A community-driven resource for genomic epidemiology and antimicrobial resistance prediction of *Neisseria gonorrhoeae* at Pathogenwatch

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

[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] World Health Organization Collaborating Centre for Gonorrhoea and Other STIs, Department of Laboratory Medicine, Faculty of Medicine and Health, Örebro University, Örebro, Sweden.
[5] National Infection Service, Public Health England, London, United Kingdom.
[6] Division of Infectious Diseases, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America.
[7] National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada.
[8] 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.
[9] Department of Microbiology and Immunology and Emory Antibiotic Resistance Center, Emory University School of Medicine, Atlanta, Georgia, United States of America.
[10] Laboratories of Bacterial Pathogenesis, Veterans Affairs Medical Center, Decatur, Georgia, United States of America.
[11] European Centre for Disease Prevention and Control, Stockholm, Sweden.
[12] Department of the Global HIV, Hepatitis and STI programmes, World Health Organization, Geneva, Switzerland.
[13] Microbiotica, Biodata Innovation Centre, Hinxton, Cambridgeshire, United Kingdom.

[^]Members of the *N. gonorrhoeae* Pathogenwatch Scientific Steering Group at the time of publication of this manuscript.
[#]Current affiliations; Leonor Sánchez Busó: Genomics and Health Area, Foundation for the Promotion of Health and Biomedical Research in the Valencian Community (FISABIO-Public Health), Valencia, Spain; Richard J. Goater: European Molecular Biology Lab, Heidelberg, Baden-Wuerttemberg, Germany.
[*]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. Whole genome sequencing (WGS) data can be used to identify new AMR clones, 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* at Pathogenwatch (https://pathogen.watch/ngonorrhoeae). An international advisory group of experts in epidemiology, public health, genetics and genomics of *N. gonorrhoeae* was convened to inform on the utility of current and future analytics in the platform. We implement backwards compatibility with MLST, NG-MAST and NG-STAR typing schemes as well as an exhaustive library of genetic AMR determinants linked to a genotypic prediction of 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: AMR prediction from genome data revealed specificity values over 99% for azithromycin, ciprofloxacin and ceftriaxone and sensitivity values around 99% for benzylpenicillin and tetracycline. A case study using the Pathogenwatch collection of *N. gonorrhoeae* public genomes showed the global expansion of an azithromycin resistant lineage carrying a mosaic *mtr* over at least the last 10 years, emphasizing the power of Pathogenwatch to explore and evaluate genomic epidemiology questions of public health concern.

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. The advisory group will assess and identify ongoing public health needs in the field of gonorrhoea, particularly regarding gonococcal AMR, in order to further enhance utility with modified or new analytic methods.

82 **Background**

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

97 Strains of *N. gonorrhoeae* resistant to every recommended treatment have rapidly emerged,
98 including resistance to penicillins, tetracyclines, fluoroquinolones, macrolides and the extended-
99 spectrum cephalosporins (ESCs) (5-8). The current recommended treatment in many countries
100 is a dual therapy with injectable ceftriaxone plus oral azithromycin, although reports of decreased
101 susceptibility to ceftriaxone as well as azithromycin resistance have increased globally (7, 8). One
102 case of failure of dual treatment was reported in 2016 in the United Kingdom (UK) (9). Additionally,
103 in 2018 a gonococcal strain with resistance to ceftriaxone combined with high-level resistance to
104 azithromycin was detected in both the UK and Australia (10). The transmission of a ceftriaxone-
105 resistant clone (FC428) has been documented internationally since 2015, raising concerns about
106 the long-term effectiveness of the current treatment in the absence of an available alternative
107 (11). In some countries such as in Japan, China and since 2019 in the UK, a single dose of
108 ceftriaxone 1 gram is the recommended treatment due to the increasing incidence of azithromycin

109 resistance in *N. gonorrhoeae* and other STI pathogens such as *Mycoplasma genitalium* (12).
110 Extensive investigations have been ongoing for years to unveil the genetic mechanisms that
111 explain most of the observed susceptibility patterns for the main classes of antimicrobials for *N.*
112 *gonorrhoeae*. For ciprofloxacin, nearly all resistant strains have the GyrA S91F amino acid
113 alteration (13-15), however, resistance prediction from genomic data is not as straightforward for
114 other antibiotics. Known resistance mechanisms often involve additive or suppressive effects as
115 well as epistatic interactions that all together explain just part of the observed phenotypic
116 resistance. For example, there is good evidence that many mosaic structures of the *penA* gene
117 are associated with decreased susceptibility to ESCs (16, 17), however, mosaics do not explain
118 all cases of ESC resistance, especially for ceftriaxone, and some mosaic *penA* alleles do not
119 cause decreased susceptibility or resistance to this antibiotic (16-19). On top of these, variants
120 that overexpress the MtrCDE efflux pump, mutations in *porB1b* that reduce drug influx and non-
121 mosaic mutations in penicillin-binding proteins also contribute to decreased susceptibility to ESCs
122 (20). Furthermore, mutations in the *rpoB* and *rpoD* genes, encoding subunits of the RNA
123 polymerase, have been recently related to resistance to ESCs in clinical *N. gonorrhoeae* isolates
124 (21). Mutations in the 23S rRNA gene (A2045G and C2597T in *N. gonorrhoeae* nomenclature,
125 coordinates from the WHO 2016 reference panel (22), A2059G and C2611T in *Escherichia coli*)
126 are frequently associated with azithromycin resistance, as do variants in *mtrR* or its promoter that
127 increase the expression of the MtrCDE efflux pump (5). Recently, epistatic interactions between
128 a mosaic *mtr* promoter region and a mosaic *mtrD* gene have also been reported to increase the
129 expression of this pump, contributing to macrolide resistance (23, 24). Mutations in *rplD* have also
130 been associated with reduced susceptibility to this antibiotic (25) and contrarily, loss-of-function
131 mutations in *mtrC* have been linked to increased susceptibility to several antibiotics including
132 azithromycin (26). Thus, we can relatively confidently predict decreased susceptibility or
133 resistance to an antimicrobial using the current known genetic mechanisms, however, phenotypic
134 testing is still necessary to detect resistant cases caused by unknown or novel mechanisms.
135 These inconsistencies with the genomic data will allow the discovery of these new mechanisms,
136 which will keep improving the resistance predictions from WGS.

