A community-driven resource for genomic surveillance of *Neisseria gonorrhoeae* at Pathogenwatch

Leonor Sánchez-Busó[1^,*], Corin A. Yeats[1], Ben Taylor[1,2], Richard Goater[2], Anthony Underwood[2], Khalil Abudahab[2], Silvia Argimón[2], Kevin C. Ma[3], Tatum D. Mortimer[3], Michelle J. Cole[4,^], Yonatan H. Grad[3,5,^], Irene Martin[6,^], Brian H. Raphael[7,^], William M. Shafer[8,9,^], Gianfranco Spiteri[10,^], Katy Town[7,^], Teodora Wi[11,^], Simon R. Harris[12], Magnus Unemo[13,^] and David M. Aanensen[1,2,^,*], representing the *N. gonorrhoeae* Pathogenwatch Scientific Steering Group.

[1] Centre for Genomic Pathogen Surveillance, Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, Oxfordshire, United Kingdom.
[2] Centre for Genomic Pathogen Surveillance, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom.
[3] Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, United States of America.
[4] National Infection Service, Public Health England, London, United Kingdom.
[5] Division of Infectious Diseases, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America.
[6] National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada.
[7] Division of STD prevention, National Center for HIV/AIDS, Viral Hepatitis, STD and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America.
[8] Department of Microbiology and Immunology and Emory Antibiotic Resistance Center, Emory University School of Medicine, Atlanta, Georgia, United States of America.
[9] Laboratories of Bacterial Pathogenesis, Veterans Affairs Medical Center, Decatur, Georgia, United States of America.
[10] European Centre for Disease Prevention and Control, Stockholm, Sweden.
[11] Department of the Global HIV, Hepatitis and STI programmes, World Health Organization, Geneva, Switzerland.
[12] Microbiotica, Biodata Innovation Centre, Hinxton, Cambridgeshire, United Kingdom.
[13] World Health Organization Collaborating Centre for Gonorrhoea and Other STIs, Department of Laboratory Medicine, Faculty of Medicine and Health, Örebro University, Örebro, Sweden.
^Current members of the *N. gonorrhoeae* Pathogenwatch Scientific Steering Group.

*Corresponding authors: Leonor Sánchez-Busó (leo.sanchez-buso@cgps.group) and David M. Aanensen (david.aanensen@cgps.group).

Keywords: *Neisseria gonorrhoeae*, Pathogenwatch, public health, genomics, epidemiology, surveillance, antimicrobial resistance.

Abstract

Background: Antimicrobial resistant (AMR) *Neisseria gonorrhoeae* is an urgent threat to public health, as strains resistant to at least one of the two last line antibiotics used in empiric therapy of gonorrhoea, ceftriaxone and azithromycin, have spread internationally. With new treatment options not yet available, this has prompted a call for collaborative action on global surveillance for this sexually transmitted pathogen. Whole genome sequencing (WGS) data can be used to identify new AMR clones, outbreaks, transmission networks and inform the development of point-of-care tests for antimicrobial susceptibility, novel antimicrobials and vaccines. Community driven tools that provide an easy access to and analysis of genomic and epidemiological data is the way forward for public health surveillance.

Methods: Here we present a public health focussed scheme for genomic epidemiology of *N. gonorrhoeae* using Pathogenwatch (https://pathogen.watch/ngonorrhoeae), which enables the processing of raw or assembled genomic data. We implement backwards compatibility with MLST, NG-MAST and NG-STAR typing schemes as well as an exhaustive library of genetic AMR determinants associated with resistance to eight antibiotics. A collection of over 12,000 *N. gonorrhoeae* genome sequences from public archives has been quality-checked, assembled and made public together with available metadata for contextualization.

Results: An international advisory group of experts in epidemiology, public health, genetics and genomics of *N. gonorrhoeae* was convened to identify public health needs in the field and inform on the utility of current and future analytics in the platform, including a customised library of genetic AMR determinants. After uploading genome data, this platform automatically provides typing information, detects genetic determinants of AMR for eight antibiotics including azithromycin and the extended-spectrum cephalosporins ceftriaxone and cefixime, and infers resistance based on the specific combination of mechanisms. Furthermore, genomes are contextualised with globally available genomic data to aid epidemiological investigation.

Conclusions: The *N. gonorrhoeae* scheme in Pathogenwatch provides customized bioinformatic pipelines guided by expert opinion that can be adapted to public health agencies and departments with little expertise in bioinformatics and lower resourced settings with internet connection but limited computational infrastructure. This advisory group will assess and identify ongoing public health needs in the field of gonococcal AMR in order to further enhance utility with modified or new analytic methods.

Background

Antimicrobial resistance (AMR) is an urgent threat to public health. *Neisseria gonorrhoeae*, the strictly human pathogen causing the sexually-transmitted infection (STI) gonorrhoea, has developed or acquired resistance to the last-line antibiotics used in empiric therapy to treat the infection, and thus has become one of the major global priorities in order to tackle AMR. In 2017, due to the increase in AMR infections and the absence of an effective vaccine, the World Health Organization (WHO) included *N. gonorrhoeae* as a high priority pathogen in need of research and development of new antimicrobials and ideally a vaccine (1). In 2019, the Centers for Disease Control and Prevention (CDC) again included the gonococcus on the list of urgent threats in the United States (2). The most recent WHO estimates from 2016 indicate an annual global incidence of 87 million cases of gonorrhoea among adults (3, 4). Untreated cases can develop complications including an increased acquisition and transmission of HIV. In women, long-term infections can cause infertility, pelvic inflammatory disease, ectopic pregnancy, miscarriage or premature labour (5). Infections during pregnancy can transmit to newborns at birth causing eye damage that can have permanent effects on vision (6).

Strains of *N. gonorrhoeae* resistant to every recommended treatment have rapidly emerged, including resistance to penicillins, tetracyclines, fluoroquinolones, macrolides and the extended-spectrum cephalosporins (ESCs) (5-7). The current recommended treatment in many countries is a dual therapy with injectable ceftriaxone plus oral azithromycin, although reports of decreased susceptibility to ceftriaxone as well as azithromycin resistance have increased globally (7, 8). One case of failure of dual treatment was reported in 2016 in the United Kingdom (UK) (9). Additionally, in 2018 a gonococcal strain with resistance to ceftriaxone combined with high-level resistance to azithromycin was detected in both the UK and Australia (10, 11). A ceftriaxone-resistant clone (FC428) has been transmitted internationally, raising concerns about the long-term effectiveness

of the current treatment in the absence of an available alternative (12). In some countries such as in Japan, China and since 2019 in the UK, a single dose of ceftriaxone 1 gram is recommended due to the increasing incidence of azithromycin resistance in *N. gonorrhoeae* and other STI pathogens such as *Mycoplasma genitalium* (13). Extensive investigations have been ongoing for years to unveil the genetic mechanisms that explain most of the observed susceptibility patterns for the main classes of antimicrobials for *N. gonorrhoeae*. For ciprofloxacin, nearly all of the resistant strains have the GyrA S91F amino acid alteration (14-16), however, resistance prediction from genomic data is not as straightforward for other antibiotics. Known resistance mechanisms often involve additive or suppressive effects as well as epistatic interactions that all together explain just part of the observed phenotypic resistance. For example, there is good evidence that many mosaic structures of the *penA* gene are associated with decreased susceptibility of ESCs (17, 18), however, mosaics do not explain all cases of ESC resistance, especially for ceftriaxone, and some mosaic *penA* alleles do not cause decreased susceptibility or resistance to this antibiotic (17-20). On top of these, variants that overexpress the MtrCDE efflux pump, mutations in *porB* that reduce drug influx and non-mosaic mutations in penicillin-binding proteins also contribute to decreased susceptibility to ESCs (21). Furthermore, mutations in the *rpoB* and *rpoD* genes, encoding subunits of the RNA polymerase, have been recently related to resistance to ESCs in clinical *N. gonorrhoeae* isolates (22). Mutations in the 23S rRNA gene (A2045G and C2597T in *N. gonorrhoeae* nomenclature, coordinates from the WHO 2016 reference panel (23), A2059G and C2611T in *Escherichia coli*) are frequently associated with azithromycin resistance, as do variants in *mtrR* or its promoter that increase the expression of the MtrCDE efflux pump (5). Recently, epistatic interactions between a mosaic *mtr* promoter region and a mosaic *mtrD* gene have also been reported to increase the expression of this pump, contributing to macrolide resistance (24, 25). Mutations in *rplD* have also been associated with reduced susceptibility to this antibiotic (26) and contrarily, loss-of-function mutations in *mtrC* have been linked to increased susceptibility to several antibiotics including azithromycin (27).

A myriad of methods have been used to discriminate among strains of *N. gonorrhoeae*, from phenotypic to DNA-based techniques (28), but whole genome sequencing (WGS) can provide the complete genome information of a bacterial strain. The cost of amplifying all loci of the different typing schemes via nucleic acid amplification and traditional Sanger sequencing can be more expensive than the cost of WGS of one bacterial genome in many settings. With WGS, multiple genetic AMR mechanisms as well as virulence and typing regions can be targeted simultaneously with the appropriate bioinformatic tools and pipelines. It also provides a significant improvement in resolution and accuracy over traditional molecular epidemiology and typing methods, allowing a genome-wide comparison of strains that can: identify AMR clones, outbreaks, transmission networks, national and international spread, known and novel resistance mechanisms as well as also inform on the development of point-of-care tests for antimicrobial susceptibility, novel antimicrobials and vaccines (29, 30). However, implementation of WGS for genomic surveillance poses practical challenges, especially for Low- and Middle-Income Countries (LMICs), due to the need of a major investment to acquire and maintain the required infrastructure. The cost of sequencing is decreasing very rapidly in well-resourced settings, especially in large sequencing centres, but it is still prohibitive for routine surveillance in many others.

WGS produces a very high volume of data that needs to be pre-processed and analysed using bioinformatics. Bioinformatics expertise is not always readily available in laboratory and public health settings, and currently there are no international standards and proficiency trials for which algorithms to use to process WGS data. There are several open source tools specialised in each step of the pipeline as well as proprietary software containing workflows that simplify the analyses. However, these are less customizable and may not be affordable for all (31, 32). Choosing the best algorithms and parameters when analysing genomic data is not straightforward as it requires a fair knowledge of the pathogen under study and its genome diversity. Multiple databases containing genetic determinants of AMR for bacterial pathogens are available (31, 32), however, choosing which one is most complete for a particular organism frequently requires an extensive literature search. Public access web-based species-specific tools and AMR databases revised

and curated by experts would be the most approachable option for both well-resourced and LMICs with a reliable internet connection. Very importantly though, the full benefits of using WGS for both molecular epidemiology and AMR prediction can only be achieved if the WGS data are linked to phenotypic data for the gonococcal isolates and, as much as feasible, epidemiological data for the patients.

Here, we present a public health focussed system to facilitate genomic epidemiology of *N. gonorrhoeae* within Pathogenwatch (https://pathogen.watch/ngonorrhoeae), which includes the latest analytics for typing, detection of genetic AMR determinants and prediction of AMR from *N. gonorrhoeae* genome data, linked to metadata where available, as well as a collection of over 12,000 gonococcal genomes from public archives for contextualization. We formed an advisory group including experts in the field of *N. gonorrhoeae* epidemiology, public health, AMR, genetics and genomics to consult on the development and design of the tool, such as the analytics and genetic AMR mechanisms to include, in order to adapt the platform for ongoing public health needs. We present this scheme as a community-steered model for genomic surveillance of other pathogens.

Methods

Generation of the N. gonorrhoeae core genome library

Pathogenwatch implements a library of core genome sequences for several supported organisms. In the case of *N. gonorrhoeae*, a core gene set was built from the 14 reference genomes that constitute the 2016 WHO reference strain panel (23) using the pangenome analysis tool Roary (33) as described in Harris *et al* (2018) (16). Briefly, the minimum percentage of identity for blastp was set to 97% and the resulting core genes were aligned individually using MAFFT. The resulting genes with a percentage of identity above 99% were post-processed as described in (34). Overlapping genes were merged into pseudocontigs and clusters representing paralogs or fragment matches were removed. Representative sequences from each cluster were selected as

177 the longest compared to a consensus obtained from the cluster alignment. The final core gene
178 set contains 1,542 sequences that span a total of 1,470,119 nucleotides. A BLAST database was
179 constructed from these core segments and used to profile new assemblies.

180 ***Profiling new assemblies***

181 New genome assemblies can be uploaded by a user (drag and drop) or calculated from high-
182 throughput short read data directly within Pathogenwatch using SPAdes (35) as described in (36).

183 A taxonomy assignment step for species identification is performed on the uploaded assemblies
184 by using Speciator (37). New assemblies are then queried against a species-specific BLAST
185 database using blastn. For *N. gonorrhoeae*, every core loci needs to match at least 80% of its
186 length to be considered as present. Further filtering steps are applied to remove loci that can be
187 problematic for tree building, such as a paralogs or loci with unusually large number of variant
188 sites compared to an estimated substitution rate on the rest of the genome, as described in (38).
189 The overall substitution rate is calculated as the number of total differences in the core library
190 divided by the total number of nucleotides. Indels are ignored to minimise the noise that could be
191 caused by assembly or sequencing errors. The expected number of substitutions per locus is
192 determined by multiplying this substitution rate by the length of the representative sequence.

