLGEN

De bacterie *Neisseria gonorrhoeae*, de gonokok, veroorzaakt de seksueel overdraagbare aandoening (SOA) gonorroe, wereldwijd een van de meest voorkomende SOA's. Behandeling en beheersing van gonorroe wordt momenteel bedreigd door antibioticaresistentie (ABR). *N. gonorrhoeae* heeft namelijk resistentie ontwikkeld tegen alle antibiotica die voorheen gebruikt werden voor de behandeling van gonorroe, en ook resistentie tegen het huidige antibioticum ceftriaxon is de laatste jaren al sporadisch gedetecteerd. Om de verspreiding van antibioticaresistente *N. gonorrhoeae* te voorkomen, is het noodzakelijk om ABR te monitoren. Dit kan worden ondersteund met genoomanalyses, welke het mogelijk maken om genetische determinanten voor ABR te detecteren en te volgen. **Deel I** van dit proefschrift beschrijft genoomanalyses voor het monitoren van ABR binnen Nederland. Vergelijkbare genoomanalyses zijn gebruikt in **deel II**, met het doel om nieuwe inzichten te verkrijgen in de biologie van de gonokok.

In **hoofdstuk 1** worden verschillende aspecten van gonorroe geïntroduceerd en dit hoofdstuk bevat een samenvatting van de tot op heden gebruikte antibiotica voor gonorroe met de bijbehorende genetische resistentiemechanismen. Ook worden de voor dit proefschrift gebruikte genoomanalyses en bioinformatica-analyses geïntroduceerd. Moleculair typeren wordt beschreven als mogelijkheid om genetische diversiteit in bacteriepopulaties te bestuderen, waarbij gebruik gemaakt wordt van sequenties van bepaalde genen of genfragementen. Volledige genoomanalyses (*whole-genome sequencing*) hebben een aanzienlijk hogere resolutie en kunnen worden gebruikt voor het vergelijken van genomen op gen-niveau of op nucleotide-niveau.

DEEL I: GENOOMANALYSES OM ANTIBIOTICARESISTENTIE TE MONITOREN

EEN MOLECULAIRE TEST OM FLUOROQUINOLONEN-GEVOELIGHEID VAN *N. GONORRHOEAE* TE BEPALEN

Behandelrichtlijnen voor bacteriële infecties worden opgesteld op basis van ABR prevalentie. In het verleden werd gonorroe effectief behandeld met ciprofloxacine, behorend tot de antibiotica klasse fluoroquinolonen. Ciprofloxacine heeft voordelen ten opzichte van andere antibiotica, zoals orale toediening en relatief weinig bijwerkingen. Toen fluoroquinolonen-resistente *N. gonorrhoeae* stammen sterk opkwamen, werd ciprofloxacine niet langer aangeraden als

behandeling. Een aanzienlijk deel van de *N. gonorrhoeae* stammen is echter nog gevoelig voor ciprofloxacine en infecties met deze gevoelige stammen zouden nog behandeld kunnen worden met ciprofloxacine. Dit vereist wel dat de gevoeligheid wordt bepaald voorafgaand aan de behandeling, wat onmogelijk is met de huidige kweek-gebaseerde gevoeligheidsbepalingen die minimaal 2 dagen duren. In **hoofdstuk 2** is een moleculaire test (*polymerase chain reaction*; PCR) geëvalueerd voor gelijktijdige detectie van *N. gonorrhoeae* en de ciprofloxacine-resistentie veroorzakende mutatie *gyrA*-S91F. We stellen voor om deze PCR-test te implementeren in de routinediagnostiek voor fluoroquinolonen- gevoeligheidsbepaling van monsters waarin *N. gonorrhoeae* reeds is gedetecteerd met de huidige detectiemethode (*transcription-mediated amplification;* TMA), aangezien deze een hogere sensitiviteit heeft voor *N. gonorrhoeae* detectie. Naast dat de snellere fluoroquinolonen-gevoeligheidsbepaling het voorschrijven van ciprofloxacine weer mogelijk maakt, voorziet de PCR-test ook in fluoroquinolonen-gevoeligheidsbepalingen van niet-kweekbare monsters en daarmee ondersteunt deze test het monitoren van ABR.

CEFTRIAXON VERMINDERD-GEVOELIGE *N. GONORRHOEAE* STAMMEN IN AMSTERDAM

Het huidige antibioticum ceftriaxon is de laatste beschikbare behandelmogelijkheid voor gonorroe en daarom is het van groot belang om ceftriaxon-resistentie continu te monitoren. Routinematige ceftriaxon-gevoeligheidsbepalingen lieten een toename zien in het aantal *N. gonorrhoeae* isolaten met verminderde gevoeligheid voor ceftriaxon tussen 2014 en 2019 in Amsterdam. Het typeren van het gen dat codeert voor het aangrijpingspunt van ceftriaxon (*penA* gen) had uitgewezen dat de verminderde ceftriaxon-gevoeligheid in de eerste instantie werd veroorzaakt door een mozaïek *penA* gen (stukken van het *penA* gen zijn hierin verwisseld met het *penA* gen van een gerelateerde bacterie). Later veranderde het *penA* gen van ceftriaxon verminderd-gevoelige isolaten naar een niet-mozaïek gen met een A501 mutatie. In **hoofdstuk 3** is deze verschuiving in ABR mechanisme in meer detail onderzocht door middel van volledige genoomanalyses. Alle isolaten met verminderde gevoeligheid voor ceftriaxon verkregen tussen 2014 en 2019 zijn genetisch gekarakteriseerd, evenals een willekeurige selectie ceftriaxon-gevoelige isolaten van dezelfde tijdsperiode. De eerdere observatie van de verschuiving in het *penA* gen bleek een verschuiving in ceftriaxon verminderd-gevoelige *N. gonorrhoeae* stammen over de tijd: van het mozaïek *penA* gen bevattende type (MLST) 1901 naar MLST 7363 en recent naar MLST 7827, beide met een niet-mozaïek *penA* gen met een A501 mutatie. Ook mutaties in het *porB1b* gen kwamen veelvuldig voor onder ceftriaxon verminderd-gevoelige stammen. De ceftriaxon verminderd-gevoelige MLST 7827 stam met *penA* A501V en *porB1b* G120K/A121D mutaties kwam op in Amsterdam tussen 2017 en 2019. Ook in Noorwegen kwam deze stam op in deze tijdsperiode en de genomen van Noorse en Nederlandse isolaten van dit type waren vrijwel identiek, wat suggereerde dat deze stam circuleerde in Europa. Dit benadrukt het belang van het continu monitoren van ABR om verschuivingen in ABR mechanismen over de tijd te detecteren.