137 A myriad of methods have been used to discriminate among strains of *N. gonorrhoeae*, from
138 phenotypic to DNA-based techniques (27), but whole genome sequencing (WGS) can provide
139 the complete genome information of a bacterial strain. The cost of amplifying all loci of the different
140 typing schemes via nucleic acid amplification and traditional Sanger sequencing can be more
141 expensive than the cost of WGS of one bacterial genome in many settings. With WGS, multiple
142 genetic AMR mechanisms as well as virulence and typing regions can be targeted simultaneously
143 with the appropriate bioinformatic tools and pipelines. It also provides a significant improvement
144 in resolution and accuracy over traditional molecular epidemiology and typing methods, allowing
145 a genome-wide comparison of strains that can: identify AMR clones, outbreaks, transmission
146 networks, national and international spread, known and novel resistance mechanisms as well as
147 also inform on the development of point-of-care tests for antimicrobial susceptibility, novel
148 antimicrobials and vaccines (28, 29). However, implementation of WGS for genomic surveillance
149 poses practical challenges, especially for Low- and Middle-Income Countries (LMICs), due to the
150 need of a major investment to acquire and maintain the required infrastructure.

151 WGS produces a very high volume of data that needs to be pre-processed and analysed using
152 bioinformatics. Bioinformatics expertise is not always readily available in laboratory and public
153 health settings, and currently there are no international standards and proficiency trials for which
154 algorithms to use to process WGS data. There are several open-source tools specialised in each
155 step of the pipeline as well as proprietary software containing workflows that simplify the analyses.
156 However, these are less customizable and may not be affordable for all (30, 31). Choosing the
157 best algorithms and parameters when analysing genomic data is not straightforward as it requires
158 a fair knowledge of the pathogen under study and its genome diversity. Multiple databases
159 containing genetic determinants of AMR for bacterial pathogens are available (30, 31), however,
160 choosing which one is most complete for a particular organism frequently requires an extensive
161 literature search. Public access web-based species-specific tools and AMR databases revised
162 and curated by experts would be the most approachable option for both well-resourced and LMICs
163 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, clinical and 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 that can be applied to other pathogens.

Methods

The Pathogenwatch platform: technical summary

Pathogenwatch is a web-based platform with several different components. The main interface is a React (32) single-page application with a style based on Material Design Lite (33). Phylogenetic trees are plotted using Phylocanvas (34), maps using Leaflet (35) and networks with Sigma (36). The back end is written in Node.js and contains an API service for the user interface and four "Runner" services for the following analyses: species prediction, single-genome analyses, tree building and core genome multi-locus sequence typing (cgMLST) clustering. Docker containers are used for queuing tasks, streaming input or result files through standard input and storing JSON data from standard output. A MongoDB cluster is used for data storage and task queuing/synchronisation. Pathogenwatch shares some visualization components with Microreact

189 (37), such as those associated with the phylogenetic tree and the map. However, Pathogenwatch
190 includes an analytical framework which is unique to this platform.

191 ***Generation of the N. gonorrhoeae core genome library***

192 Pathogenwatch implements a library of core genome sequences for several supported organisms.
193 In the case of *N. gonorrhoeae*, a core gene set was built from the 14 finished reference genomes
194 that constitute the 2016 WHO reference strain panel (22) using the pangenome analysis tool
195 Roary (38) as described in Harris *et al* (2018) (15). Briefly, the minimum percentage of identity for
196 blastp was set to 97% and the resulting core genes were aligned individually using MAFFT. The
197 resulting genes with a percentage of identity above 99% were post-processed as described in
198 (39). Representatives for each family were selected by choosing the sequence with the fewest
199 differences to the others on average and searched using tblastn (percentage of identity >= 80%,
200 E-value <= 1e-35) against the 14 high quality reference genomes. Families without a complete
201 match in every reference (100% coverage) or had multiple matches were removed. Overlapping
202 genes from each reference were merged into pseudocontigs and grouped by gene composition.
203 For each family, a representative was selected as before and searched/filtered using the
204 references as before. The final core gene set contains 1,542 sequences that span a total of
205 1,470,119 nucleotides (approximately 67% of a typical *N. gonorrhoeae* genome length, 2.2Mb).
206 A BLAST database was constructed from these core segments and used to profile new
207 assemblies.

208 ***Profiling new assemblies***

209 New genome assemblies can be uploaded by a user (drag and drop) or calculated from high-
210 throughput short read data directly within Pathogenwatch using SPAdes (40) as described in (41).

211 A taxonomy assignment step for species identification is performed on the uploaded assemblies
212 by using Speciator (42). New assemblies are then queried against a species-specific BLAST
213 database using blastn. For *N. gonorrhoeae*, every core loci needs to match at least 80% of its
214 length to be considered as present. Further filtering steps are applied to remove loci that can be

problematic for tree building, such as paralogs or loci with unusually large number of variant sites compared to an estimated substitution rate on the rest of the genome, as described in (43). The overall substitution rate is calculated as the number of total differences in the core library divided by the total number of nucleotides. Indels are ignored to minimise the noise that could be caused by assembly or sequencing errors. The expected number of substitutions per locus is determined by multiplying this substitution rate by the length of the representative sequence.

The number of substitutions observed for each locus between the new assembly and the reference sequence are scaled to the total number of nucleotides that match the core library, creating a pairwise score that is saved on a distance matrix and is used for Neighbour-Joining tree construction, as described in (44).

Algorithms for sequence typing and cgMLST clustering

Alleles and sequence types (STs) for Multi-Locus Sequence Typing (MLST) (45) and cgMLST (core genome MLST, *N. gonorrhoeae* cgMLST v1.0) (46) were obtained from PubMLST (47, 48), for *N. gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) (49) from (50) and for *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) (51) from (52) (Table 1). A search tool implemented as part of Pathogenwatch is used to make the assignments for MLST, cgMLST and NG-STAR, while NGMASTER (53) 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 (54). 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 using single linkage clustering as described in (55). Users can select the clustering threshold (i.e. number of loci with differing alleles) and a network graph based on the SLINK (56) algorithm is calculated within individual genome reports.