193 The number of substitutions observed for each locus between the new assembly and the
194 reference sequence are scaled to the total number of nucleotides that match the core library,
195 creating a pairwise score that it is saved on a distance matrix and is used for tree construction,
196 as described in (39).

197 ***Algorithms for sequence typing and cgMLST clustering***

198 Alleles and sequence types (STs) for Multi-Locus Sequence Typing (MLST) (40) and cgMLST
199 (core genome MLST, *N. gonorrhoeae* cgMLST v1.0) (41) were obtained from PubMLST (42, 43),
200 for *N. gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) (44) from (45) and for *N.*
201 *gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) (46) from (47). A search

tool implemented as part of Pathogenwatch is used to make the assignments for MLST, cgMLST and NG-STAR, while NGMASTER (48) is used for NG-MAST. Briefly, exact matches to known alleles are searched for, while novel sequences are assigned a unique identifier. The combination of alleles is used to assign a ST as described in (49). Databases are regularly updated and novel alleles and STs should be submitted by the user to the corresponding schemes for designation.

cgMLST typing information is used for clustering individual genomes with others in the Pathogenwatch database as described in (50). Users can select the clustering threshold (i.e. number of loci with differing alleles) and a network graph is calculated within individual genome reports.

AMR library and detection of genetic AMR determinants

Genes and point mutations (single nucleotide polymorphisms (SNPs) and indels) were detected using PAARSNP v2.4.9 (51). PAARSNP also provides a prediction of AMR phenotype inferred from the combination of identified mechanisms. Genetic determinants described in the literature as involved in AMR in *N. gonorrhoeae* were collated into a library in TOML format (version 0.0.11). A test dataset containing 3,987 isolates from 13 studies (16, 19, 23, 52-61) (Additional file 1: Table S1) providing minimum inhibitory concentration (MIC) information for six antibiotics (benzylpenicillin, tetracycline, ciprofloxacin, cefixime, ceftriaxone and azithromycin) was used to benchmark and to curate this library. A validation benchmark was posteriorly run with a dataset of 1,607 isolates from 3 other publications (62-64) with MIC information for the same six antibiotics plus spectinomycin (Additional file 1: Table S1). EUCAST clinical breakpoints v9.0 (65) were used for S (susceptibility), I (intermediate resistance/decreased susceptibility) or R (resistance) (SIR) categorical interpretation of MICs for all antibiotics except for azithromycin, for which the epidemiological cut-off (ECOFF) was used. As a result of the benchmark analyses, sensitivity, specificity and positive/negative predictive values (PPV/NPV) were obtained for the AMR mechanisms implemented in the library and, globally, for each of the antibiotics. Confidence intervals for these statistics were calculated using the *epi.tests* function in the *epiR* R package

v1.0-14 (66). Individual or combined AMR mechanisms with a PPV below 15% were discarded from the library to optimise the overall predictive values. Visual representations of the observed ranges of MIC values for a particular antibiotic for each of the observed combinations of genetic AMR mechanisms on the test dataset were used to identify and assess combinations of mechanisms that have an additive or suppressive effect on AMR. These were included in the library.

As part of the quality assessment of the AMR library, we ran the 2016 WHO *N. gonorrhoeae* reference genomes 2016 panel (n=14) through Pathogenwatch and compared the detected list of genetic AMR mechanisms with the list published in the original study (23). For the WHO U strain, a discrepancy on a mutation in *parC* was further investigated by mapping the original raw Illumina data (European Nucleotide Archive (ENA) run accession ERR449479) to the reference genome assembly (ENA genome accession LT592159.1) and visualized using Artemis (67).

In short-read assemblies, the four copies of the 23S rDNA gene are collapsed into one, thus the detection of the A2045G and C2597T mutations is dependent on the consensus bases resulting from the number of mutated copies (57, 60, 68).

Quality check and assembly of public sequencing data

Public *N. gonorrhoeae* genomes with geolocation data were obtained from the ENA in November 2019. This list was complemented by an exhaustive literature search of studies on *N. gonorrhoeae* genomics without metadata submitted to the ENA but instead made available as supplementary information in the corresponding publications. Raw paired-end short read data from a list of 12,192 isolates was processed with the GHRU assembly pipeline v1.5.4 (69). This pipeline runs a Nextflow workflow to quality-check (QC) paired-end short read fastq files before and after filtering and trimming, assembles the data and quality-checks the resulting assembly. In this pipeline, QC of short reads was performed using FastQC v0.11.8 (70). Trimming was done with Trimmomatic v0.38 (71) by cutting bases from the start and end of reads if they were below a Phred score of 25, trimming using a sliding window of size 4 and cutting once the average quality

within the window fell below a Phred score of 20. Only reads with length above a third of the original minimum read length were kept for further analyses. After trimming, reads were corrected using the kmer-based approach implemented in Lighter v1.1.1 (72) with a kmer length of 32 bp and a maximum number of corrections allowed within a 20 bp window of 1. ConFindr v0.7.2 was used to assess intra- and inter-species contamination (73). Mash v2.1 (74) was applied to estimate genome size using a kmer size of 32 bp and Seqtk v1.3 (75) to down sample fastq files if the depth of coverage was above 100x. Flash v1.2.11 (76) was used to merge reads with a minimum overlap length of 20 bp and a maximum overlap of 100 bp to facilitate the subsequent assembly process. SPAdes v3.12 (35) was used for genome assembly with the --careful option selected to reduce the number of mismatches and short indels with a range of kmer lengths depending on the minimum read length. The final assemblies were quality-checked using Quast v5.0.2 (77) and ran through the species identification tool Bactinspector (78). QC conditions were assessed and summarised using Qualifyr (79).

Fastq files with poor quality in which the trimming step discarded all reads from either one or both pairs were excluded from the analyses. Assemblies with an N50 below 25,000 bp, a number of contigs above 300, a total assembly length above 2.5 Mb or a percentage of contamination above 5% were also excluded.

Metadata for public genomes

Geolocation data (mainly country), collection dates (day, month and year when available), ENA project accession and associated Pubmed ID were obtained from the ENA API for all the genomes in the pipeline (80). A manual extensive literature search was performed to identify the publications containing the selected genomes. In order to complete published studies as much as possible, extra genomes were downloaded and added to the dataset. Metadata for the final set was completed with the information contained in supplementary tables on the corresponding publications, including MIC data. Submission date was considered instead of collection date when the latter was not available, however, this occurred in only a few cases (<0.5%).

Results

Upload and analyse N. gonorrhoeae genome data

Data can be uploaded in the form of assemblies or raw data (fastq format) into Pathogenwatch, which allows users to run different analytics on genomic data simultaneously (Figure 1). If raw data is provided, an assembly is calculated before running the analyses. These analytics include four typing schemes for *N. gonorrhoeae* as well as a genotypic AMR prediction using a customized AMR library that includes known genetic mechanisms of resistance for 8 antimicrobials: ceftriaxone, cefixime, azithromycin, ciprofloxacin, spectinomycin, tetracycline, benzylpenicillin and sulfonamides. Statistics on the quality of the assemblies are also provided in the form of matches to the core genome, total genome length, N50, number of contigs, number of non-ATCG bases and GC content (Additional file 2: Figure S1).

Genomes from one or multiple studies can be grouped into collections (Figure 2 and Additional file 2: Figure S2), and the genomic data are automatically processed by comparing with a core *N. gonorrhoeae* genome built from WHO reference strain genomes (16, 23). A phylogenetic tree, inferred using the Neighbour-Joining algorithm on core SNPs, is obtained as a result, representing the genetic relationship among the isolates in the collection. Metadata can be uploaded at the same time as the genome data, and if the collection location coordinates for an isolate are provided, this information is plotted into a map (Additional file 2: Figure S1). If date or year of isolation is also provided, this information is represented in a timeline. The three panels on the main collection layout - the tree, the map and a table or timeline – are functionally integrated so filters and selections made by the user update all of them simultaneously. Users can also easily switch among the metadata and the results of the main analytics: typing, genome assembly statistics, genotypic AMR prediction, AMR-associated SNPs, AMR-associated genes and the timeline (Additional file 2: Figure S1). A video demonstrating the usage and main features of Pathogenwatch is available (81).

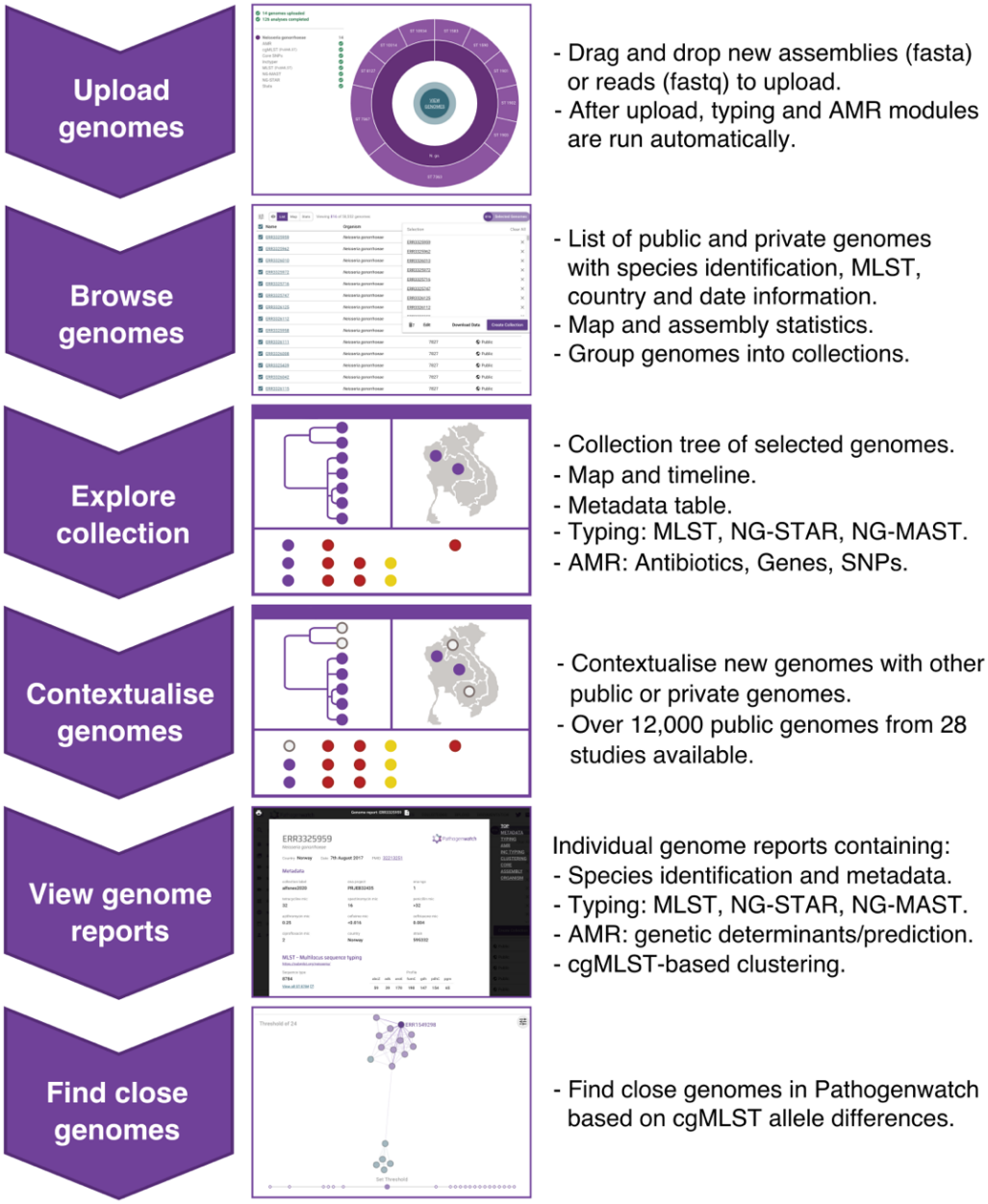

Figure 1. Main workflow in Pathogenwatch. New genomes can be uploaded and combined with public data for contextualisation. The collection view allows data exploration through a combined phylogenetic tree, a map, a timeline and the metadata table, which can be switched to show typing information (Multi-Locus Sequence Typing, MLST; *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance, NG-STAR; and *N. gonorrhoeae* Multi-Antigen Sequence Typing, NG-MAST) as well as known genetic AMR mechanisms for eight antibiotics. Genome reports summarise the metadata, typing and AMR marker results for individual isolates and allow finding other close genomes in Pathogenwatch based on core genome MLST (cgMLST). SNPs: single nucleotide polymorphisms.