DE DYNAMIEK VAN DE GONOKOKKENPOPULATIE IN AMSTERDAM TIJDENS DE EERSTE COVID-19 LOCKDOWN

Een update van het monitoren van *N. gonorrhoeae* is gegeven in **hoofdstuk 4**, waarin de dynamiek van *N. gonorrhoeae* stammen wordt beschreven tijdens de eerste COVID-19 lockdown in Amsterdam in 2020. Een eerdere studie had laten zien dat bezoekers van het Centrum voor Seksuele Gezondheid (CSG) in Amsterdam tijdens de lockdown tijdelijk minder losse sekspartners rapporteerden. Deze gedragsverandering heeft mogelijk ook invloed gehad op de verspreiding van SOA's en daardoor mogelijk ook op populaties seksueel overdraagbare bacteriën. In dit hoofdstuk vergeleken we de genotypische en fenotypische karakteristieken van *N. gonorrhoeae* stammen van voor (tussen 15 januari en 29 februari 2020) en tijdens (15 mei en 30 juni 2020) de eerste lockdown. De samenstelling van de gonokokkenpopulatie veranderde: terwijl de MLST 8156 stam licht dominant was voor de lockdown, was de MLST 9362 stam dat overduidelijk tijdens de lockdown. Deze MLST 9362 stam droeg mutaties in *mtr* genen, welke geassocieerd zijn met azitromycine-resistentie. Mutaties die verband houden met verminderde gevoeligheid voor ceftriaxon werden niet gevonden in deze stam. De opkomst van deze stam resulteerde daarom in meer azitromycine-resistentie en een hogere ceftriaxon gevoeligheid tijdens de lockdown. De volledige genomen van de MLST 9362 isolaten waren vrijwel identiek, hetgeen erop wees dat deze stam klonaal is verspreid binnen Amsterdam tijdens de lockdown, mogelijk door lokale transmissie. Dit hoofdstuk laat zien dat publieke gezondheidsmaatregelen gevolgen kunnen hebben voor de dynamiek van infectieziekten zoals gonorroe, en dat hiermee rekening gehouden moet worden bij het monitoren ervan.

HET MONITOREN VAN ANTIBIOTICARESISTENTE *ESCHERICHIA COLI*

Vergelijkbare genoomanalyses zijn toegepast op het monitoren van *extended-spectrum* β-lactamases (ESBL) producerende *Escherichia coli* bacteriën, welke resistent zijn tegen *extended-spectrum* cefalosporines. ESBL-dragende *E. coli* (ESBL-Ec) is een ziekteverwekker die blaasontsteking of darmklachten kan veroorzaken, maar vaak in de darmen wordt gedragen zonder dat er symptomen optreden. Omdat ESBL genen kunnen worden uitgewisseld tussen verschillende bacteriesoorten en daardoor breed ABR kunnen veroorzaken (hoewel niet in de Neisseriaceae familie), is het uiterst relevant om de circulatie van deze genen te monitoren. Een eerdere Nederlandse studie had meer ESBL-Ec gevonden in een groep mannen die seks hebben met mannen (MSM) die deelnemen aan de Amsterdam Cohort Studies (16.3%) vergeleken met de algemene Nederlandse populatie (8.6%). In **hoofdstuk 5** zijn de genetische karakteristieken van de ESBL-Ec isolaten van MSM in kaart gebracht en is onderzocht of de verhoogde ESBL-Ec prevalentie mogelijk verklaard kon worden door transmissie tussen cohort deelnemers. Volledige genoomanalyses lieten zien dat de meeste ESBL-Ec isolaten behoorden tot de wereldwijd verspreide MLST 131 stam. Indien volledige genomen van ESBL-Ec isolaten vrijwel identiek waren, werd dit gedefinieerd als mogelijke transmissie. Onder andere waren de MLST 14 en MLST 394 stammen betrokken bij mogelijke transmissie en deze stammen zijn nauwelijks

gevonden in publieke data. Het feit dat wij deze stammen hebben gevonden in meerdere deelnemers en dat de isolaten genetisch vrijwel identiek waren, onderbouwt de mogelijkheid van ESBL-Ec transmissie tussen deelnemers. Dit betekent dat ESBL-Ec infectie overwogen zou moeten worden in het geval dat seksueel actieve MSM symptomen hebben die mogelijk verband hebben met een *E. coli* infectie.

DEEL II: GENOOMANALYSES OM DE BIOLOGIE VAN DE GONOKOK BETER TE BEGRIJPEN

HET VERLOOP VAN EEN *N. GONORRHOEAE* INFECTIE BETER BEGRIJPEN

Het in kaart brengen van de genetische variatie die optreedt in *N. gonorrhoeae* tijdens een infectie kan inzicht geven in de processen die een rol spelen in het verloop van de infectie. In **hoofdstuk 6** zijn de genomen vergeleken van *N. gonorrhoeae* isolaten die verkregen zijn van eenzelfde lichaamslocatie van een patiënt op meerdere tijdstippen (verkregen in het geval van behandelfalen tijdens een klinische studie), of van verschillende lichaamslocaties op één tijdstip. Van deze isolaten zijn gensequenties behorende tot veelgebruikte typeringsschema's vergeleken, evenals de volledige genomen op nucleotide-niveau, zowel met als zonder het maskeren van recombinatie (recombinatie is het uitwisselen van genetisch materiaal tussen bacteriën). Een groot aantal nucleotiden verschilden tussen isolaten met verschillende typeringsprofielen, terwijl isolaten met dezelfde typeringsprofielen nauwelijks verschilden op nucleotide-niveau. Isolaten van één patiënt waren genetisch vrijwel identiek, met minder dan 10 nucleotides verschil wanneer recombinatie werd gemaskeerd. De genetische variatie die optrad tijdens de infectie werd voornamelijk gevonden in genen die hypervariabel zijn onder immuundruk, wat aanduidde dat de humane immuunrespons de belangrijkste veroorzaker is van genetische variatie in *N. gonorrhoeae* tijdens een infectie. Isolaten van verschillende patiënten verschilden soms ook minder dan 10 nucleotides, wat mogelijk wijst op *N. gonorrhoeae* transmissie tussen patiënten. Dit liet zien dat alleen genetische variatie niet altijd onderscheid kan maken tussen een persisterende infectie en een overgedragen infectie, en dat hiervoor aanvullende patiëntgegevens nodig zijn. De kennis over de mate waarin *N. gonorrhoeae* isolaten van een patiënt of van meerdere patiënten genetisch kunnen verschillen, kan worden toegepast in de publieke gezondheidszorg, bijvoorbeeld bij het gebruik van genoomanalyses om transmissienetwerken te identificeren.