241 ***AMR library and detection of genetic AMR determinants***

242 Genes and point mutations (single nucleotide polymorphisms (SNPs) and indels) were detected
243 using Pathogenwatch AMR v2.4.9 (57). Pathogenwatch AMR also provides a prediction of AMR
244 phenotype inferred from the combination of identified mechanisms. Genetic determinants
245 described in the literature as involved in AMR in *N. gonorrhoeae* were collated into a library in
246 TOML format (version 0.0.11). A test dataset containing 3,987 isolates from 13 studies (15, 18,
247 22, 58-67) (Additional file 1: Table S1) providing minimum inhibitory concentration (MIC)
248 information for six antibiotics (benzylpenicillin, tetracycline, ciprofloxacin, cefixime, ceftriaxone
249 and azithromycin) was used to benchmark and to curate this library. A validation benchmark was
250 posteriorly run with a dataset of 1,607 isolates from 3 other publications (68-70) with MIC
251 information for the same six antibiotics plus spectinomycin (Additional file 1: Table S1). EUCAST
252 clinical breakpoints v9.0 (71) were used to define susceptibility (S), susceptibility with an
253 increased exposure (I) or resistance (R) (SIR) categorical interpretations of MICs for all antibiotics
254 except for azithromycin, for which the EUCAST epidemiological cut-off (ECOFF) was used to
255 define non-susceptibility/resistance (ECOFF>1mg/L). As a result of the benchmark analyses,
256 sensitivity, specificity and positive/negative predictive values (PPV/NPV) were obtained for the
257 AMR mechanisms implemented in the library and, globally, for each of the antibiotics. Confidence
258 intervals (95%) for these statistics were calculated using the *epi.tests* function in the *epiR* R
259 package v1.0-14 (72). Individual or combined AMR mechanisms with a PPV below 15% were
260 discarded from the library to optimise the overall predictive values. Visual representations of the
261 observed ranges of MIC values for a particular antibiotic for each of the observed combinations
262 of genetic AMR mechanisms on the test dataset were used to identify and assess combinations
263 of mechanisms that have an additive or suppressive effect on AMR. These were included in the
264 library.

265 As part of the accuracy testing of the AMR library, we ran the 2016 WHO *N. gonorrhoeae*
266 reference genomes 2016 panel (n=14) through Pathogenwatch and compared the detected list
267 of genetic AMR mechanisms with the list published in the original study (22). For the WHO U

268 strain, a discrepancy on a mutation in *parC* was further investigated by mapping the original raw
269 Illumina data (European Nucleotide Archive (ENA) run accession ERR449479) to the reference
270 genome assembly (ENA genome accession LT592159.1) and visualized using Artemis (73).

271 In short-read assemblies, the four copies of the 23S rRNA gene are collapsed into one, thus the
272 detection of the A2045G and C2597T mutations is dependent on the consensus bases resulting
273 from the number of mutated copies (63, 66, 74).

274 ***Quality check and assembly of public sequencing data***

275 Public *N. gonorrhoeae* genomes with geolocation data were obtained from the ENA in November
276 2019. This list was complemented by an exhaustive literature search of studies on *N. gonorrhoeae*
277 genomics without metadata submitted to the ENA but instead made available as supplementary
278 information in the corresponding publications. Raw paired-end short read data from a list of
279 12,192 isolates was processed with the GHRU assembly pipeline v1.5.4 (75). This pipeline runs
280 a Nextflow workflow to quality-check (QC) paired-end short read fastq files before and after
281 filtering and trimming, assembles the data and quality-checks the resulting assembly. Results
282 from the pipeline are provided in Additional file 2. In this pipeline, QC of short reads was performed
283 using FastQC v0.11.8 (76). Trimming was done with Trimmomatic v0.38 (77) by cutting bases
284 from the start and end of reads if they were below a Phred score of 25, trimming using a sliding
285 window of size 4 and cutting once the average quality within the window fell below a Phred score
286 of 20. Only reads with length above a third of the original minimum read length were kept for
287 further analyses. After trimming, reads were corrected using the kmer-based approach
288 implemented in Lighter v1.1.1 (78) with a kmer length of 32 bp and a maximum number of
289 corrections allowed within a 20 bp window of 1. ConFindr v0.7.2 was used to assess intra- and
290 inter-species contamination (79). Mash v2.1 (80) was applied to estimate genome size using a
291 kmer size of 32 bp and Seqtk v1.3 (81) to down sample fastq files if the depth of coverage was
292 above 100x. Flash v1.2.11 (82) was used to merge reads with a minimum overlap length of 20 bp
293 and a maximum overlap of 100 bp to facilitate the subsequent assembly process. SPAdes v3.12

294 (40) was used for genome assembly with the --careful option selected to reduce the number of
295 mismatches and short indels with a range of kmer lengths depending on the minimum read length.
296 The final assemblies were quality-checked using Quast v5.0.2 (83) and ran through the species
297 identification tool Bactinspector (84). QC conditions were assessed and summarised using
298 Qualifyr (85).

299 Fastq files with poor quality in which the trimming and filtering step discarded all reads from either
300 one or both pairs were excluded from the analyses because the assembly pipeline is optimised
301 for paired-end data. Assemblies with an N50 below 25,000 bp, a number of contigs above 300, a
302 total assembly length above 2.5 Mb or a percentage of contamination above 5% were also
303 excluded.

304 ***Metadata for public genomes***

305 Geolocation data (mainly country), collection dates (day, month and year when available), ENA
306 project accession and associated Pubmed ID were obtained from the ENA API for all the genomes
307 in the pipeline (86). A manual extensive literature search was performed to identify the
308 publications containing the selected genomes. In order to complete published studies as much
309 as possible, extra genomes were downloaded and added to the dataset. Metadata for the final
310 set was completed with the information contained in supplementary tables on the corresponding
311 publications, including phenotypic antimicrobial susceptibility data. Submission date was
312 considered instead of collection date when the latter was not available, however, this occurred in
313 only a few cases (<0.5%).

314 ***Creation of the N. gonorrhoeae Pathogenwatch Scientific Steering Group***

315 International experts in the field of *N. gonorrhoeae* AMR, microbiology, genetics, genomics,
316 epidemiology and public health were approached and agreed to participate as members of the
317 '*N. gonorrhoeae* Pathogenwatch Scientific Steering Group' in order to discuss the analytics in
318 Pathogenwatch and make sure they met the current needs of the public health and scientific
319 community. During the updates made to the platform and the preparation of this manuscript, these

320 experts participated in virtual sessions to discuss the list of genetic AMR determinants and their
321 association with SIR categories (Table 2) based on experimental and/or computational evidence.
322 Some of the members of the group had previously been directly involved in many of these studies.
323 Other current and future updates were also discussed, such as the inclusion of the NG-STAR
324 typing scheme (51) and the organization of published genomes into public collections, data
325 sharing, privacy and the interconnectivity of Pathogenwatch with other platforms, such as
326 PubMLST (48) or the ENA. The group will regularly discuss new updates to the platform.