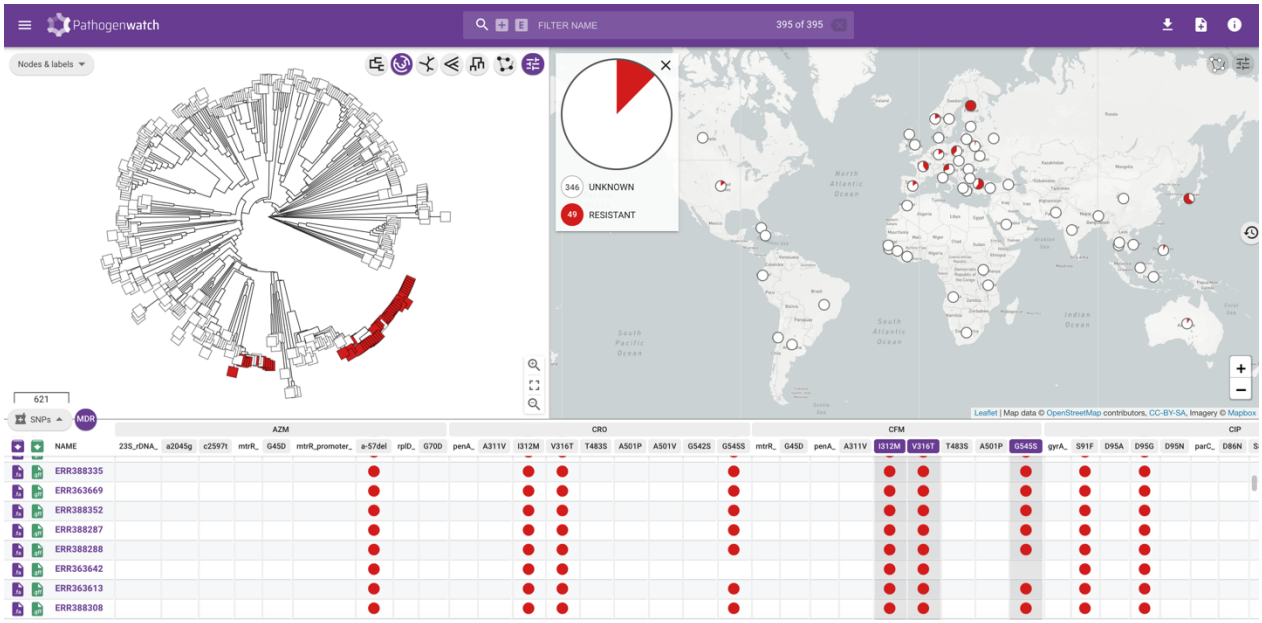

Figure 2. Main display of a Pathogenwatch collection, showing a phylogenetic tree, a map and a table of SNPs associated to AMR of 395 *N. gonorrhoeae* genomes from a global study (58, 82). Isolates carrying three mosaic *penA* marker mutations are marked in red in the tree and the map. The table can be switched to show the metadata, a timeline, typing results (Multi-Locus Sequence Typing, MLST; *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance, NG-STAR and *N. gonorrhoeae* Multi-Antigen Sequence Typing, NG-MAST) as well as AMR analytics (known genetic mechanisms and genotypic AMR prediction) implemented in the platform. Further detail is shown in Additional file 2: Figure S1

Sequence typing schemes: cgMLST, MLST, NG-MAST and NG-STAR

Pathogenwatch implements four sequence typing schemes for *N. gonorrhoeae*: cgMLST (41), MLST (40), NG-MAST (44) and NG-STAR (46) (Table 1). Each of the schemes is based on a group of loci for which individual allele numbers are assigned relying on an existing database of allele sequences. A unique ST is generated from the combination of allele numbers to represent each isolate. The cgMLST scheme includes 1,649 loci from the *N. gonorrhoeae* cgMLST v1.0 scheme in PubMLST (43) and it is used for clustering individual genomes with others in the database based on allele differences (Additional file 2: Figure S3). The MLST scheme, also hosted in PubMLST, includes 7 housekeeping genes and gene fragments more conserved and slowly evolving in the *Neisseria* genus. NG-MAST includes internal fragments from two highly polymorphic and rapidly evolving outer membrane protein genes, *porB* and *tbpB.* NG-STAR was developed more recently with the aim of standardizing the nomenclature associated with AMR

352 determinants as well as having a typing scheme that would distinguish among lineages with
353 different AMR mechanisms. It includes 7 genes associated with resistance to β-lactams,
354 macrolides and fluoroquinolones (Table 1).

355 Table 1. *N. gonorrhoeae* sequence typing schemes implemented in Pathogenwatch.
356

Typing scheme*	Loci (number)	Note	Pathogenwatch implementation	References
cgMLST	(N=1,649)	*N. gonorrhoeae* cgMLST v1.0	Typing algorithm, database from PubMLST	(41-43, 83)
MLST	*abcZ, adk, aroE, fumC, gdh, pdhC, pgm* (N=7)	Housekeeping genes in *Neisseria* spp.	In-house MLST tool, database from PubMLST	(40, 42, 43, 83)
NG-MAST	*porB, tbpB* (N=2)	Genes encoding highly-variable membrane proteins	NG-MASTER, database from NG-MAST website	(44, 45, 48)
NG-STAR	*penA, mtrR, porB, ponA, gyrA, parC, 23S rDNA* (N=7)	Genes involved in antimicrobial resistance	In-house MLST tool, database from NG-STAR website	(46, 47, 83)

357 * Typing scheme: cgMLST = core genome Multi-Locus Sequence Typing, MLST = Multi-Locus Sequence Typing, NG-
358 MAST = *N. gonorrhoeae* Multi-Antigen Sequence Typing, NG-STAR = *N. gonorrhoeae* Sequence Typing for
359 Antimicrobial Resistance.
360
361

362 ***Library of genetic AMR mechanisms: test and validation***

363 We compiled described genetic AMR mechanisms previously reported for *N. gonorrhoeae* up to
364 the writing of this manuscript into the AMR library in Pathogenwatch (Table 2).

365

366

367

368

369

370

Table 2. List of *N. gonorrhoeae* genetic antimicrobial resistance (AMR) determinants in Pathogenwatch. References that report evidence of association of each mechanism to AMR in clinical isolates and/or where their role on AMR has been confirmed in the laboratory through, i.e. transformation experiments, are included in the table. Effect: R = Resistance, I = Intermediate resistance (decreased susceptibility), A = Additive effect, N = Negative effect. R and I follow the EUCAST clinical breakpoints except for azithromycin, for which the epidemiological cut-off (ECOFF) is reported and used instead.

Antibiotic (MIC breakpoint mg/L)	Genetic AMR determinants	Effect	Evidence (References)
Azithromycin (R: MIC>1, ECOFF)	*23S* rDNA 2045A>G substitution (2059A>G in *E. coli*)	R	(68)
	23S rDNA 2597C>T substitution (2611C>T in *E. coli*)	R	(84)
	ermA, ermB, ermC, ermF genes	R	(85, 86)
	ereA, ereB genes	R	(23)
	mefA gene	R	(86, 87)
	macAB promoter -48G>T substitution*	R	(88)
	mtr mosaic**		
	N. meningitidis-like mosaic (n=1)	R	(24)
	N. lactamica-like mosaic (n=2)	R	(24)
	mtrD mosaic**		
	N. meningitidis-like mosaic (n=1)	R	(24)
	N. lactamica-like mosaic (n=2)	R	(24)
	mtrR promoter -57delA*	A	(89, 90)
	mtrR G45D	A	(91, 92)
	mtrC loss-of-function	N	(27)
	rplV ARAK tandem duplication (position 90)	R	(19)
	rplV KGPSLK tandem duplication (position 83)	R	(19)
	rplD G70D	A	(26)
Ceftriaxone*** (R: MIC>0.125)	*penA* mosaic (A311V, I312M, V316P/T, T483S and G545S)	R	(93-95)
	penA V316P, T483S, A501P/V, G542S	R	(93, 94)
	rpoB P157L, G158V, R201H	R	(22)
	rpoD D92-95 deletion, E98K	I	(22)
Cefixime*** (R: MIC>0.125)	*mtrR* G45D	A	(91, 92)
	penA mosaic (I312M, V316T, G545S)	R	(93-95)
	penA mosaic (A311V, I312M, V316P/T, T483S and G545S)	R	(93-95)
	penA V316P, T483S, A501P	I	(93, 94)
	rpoB P157L, G158V, R201H	I	(22)
	rpoD D92-95 deletion, E98K	I	(22)
Ciprofloxacin (I: 0.03<MIC≤0.06; R: MIC>0.06)	*gyrA* S91F, D95A/N	R	(96)
	gyrA D95G	I	(96)
	norM promoter -7A>G, -104C>T substitutions*	I	(97)
	parC D86N, S87R	R	(96)
	parC S87I/N, S88P, E91K	I	(96)
	parE G410V	I	(98)
Tetracycline**** (I: 0.5<MIC≤1; R: MIC>1)	*mtrR* A39T, G45D	A	(91, 92)
	mtrR loss-of-function	I	(23)
	mtrR promoter -56A>C substitution, -57delA deletion*	I	(24, 89, 90)
	mtrR promoter -131G>A (*mtrC* -120G>A substitution, *mtr120)**	I	(91)
	rpsJ V57M	I	(99)
	tetM gene	R	(100)

Penicillins (I: 0.06<MIC≤1; R: MIC>1)	*blaTEM* gene	R	(101)
	mtrR G45D	I	(91, 92)
	mtrR A39T	A	(91)
	mtrR loss-of-function	I	(23)
	mtrR promoter -56A>C, -57delA*	I	(24, 90)
	mtrR promoter -131G>A (*mtrC* -120G>A substitution, *mtr120)**	I	(91)
	penA I312M, V316P/T, ins346D, T483S, A501P/T/V, G542S, G545S, P551S	I	(93, 94)
	penA mosaic (I312M, V316T, G545S)	A	(93-95)
	ponA1 L421P	I	(102)
	porB1b G120K, A121N/D	I	(103)
Spectinomycin (R: MIC>64)	*16S* rDNA 1184C>T (1192C>T in *E. coli*)	R	(104)
	rpsE T24P	R	(105)
	rpsE V27- deletion, K28E	R/A	(105)
Sulfonamides *****	*folP* R228S	R	(23, 106)

*Nomenclature of the mutations on the *macAB, mtrR and norM* promoter regions is based on *N. gonorrhoeae* coordinates considering the distance from the start of the *macAB, mtrR* and *norM* genes, respectively. **Note that mosaics are caused by recombination events, which can have variable breakpoints with different effects on azithromycin MIC if any. In this version, we have included the three mosaics described by Wadsworth *et al.* (24), but the list will be expanded as new mosaic *mtr* (intergenic region between *mtrR* and *mtrC*) and *mtrD* alleles having an effect on azithromycin MICs are published. ***The list of genetic AMR mechanisms for the ESCs ceftriaxone and cefixime do not include all known *porB1b* or *mtrR*-associated variants as their effect was found not to be relevant in increasing MIC on the benchmark analyses for phenotypic AMR prediction purposes despite the experimental evidence reported in Zhao *et al.* (107). In case of strains carrying *penA*-associated mutations, their immediate predicted phenotype is that of those carrying *penA*-associated variants. ****The list of genetic AMR mechanisms for tetracycline does not include *porB1b* mutations as their effect was found not to be relevant in increasing MIC on the benchmark analyses for phenotypic AMR prediction purposes. *****Sulfonamides are not a treatment alternative for gonorrhoea, however the *folP* R228S mutation is kept in this version of the library for surveillance purposes.

This list was benchmarked using a test dataset of 3,987 *N. gonorrhoeae* isolates from 13 different studies containing MIC information for at least part of the following six antibiotics: ceftriaxone, cefixime, azithromycin, ciprofloxacin, benzylpenicillin and tetracycline (Additional file 1: Table S1). EUCAST clinical breakpoints were applied for five of the antimicrobials except for azithromycin, for which the adoption of an ECOFF>1 mg/L is now recommended to distinguish isolates with azithromycin resistance determinants, instead of a clinical resistance breakpoint (108, 109). A visualization of the range of MICs on each particular combination of genetic AMR mechanisms observed on the isolates from the benchmark test dataset (Figure 3a-b and Additional file 2: Figures S4-S9) revealed combinations that show an additive effect on AMR. These combinations were included in the AMR library to improve the accuracy of the genotypic prediction. For example, *rpsJ* V57M and some *mtrR*-associated mutations individually cause decreased susceptibility or intermediate resistance to tetracycline (MICs between 0.5-1 mg/L), however, a

combination of these variants can increase MICs above the EUCAST resistance breakpoint for tetracycline (MICs>1 mg/L) (Additional file 2: Figure S8). This is the case of the combination of *rpsJ* V57M with the *mtrR* promoter -57delA mutation (N=681 isolates, 94.9% positive predictive value, PPV) or with *mtrR* promoter -57delA and *mtrR* G45D (N=83 isolates, 93.9% PPV). Several combinations of *penA, ponA1, mtrR* and *porB1b* mutations were observed to be able to increase the benzylpenicillin MIC above the resistant threshold in most of the cases (Additional file 2: Figure S9). This is the case of the *porB1b* mutations combined with *mtrR* A39T (N=31 isolates, 100% PPV), with the *mtrR* promoter -57delA deletion (N=286 isolates, 96.5% PPV) or with *mtrR* promoter -57delA and *ponA1* L421P (N=269 isolates, 96.3%). Despite mosaic *penA* not being a main driver of resistance to penicillins, a combination of the *porB1b* mutations with the three main mosaic *penA* mutations (G545S, I312M and V316T) was also observed to produce a resistant phenotype in all cases (N=17 isolates, 100% PPV). A recent publication showed that loss-of-function mutations in *mtrC* increased susceptibility to azithromycin and are associated with isolates from the cervical environment (27). We included the presence of a disrupted *mtrC* as a modifier of antimicrobial susceptibility in the presence of an *mtr* mosaic, as it did not show a significant effect in the presence of 23S rDNA A2045G and C2597T mutations.