DE POPULATIEBIOLOGIE VAN DE GONOKOK BETER BEGRIJPEN

Het bacteriële genoom bevat genen die essentieel zijn om te overleven en daarom aanwezig zijn in alle isolaten van een bacteriesoort (*core genome*). Daarnaast zijn er ook genen die slechts in een deel van de isolaten aanwezig zijn omdat deze enkel benodigd zijn voor aanpassingen aan een bepaalde omgeving of gastheer (*accessory genome*). Zo draagt slechts een deel van de

gonokokken plasmiden die ABR kunnen veroorzaken. Naast de plasmiden, wat circulaire stukken DNA zijn die zich buiten het chromosoom bevinden, kunnen er in het chromosoom mogelijk ook genen zijn die slechts in een deel van de gonokokken aanwezig zijn en die mogelijk verband houden met ABR. In **hoofdstuk 7** zijn deze genen geïdentificeerd door publiekelijk beschikbare *N. gonorrhoeae* genomen in de online PubMLST database te vergelijken op gen-niveau. In totaal bleken 247 genen aanwezig te zijn in slechts een deel van de gonokokken. Deze genen waren vooral gelokaliseerd op de alreeds bekende mobiele genetische elementen, zoals de plasmiden. Opvallend was dat de aanwezigheid van deze genen veelal samenhing met de samenstelling en sequenties in de rest van het genoom. Dit liet zien dat de populatiestructuur van de gonokok ordelijk te verdelen is in groepen, welke al dan niet een karakteristiek hebben dat leidt tot aanwezigheid van deze genen. Het is ook mogelijk dat een stam met bepaalde genen verder is verspreid en dat de samenstelling van het genoom daarbij intact is gebleven. Aanwezigheid van bepaalde genen die coderen voor eiwitten die betrokken zijn bij DNA uitwisseling bleek samen te hangen met ABR, wat doet vermoeden dat de uitwisseling van genetisch materiaal een belangrijke rol speelt in de ontwikkeling of het behoud van ABR.

In **hoofdstuk 8** is een overzicht gegeven van de inzichten die verkregen zijn in dit proefschrift en worden aanbevelingen gedaan voor toekomstig onderzoek. Op basis van de resultaten in dit proefschrift en uit recente literatuur wordt verwacht dat ABR in *N. gonorrhoeae* zal blijven ontwikkelen in de toekomst. Om deze ontwikkeling af te remmen, zou het gebruik van huidige antibiotica verminderd kunnen worden en zou voormalige antibiotica weer voorgeschreven kunnen worden voor infecties met daarvoor gevoelige stammen. Toekomstig onderzoek moet zich daarom richten op de ontwikkeling van moleculaire testen die gevoeligheidsbepaling voorafgaand aan de behandeling mogelijk maakt en deze testen moeten continu worden afgestemd op ABR monitoringsdata. Onderzoek naar gonorroe vaccins en naar nieuwe antibiotica voor gonorroe behandeling is nog steeds noodzakelijk, in het bijzonder voor gonorroe in de keel. Huidige antibiotica doseringen zijn mogelijk suboptimaal voor de behandeling van gonorroe in de keel, wat ABR ontwikkeling in de hand werkt. Toekomstig onderzoek moet daarom gericht zijn op het optimaliseren van antibiotica dosering voor gonorroe in de keel, gebaseerd op farmacokinetische en farmacodynamische parameters. Onderzoek naar ABR kan worden aangevuld met genoomanalyses om genetische determinanten voor ABR te monitoren en om begrip over ABR ontwikkeling en verspreiding te vergroten. De keuze voor methodologie van de genoomanalyses hangt sterk af van de onderzoeksvraag en daarom is transparantie over de gebruikte parameters hierbij erg belangrijk. Wetenschappers zouden hun ruwe sequentiedata en hun bioinformatica analyseprotocollen daarom altijd beschikbaar moeten maken in publieke databases, waarvoor tevens meer financiering beschikbaar moet komen om ze goed te onderhouden.

ABOUT THE AUTHOR

Jolinda Elenbaas was born in Goes on 25 September 1995 and grew up in a town called 's-Gravenpolder. After graduating from secondary school (Calvijn College, Goes) in 2013, she moved to Katwijk aan Zee, and one year later to Leiden, for her studies. She graduated from Leiden University with a Bachelor in Bio-Pharmaceutical Sciences in 2016 and with a Master in Bio-Pharmaceutical Sciences in 2019 (cum laude). During her master's study, she completed a 9-month research internship at the Leiden Academic Centre of Drug Research (LACDR), where she studied the role of metabolites in pain after abdominal surgery. She learned to analyse large-scale metabolomics datasets and how to implement metabolomics data in a pharmacodynamic model. This sparked her interest in performing computational data analyses to answer biological research questions. After her internship, she continued working at the Biomedical Metabolomics Facility in Leiden (BMFL) as research technician for 3 months, during which she further improved her laboratory skills. Afterwards, she completed another 6-month internship at the organ-on-a-chip company MIMETAS, where she developed a 3D cell platform to study vascular growth towards tumour tissue.

During her last internship she realised that her interest lay more in computational experimental work (dry-lab) than in laboratory experimental work (wet-lab). She therefore decided to start her PhD at the Public Health Laboratory of Amsterdam, where she used genomics analyses to study the epidemiology and antimicrobial resistance mechanisms of the sexually transmittable bacterium *Neisseria gonorrhoeae*, under the supervision of Prof. dr. Henry de Vries, Dr. Alje van Dam and Dr. Sylvia Bruisten. In 2020, a collaboration was initiated with Dr. Odile Harrison from the Maiden Lab at Oxford University, and Jolinda got to spend 3 weeks at the Maiden Lab in 2022. She worked there on a database for genomic data of the sexually transmittable bacterium *Mycoplasma genitalium*.