327 ***Data sharing and privacy***

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

336

337 **Results**

338 ***N. gonorrhoeae genome analytics in Pathogenwatch***

339 Pathogenwatch is a web-based platform for epidemiological surveillance using genome
340 sequencing data. After upload, different analytics are run simultaneously (Figure 1): cgMLST (46),
341 MLST (45), NG-MAST (49) and NG-STAR (51) typing schemes (Table 1), a genotypic prediction
342 of phenotypic resistance using a customized AMR library (Table 2) that includes known genetic
343 AMR mechanisms for 8 antimicrobials, as well as statistics on the quality of the assemblies
344 (Additional file 3: Figure S1). These analytical features differentiate Pathogenwatch from a parallel

345 platform from the same group, Microreact (37), which shares one of the main layouts with
346 Pathogenwatch (a phylogenetic tree, a map and a table or timeline), but it is intended for
347 visualization of pre-computed phylogenetic trees with accompanying metadata, while
348 Pathogenwatch also includes analytical tools.

349 Genomes from one or multiple studies can be grouped into collections (Figure 2 and Additional
350 file 3: Figure S2), and the genomic data are automatically processed by comparing with a core *N.*
351 *gonorrhoeae* genome built from WHO reference strain genomes (15, 22). A phylogenetic tree is
352 obtained as a result, representing the genetic relationship among the isolates in the collection.
353 Metadata can be uploaded at the same time as the genome data, and if the collection location
354 coordinates for an isolate are provided, this information is plotted into a map (Additional file 3:
355 Figure S1). If date or year of isolation is also provided, this information is represented in a timeline.
356 The three panels on the main collection layout - the tree, the map and a table or timeline – are
357 functionally integrated so filters and selections made by the user update all of them
358 simultaneously. Users can also easily switch among the metadata and the results of the main
359 analytics: typing, genome assembly statistics, genotypic AMR prediction, AMR-associated SNPs,
360 AMR-associated genes and the timeline (Additional file 3: Figure S1). cgMLST is used for finding
361 close genomes in the database based on allele differences to one individual isolate (Additional
362 file 3: Figure S3). A video demonstrating the usage and main features of Pathogenwatch is
363 available (87). Notes on data sharing and privacy are available in the Methods section.

364

365

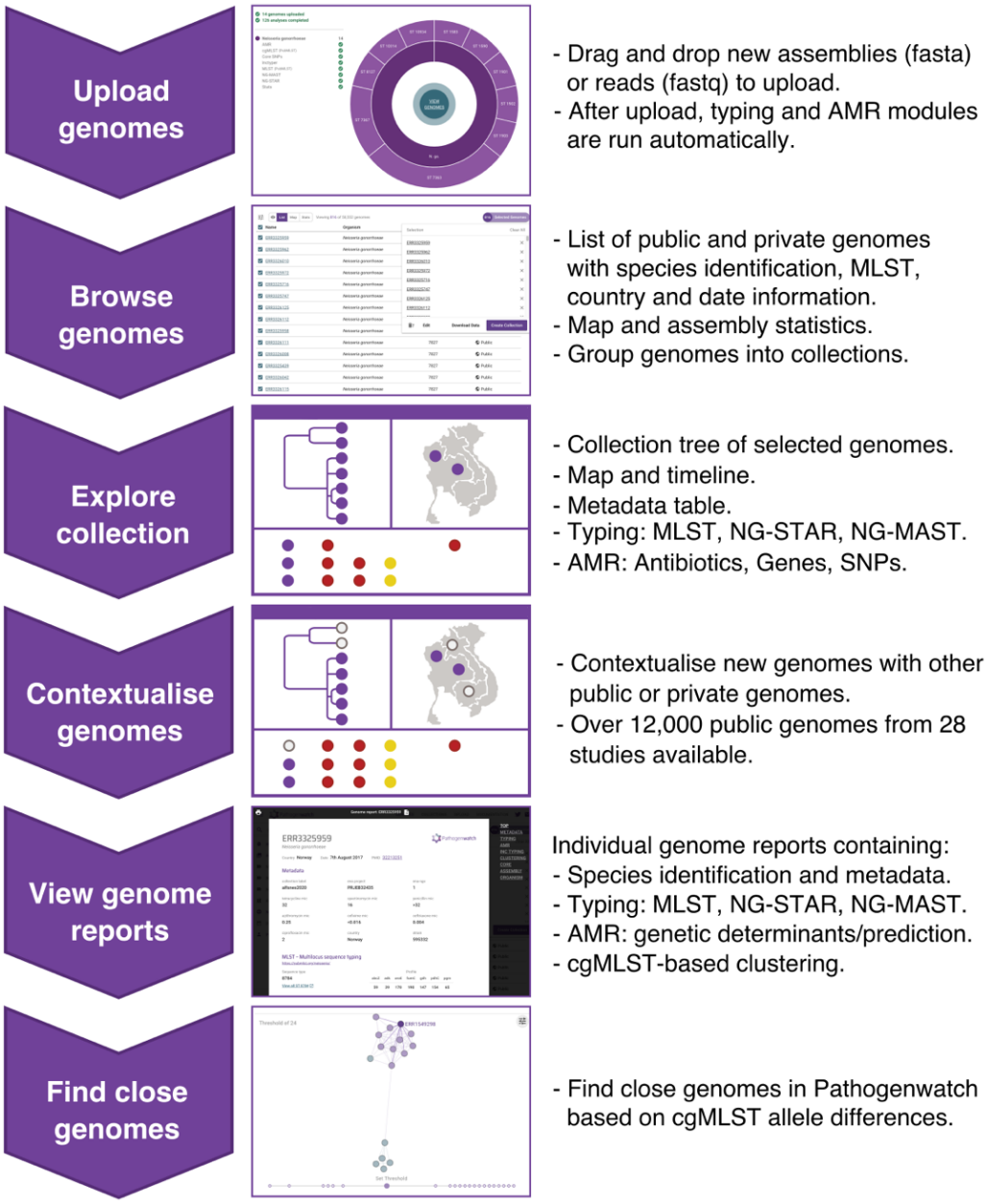


Figure 1. Main workflow in Pathogenwatch. New genomes can be uploaded and combined with public data for contextualisation. The collection view allows data exploration through a combined phylogenetic tree, a map, a timeline and the metadata table, which can be switched to show typing information (Multi-Locus 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.