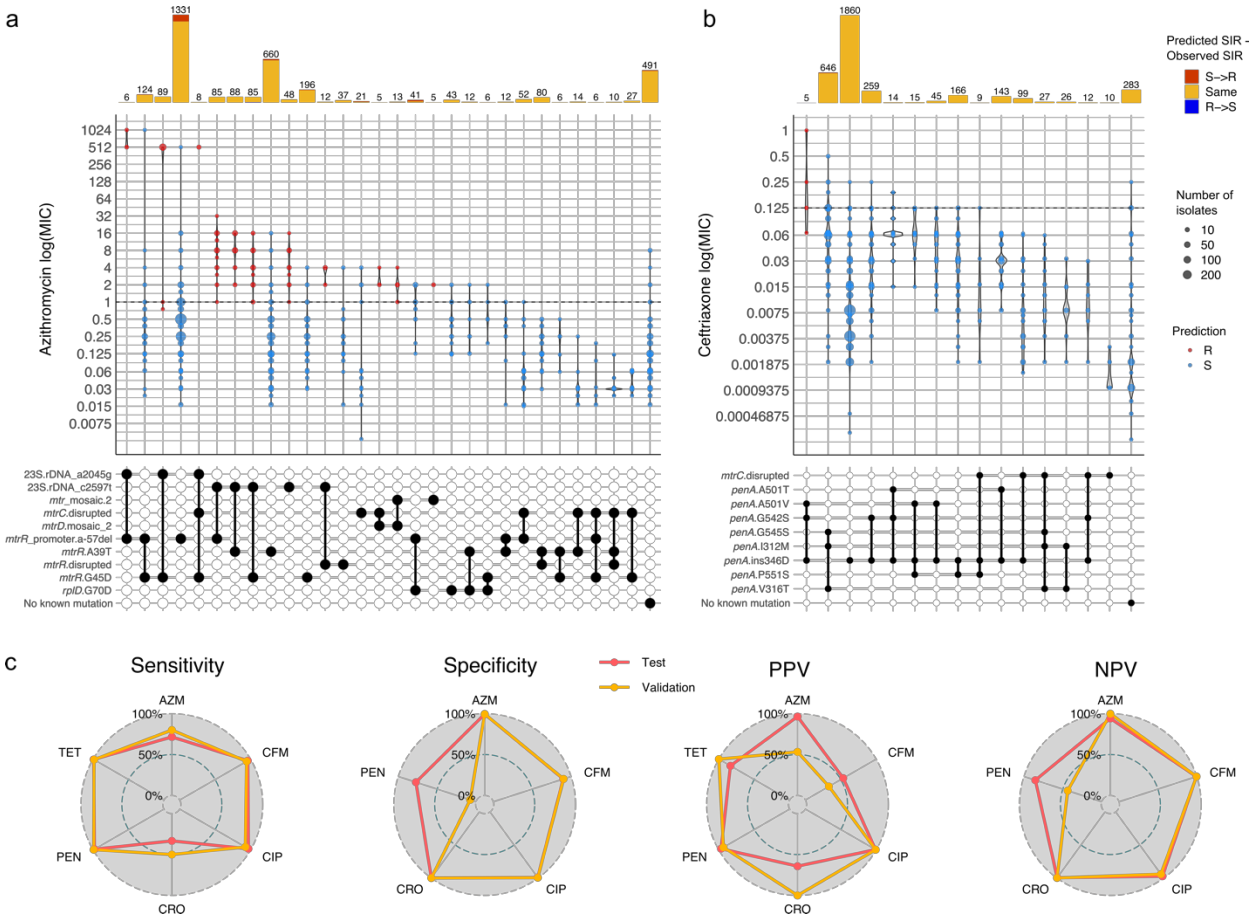

Figure 3. Distribution of minimum inhibitory concentration (MIC) values (mg/L) for the last-line antibiotics for *N. gonorrhoeae* azithromycin (a) and ceftriaxone (b) in a collection of 3,987 *N. gonorrhoeae* isolates with different combinations of genetic antimicrobial resistance (AMR) mechanisms. Only combinations observed in at least 5 isolates are shown (see Additional file 2: Figure S4-S9 for expanded plots for six antibiotics). Dashed horizontal lines on the violin plots mark the EUCAST epidemiological cut-off (ECOFF) for azithromycin and EUCAST clinical breakpoint for ceftriaxone. Point colours inside violins represent the genotypic AMR prediction by Pathogenwatch on each combination of mechanisms (indicated by black circles connected vertically; horizontal thick grey lines connect combinations of mechanisms that share an individual determinant). Barplots on the top show the abundance of isolates with each combination of mechanisms. Bar colours represent the differences between the predicted and the observed SIR (i.e. red for a predicted susceptible mechanism when the observed phenotype is resistant). (c) Radar plots comparing the sensitivity, specificity, positive and negative predictive values (PPV/NPV) for six antibiotics for the test and validation benchmark analyses. AZM = Azithromycin, CFM = Cefixime, CIP = Ciprofloxacin, CRO = Ceftriaxone, PEN = Benzylpenicillin, TET = Tetracycline.

Results from the benchmark (Additional file 1: Table S2) show sensitivity values (true positive rates, TP/(TP+FN); TP=True Positives, FN=False Negatives) above 96% for tetracycline (99.2%), benzylpenicillin (98.1%), ciprofloxacin (97.1%) and cefixime (96.1%), followed by azithromycin (71.6%) and ceftriaxone (33.3%). These results reflect the complexity of the resistance

mechanisms for azithromycin and ceftriaxone, where the known genetic determinants explain only part of the antimicrobial susceptibility. However, specificity values (true negative rates, TN/(TN+FP); TN=True Negatives, FP=False Positives) for these two antibiotics as well as ciprofloxacin were above 99% (Additional file 1: Table S2), demonstrating that the genetic mechanisms included in the database have a role in AMR. The specificity value for cefixime was lower but nearly 90%, mainly due to the high number of isolates with an MIC below the threshold but with three mutations characterising a mosaic *penA* allele (G545S, I312M and V316T, TP=367, TN=323, PPV=53.2%; Additional file 1: Table S3). Benzylpenicillin and tetracycline showed specificity values of 77.3% and 61.3%, respectively. In the first case, all the mechanisms included in the library showed a PPV value above 94%. For tetracycline, a considerable number of false positive results are mainly caused by the presence of *rpsJ* V57M, for which PPV=83.8% (TP=1083, FP=209; Additional file 1: Table S3). However, this mutation was kept in the AMR library because it can cause intermediate resistance to tetracycline on its own (Additional file 2: Figure S8).

Results from the benchmark analysis on the 3,987-isolates dataset were used to curate and optimize the AMR library. Thus, in order to objectively validate it, the benchmark analysis was also run on a combination of three different collections (N=1,607, Additional file 1: Table S1) with available MIC information for seven antibiotics including spectinomycin (Additional file 1: Table S4) (63, 64, 110). Results from the test and validation benchmark runs were compared, showing that sensitivity values on the six overlapping antibiotics were very similar, with the validation set performing even better for azithromycin and ceftriaxone (Figure 3c). In terms of specificity, both datasets performed equally well for all antibiotics except for benzylpenicillin, in which specificity drops in the validation dataset. This is due to the *penA*_ins346D mutation (TP=1125, FP=83) and the *blaTEM* genes (TP=525, FP=36), which despite showing false positives, have a PPV above 93% (Additional file 1: Table S4). In general, discrepancies found between the test and the validation datasets can be explained by particular mechanisms that on their own show high

predictive values and affect antibiotics for which we do not currently understand all the factors involved in resistance, such as azithromycin and the ESCs (Additional file 1: Table S4).

An additional quality assessment of the AMR library was performed using the 14 *N. gonorrhoeae* reference genomes from the WHO 2016 panel (23), which were uploaded into Pathogenwatch. All the genetic AMR determinants described as present in these isolates and implemented in the Pathogenwatch AMR library were obtained as a result (Additional file 1: Table S5). Only one discrepancy was found when compared to the original publication. The WHO U strain was reported as carrying a *parC* S87W mutation. However, mapping the original Illumina data from this isolate with the final genome assembly revealed that this strain carries a wild type allele (Additional file 2: Figure S10). MLST and NG-MAST types were the same as those reported in the original publication (note that NG-STAR was not available at that time) and the *porA* mutant gene was found in WHO U as previously described. This mutant *porA* has nearly a 95% nucleotide identity to *N. meningitidis* and 89% to *N. gonorrhoeae*, and it is included as screening because it has previously been shown to cause false negative results in some molecular detection tests for *N. gonorrhoeae* (111).

Over 12,000 public genomes available

All *N. gonorrhoeae* short-read sequencing raw data with geolocation data (minimum of country and preferably also year) and associated to a scientific publication was downloaded from the ENA. This collection was expanded after an exhaustive literature search on studies that did not upload geolocation data to the ENA but released as a part of scientific publication(s). Over 12,000 genomes were assessed for sequencing quality data and contamination, assembled using a common pipeline and thresholds as well as post-assembly quality check (Additional file 3). Data for 11,461 isolates were successfully assembled and passed all quality cut-offs, providing 12,515 isolates after including the previously-available Euro-GASP 2013 dataset (16). New assemblies were uploaded and made public on Pathogenwatch, which now constitutes the largest repository of curated *N. gonorrhoeae* genomic data with associated metadata, typing and AMR information

491 at the time of submission of this manuscript. Updated data spans 27 different publications (19,
492 44, 48, 52-55, 57-59, 61-64, 110, 112-125) and is organized into individual collections associated
493 with the different studies (Additional file 1: Table S6). Available metadata was added for the
494 genomes from these publications while basic metadata fields were kept for others (country,
495 year/date and ENA project number).

496 The *N. gonorrhoeae* public data available on Pathogenwatch spans nearly a century (1928-2018)
497 and almost 70 different countries (Additional file 2: Figure S11). However, sequencing efforts are
498 unevenly distributed around the world, and over 90% of the published isolates were isolated in
499 only 10 countries, headed by the United Kingdom (N=3,476), the United States (N=2,774) and
500 Australia (N=2,388) (Additional file 1: Table S7, Figure 4). A total of 554 MLST, 1,670 NG-MAST
501 and 1,769 NG-STAR different STs were found in the whole dataset, from which a considerable
502 number were new profiles caused by previously undetected alleles or new combinations of known
503 alleles (N=92 new MLST STs, N=769 new NG-STAR STs and N=2,289 isolates with new NG-
504 MAST *porB* and/or *tbpB* alleles). These new alleles and profiles were submitted to the
505 corresponding scheme servers.