During the first period of the COVID-19 pandemic, which occurred during Jolinda's PhD trajectory, she temporarily interrupted her research to support the Public Health Laboratory with the SARS-CoV-2 diagnostics. She resumed her research 6 months later. The results of her PhD studies are presented in this thesis.

Jolinda is married to Frank de Korne and together they live in Leiden. They love to travel and can often be found kitesurfing on the North Sea.

LIST OF PUBLICATIONS

PART OF THIS THESIS

De Korne-Elenbaas J, Pol A, Vet J, Dierdorp M, van Dam AP, Bruisten SM. Simultaneous detection of *Neisseria gonorrhoeae* and fluoroquinolone resistance mutations to enable rapid prescription of oral antibiotics. **Sex Transm Dis. 2020,** 47(4):238-242.

Author contributions: SMB, AP, JV and APD conceptualised the study. AP and JV designed and provided the diagnostic kits. MD and JKE performed the laboratory work and initial analyses. JKE conducted the final data analyses and drafted the manuscript. SMB, AP, JV, MD and APD contributed to data interpretation and project discussions and revised the manuscript. All authors read and approved the final manuscript.

De Korne-Elenbaas J, Bruisten SM, de Vries HJC, van Dam AP. Emergence of a cephalosporin reduced susceptible *Neisseria gonorrhoeae* clone between 2014-2019 in Amsterdam, the Netherlands, revealed by a genomic population analysis. **J Antimicrob Chemother. 2021,** 18;76(7):1759-1768.

Author contributions: APD, SMB and JKE designed the study. JKE prepared the samples for whole-genome sequencing, conducted the analyses and drafted the manuscript. APD, SMB and HJCV contributed to the data interpretation and project discussions, and revised the manuscript. All authors read and approved the final manuscript.

Zondag HCA*, de Korne-Elenbaas J*, Bruisten SM, de Vries HJC, van Dam AP. Increased clonality among *Neisseria gonorrhoeae* isolates during the COVID-19 pandemic in Amsterdam, the Netherlands. **Submitted for publication.**
*contributed equally

Author contributions: APD, HCAZ and JKE designed the study. HCAZ and JKE conducted the analyses and drafted the manuscript. APD, SMB and HJCV contributed to the data interpretation and project discussions, and revised the manuscript. All authors read and approved the final manuscript.

De Korne-Elenbaas J*, van der Putten BCL*, Boek NDM, Matser A, Schultsz C, Bruisten SM, van Dam AP. Putative transmission of extended-spectrum β-lactamase-producing *Escherichia coli* among men who have sex with men in Amsterdam, the Netherlands. **Submitted for publication.**
*contributed equally

Author contributions: APD and JKE conceived the study and designed it together with BCLP. The samples were obtained during a previous study coordinated by AM. JKE prepared the samples for whole-genome sequencing. NDMB performed the initial analyses under supervision of JKE and BCLP. JKE and BCLP performed the final analyses and drafted the manuscript. APD, SMB and CS contributed to the data interpretation and project discussions, and revised the manuscript. All authors read and approved the final manuscript.

De Korne-Elenbaas J, Bruisten SM, de Vries HJC, van Dam AP. Within-host genetic variation in *Neisseria gonorrhoeae* over the course of infection. **Microbiology Spectrum. 2022,** 10(3):e0031322.

Author contributions: APD and JKE conceptualised the study. JKE conducted the analyses and drafted the manuscript. APD, SMB and HJCV contributed to the data interpretation and project discussions, and revised the manuscript. All authors read and approved the final manuscript.

De Korne-Elenbaas J, Bruisten SM, van Dam AP, Maiden MCJ, Harrison OB. The *Neisseria gonorrhoeae* accessory genome and its association with the core genome and antimicrobial resistance. **Microbiology Spectrum. 2022,** 10(3):e0265421.

Author contributions: OBH and JKE designed the study, conducted the analyses, and wrote the manuscript; SMB, APD, and MCJM contributed to project discussions and revised the manuscript. All authors read and approved the final manuscript.

OTHER

Osnes MN, Didelot X, de Korne-Elenbaas J, Alfsnes K, Brynildsrud OB, Syversen G, Nilsen ØJ, Freiesleben De Blasio B, Caugant DA, Eldholm V. Sudden emergence of a *Neisseria gonorrhoeae* clade with reduced susceptibility to extended-spectrum cephalosporins, Norway. **Microbial Genomics. 2020**, 6(12):mgen000480.

Van der Putten BCL, Mendes CI, Talbot BM, de Korne-Elenbaas J, Mamede R, Vila-Cerqueira P, Coelho LP, Gulvik CA, Katz L. Software testing in microbial bioinformatics: a call to action. **Microbial genomics. 2021,** 8(3):000790.

De Vries HJC, de Laat M, Jongen VW, Heijman T, Wind CM, Boyd A, de Korne-Elenbaas J, van Dam AP, Schim van der Loeff M. Efficacy of ertapenem, gentamicin, fosfomycin, and ceftriaxone for the treatment of anogenital gonorrhoea (NABOGO): a randomised, non-inferiority trial. **Lancet Infectious Diseases. 2022,** 22(5):706-717.

LIST OF CONTRIBUTING AUTHORS

Name	Affiliation(s)
Dr. A.P. van Dam	Public Health Laboratory, Department of Infectious Diseases, Public Health Service of Amsterdam, the Netherlands; Amsterdam UMC, University of Amsterdam, Department of Medical Microbiology, Amsterdam Institute for Infection & Immunity, Amsterdam, The Netherlands
Dr. S.M. Bruisten	Public Health Laboratory, Department of Infectious Diseases, Public Health Service of Amsterdam, the Netherlands; Amsterdam UMC, University of Amsterdam, Amsterdam Institute for Infection & Immunity, Amsterdam, The Netherlands
Prof. dr. H.J.C. de Vries	Centre for Sexual Health, Department of Infectious Diseases, Public Health Service Amsterdam, The Netherlands; Amsterdam UMC, University of Amsterdam, Department of Dermatology, Amsterdam Institute for Infection & Immunity, Amsterdam, The Netherlands
Dr. A. Pol	NYtor BV, Nijmegen, the Netherlands
Dr. J. Vet	NYtor BV, Nijmegen, the Netherlands
M. Dierdorp, BSc	Public Health Laboratory, Department of Infectious Diseases, Public Health Service of Amsterdam, the Netherlands
H.C.A. Zondag, MSc	Public Health Laboratory, Department of Infectious Diseases, Public Health Service of Amsterdam, the Netherlands; Amsterdam UMC, University of Amsterdam, Department of Medical Microbiology, Amsterdam Institute for Infection & Immunity, Amsterdam, The Netherlands
Dr. B.C.L. van der Putten	Department of Global Health, Amsterdam Institute for Global Health and Development, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; Amsterdam UMC, University of Amsterdam, Department of Medical Microbiology, Amsterdam Institute for Infection and Immunity, Amsterdam, the Netherlands