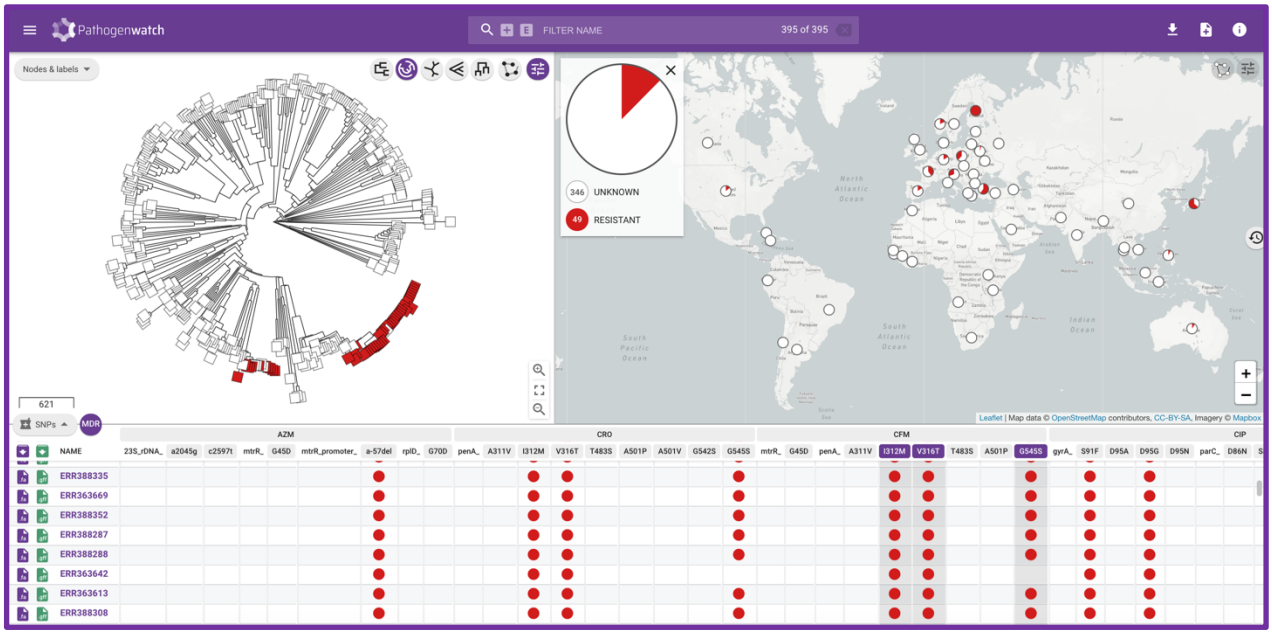


Figure 2. Main display of a Pathogenwatch collection, showing a phylogenetic tree, a map and a table of SNPs associated with AMR of 395 *N. gonorrhoeae* genomes from a global study (64, 88). Isolates carrying three mosaic *penA* marker mutations are marked in red in the tree and the map. The table can 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 3: Figure S1.

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

Typing scheme*	Loci (number)	Note	Pathogenwatch implementation	References
cgMLST	(N=1,649)	*N. gonorrhoeae* cgMLST v1.0	Typing algorithm, database from PubMLST	(46-48, 89)
MLST	*abcZ, adk, aroE, fumC, gdh, pdhC, pgm* (N=7)	Housekeeping genes in *Neisseria* spp.	In-house typing tool, database from PubMLST	(45, 47, 48, 89)
NG-MAST	*porB, tbpB* (N=2)	Genes encoding highly-variable membrane proteins	NG-MASTER, database from NG-MAST website	(49, 50, 53)
NG-STAR	*penA, mtrR, porB, ponA, gyrA, parC, 23S rDNA* (N=7)	Genes involved in antimicrobial resistance	In-house typing tool, database from NG-STAR website	(51, 52, 89)

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

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

408

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	(74)
	23S rDNA 2597C>T substitution (2611C>T in *E. coli*)	R	(90)
	ermA, ermB, ermC, ermF genes	R	(91, 92)
	ereA, ereB genes	R	(22)
	mefA gene	R	(92, 93)
	macAB promoter -48G>T substitution*	R	(94)
	mtrR promoter mosaic**		
	N. meningitidis-like mosaic (n=1)	R	(23)
	N. lactamica-like mosaic (n=2)	R	(23)
	mtrD mosaic**		
	N. meningitidis-like mosaic (n=1)	R	(23)
	N. lactamica-like mosaic (n=2)	R	(23)
	mtrR promoter -57delA*	A	(95, 96)
	mtrR G45D	A	(97, 98)
	mtrC loss-of-function	N	(26)
	rplV ARAK tandem duplication (position 90)	R	(18)
	rplV KGPSLK tandem duplication (position 83)	R	(18)
	rplD G70D	A	(25)
Ceftriaxone*** (R: MIC>0.125)	*penA* mosaic (A311V, I312M, V316P/T, T483S and G545S)	R	(99-101)
	penA V316P, T483S, A501P/V, G542S	R	(99, 100)
	rpoB P157L, G158V, R201H	R	(21)
	rpoD D92-95 deletion, E98K	I	(21)
Cefixime*** (R: MIC>0.125)	*mtrR* G45D	A	(97, 98)
	penA mosaic (I312M, V316T, G545S)	R	(99-101)
	penA mosaic (A311V, I312M, V316P/T, T483S and G545S)	R	(99-101)
	penA V316P, T483S, A501P	I	(99, 100)
	rpoB P157L, G158V, R201H	I	(21)
	rpoD D92-95 deletion, E98K	I	(21)
Ciprofloxacin (I: 0.03<MIC≤0.06; R: MIC>0.06)	*gyrA* S91F, D95A/N	R	(102)
	gyrA D95G	I	(102)
	norM promoter -7A>G, -104C>T substitutions*	I	(103)
	parC D86N, S87R	R	(102)
	parC S87I/N, S88P, E91K	I	(102)
	parE G410V	I	(104)
Tetracycline**** (I: 0.5<MIC≤1; R: MIC>1)	*mtrR* A39T, G45D	A	(97, 98)
	mtrR loss-of-function	I	(22)
	mtrR promoter -56A>C substitution, -57delA deletion*	I	(23, 95, 96)
	mtrR promoter -131G>A (*mtrC* -120G>A substitution, *mtr120)**	I	(97)
	rpsJ V57M	I	(105)
	tetM gene	R	(106)

Penicillins (I: 0.06<MIC≤1; R: MIC>1)	*blaTEM* gene	R	(107)
	mtrR G45D	I	(97, 98)
	mtrR A39T	A	(97)
	mtrR loss-of-function	I	(22)
	mtrR promoter -56A>C, -57delA*	I	(23, 96)
	mtrR promoter -131G>A (*mtrC* -120G>A substitution, *mtr120)**	I	(97)
	penA I312M, V316P/T, ins346D, T483S, A501P/T/V, G542S, G545S, P551S	I	(99, 100)
	penA mosaic (I312M, V316T, G545S)	A	(99-101)
	ponA1 L421P	I	(108)
	porB1b G120K, A121N/D	I	(109)
Spectinomycin (R: MIC>64)	*16S* rDNA 1184C>T (1192C>T in *E. coli*)	R	(110)
	rpsE T24P	R	(111)
	rpsE V27- deletion, K28E	R/A	(111)
Sulfonamides *****	*folP* R228S	R	(22, 112)

*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.* (23), 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.* (113). 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.