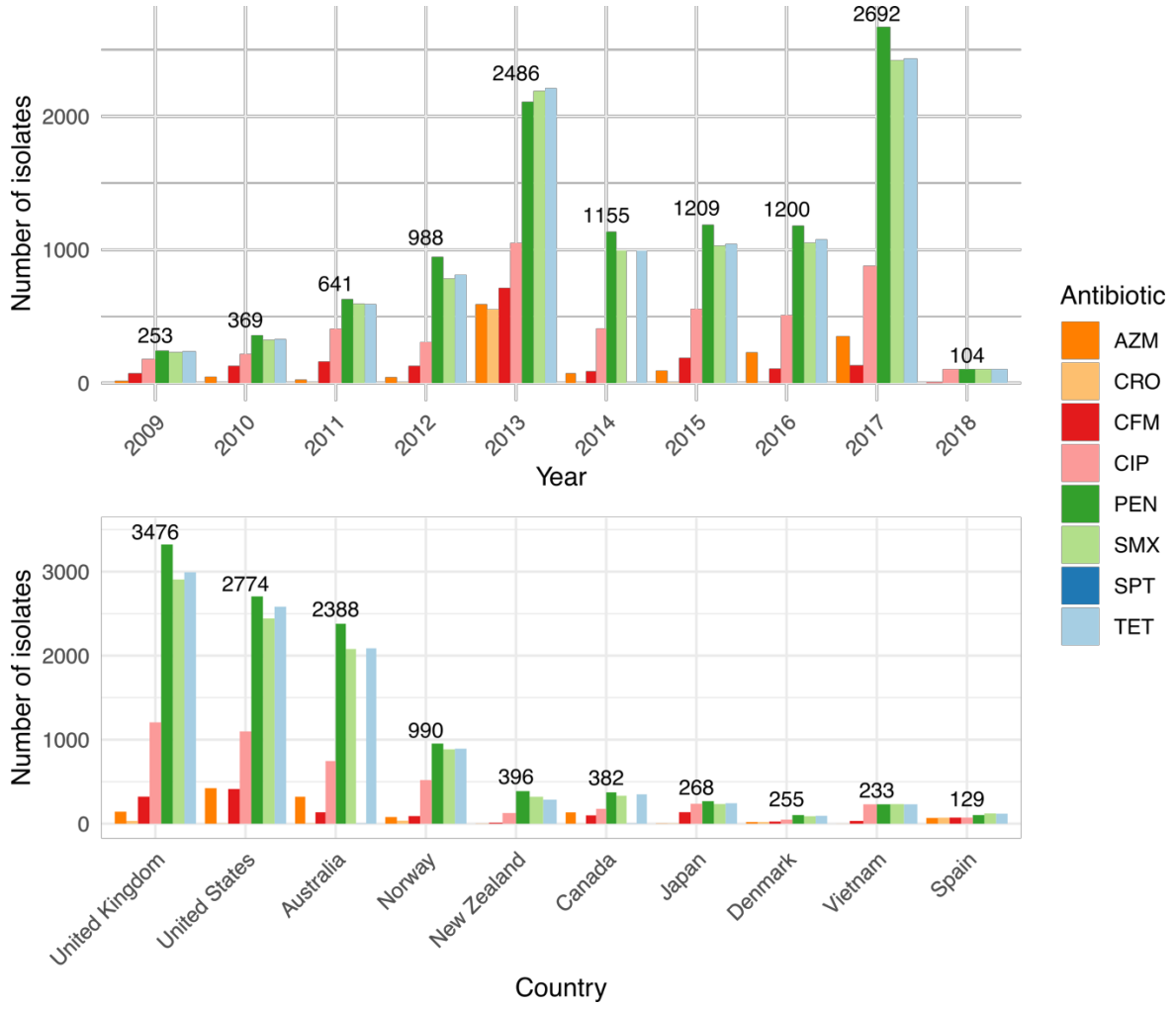

Figure 4. Summary of the geolocalization and collection date of 12,515 public *N. gonorrhoeae* genomes in Pathogenwatch. Coloured bars represent the genotypic antimicrobial resistance (AMR) prediction based on the mechanisms included in the library. AZM = Azithromycin, CFM = Cefixime, CIP = Ciprofloxacin, CRO = Ceftriaxone, PEN = Benzylpenicillin, TET = Tetracycline.

506

507 Genomic studies are often biased towards AMR isolates, and this is reflected in the most
508 abundant STs found for the three typing schemes within the public data. Isolates with MLST
509 ST1901, ST9363 and ST7363, which contain resistance mechanisms to almost every antibiotic
510 included in the study, represent over 25% of the data (Figure 5). Isolates with MLST ST1901 and
511 ST7363 are almost always resistant to tetracycline, sulfonamides, benzylpenicillin and
512 ciprofloxacin and nearly 50% of isolates from these two types harbour resistance mechanisms to
513 cefixime. Ciprofloxacin resistance is not widespread among ST9363 isolates, in which
514 azithromycin resistance can approach to nearly 50% of the isolates for this ST (Figure 5). NG-

STAR ST63 (carrying the non-mosaic *penA-2* allele, *penA* A517G and *mtrR* A39T mutations as described in (47)) is the most represented in the dataset and carries resistance mechanisms to tetracycline, sulfonamides, and benzylpenicillin, but is largely susceptible to spectinomycin, ciprofloxacin, the ESCs cefixime and ceftriaxone and azithromycin. NG-STAR ST90 isolates, conversely, are largely resistant to cefixime, ciprofloxacin and benzylpenicillin as they carry the key resistance mutations in mosaic *penA-34*, as well as in the *mtrR* promoter, *porB1b*, *ponA*, *gyrA* and *parC* (as described in (47)). NG-MAST ST1407 is commonly associated with MLST ST1901 and is the second most represented ST in the dataset following NG-MAST ST2992, which mainly harbours resistance to tetracycline, benzylpenicillin and sulfonamides (Figure 5).

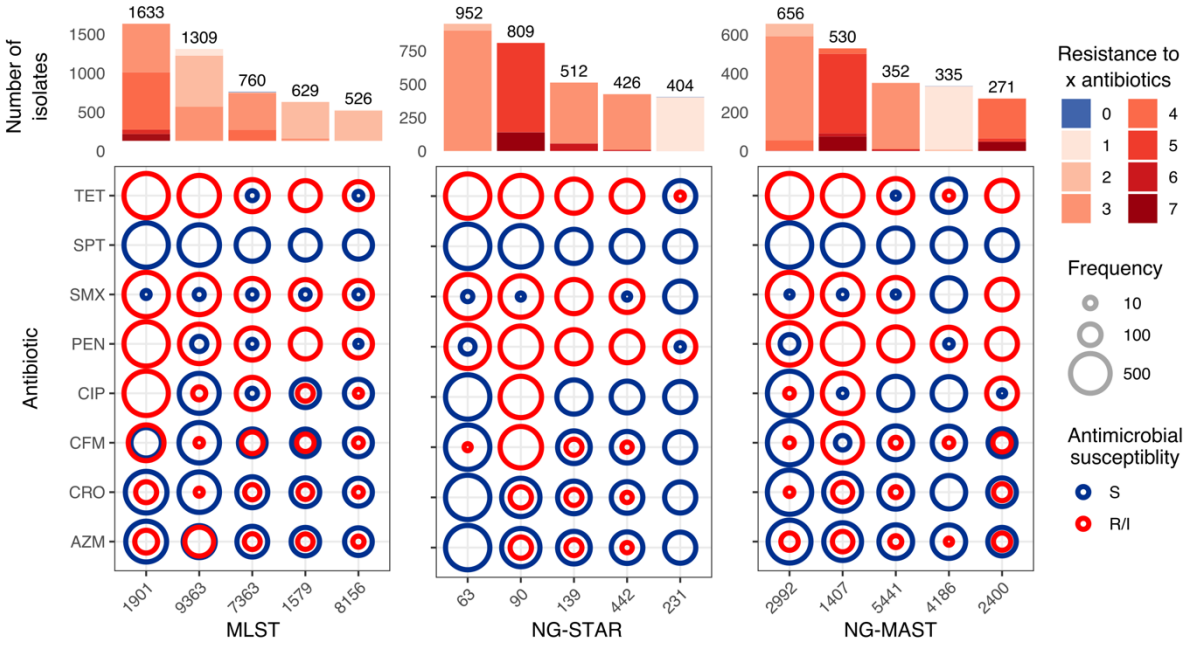

Figure 5. Predicted antimicrobial resistance (AMR) profiles of the top five Multi-Locus Sequence Typing (MLST), *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) and *N. gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) types in the *N. gonorrhoeae* public data in Pathogenwatch. The main graph shows the proportion of resistant (including intermediate phenotypes, in red) versus susceptible genomes (in dark blue) from each sequence type (ST) and antibiotic. Bars on the top show the number of isolates from each ST coloured by the number of antibiotics the genomes are predicted to be resistant to.

Data sharing and privacy

Sequencing data and metadata files uploaded by the user are kept within the user's private account. Genomes can be grouped into collections and these can be toggled between private and accessible to collaborators via a URL. Collection URLs include a 12-letter random string to secure them against brute force searching. Setting a collection to 'off-line mode' allows users to work in challenging network conditions, which may be beneficial in LMICs – all data are held within the browser. Users can also integrate private and potentially confidential metadata into the display without uploading it to the Pathogenwatch servers (locally within the browser on a user's machine).

Discussion

We present a public health focussed *N. gonorrhoeae* framework within Pathogenwatch, an open access platform for genomic surveillance supported by an expert group that can be adapted to any public health or microbiology laboratory. Little bioinformatics expertise is required, and users can choose to either upload raw short read data or assembled genomes. In both cases, the upload of high-quality data is encouraged in the form of quality-checked reads and/or quality-checked assemblies. Recent benchmark analyses show particular recommendations for long-read or hybrid data (126) as well as short read-only data (35, 127). On upload, several analyses are run on the genomes, and results for the three main typing schemes (MLST, NG-MAST and NG-STAR) as well as the detection of genetic determinants of AMR and a prediction of phenotypic resistance using these mechanisms can be obtained simultaneously. The library of AMR determinants contained in Pathogenwatch for *N. gonorrhoeae* has been revised and extended to include the latest mechanisms and epistatic interactions with experimental evidence of decreasing susceptibility or increasing resistance to at least one of eight antibiotics (Tables 2 and 3). A benchmark analysis on a test and validation datasets revealed sensitivity and/or specificity values >90% for most of the tested antibiotics (Additional file 1: Table S2).

553 The continuous increase in reporting of *N. gonorrhoeae* AMR isolates worldwide led to a call for
554 international collaborative action in 2017 to join efforts towards a global surveillance scheme. This
555 was part of the WHO global health sector strategy on STIs (2016-2021), which set the goal of
556 ending STI epidemics as a public health concern by year 2030 (7, 8). Several programmes are
557 currently in place at different global, regional or national levels to monitor gonorrhoea AMR trends,
558 emerging resistances and refine treatment guidelines and public health policies. This is the case
559 of, for example, the WHO Global Gonococcal Antimicrobial Surveillance Programme (WHO
560 GASP) (8), the Euro-GASP in Europe (6, 16, 128), the Gonococcal Isolate Surveillance Project
561 (GISP) in the United States (129), the Canadian Gonococcal Antimicrobial Surveillance
562 Programme (130), the Gonococcal Surveillance Programme (AGSP) in Australia (131) or the
563 Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP) in England and
564 Wales (132). The WHO in collaboration with CDC has recently started an enhanced GASP
565 (EGASP) (133) in some sentinel countries such as the Philippines and Thailand (134), aimed at
566 collecting standardized and quality-assured epidemiological, clinical, microbiological and AMR
567 data. On top of these programs, WHO launched the Global AMR Surveillance System (GLASS)
568 in 2015 to foster national surveillance systems and enable standardized, comparable and
569 validated AMR data on priority human bacterial pathogens (135). Efforts are now underway to link
570 GASP to GLASS. However, gonococcal AMR surveillance is still suboptimal or even lacking in
571 many locations, especially in LMICs, such as some parts of Asia, Central and Latin America,
572 Eastern Europe and Africa, which worryingly have the greatest incidence of gonorrhoea (3).
573 LMICs often have access to antimicrobials without prescription, have limited access to an optimal
574 treatment, lack the capacity needed to perform a laboratory diagnosis due to limited or non-
575 existent quality-assured laboratories, microbiological and bioinformatics expertise or training,
576 insufficient availability and exorbitant prices of some reagents on top of a lack of funding, which
577 altogether compromises infection control.

578 High throughput sequencing approaches have proved invaluable over traditional molecular
579 methods to identify AMR clones of bacterial pathogens, outbreaks, transmission networks and

national and international spread among others (29, 30). Genomic surveillance efforts to capture the local and international spread of *N. gonorrhoeae* have resulted in several publications within the last decade involving high throughput sequence data of thousands of isolates from many locations across the world. The analysis of this data requires expertise, not always completely available, in bioinformatics, genomics, genetics, AMR, phylogenetics, epidemiology, etc. For lower-resourced settings, initiatives such as the NIHR Global Health Research Unit, Genomic Surveillance of Antimicrobial Resistance (136) are essential to build genomic surveillance capacity and provide the necessary microbiology and bioinformatics training for quality-assured genomic surveillance of AMR.

One of the strengths of genomic epidemiology is being able to compare new genomes with existing data from a broader geographical level, which provides additional information on, i.e. if new cases are part of a single clonal expansion or multiple introductions from outside a specific location. Currently, over 12,000 isolates of *N. gonorrhoeae* have been sequenced using high throughput approaches and publicly deposited on the ENA linked to a scientific publication. We have quality-checked and assembled these data using a common pipeline and we make it available through Pathogenwatch, with the aim of representing as much genomic diversity of this pathogen as possible to serve as background for new analyses. These public genomes are associated with at least 27 different scientific publications, and have been organized in Pathogenwatch as individual collections (Additional file 1: Table S6).

In this study, we have gathered an advisory group of *N. gonorrhoeae* experts in different fields such as AMR, microbiology, genetics, genomics, epidemiology and public health who will consult and discuss current and future analytics to be included to address the global public health needs of the community. We suggest this strategy as a role model for other pathogens in this and other genomic surveillance platforms, so the end user, who may not have full computational experience in some cases, can be confident that the analytics and databases underlying this tool are appropriate, and can have access to all the results provided by Pathogenwatch through uploading the data via a web browser. We are aware that this is a constantly moving field and analytics will

607 be expanded and updated in the future. These updates will be discussed within an advisory group
608 to make sure they are useful in the field and the way results are reported is of use to different
609 profiles (microbiologists, epidemiologists, public health professionals, etc.).