N.D.M. Boek, MSc	Amsterdam UMC, University of Amsterdam, Department of Medical Microbiology, Amsterdam Institute for Infection and Immunity, Amsterdam, the Netherlands
Dr. A. Matser	Department of Infectious Diseases, Public Health Service of Amsterdam, the Netherlands
Prof. dr. C. Schultsz	Department of Global Health, Amsterdam Institute for Global Health and Development, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; Amsterdam UMC, University of Amsterdam, Department of Medical Microbiology, Amsterdam Institute for Infection and Immunity, Amsterdam, the Netherlands
Prof. dr. M.C.J. Maiden	Department of Biology, Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom
Dr. O.B. Harrison	Department of Biology, Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom

PHD PORTFOLIO

PhD TRAINING

YEAR		WORKLOAD (ECTS)
	GENERAL COURSES	
2019	Infectious Diseases, AMC Graduate School	1.3
2019	e-BROK: Basiscursus Regelgeving en Organisatie voor Klinische Onderzoekers, AMC Graduate School	1.5
2020	Research Data Management, AMC Graduate School	0.9
2020	Genetic Epidemiology, AMC Graduate School	1.1
2022	Talent Ontwikkelings Programma, De Baak, Driebergen	1.1
	SPECIFIC COURSES	
2019	Bioinformatics Sequence Analysis, AMC Graduate School	1.1
2019	Clinical bioinformatics for microbial genomics and metagenomics, European Society of Clinical Microbiology and Infectious Diseases, Lausanne, Switzerland	1.5
2021 - 2022	Data Analyst with Python (virtual), DataCamp	1.1
2019 - 2023	Weekly PhD educational hour (Journal Club, Peer education, Epidemiology education), Department of Infectious Diseases, Public Health Service of Amsterdam	14
	SEMINARS, WORKSHOPS AND MASTER CLASSES	
2019	GGD Onderzoeksdag 2019	0.2
2019	SOA & HIV expert meeting, RIVM, Bilthoven	0.2
2019	Special Interest Group for Bioinformatics meeting, Nederlandse Vereniging voor Medische Microbiologie (NVMM), Utrecht	0.2
2020	Seminar Genomic Epidemiology of Infectious Diseases, Koninklijke Nederlandse Vereniging voor Microbiologie (KNVM), virtual	0.2
2021	GGD Onderzoeksdag 2021, virtual	0.2
2021	Workshop Presenteren, Association of Amsterdam UMC PhD Candidates (ASAP), virtual	0.1
2021	Research Meeting 'Genomic population structures of microbial pathogens', Royal Society London, virtual	0.2
2021	ASM-NGS pre-conference hackathon 'Software testing in microbial bioinformatics', virtual	0.4

YEAR		WORKLOAD (ECTS)
2019 - 2023	Monthly seminar of the Public Health Laboratory, Public Health Service of Amsterdam	1
2021 - 2023	Weekly seminar of the Laboratory of Experimental Biology, Amsterdam UMC	1
	PRESENTATIONS	
2019	Simultaneous detection of *Neisseria gonorrhoeae* and fluoroquinolone resistance mutations to enable rapid prescription of oral antibiotics (oral), IUSTI-Europe Congress, Tallinn, Estonia	0.5
2020	Emergence of *Neisseria gonorrhoeae* cluster (MLST7827) with reduced susceptibility to ceftriaxone in Amsterdam, The Netherlands (oral), IUSTI-Europe Congress, virtual	0.5
2021	Emergence of a cephalosporin reduced susceptible *Neisseria gonorrhoeae* clone between 2014-2019 in Amsterdam (poster), KNVM & NVMM Scientific Spring Meeting, virtual	0.5
2021	The *Neisseria gonorrhoeae* accessory genome and its association with the core genome (oral), PubMLST forum on STDs, virtual	0.5
2021	Assessing within-host genetic variation in *Neisseria gonorrhoeae* at different anatomical locations and over time, STI & HIV World Congress 2021, virtual	0.5
2022	Within-host genetic variation in *Neisseria gonorrhoeae* over the course of infection (oral), NgoRS Conference, virtual	0.5
2022	Within-host genetic variation in *Neisseria gonorrhoeae* over the course of infection (oral), Department of Dermatology of Amsterdam UMC, virtual	0.5
2022	The *Neisseria gonorrhoeae* accessory genome and its association with the core genome and antimicrobial resistance (oral), ECCMID 2022, Lisbon, Portugal	0.5
2022	Within-host genetic variation in *Neisseria gonorrhoeae* over the course of infection (oral), ECCMID 2022, Lisbon, Portugal	0.5
2022	Moleculaire epidemiologie van *Neisseria gonorrhoeae* tijdens de eerste lockdown in Amsterdam (oral), SOA & HIV Expert Meeting 2022, RIVM, Bilthoven	0.5