Library of genetic AMR mechanisms: genotypic and phenotypic benchmarks

We compiled described genetic AMR mechanisms previously reported for *N. gonorrhoeae* up to the writing of this manuscript into the AMR library in Pathogenwatch (Table 2).A genotypic accuracy testing of the AMR library was performed using the 14 *N. gonorrhoeae* reference genomes from the WHO 2016 panel (22), 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 S2). 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 3: Figure S4). 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

435 identity to *N. meningitidis* and 89% to *N. gonorrhoeae*, and it is included as screening because it
436 has previously been shown to cause false negative results in some molecular detection tests for
437 *N. gonorrhoeae* (114).

438 Then, we also performed a genotypic-phenotypic benchmark using a test dataset of 3,987 *N.*
439 *gonorrhoeae* isolates from 13 different studies containing MIC information for at least part of the
440 following six antibiotics: ceftriaxone, cefixime, azithromycin, ciprofloxacin, benzylpenicillin and
441 tetracycline (Additional file 1: Table S1). EUCAST clinical breakpoints were applied for five of the
442 antimicrobials except for azithromycin, for which the adoption of an ECOFF>1 mg/L is now
443 recommended to distinguish isolates with azithromycin resistance determinants, instead of a
444 clinical resistance breakpoint (115, 116). A visualization of the range of MICs on each particular
445 combination of genetic AMR mechanisms observed on the isolates from the benchmark test
446 dataset (Figure 3a-b and Additional file 3: Figures S5-S10) revealed combinations that show an
447 additive effect on AMR. These combinations were included in the AMR library to improve the
448 accuracy of the genotypic prediction. For example, *rpsJ* V57M and some *mtrR*-associated
449 mutations individually are associated with a decreased susceptibility or intermediate resistance
450 to tetracycline (MICs of 0.5-1 mg/L), however, a combination of these variants can increase MICs
451 above the EUCAST resistance breakpoint for tetracycline (MICs>1 mg/L) (Additional file 3: Figure
452 S9). This is the case of the combination of *rpsJ* V57M with the *mtrR* promoter -57delA mutation
453 (N=681 isolates, 94.9% positive predictive value, PPV) or with *mtrR* promoter -57delA and *mtrR*
454 G45D (N=83 isolates, 93.9% PPV). Several combinations of *penA, ponA1, mtrR* and *porB1b*
455 mutations were observed to be able to increase the benzylpenicillin MIC above the resistant
456 threshold in most of the cases (Additional file 3: Figure S10). This is the case of the *porB1b*
457 mutations combined with *mtrR* A39T (N=31 isolates, 100% PPV), with the *mtrR* promoter -57delA
458 deletion (N=286 isolates, 96.5% PPV) or with *mtrR* promoter -57delA and *ponA1* L421P (N=269
459 isolates, 96.3%). Despite mosaic *penA* not being a main driver of resistance to penicillins, a
460 combination of the *porB1b* mutations with the three main mosaic *penA* mutations (G545S, I312M
461 and V316T) was also associated with 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 (26). We included the presence of a disrupted *mtrC* as a modifier of antimicrobial susceptibility in the presence of an *mtr* mosaic, as we did not have enough evidence from the test dataset to assess the MIC ranges of isolates with the 23S rDNA A2045G and C2597T mutations with and without a disrupted *mtrC* gene.