610 Future analytics that are under discussion include the automatic submission of new MLST, NG-
611 STAR and NG-MAST STs and alleles to the corresponding servers and the automatic submission
612 of data to public archives such as the ENA. Including a separate library to automatically screen
613 targets of potential interest for vaccine design (137-139) as well as targets of new antibiotics on
614 phase II or III clinical trials (i.e. zoliflodacin (140) or gepotidacin (141)) can also be an interesting
615 addition to the scheme. Regarding AMR, new methods for phenotypic prediction using genetic
616 data are continuously being reported (56, 142, 143), especially those based on machine learning
617 algorithms (144), and will be considered for future versions of the platform.

618 **Conclusions**

619 In summary, we present a genomic surveillance platform adapted to *N. gonorrhoeae*, one of the
620 main public health priorities compromising the control of AMR infections, where decisions on
621 existing and updated databases and analytics as well as how results are reported will be
622 discussed with an advisory board of experts in different public health areas. This will allow
623 scientists from both higher or lower resourced settings with different capacities regarding high
624 throughput sequencing, bioinformatics and data interpretation, to be able to use a reproducible
625 and quality-assured platform where analyse and contextualise genomic data resulting from the
626 investigation of treatment failures, outbreaks, transmission chains and networks at different
627 regional scales. This open access and reproducible platform constitutes one step further into an
628 international collaborative effort where countries can keep ownership of their data in line with
629 national STI and AMR surveillance and control programs while aligning with global strategies for
630 a joint action towards battling AMR *N. gonorrhoeae*.
631

632

List of abbreviations

AGSP: Australian Gonococcal Surveillance Programme

AMR: Antimicrobial Resistance

AZM: Azithromycin

CDC: Centers for Disease Control and Prevention

CFM: Cefixime

cgMLST: Core Genome Multi-Locus Sequence Typing

CIP: Ciprofloxacin

CRO: Ceftriaxone

ECOFF: Epidemiological Cut-Off

EGASP: Enhanced Gonococcal Antimicrobial Surveillance Programme

ENA: European Nucleotide Archive

ESCs: Extended Spectrum Cephalosporins

EUCAST: European Committee on Antimicrobial Susceptibility Testing

Euro-GASP: European Global Antimicrobial Surveillance Programme

FN: False Negative

FP: False Positive

GASP: Global Gonococcal Antimicrobial Surveillance Programme

GISP: Gonococcal Isolate Surveillance Project

GRASP: Gonococcal Resistance to Antimicrobials Surveillance Programme

HIV: Human Immunodeficiency Virus

LMICs: Low and Middle-Income Countries

MIC: Minimum Inhibitory Concentration

MLST: Multi-Locus Sequence Typing

NG-MAST: *N. gonorrhoeae* Multi-Antigen Sequence Typing

NG-STAR: *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance

NPV: Negative Predictive Value

PEN: Benzylpenicillin

PPV: Positive Predictive Value

SNPs: Single Nucleotide Polymorphisms

ST: Sequence Type

STI: Sexually-Transmitted Infection

TET: Tetracycline

TN: True Negative

TP: True Positive

UK: United Kingdom

WGS: Whole Genome Sequencing

WHO: World Health Organization

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The assemblies included in the current version of the *N. gonorrhoeae* Pathogenwatch scheme and used for the AMR benchmark analyses were generated from raw sequencing data stored in the ENA. Project accession numbers are included in Additional File 1: Tables S1 and S6. The generated assemblies can be downloaded from Pathogenwatch. The AMR library can be accessed from: https://gitlab.com/cgps/pathogenwatch/amr-libraries/-/blob/master/485.toml. The code to reproduce the figures and analyses in this manuscript can be found in https://gitlab.com/cgps/pathogenwatch/publications.

Competing interests

The authors declare that they have no competing interests.

Funding

Pathogenwatch is developed with support from Li Ka Shing Foundation (Big Data Institute, University of Oxford) and Wellcome (099202). LSB and DMA are supported by the Li Ka Shing Foundation (Big Data Institute, University of Oxford) and the Centre for Genomic Pathogen Surveillance (CGPS, http://pathogensurveillance.net). DMA and SA are supported by the National Institute for Health Research (UK) Global Health Research Unit on Genomic Surveillance of AMR (16_136_111). The department of MJC receives funding from the European Centre for Disease Prevention and Control and the National Institute for Health Research (Health Protection Research Unit) for gonococcal whole-genome sequencing. YHG was supported by the NIH/NIAID grant R01 AI132606. KCM is supported by the NSF GRFP grant number DGE1745303. TDM is supported by the National Institute of Allergy and Infectious Diseases at the National Institutes of Health [1 F32 AI145157-01]. WMS is a recipient of a Senior Research Career Scientist Award from the Biomedical Laboratory Research and Development Service of the Department of Veterans. Work on antibiotic resistance in his laboratory is supported by NIH grants R37 AI-021150 and R01 AI-147609. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Veterans Affairs, The National Institutes of Health or the United States Government. The findings and conclusions in this article are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention. The WHO Collaborating Centre for Gonorrhoea and other STIs directed by MU receives funding from the European Centre for Disease Prevention and Control and the World Health Organization. This publication made use of the Neisseria Multi-Locus Sequence Typing website (https://pubmlst.org/neisseria/) sited at the University of Oxford (43) and funded by Wellcome and European Union.

Authors' contributions

DMA conceived the Pathogenwatch application. CY, RG, KA, BT, AU and DMA developed the Pathogenwatch application. LSB and DMA contributed to the conception and design of the work. CY and LSB generated, updated and benchmarked the *N. gonorrhoeae* AMR library. BT, CY, AU and LSB obtained, quality-checked and reassembled the raw data from the ENA. LSB revised the

716 assembled data, obtained all metadata available from the corresponding scientific publications
717 and created collections. LSB drafted the manuscript. LSB, DMA, CY, SA, KCM, TDM, MJC, YHG,
718 IM, BHR, WMS, GS, KT, TW and MU contributed to the acquisition, interpretation and discussion
719 of the data. LSB, CY and LSB analysed the data. All authors read and approved the final
720 manuscript.

721 ***Acknowledgements***

722 We would like to thank MJC, YHG, IM, BHR, WMS, GS, KT, TW and MU for their support on the
723 development of the *N. gonorrhoeae* Pathogenwatch scheme and the creation of the *N.*
724 *gonorrhoeae* Pathogenwatch Scientific Steering Group.

725

726 **References**

727 1. World Health Organization. Global priority list of antibiotic-resistant bacteria to guide
728 research, discovery, and development of new antibiotics.
729 http://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-
730 ET_NM_WHO.pdf. 2017.
731 2. Centres for Disease Prevention and Control. U.S. Department of Health and Human
732 Services. Atlanta, GA. Antibiotic Resistance Threats in the United States, 2019.
733 https://www.cdc.gov/drugresistance/pdf/threats-report/2019-ar-threats-report-508.pdf. 2019.
734 3. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al.
735 Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates,
736 2016. Bull World Health Organ. 2019;97(8):548-62P.
737 4. World Health Organization. Report on global sexually transmitted infection surveillance
738 2018. http://apps.who.int/iris/bitstream/handle/10665/277258/9789241565691-eng.pdf?ua=1.
739 2019.
740 5. Unemo M, Shafer WM. Antimicrobial resistance in Neisseria gonorrhoeae in the 21st
741 century: past, evolution, and future. Clin Microbiol Rev. 2014;27(3):587-613.
742 6. Cole MJ, Spiteri G, Town K, Unemo M, Hoffmann S, Chisholm SA, et al. Risk factors for
743 antimicrobial-resistant Neisseria gonorrhoeae in Europe. Sex Transm Dis. 2014;41(12):723-9.
744 7. Wi T, Lahra MM, Ndowa F, Bala M, Dillon JR, Ramon-Pardo P, et al. Antimicrobial
745 resistance in Neisseria gonorrhoeae: Global surveillance and a call for international
746 collaborative action. PLoS Med. 2017;14(7):e1002344.
747 8. Unemo M, Lahra MM, Cole M, Galarza P, Ndowa F, Martin I, et al. World Health
748 Organization Global Gonococcal Antimicrobial Surveillance Program (WHO GASP): review of
749 new data and evidence to inform international collaborative actions and research efforts. Sex
750 Health. 2019;16(5):412-25.
751 9. Fifer H, Natarajan U, Jones L, Alexander S, Hughes G, Golparian D, et al. Failure of
752 Dual Antimicrobial Therapy in Treatment of Gonorrhea. N Engl J Med. 2016;374(25):2504-6.
753 10. Eyre DW, Sanderson ND, Lord E, Regisford-Reimmer N, Chau K, Barker L, et al.
754 Gonorrhoea treatment failure caused by a Neisseria gonorrhoeae strain with combined
755 ceftriaxone and high-level azithromycin resistance, England, February 2018. Euro Surveill.
756 2018;23(27).

11. Whiley DM, Jennison A, Pearson J, Lahra MM. Genetic characterisation of Neisseria gonorrhoeae resistant to both ceftriaxone and azithromycin. Lancet Infect Dis. 2018;18(7):717-8.
12. Lahra MM, Martin I, Demczuk W, Jennison AV, Lee KI, Nakayama SI, et al. Cooperative Recognition of Internationally Disseminated Ceftriaxone-Resistant Neisseria gonorrhoeae Strain. Emerg Infect Dis. 2018;24(4).
13. Fifer H, Saunders J, Soni S, Sadiq ST, FitzGerald M. 2018 UK national guideline for the management of infection with Neisseria gonorrhoeae. Int J STD AIDS. 2020;31(1):4-15.
14. Deguchi T, Yasuda M, Asano M, Tada K, Iwata H, Komeda H, et al. DNA gyrase mutations in quinolone-resistant clinical isolates of Neisseria gonorrhoeae. Antimicrob Agents Chemother. 1995;39(2):561-3.
15. Tanaka M, Takahashi K, Saika T, Kobayashi I, Ueno T, Kumazawa J. Development of fluoroquinolone resistance and mutations involving GyrA and ParC proteins among Neisseria gonorrhoeae isolates in Japan. J Urol. 1998;159(6):2215-9.
16. Harris SR, Cole MJ, Spiteri G, Sanchez-Buso 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):758-68.
17. Spratt BG. Hybrid penicillin-binding proteins in penicillin-resistant strains of Neisseria gonorrhoeae. Nature. 1988;332(6160):173-6.
18. Ohnishi M, Golparian D, Shimuta K, Saika T, Hoshina S, Iwasaku 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):3538-45.
19. Grad YH, Harris SR, Kirkcaldy RD, Green AG, Marks DS, Bentley SD, et al. Genomic epidemiology of gonococcal resistance to extended spectrum cephalosporins, macrolides, and fluoroquinolones in the US, 2000-2013. J Infect Dis. 2016;214:1579-87.
20. Abrams AJ, Kirkcaldy RD, Pettus K, Fox JL, Kubin G, Trees DL. 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(8):492-4.
21. Lindberg R, Fredlund H, Nicholas R, Unemo M. Neisseria gonorrhoeae isolates with reduced susceptibility to cefixime and ceftriaxone: association with genetic polymorphisms in penA, mtrR, porB1b, and ponA. Antimicrob Agents Chemother. 2007;51(6):2117-22.
22. Palace SG, Wang Y, Rubin DH, Welsh MA, Mortimer TD, Cole K, et al. RNA polymerase mutations cause cephalosporin resistance in clinical Neisseria gonorrhoeae isolates. Elife. 2020;9.
23. Unemo M, Golparian D, Sanchez-Buso L, Grad Y, Jacobsson S, Ohnishi M, et al. The novel 2016 WHO Neisseria gonorrhoeae reference strains for global quality assurance of laboratory investigations: phenotypic, genetic and reference genome characterization. J Antimicrob Chemother. 2016;71(11):3096-108.
24. Wadsworth CB, Arnold BJ, Sater MRA, Grad YH. Azithromycin Resistance through Interspecific Acquisition of an Epistasis-Dependent Efflux Pump Component and Transcriptional Regulator in Neisseria gonorrhoeae. mBio. 2018;9(4).
25. Rouquette-Loughlin CE, Reimche JL, Balthazar JT, Dhulipala V, Gernert KM, Kersh EN, et al. Mechanistic Basis for Decreased Antimicrobial Susceptibility in a Clinical Isolate of Neisseria gonorrhoeae Possessing a Mosaic-Like mtr Efflux Pump Locus. mBio. 2018;9(6).
26. Ma KC, Mortimer TD, Duckett MA, Hicks AL, Wheeler NE, Sánchez-Busó L, et al. Increased power from bacterial genome-wide association conditional on known effects identifies Neisseria gonorrhoeae macrolide resistance mutations in the 50S ribosomal protein L4. bioRxiv. 2020.
27. Ma KC, Mortimer TD, Hicks AL, Wheeler NE, Sánchez-Busó L, Golparian D, et al. Increased antibiotic susceptibility in Neisseria gonorrhoeae through adaptation to the cervical environment. bioRxiv. 2020.
28. Unemo M, Dillon JA. Review and international recommendation of methods for typing neisseria gonorrhoeae isolates and their implications for improved knowledge of gonococcal epidemiology, treatment, and biology. Clin Microbiol Rev. 2011;24(3):447-58.