YEAR		WORKLOAD (ECTS)
2022	Tracking *Mycoplasma genitalium* on a molecular levelin the Netherlands and globally, symposium 'Time to test and treat Mycoplasma genitalium infection?' (oral), Public Health Laboratory of Amsterdam	0.5
2022	Increased clonality among *Neisseria gonorrhoeae* isolates during the COVID-19 pandemic in Amsterdam (poster), IMMEM XIII, Bath, United Kingdom	0.5
2022	Emergence of a cephalosporin reduced susceptible *Neisseria gonorrhoeae* clone between 2014-2019 in Amsterdam (oral), IPNC 2022, Cape Town, South Africa	0.5
2022	Increased clonality among *Neisseria gonorrhoeae* isolates during the COVID-19 pandemic in Amsterdam (poster), IPNC 2022, Cape Town, South Africa	0.5
	(INTER)NATIONAL CONFERENCES	
2019	European Congress of Microbiology and Infectious Diseases (ECCMID), Amsterdam, The Netherlands	1.1
2019	International Union against STIs (IUSTI)-Europe, Tallinn, Estonia	1.1
2020	IUSTI-Europe, virtual	0.5
2020	ASM Conference on Rapid Applied Microbial Next-Generation Sequencing and Bioinformatic Pipelines (ASM-NGS), virtual	0.5
2021	KNVM & NVMM Scientific Spring Meeting, virtual	0.6
2021	STI & HIV World Congress, virtual	0.5
2022	*Neisseria gonorrhoeae* Research Society (NgoRS) Conference, virtual	0.9
2022	ECCMID, Lisbon, Portugal	1.1
2022	13th International Meeting on Microbial Epidemiological Markers (IMMEM XIII), Bath, United Kingdom	1.1
2022	International Pathogenic *Neisseria* Conference (IPNC), Cape Town, South Africa	1.1

TEACHING

YEAR		WORKLOAD (ECTS)
	LECTURING	
2021	Guest lecture on STIs, Medical Microbiology Course, BSc Biomedical Sciences, Amsterdam UMC	0.5
	TUTORING, MENTORING	
2020 & 2021	Assisting Medical Microbiology Computer Practicals in Galaxy	1.1
	SUPERVISING	
2021	Internship Niels Boek, MSc student Biomedical Sciences at University of Amsterdam	2
2022	Internship Serife Öcal, BSc student Gezondheid & Leven at Vrije Universiteit Amsterdam	2

PARAMETERS OF ESTEEM

GRANTS

2020 Short-term Fellowship from the European Molecular Biology Organization (EMBO) for the research visit to the Maiden Lab at Oxford University

PUBLICATIONS

PEER REVIEWED

2020 **De Korne-Elenbaas J**, Pol A, Vet J, Dierdorp M, van Dam AP, Bruisten SM. (2020). Simultaneous detection of *Neisseria gonorrhoeae* and fluoroquinolone resistance mutations to enable rapid prescription of oral antibiotics. *Sexually Transmitted Diseases*, 47(4):238-242. doi: 10.1097/OLQ.0000000000001141

2020 Osnes MN, Didelot X, **de Korne-Elenbaas J**, Alfsnes K, Brynildsrud OB, Syversen G, Nilsen ØJ, De Blasio BF, Caugant DA, Eldholm V. (2020). Sudden emergence of a *Neisseria gonorrhoeae* clade with reduced susceptibility to extended-spectrum cephalosporins, Norway. *Microbial Genomics,* 6(12):mgen000480. doi: 10.1099/mgen.0.000480.

2021 **De Korne-Elenbaas J**, Bruisten SM, de Vries HJC, Van Dam AP. (2021). Emergence of a *Neisseria gonorrhoeae* clone with reduced cephalosporin susceptibility between 2014 and 2019 in Amsterdam, The Netherlands, revealed by genomic population analysis. *Journal of Antimicrobial Chemotherapy,* 18;76(7):1759-1768. doi: 10.1093/jac/dkab082

2021 Van der Putten BCL, Mendes CI, Talbot BM, **de Korne-Elenbaas J,** Mamede R, Vila-Cerqueira P, Coelho LP, Gulvik CA, Katz L. (2021). Software testing in microbial bioinformatics: a call to action. *Microbial Genomics,* 8(3):000790. doi: 10.1099/mgen.0.000790

2022 De Vries HCJ, de Laat M, Jongen VW, Heijman T, Wind CM, Boyd A, **de Korne-Elenbaas J**, van Dam AP, Schim van der Loeff M. (2022). Efficacy of ertapenem, gentamicin, fosfomycin, and ceftriaxone for the treatment of anogenital gonorrhoea (NABOGO): a randomised, non-inferiority trial. *Lancet Infectious Diseases*, 22(5):706-717. doi: 10.1016/S1473-3099(21)00625-3

2022 **De Korne-Elenbaas J**, Bruisten SM, de Vries HJC, van Dam AP. (2022) Within-host genetic variation in *Neisseria gonorrhoeae* over the course of infection. *Microbiology Spectrum*, 10(3):e0031322. doi: 10.1128/spectrum.00313-22

2022 **De Korne-Elenbaas J**, Bruisten SM, van Dam AP, Maiden MCJ, Harrison OB. (2022). The *Neisseria gonorrhoeae* accessory genome and its association with the core genome and antimicrobial resistance. *Microbiology Spectrum,* 10(3):e0265421. doi: 10.1128/spectrum.02654-21

OTHER

2022 **De Korne-Elenbaas J*,** van der Putten BCL*, Boek NDM, Matser A, Schultsz C, Bruisten SM, van Dam AP. (2022). Putative transmission of Extended-Spectrum β-Lactamase-producing *Escherichia coli* among men who have sex with men in Amsterdam, the Netherlands. *Submitted.*

2022 Zondag HCA*, **de Korne-Elenbaas J*,** de Vries HJC, Bruisten SM, van Dam AP. (2022). Increased clonality among *Neisseria gonorrhoeae* isolates during the COVID-19 pandemic in Amsterdam, the Netherlands. *Submitted.*

*equal contribution

DANKWOORD

God dank ligt dit proefschrift er! Zonder kracht van boven was het me nooit gelukt. Ik ben ontzettend dankbaar voor alles wat ik afgelopen jaren heb kunnen ontdekken, beleven en leren op het gebied van de wetenschap, maar zeker ook op persoonlijk vlak. Hier zijn heel veel mensen onmisbaar voor geweest.

Allereerst wil ik mijn promotor Henry bedanken. De gesprekken met jou voelde altijd als aanmoediging. Je gaf ruimte voor mijn eigen interesses, maar toch hield je altijd het grotere plaatje voor ogen. In plaats van de vraag 'Hoe gaat het?', vroeg jij bij een van mijn voortgangsgesprekken: 'Slaap je goed?'. Voor mij het toonbeeld van jouw klinische blik, waar ik veel van heb geleerd. Dank ook voor het vertrouwen dat je me altijd gaf, mede daardoor leerde ik ook meer op mezelf vertrouwen.