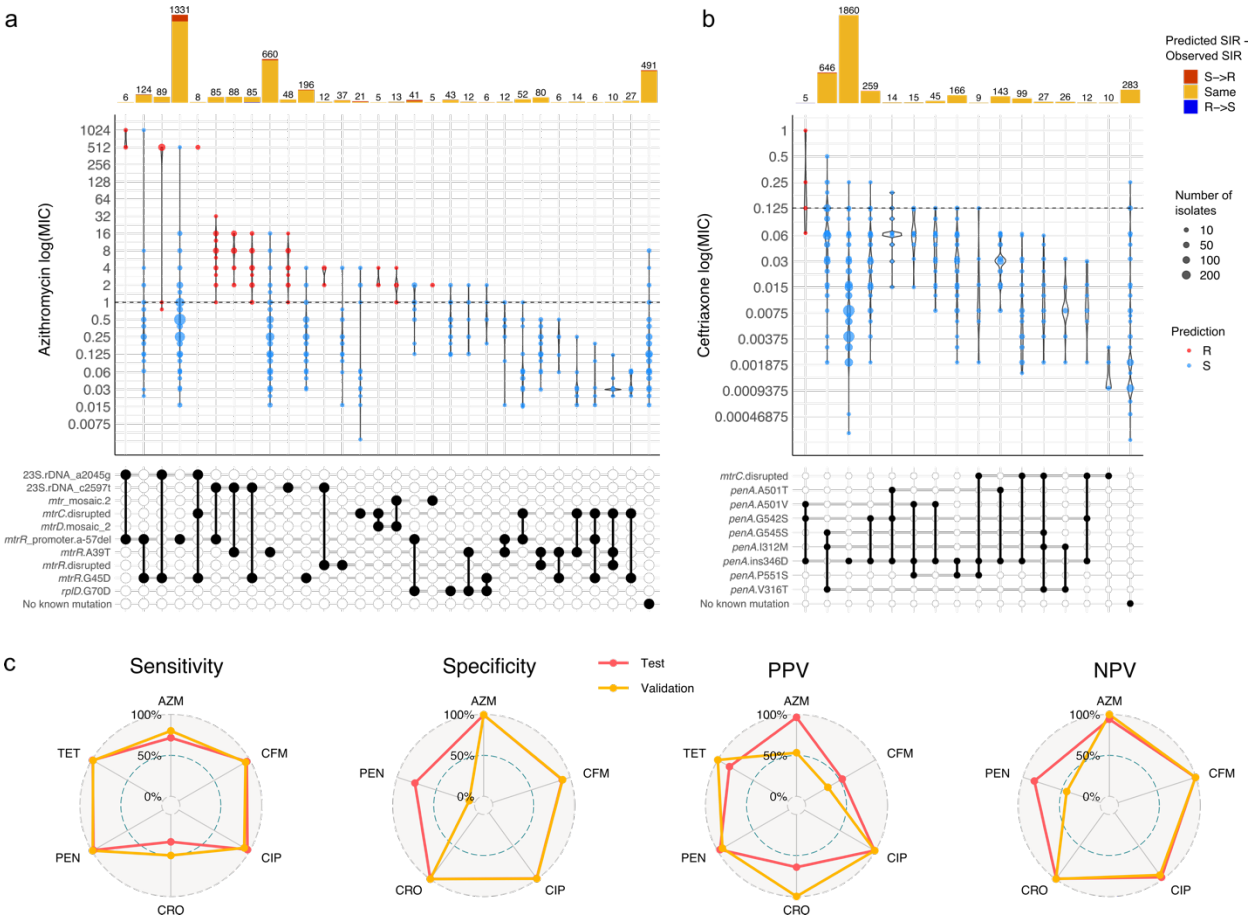


Figure 3. Distribution of minimum inhibitory concentration (MIC) values (mg/L) for the last-line antibiotics for *N. gonorrhoeae* azithromycin (a) and ceftriaxone (b) in a collection of 3,987 *N. gonorrhoeae* isolates with different combinations of genetic antimicrobial resistance (AMR) mechanisms. Only combinations observed in at least 5 isolates are shown (see Additional file 3: Figure S5-S10 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

481 and validation benchmark analyses. AZM = Azithromycin, CFM = Cefixime, CIP = Ciprofloxacin, CRO = Ceftriaxone,
482 PEN = Benzylpenicillin, TET = Tetracycline.
483

484 Results from the benchmark (Additional file 1: Table S3) show sensitivity values (true positive
485 rates, TP/(TP+FN); TP=True Positives, FN=False Negatives) above 96% for tetracycline (99.2%),
486 benzylpenicillin (98.1%), ciprofloxacin (97.1%) and cefixime (96.1%), followed by azithromycin
487 (71.6%) and ceftriaxone (33.3%). These results reflect the complexity of the resistance
488 mechanisms for azithromycin and ceftriaxone, where the known genetic determinants explain
489 only part of the antimicrobial susceptibility. However, specificity values (true negative rates,
490 TN/(TN+FP); TN=True Negatives, FP=False Positives) for these two antibiotics as well as
491 ciprofloxacin were above 99% (Additional file 1: Table S3), demonstrating that the genetic
492 mechanisms included in the database have a role in AMR. The specificity value for cefixime was
493 lower but nearly 90%, mainly due to the high number of isolates with an MIC below the threshold
494 but with three mutations characterising a mosaic *penA* allele (G545S, I312M and V316T, TP=367,
495 TN=323, PPV=53.2%; Additional file 1: Table S4). Benzylpenicillin and tetracycline showed
496 specificity values of 77.3% and 61.3%, respectively. In the first case, all the mechanisms included
497 in the library showed a PPV value above 94%. For tetracycline, a considerable number of false
498 positive results are mainly caused by the presence of *rpsJ* V57M, for which PPV=83.8%
499 (TP=1083, FP=209; Additional file 1: Table S4). However, this mutation was kept in the AMR
500 library because it can cause intermediate resistance to tetracycline on its own (Additional file 3:
501 Figure S9).

502 Results from the benchmark analysis on the 3,987-isolates dataset were used to curate and
503 optimize the AMR library. Thus, in order to objectively validate it, the benchmark analysis was
504 also run on a combination of three different collections (N=1,607, Additional file 1: Table S1) with
505 available MIC information for seven antibiotics including spectinomycin (Additional file 1: Table
506 S3) (69, 70, 117). Results from the test and validation benchmark runs were compared, showing
507 that sensitivity values on the six overlapping antibiotics were very similar, with the validation
508 benchmark performing even better for azithromycin and ceftriaxone (Figure 3c). In terms of

509 specificity, both datasets performed equally well for all antibiotics except for benzylpenicillin, in
510 which specificity drops in the validation benchmark. This is due to the *penA*_ins346D mutation
511 (TP=1125, FP=83) and the *blaTEM* genes (TP=525, FP=36), which despite showing false
512 positives, have a PPV above 93% (Additional file 1: Table S5). In general, discrepancies found
513 between the test and the validation benchmarks can be explained by particular mechanisms that
514 on their own show high predictive values and affect antibiotics for which we do not currently
515 understand all the factors involved in resistance, such as azithromycin and the ESCs (Additional
516 file 1: Table S5).

517 ***Over 12,000 public genomes available***

518 Data for 11,461 isolates were successfully assembled and passed all quality cut-offs, resulting in
519 12,515 isolates after including the previously-available Euro-GASP 2013 dataset (15). New
520 assemblies were uploaded and made public on Pathogenwatch, which now constitutes the largest
521 repository of curated *N. gonorrhoeae* genomic data with associated metadata, typing and AMR
522 information at the time of submission of this manuscript. Updated data spans 27 different
523 publications (18, 53, 58-61, 63-65, 67-70, 117-131) and is organized into individual collections
524 associated with the different studies (Additional file 1: Table S6). Available metadata was added
525 for the genomes from these publications while basic metadata fields were kept for others (country,
526 year/date and ENA project number).

527 We cross-checked that the main clusters found in the phylogenetic trees obtained after creating
528 the public collections in Pathogenwatch were consistent with those observed in the trees in the
529 corresponding publications. For example, recent works defined two major clusters of *N.*
530 *gonorrhoeae*, termed Lineages A and B, which were found to be consistent with the corresponding
531 Pathogenwatch trees as exemplified for isolates from England in Town *et al* (2020) (68) (Figure
532 S11a). We were also able to differentiate the cefixime-resistant *penA10* and *penA34*-carrying
533 clones from Vietnam from Lan *et al* (2020) (124) (Figure S11b) as well as the 10 major clusters
534 defined in the *N. gonorrhoeae* population circulating in New York City (NYC) as described in

535 Mortimer *et al* (2020) (120) (Figure S11c). In the last case, we also liked to emphasize the
536 usefulness of Microreact (37) as a parallel tool to Pathogenwatch for more complex visualization
537 purposes, such as showing the 10 major clusters in NYC as metadata blocks of different colours.