812 29. Loman NJ, Pallen MJ. Twenty years of bacterial genome sequencing. Nat Rev Microbiol.
813 2015;13(12):787-94.
814 30. Balloux F, Bronstad Brynildsrud O, van Dorp L, Shaw LP, Chen H, Harris KA, et al. From
815 Theory to Practice: Translating Whole-Genome Sequencing (WGS) into the Clinic. Trends
816 Microbiol. 2018;26(12):1035-48.
817 31. Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to
818 study antimicrobial resistance. Nat Rev Genet. 2019;20(6):356-70.
819 32. Hendriksen RS, Bortolaia V, Tate H, Tyson GH, Aarestrup FM, McDermott PF. Using
820 Genomics to Track Global Antimicrobial Resistance. Front Public Health. 2019;7:242.
821 33. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid
822 large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691-3.
823 34. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions. Core
824 assignment. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/core-genome-
825 tree/core-assignment. Accessed 30 June 2020.
826 35. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a
827 new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol.
828 2012;19(5):455-77.
829 36. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions. Short
830 read assembly. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/short-read-
831 assembly. Accessed 30 June 2020.
832 37. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions.
833 Speciator. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/species-
834 assignment/speciator. Accessed 30 June 2020.
835 38. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions. Core
836 filter. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/core-genome-tree/core-filter.
837 Accessed 30 June 2020.
838 39. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions. Tree
839 construction. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/core-genome-
840 tree/tree-construction. Accessed 30 June 2020.
841 40. Bennett JS, Jolley KA, Sparling PF, Saunders NJ, Hart CA, Feavers IM, et al. Species
842 status of Neisseria gonorrhoeae: evolutionary and epidemiological inferences from multilocus
843 sequence typing. BMC Biol. 2007;5:35.
844 41. Harrison OB, Cole K, Peters J, Cresswell F, Dean G, Eyre DW, et al. Genomic analysis
845 of urogenital and rectal Neisseria meningitidis isolates reveals encapsulated hyperinvasive
846 meningococci and coincident multidrug-resistant gonococci. Sex Transm Infect. 2017;93(6):445-
847 51.
848 42. PubMLST. Neisseria Multi-Locus Sequence Typing website.
849 https://pubmlst.org/neisseria/. Accessed 30 June 2020.
850 43. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb
851 software, the PubMLST.org website and their applications. Wellcome Open Res. 2018;3:124.
852 44. Martin IM, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Rapid sequence-based
853 identification of gonococcal transmission clusters in a large metropolitan area. J Infect Dis.
854 2004;189(8):1497-505.
855 45. NG-MAST. Neisseria gonorrhoeae Multi-Antigen Sequence Typing database.
856 http://www.ng-mast.net/. Accessed 30 June 2020.
857 46. Demczuk W, Sidhu S, Unemo M, Whiley DM, Allen VG, Dillon JR, et al. Neisseria
858 gonorrhoeae Sequence Typing for Antimicrobial Resistance, a Novel Antimicrobial Resistance
859 Multilocus Typing Scheme for Tracking Global Dissemination of N. gonorrhoeae Strains. J Clin
860 Microbiol. 2017;55(5):1454-68.
861 47. NG-STAR. Neisseria gonorrhoeae Sequence Typing for Antimicrobial Resistance.
862 https://ngstar.canada.ca/. Accessed 30 June 2020.
863 48. Kwong JC, Gonçalves da Silva A, Dyet K, Williamson DA, Stinear TP, Howden BP, et al.
864 NGMASTER: in silico Multi-Antigen Sequence Typing for Neisseria gonorrhoeae. Microb
865 Genom. 2016;2(8):e000076.

49. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions. MLST. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/typing-methods/mlst. Accessed 30 June 2020.
50. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions. cgMLST clustering. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/cgmlst-clusters. Accessed 30 June 2020.
51. Centre for Genomic Pathogen Surveillance. Pathogenwatch technical descriptions. PAARSNP. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/antimicrobial-resistance-prediction/paarsnp. Accessed 30 June 2020.
52. Chisholm SA, Wilson J, Alexander S, Tripodo F, Al-Shahib A, Schaefer U, et al. An outbreak of high-level azithromycin resistant Neisseria gonorrhoeae in England. Sex Transm Infect. 2015;0:1-3.
53. Golparian D, Harris SR, Sanchez-Buso 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(1):116.
54. Demczuk W, Lynch T, Martin I, Van Domselaar G, Graham M, Bharat A, et al. Whole-genome phylogenomic heterogeneity of Neisseria gonorrhoeae isolates with decreased cephalosporin susceptibility collected in Canada between 1989 and 2013. J Clin Microbiol. 2015;53(1):191-200.
55. Demczuk W, Martin I, Peterson S, Bharat A, Van Domselaar G, Graham M, et al. Genomic Epidemiology and Molecular Resistance Mechanisms of Azithromycin-Resistant Neisseria gonorrhoeae in Canada from 1997 to 2014. J Clin Microbiol. 2016;54(5):1304-13.
56. Eyre DW, De Silva D, Cole K, Peters J, Cole MJ, Grad YH, et al. WGS to predict antibiotic MICs for Neisseria gonorrhoeae. J Antimicrob Chemother. 2017;72(7):1937-47.
57. Fifer H, Cole M, Hughes G, Padfield S, Smolarchuk C, Woodford N, et al. Sustained transmission of high-level azithromycin-resistant Neisseria gonorrhoeae in England: an observational study. Lancet Infect Dis. 2018;18(5):573-81.
58. Sanchez-Buso L, Golparian D, Corander J, Grad YH, Ohnishi M, Flemming R, et al. The impact of antimicrobials on gonococcal evolution. Nat Microbiol. 2019.
59. Grad YH, Kirkcaldy RD, Trees D, Dordel J, Harris SR, Goldstein E, et al. Genomic epidemiology of Neisseria gonorrhoeae with reduced susceptibility to cefixime in the USA: a retrospective observational study. Lancet Infect Dis. 2014;14(3):220-6.
60. Jacobsson S, Golparian D, Cole M, Spiteri G, Martin I, Bergheim T, et al. WGS analysis and molecular resistance mechanisms of azithromycin-resistant (MIC >2 mg/L) Neisseria gonorrhoeae isolates in Europe from 2009 to 2014. J Antimicrob Chemother. 2016.
61. Lee RS, Seemann T, Heffernan H, Kwong JC, Goncalves da Silva A, Carter GP, et al. Genomic epidemiology and antimicrobial resistance of Neisseria gonorrhoeae in New Zealand. J Antimicrob Chemother. 2018;73(2):353-64.
62. Town K, Harris S, Sanchez-Buso L, Cole MJ, Pitt R, Fifer H, et al. Genomic and Phenotypic Variability in Neisseria gonorrhoeae Antimicrobial Susceptibility, England. Emerg Infect Dis. 2020;26(3):505-15.
63. Yahara K, Nakayama SI, Shimuta K, Lee KI, Morita M, Kawahata T, et al. Genomic surveillance of Neisseria gonorrhoeae to investigate the distribution and evolution of antimicrobial-resistance determinants and lineages. Microb Genom. 2018;4(8).
64. 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(2):151-7.
65. European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 9.0. https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_9.0_Breakpoint_Tables.pdf. Accessed 30 June 2020.
66. Stevenson M, Nunes T, Heuer C, Marshall J, Sanchez J, Thornton R, et al. epiR: Tools for the Analysis of Epidemiological Data. R package version 1.0-14. https://CRAN.R-project.org/package=epiR. 2020.

67. 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. Bioinformatics. 2012;28(4):464-9.
68. 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):3812-6.
69. Underwood A. Global Health Research Unit (GHRU) assembly pipeline. https://gitlab.com/cgps/ghru/pipelines/assembly. Accessed 30 June 2020.
70. Babraham Bioinformatics. FastQC. https://bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 30 June 2020.
71. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
72. Song L, Florea L, Langmead B. Lighter: fast and memory-efficient sequencing error correction without counting. Genome Biol. 2014;15(11):509.
73. Low AJ, Koziol AG, Manninger PA, Blais B, Carrillo CD. ConFindr: rapid detection of intraspecies and cross-species contamination in bacterial whole-genome sequence data. PeerJ. 2019;7:e6995.
74. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 2016;17(1):132.
75. Li H. Seqtk. https://github.com/lh3/seqtk. Accessed 30 June 2020.
76. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011;27(21):2957-63.
77. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5.
78. Underwood A. Bactinspector. https://gitlab.com/antunderwood/bactinspector. Accessed 30 June 2020.
79. Underwood A. Qualifyr. https://gitlab.com/cgps/qualifyr. Accessed 30 June 2020.
80. European Nucleotide Archive API portal. https://www.ebi.ac.uk/ena/portal/api/. Accessed 30 June 2020.
81. Sánchez-Busó L. N. gonorrhoeae Pathogenwatch video demo. https://vimeo.com/434706201. 2020.
82. Centre for Genomic Pathogen Surveillance. Pathogenwatch collection. Sánchez-Busó et al. (2019). https://pathogen.watch/collection/9stz0m94hin7-sanchez-buso-et-al-2019.
83. Page A, Taylor B, Keane J. Multilocus sequence typing by blast from de novo assemblies against PubMLST. Journal of Open Source Software. 2016;1(8).
84. Ng LK, Martin I, Liu G, Bryden L. Mutation in 23S rRNA associated with macrolide resistance in Neisseria gonorrhoeae. Antimicrob Agents Chemother. 2002;46(9):3020-5.
85. Roberts MC, Chung WO, Roe D, Xia M, Marquez C, Borthagaray G, et al. Erythromycin-resistant Neisseria gonorrhoeae and oral commensal Neisseria spp. carry known rRNA methylase genes. Antimicrob Agents Chemother. 1999;43(6):1367-72.
86. Cousin S, Jr., Whittington WL, Roberts MC. Acquired macrolide resistance genes in pathogenic Neisseria spp. isolated between 1940 and 1987. Antimicrob Agents Chemother. 2003;47(12):3877-80.
87. Luna VA, Cousin S, Jr., Whittington WL, Roberts MC. Identification of the conjugative mef gene in clinical Acinetobacter junii and Neisseria gonorrhoeae isolates. Antimicrob Agents Chemother. 2000;44(9):2503-6.
88. Rouquette-Loughlin CE, Balthazar JT, Shafer WM. Characterization of the MacA-MacB efflux system in Neisseria gonorrhoeae. J Antimicrob Chemother. 2005;56(5):856-60.
89. Veal WL, Nicholas RA, Shafer WM. Overexpression of the MtrC-MtrD-MtrE efflux pump due to an mtrR mutation is required for chromosomally mediated penicillin resistance in Neisseria gonorrhoeae. J Bacteriol. 2002;184(20):5619-24.