Ook heel veel dank aan mijn copromotoren. Sylvia, jij had altijd tijd voor me en leefde helemaal met me mee de afgelopen jaren. Tegelijk gaf je me vrijheid als je merkte dat het me wel lukte. Jij blinkt uit in het halen van deadlines, sterker nog, jouw feedback ontving ik vrijwel altijd ruim voor de deadline. Voor vragen kon ik altijd bij je binnenlopen en meestal wist je de antwoorden precies te vinden tussen de stapels geprinte emails, zelfs tussen die van jaren geleden! Dat we allebei van aanpakken houden, maakte onze samenwerking fijn. Ik ben je dankbaar voor alle kansen die je me hebt gegeven om mezelf te ontwikkelen en voor jouw aanmoediging daarbij.

Alje, wat heb ik ontzettend veel van jou geleerd. Vaak hanteerde ik de regel 'Als Alje het zegt, dan klopt het'. Jouw kennis over microbiologie is onuitputtelijk, waardoor brainstormen met jou me altijd weer inspireerde. Ik ben blij dat ik jou toch ook wat kon leren over sequencen en bio-informatica. Ik waardeer enorm hoe je altijd tijd vrij wist te maken om te vragen hoe het met me ging, zelfs toen je alleen nog op donderdag voor de GGD werkte. Heel veel dank voor je eerlijkheid, betrokkenheid, realisme en vertrouwen. Ik ben blij dat we toch nog samen naar de IPNC in Kaapstad konden!

Dear members of my PhD committee, thank you for joining the committee and for reading my thesis. I look forward to the discussions with you.

Onnoemelijk veel steun heb ik tijdens mijn promotietraject gekregen van mijn lieve mede lab-PhD'ers en kamergenoten Joyce en Hélène. Fantastisch dat jullie mijn paranimfen zijn! We liepen tijdens onze promotietrajecten veel samen op, en dat maakt het des te bijzonder dat jullie tijdens mijn verdediging daadwerkelijk naast mij staan.

Joyce, wij hadden beide onze eerste werkdag tijdens de GGD Onderzoeksdag in 2019 en 'groeiden samen op' bij de GGD. Dank voor alle gezellige uren kletsen, over PhD issues, maar nog veel meer over baksels, reizen, trouwen, duurzaamheid en sporten. Dank dat ik altijd m'n hart mocht luchten en dat je altijd bereid was om met me mee te denken. Dank ook dat je in Istanbul getrouwd bent, zodat ik daar met een goede reden op vakantie kon ;). Ik bewonder jouw doelgerichte houding en je doorzettingsvermogen, welke ervoor zorgden dat je netjes op tijd je promotie afrondde en dat je een leuke nieuwe baan hebt gevonden. Dat ik jouw paranimf mocht zijn, en jij nu de mijne bent, betekent veel voor me.

Hélène, wat hebben wij veel leuke momenten samen beleefd en wat hebben we ontzettend veel gelachen! Dank voor de vele uren gezelligheid op kantoor en daarbuiten, voor je onuitputtelijke humor, voor alle keren dat ik m'n hart bij je mocht luchten en voor je medeleven op die momenten. Dank ook voor alle keren dat je me hebt opgevrolijkt met een toepasselijk poezen-gifje of een foto van Odin. Jij hebt altijd oog voor de mensen om je heen en tegelijk ben je ook heel toegewijd aan werk, mede daardoor heb ik met veel plezier aan ons DINGL project gewerkt. Je bent echt een enorme steun voor me geweest in mijn hele promotietraject en ik had me geen betere PhD-collega kunnen wensen.

Heel veel dank aan al mijn collega's van het Streeklab. Ik heb veel bewondering voor de manier waarop jullie het lab, dag in dag uit, draaiende houden. Een aantal mensen wil ik in het bijzonder bedanken. Michelle, voor jou is geen vraag te gek. Jouw behulpzaamheid is onuitputtelijk en bewonderingswaardig. Heel erg bedankt voor al je hulp, en ook voor de gezelligheid en de vele wijntjes (zelfs uit papieren bekertjes op de 7e), tijdens welke we hele beschouwingen konden houden over het lab, het leven of katten. Sahare, de warmte die jij geeft aan mensen om je heen is echt uniek. Dank voor je goede zorg en voor alle knuffels. Dank ook voor je hulp op het lab, in het bijzonder bij de stage van Serife. Nadia, dankzij jou heb ik kennis gemaakt met de Iraanse cultuur, en ik ben vooral blij dat ik nu ook weet dat er pittige augurken bestaan. Dominique, dank dat ik samen met jou onbeschaamd mee kon zingen met foute Nederlandse hits. Jeffrey, als iemand oog heeft voor collega's, ben jij het. Jij wist altijd de weg te vinden naar de 6e verdieping om even te komen kletsen. Dank voor je oprechte interesse en voor de leuke feestjes bij jou thuis (met meer dan genoeg GT!). Dank dat je me gepassioneerd meer leerde over Amsterdam. Hopelijk heb ik jou ook nog wat religieuze kennis kunnen bijbrengen ;). Mirjam, dank voor je trouwe hulp bij het NYtor project en voor alle antwoorden op mijn vragen over projecten die mijn voorgangers hadden gedaan. Dankzij jouw hulp had ik een vliegende start! Martine, René, Margreet, Brenda, Akke, Monique en Nikki, ook jullie maakten kantoordagen een stuk gezelliger, dank jullie wel!

Ik wil ook mijn collega's van de bacteriologie enorm bedanken voor hun hulp bij het kweken van alle GO's, in het bijzonder Ineke, Kaoutar, Sandra, Neeltje, Jeffrey, Ursa, Fallon en Kawtar. Dankzij jullie ging alles met een sneltreinvaart! Ursa, aan jou kon ik de moeilijkste experimenten met

een gerust hart overlaten. Ook al hebben de resultaten de publicatie niet gehaald, het was zeker niet voor niets. Heel veel dank voor je hulp hierbij!

Een van de hoogtepunten van afgelopen jaren was de tijd dat ik hielp met de logistiek van de COVID-19 testen. Ik kon zo van heel dichtbij meemaken wat er gebeurt bij een crisis. Dank aan alle collega's met wie ik zo fijn heb samengewerkt in die periode. Het zorgde ervoor dat ik veel mensen beter heb leren kennen en dat ik me nog veel meer een onderdeel van het lab ben gaan voelen. In het bijzonder, Douwe, dank voor onze samenwerking. We hebben samen toen een heleboel gefixt! Ook voor jouw hulp bij alle administratieve zaken rondom mijn onderzoek wil ik je enorm bedanken. Mariken, dank dat je me de mogelijkheid gaf om mijn steentje bij te dragen. Dank dat je mijn brede interesse opmerkt en voor de kansen die je me daarom bood.