973 90. Cousin SL, Jr., Whittington WL, Roberts MC. Acquired macrolide resistance genes and
974 the 1 bp deletion in the mtrR promoter in Neisseria gonorrhoeae. J Antimicrob Chemother.
975 2003;51(1):131-3.
976 91. Warner DM, Shafer WM, Jerse AE. Clinically relevant mutations that cause derepression
977 of the Neisseria gonorrhoeae MtrC-MtrD-MtrE Efflux pump system confer different levels of
978 antimicrobial resistance and in vivo fitness. Mol Microbiol. 2008;70(2):462-78.
979 92. Shafer WM, Balthazar JT, Hagman KE, Morse SA. Missense mutations that alter the
980 DNA-binding domain of the MtrR protein occur frequently in rectal isolates of Neisseria
981 gonorrhoeae that are resistant to faecal lipids. Microbiology. 1995;141 (Pt 4):907-11.
982 93. Tomberg J, Unemo M, Davies C, Nicholas RA. Molecular and structural analysis of
983 mosaic variants of penicillin-binding protein 2 conferring decreased susceptibility to expanded-
984 spectrum cephalosporins in Neisseria gonorrhoeae: role of epistatic mutations. Biochemistry.
985 2010;49(37):8062-70.
986 94. Tomberg J, Unemo M, Ohnishi M, Davies C, Nicholas RA. Identification of amino acids
987 conferring high-level resistance to expanded-spectrum cephalosporins in the penA gene from
988 Neisseria gonorrhoeae strain H041. Antimicrob Agents Chemother. 2013;57(7):3029-36.
989 95. Unemo M, Golparian D, Nicholas R, Ohnishi M, Gallay A, Sednaoui P. High-level
990 cefixime- and ceftriaxone-resistant Neisseria gonorrhoeae in France: novel penA mosaic allele
991 in a successful international clone causes treatment failure. Antimicrob Agents Chemother.
992 2012;56(3):1273-80.
993 96. Belland RJ, Morrison SG, Ison C, Huang WM. Neisseria gonorrhoeae acquires
994 mutations in analogous regions of gyrA and parC in fluoroquinolone-resistant isolates. Mol
995 Microbiol. 1994;14(2):371-80.
996 97. Rouquette-Loughlin C, Dunham SA, Kuhn M, Balthazar JT, Shafer WM. The NorM efflux
997 pump of Neisseria gonorrhoeae and Neisseria meningitidis recognizes antimicrobial cationic
998 compounds. J Bacteriol. 2003;185(3):1101-6.
999 98. Lindback E, Rahman M, Jalal S, Wretlind B. Mutations in gyrA, gyrB, parC, and parE in
1000 quinolone-resistant strains of Neisseria gonorrhoeae. APMIS. 2002;110(9):651-7.
1001 99. Hu M, Nandi S, Davies C, Nicholas RA. High-level chromosomally mediated tetracycline
1002 resistance in Neisseria gonorrhoeae results from a point mutation in the rpsJ gene encoding
1003 ribosomal protein S10 in combination with the mtrR and penB resistance determinants.
1004 Antimicrob Agents Chemother. 2005;49(10):4327-34.
1005 100. Morse SA, Johnson SR, Biddle JW, Roberts MC. High-level tetracycline resistance in
1006 Neisseria gonorrhoeae is result of acquisition of streptococcal tetM determinant. Antimicrob
1007 Agents Chemother. 1986;30(5):664-70.
1008 101. Ashford WA, Golash RG, Hemming VG. Penicillinase-producing Neisseria gonorrhoeae.
1009 Lancet. 1976;2(7987):657-8.
1010 102. Ropp PA, Hu M, Olesky M, Nicholas RA. Mutations in ponA, the gene encoding
1011 penicillin-binding protein 1, and a novel locus, penC, are required for high-level chromosomally
1012 mediated penicillin resistance in Neisseria gonorrhoeae. Antimicrob Agents Chemother.
1013 2002;46(3):769-77.
1014 103. Olesky M, Hobbs M, Nicholas RA. Identification and analysis of amino acid mutations in
1015 porin IB that mediate intermediate-level resistance to penicillin and tetracycline in Neisseria
1016 gonorrhoeae. Antimicrob Agents Chemother. 2002;46(9):2811-20.
1017 104. Bilgin N, Richter AA, Ehrenberg M, Dahlberg AE, Kurland CG. Ribosomal RNA and
1018 protein mutants resistant to spectinomycin. EMBO J. 1990;9(3):735-9.
1019 105. Unemo M, Golparian D, Skogen V, Olsen AO, Moi H, Syversen G, et al. Neisseria
1020 gonorrhoeae strain with high-level resistance to spectinomycin due to a novel resistance
1021 mechanism (mutated ribosomal protein S5) verified in Norway. Antimicrob Agents Chemother.
1022 2013;57(2):1057-61.
1023 106. Fiebelkorn KR, Crawford SA, Jorgensen JH. Mutations in folP associated with elevated
1024 sulfonamide MICs for Neisseria meningitidis clinical isolates from five continents. Antimicrob
1025 Agents Chemother. 2005;49(2):536-40.

107. Zhao S, Duncan M, Tomberg J, Davies C, Unemo M, Nicholas RA. Genetics of chromosomally mediated intermediate resistance to ceftriaxone and cefixime in Neisseria gonorrhoeae. Antimicrob Agents Chemother. 2009;53(9):3744-51.
108. Kersh EN, Allen V, Ransom E, Schmerer M, Cyr S, Workowski K, et al. Rationale for a Neisseria gonorrhoeae Susceptible-only Interpretive Breakpoint for Azithromycin. Clin Infect Dis. 2020;70(5):798-804.
109. Cole MJ, Tan W, Fifer H, Brittain C, Duley L, Hepburn T, et al. Gentamicin, azithromycin and ceftriaxone in the treatment of gonorrhoea: the relationship between antibiotic MIC and clinical outcome. J Antimicrob Chemother. 2020;75(2):449-57.
110. Town K, Field N, Harris SR, Sanchez-Buso L, Cole MJ, Pitt R, et al. Phylogenomic analysis of Neisseria gonorrhoeae transmission to assess sexual mixing and HIV transmission risk in England: a cross-sectional, observational, whole-genome sequencing study. Lancet Infect Dis. 2020;20(4):478-86.
111. Ison CA, Golparian D, Saunders P, Chisholm S, Unemo M. Evolution of Neisseria gonorrhoeae is a continuing challenge for molecular detection of gonorrhoea: false negative gonococcal porA mutants are spreading internationally. Sex Transm Infect. 2013;89(3):197-201.
112. 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(1):3988.
113. De Silva D, Peters J, Cole K, Cole MJ, Cresswell F, Dean G, et al. Whole-genome sequencing to determine transmission of Neisseria gonorrhoeae: an observational study. Lancet Infect Dis. 2016;16(11):1295-303.
114. Mortimer TD, Pathela P, Crawley A, Rakeman JL, Lin Y, Harris SR, et al. The distribution and spread of susceptible and resistant Neisseria gonorrhoeae across demographic groups in a major metropolitan center. medRxiv. 2020;2020.04.30.20086413.
115. Alfsnes K, Eldholm V, Olsen AO, Brynildsrud OB, Bohlin J, Steinbakk M, et al. Genomic epidemiology and population structure of Neisseria gonorrhoeae in Norway, 2016-2017. Microb Genom. 2020.
116. Thomas JC, Seby S, Abrams AJ, Cartee J, Lucking S, Vidyaprakash E, 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.
117. Schmerer MW, Abrams AJ, Seby S, Thomas JCt, Cartee J, Lucking S, et al. Genomic Characterization of Neisseria gonorrhoeae Strains from 2016 U.S. Sentinel Surveillance Displaying Reduced Susceptibility to Azithromycin. Antimicrob Agents Chemother. 2020;64(5).
118. Lan PT, Golparian D, Ringlander J, Van Hung L, Van Thuong N, Unemo M. Genomic analysis and antimicrobial resistance of Neisseria gonorrhoeae isolates from Vietnam in 2011 and 2015-16. J Antimicrob Chemother. 2020;75(6):1432-8.
119. Didelot X, Dordel J, Whittles LK, Collins C, Bilek N, Bishop CJ, et al. Genomic Analysis and Comparison of Two Gonorrhea Outbreaks. MBio. 2016;7(3).
120. Osnes MN, Didelot X, de Korne-Elenbaas J, Alfsnes K, Brynildsrud OB, Syversen G, et al. The sudden emergence of a Neisseria gonorrhoeae strain with reduced susceptibility to extended-spectrum cephalosporins, Norway. bioRxiv. 2020;2020.02.07.935825.
121. Cehovin A, Harrison OB, Lewis SB, Ward PN, Ngetsa C, Graham SM, et al. Identification of Novel Neisseria gonorrhoeae Lineages Harboring Resistance Plasmids in Coastal Kenya. J Infect Dis. 2018;218(5):801-8.
122. Buckley C, Forde BM, Trembizki E, Lahra MM, Beatson SA, Whiley DM. Use of whole genome sequencing to investigate an increase in Neisseria gonorrhoeae infection among women in urban areas of Australia. Sci Rep. 2018;8(1):1503.
123. Ryan L, Golparian D, Fennelly N, Rose L, Walsh P, Lawlor B, et al. Antimicrobial resistance and molecular epidemiology using whole-genome sequencing of Neisseria gonorrhoeae in Ireland, 2014-2016: focus on extended-spectrum cephalosporins and azithromycin. Eur J Clin Microbiol Infect Dis. 2018;37(9):1661-72.

124. Wind CM, de Vries E, Schim van der Loeff MF, van Rooijen MS, van Dam AP, Demczuk WHB, et al. Decreased Azithromycin Susceptibility of Neisseria gonorrhoeae Isolates in Patients Recently Treated with Azithromycin. Clin Infect Dis. 2017;65(1):37-45.
125. Ezewudo MN, Joseph SJ, Castillo-Ramirez S, Dean D, Del Rio C, Didelot X, et al. Population structure of Neisseria gonorrhoeae based on whole genome data and its relationship with antibiotic resistance. PeerJ. 2015;3:e806.
126. Wick RR, Holt KE. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. F1000Res. 2019;8:2138.
127. Souvorov A, Agarwala R, Lipman DJ. SKESA: strategic k-mer extension for scrupulous assemblies. Genome Biol. 2018;19(1):153.
128. Spiteri G, Cole M, Unemo M, Hoffmann S, Ison C, van de Laar M. The European Gonococcal Antimicrobial Surveillance Programme (Euro-GASP)--a sentinel approach in the European Union (EU)/European Economic Area (EEA). Sex Transm Infect. 2013;89 Suppl 4:iv16-8.
129. Schwarcz SK, Zenilman JM, Schnell D, Knapp JS, Hook EW, 3rd, Thompson S, et al. National surveillance of antimicrobial resistance in Neisseria gonorrhoeae. The Gonococcal Isolate Surveillance Project. JAMA. 1990;264(11):1413-7.
130. Public Health Agency of Canada National Microbiology Laboratory. National surveillance of antimicrobial susceptibilities of Neisseria gonorrhoeae. Annual summary 2017. https://www.canada.ca/en/public-health/services/publications/drugs-health-products/national-surveillance-antimicrobial-susceptibilities-neisseria-gonorrhoeae-annual-summary-2017.html. 2017.
131. Lahra MM, Enriquez R, George CRR. Australian Gonococcal Surveillance Programme Annual Report, 2017. Commun Dis Intell (2018). 2019;43.
132. Paine TC, Fenton KA, Herring A, Turner A, Ison C, Martin I, et al. GRASP: a new national sentinel surveillance initiative for monitoring gonococcal antimicrobial resistance in England and Wales. Sex Transm Infect. 2001;77(6):398-401.
133. Weston EJ, Wi T, Papp J. Strengthening Global Surveillance for Antimicrobial Drug-Resistant Neisseria gonorrhoeae through the Enhanced Gonococcal Antimicrobial Surveillance Program. Emerg Infect Dis. 2017;23(13).
134. Sirivongrangson P, Girdthep N, Sukwicha W, Buasakul P, Tongtoyai J, Weston E, et al. The first year of the global Enhanced Gonococcal Antimicrobial Surveillance Programme (EGASP) in Bangkok, Thailand, 2015-2016. PLoS One. 2018;13(11):e0206419.
135. Seale AC, Gordon NC, Islam J, Peacock SJ, Scott JAG. AMR Surveillance in low and middle-income settings - A roadmap for participation in the Global Antimicrobial Surveillance System (GLASS). Wellcome Open Res. 2017;2:92.
136. NIHR-GHRU. Global Health Research Unit, Genomic Surveillance of Antimicrobial Resistance. https://ghru.pathogensurveillance.net/. Accessed 30 June 2020.
137. Gottlieb SL, Ndowa F, Hook EW, 3rd, Deal C, Bachmann L, Abu-Raddad L, et al. Gonococcal vaccines: Public health value and preferred product characteristics; report of a WHO global stakeholder consultation, January 2019. Vaccine. 2020.
138. Marjuki H, Topaz N, Joseph SJ, Gernert KM, Kersh EN, Antimicrobial-Resistant Neisseria gonorrhoeae Working G, et al. Genetic Similarity of Gonococcal Homologs to Meningococcal Outer Membrane Proteins of Serogroup B Vaccine. mBio. 2019;10(5).
139. Russell MW, Jerse AE, Gray-Owen SD. Progress Toward a Gonococcal Vaccine: The Way Forward. Front Immunol. 2019;10:2417.
140. Taylor SN, Marrazzo J, Batteiger BE, Hook EW, 3rd, Sena AC, Long J, et al. Single-Dose Zoliflodacin (ETX0914) for Treatment of Urogenital Gonorrhea. N Engl J Med. 2018;379(19):1835-45.
141. Taylor SN, Morris DH, Avery AK, Workowski KA, Batteiger BE, Tiffany CA, et al. Gepotidacin for the Treatment of Uncomplicated Urogenital Gonorrhea: A Phase 2, Randomized, Dose-Ranging, Single-Oral Dose Evaluation. Clin Infect Dis. 2018;67(4):504-12.
142. Institute of Medicine BoGH, Forum on Microbial Threats. Antibiotic Resistance: Implications for Global Health and Novel Intervention Strategies: Workshop Summary. The

National Academies Collection: Reports funded by National Institutes of Health. Washington (DC)2010.

143. Demczuk W, Martin I, Sawatzky P, Allen V, Lefebvre B, Hoang L, et al. Equations To Predict Antimicrobial MICs in Neisseria gonorrhoeae Using Molecular Antimicrobial Resistance Determinants. Antimicrob Agents Chemother. 2020;64(3).

144. Hicks AL, Wheeler N, Sanchez-Buso L, Rakeman JL, Harris SR, Grad YH. Evaluation of parameters affecting performance and reliability of machine learning-based antibiotic susceptibility testing from whole genome sequencing data. PLoS Comput Biol. 2019;15(9):e1007349.