Dank aan de GLIMS collega's, John, Carla, Frans, Nadia en Michelle H, die altijd voor me klaarstonden als ik labgegevens nodig had. Dank aan de datamanagers, Michelle K en Steffen, zonder wie mijn onderzoek niet mogelijk was geweest. Caspar, dank voor jouw hulp bij vragen waar niemand anders antwoord op wist.

Ook dank aan mijn collega's van Team Onderzoek. Maarten, van jou leerde ik dat de wereld te vangen is in een 2x2 tabel. Dank voor alle leerzame onderwijsuren en dat je altijd klaarstond om mijn epidemiologische vragen te beantwoorden. Dank aan alle PhD-collega's, jullie zijn belangrijk geweest tijdens mijn promotietraject doordat we van elkaar konden leren, PhD issues met elkaar konden bespreken en gezellige dingen konden doen. Het uitje naar het ping-pong cafe zal ik nooit vergeten (vooral omdat ik toen volledig ingemaakt ben door Jeffrey).

Ik heb ook ontzettend veel gehad aan de samenwerking met collega's van het AMC. Constance, dank dat je me 'adopteerde' in jouw onderzoeksgroep. Het heeft me een schat aan kennis en contacten opgeleverd. Boas, jou kan ik niet genoeg bedanken. Vanaf onze eerste meeting deelde jij openlijk je kennis met me en zorgde je ervoor dat ik vliegend van start kon gaan. Jij was mijn bioinformatica hulplijn, waarvan ik altijd antwoord kreeg. Ik heb jouw onuitputtelijke kennis altijd bewonderd, evenals je bereidheid om kennis, scripts en data te delen. Mede dankzij jou heb ik nu een heel netwerk van bioinformatici om me heen en ik weet zeker dat ik daar in de toekomst nog veel aan ga hebben. Fijn dat één daarvan nu bij het RIVM zit ;). Ik hoop dat we nog vaak samen koffie kunnen drinken en de wondere wereld van de bioinformatica kunnen beschouwen. Kees en Thomas, dank dat jullie met me meedachten over een mogelijk RNA sequencing project. Ik heb veel aan jullie kennis en ervaring gehad. Rik, dank voor de brainstormsessies die we hadden over *Neisseria gonorrhoeae*, dat gaf mij ook weer nieuwe inzichten. Many thanks also to the other (ex-) colleagues of the Schultsz lab, who always made me feel welcome during the Monday meetings.

A special thanks to Odile, who taught me so much about *Neisseria* genomics. I really enjoyed our collaboration and all the Teams meetings we had. Many thanks that I could discuss everything with you about our project, but also about a career in science and our lives. I am very honoured that you are physically present during my PhD defence as an opponent! Also thanks to the other colleagues from the Maiden Lab in Oxford, I am very happy that I got to spend some time with you.

Ik heb ook veel te danken aan mijn lieve vrienden en vriendinnen. Lieve Lis en Mies, de dates met jullie zorgden er altijd voor dat ik mijn werk kon relativeren. Met jullie kan ik lachen, dansen, slapen, klagen, eten, huilen en wijnen. Lis, dank dat jij altijd laat zien hoe je van het leven moet genieten. Mies, dank voor de zetjes die jij kan geven om dingen gewoon te doen. Peet, ook al zien we elkaar niet vaak, we blijven zusjes for life! Dank dat je altijd interesse had in mijn onderzoek en dat je zelfs mijn artikel hebt proberen te lezen. Tir, dank dat je altijd luistert en meedenkt. Jouw doorzettingsvermogen inspireert mij ook. San en Aad, wat ben ik blij dat jullie nog steeds in Leiden wonen. Dank voor de heerlijke wijnen en maaltijden die problemen als sneeuw voor de zon laten verdwijnen. Toon en Jant, dank dat alles open op tafel mag liggen bij jullie. Jullie wijze raad heeft me vaak verder geholpen. Jant, heel veel dank voor je creativiteit waarmee je dit proefschrift zo prachtig hebt gemaakt. Robin, dank voor alle keren slap ouwehoeren. Hoe jij ambitie en humor combineert is een inspiratie voor mij. Lieve mosterdzaadjes, onmisbaar waren de afgelopen jaren ook jullie bijdragen, in de vorm van heerlijke maaltijden, prachtige taartjes, goede gesprekken, nachtjes weg, muziek, snowboarden, surfen en cocktails.

Lieve schoonfamilie, dank voor jullie warmte en dat jullie altijd begaan waren met de voortgang van m'n onderzoek. Lieve oma, wat bijzonder dat je bij mijn promotie bent. Dank voor je medeleven waarmee ik opgroeide, mede daardoor kon ik worden wie ik nu ben. Lieve schoon(zussen), broers en zwagers, dank voor de nodige (Zeeuwse) nuchterheid die jullie in mijn leven brengen. Patrick en Menno, jullie als grote broers leerden me hoe ik m'n mannetje moest staan. Elselien, jij liet me zien dat je niet altijd de gebaande wegen hoeft te gaan. Ilona, van jou leerde ik dat dingen loslaten heel bevrijdend kan zijn. Pa en ma, jullie laten me mijn eigen pad volgen en geven tegelijkertijd onvoorwaardelijke steun, wat van onbeschrijfelijke waarde voor me is. Pa, dank dat jij werkelijk altijd klaar staat om te helpen, of het nu een fietsreparatie is of een verhuizing. Met jou kan ik problemen ontleden tot op het bot en daar hebben we dan nog plezier in ook. Mams, doordat jij me het beste aanvoelt, kun je me ook de beste adviezen geven. Dank dat 'hotel mama' 24/7 geopend is en dat het daar nog altijd voelt als thuis.

Frank, wat ben ik blij met jou in mijn leven. Jouw steun heeft me gebracht waar ik nu ben. Jij remt me af als ik te hard van stapel loop en moedigt me aan als ik even niet meer kan. Dank voor je optimisme en dat voor jou niets te gek is. Dank ook voor je onvoorwaardelijke liefde, wat mij zoveel sterker maakt. Ik ben klaar voor een nieuw avontuur met jou!

Is dit nu later als je groot bent?

Stef Bos