538 The *N. gonorrhoeae* public data available on Pathogenwatch spans nearly a century (1928-2018)
539 and almost 70 different countries (Additional file 3: Figure S12). However, sequencing efforts are
540 unevenly distributed around the world, and over 90% of the published isolates were isolated in
541 only 10 countries, headed by the United Kingdom (N=3,476), the United States (N=2,774) and
542 Australia (N=2,388) (Additional file 1: Table S7, Figure 4). A total of 554 MLST, 1,670 NG-MAST
543 and 1,769 NG-STAR different STs were found in the whole dataset, from which a considerable
544 number were new profiles caused by previously undetected alleles or new combinations of known
545 alleles (N=92 new MLST STs, N=769 new NG-STAR STs and N=2,289 isolates with new NG-
546 MAST *porB* and/or *tbpB* alleles). These new alleles and profiles were submitted to the
547 corresponding scheme servers.

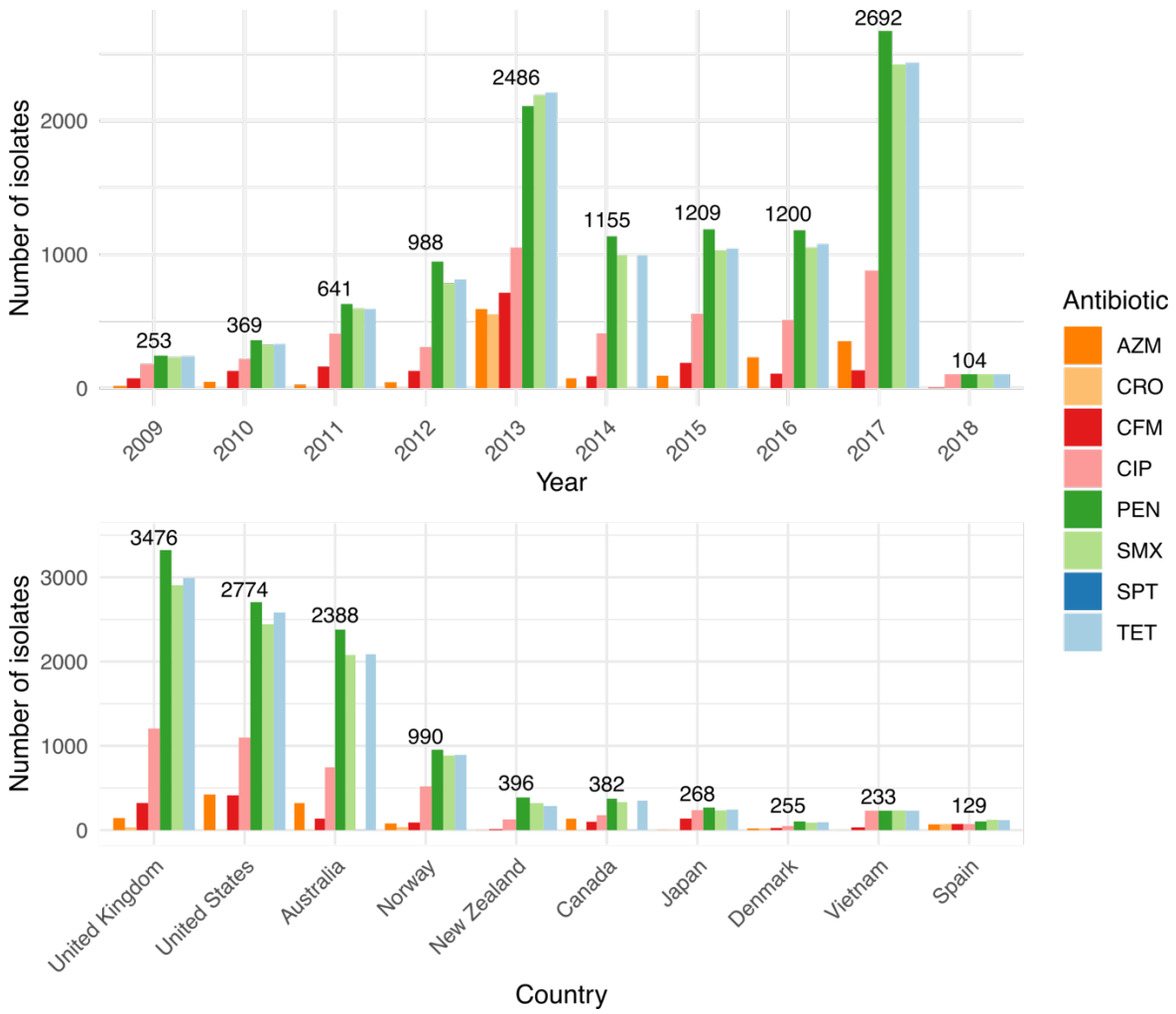


Figure 4. Summary of the geolocalization and collection date of 12,515 public *N. gonorrhoeae* genomes in Pathogenwatch. Coloured bars represent the genotypic antimicrobial resistance (AMR) prediction based on the mechanisms included in the library. AZM = Azithromycin, CFM = Cefixime, CIP = Ciprofloxacin,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

548

549 Genomic studies are often biased towards AMR isolates, and this is reflected in the most
550 abundant STs found for the three typing schemes within the public data. Isolates with MLST
551 ST1901, ST9363 and ST7363, which contain resistance mechanisms to almost every antibiotic
552 included in the study, represent over 25% of the data (Figure 5). Isolates with MLST ST1901 and
553 ST7363 are almost always associated with resistance to tetracycline, sulfonamides,
554 benzylpenicillin and ciprofloxacin and nearly 50% of isolates from these two types harbour
555 resistance mechanisms to cefixime. Ciprofloxacin resistance is not widespread among ST9363
556 isolates, which are associated with azithromycin resistance in nearly 50% of the isolates for this

557 ST (Figure 5). NG-STAR ST63 (carrying the non-mosaic *penA-2* allele, *penA* A517G and *mtrR*
558 A39T mutations as described in (52)) is the most represented in the dataset and carries resistance
559 mechanisms to tetracycline, sulfonamides, and benzylpenicillin, but is largely susceptible to
560 spectinomycin, ciprofloxacin, the ESCs cefixime and ceftriaxone and azithromycin. NG-STAR
561 ST90 isolates, conversely, are largely associated with resistance to cefixime, ciprofloxacin and
562 benzylpenicillin as they carry the key resistance mutations in mosaic *penA-34*, as well as in the
563 *mtrR* promoter, *porB1b*, *ponA*, *gyrA* and *parC* (as described in (52)). NG-MAST ST1407 is
564 commonly associated with MLST ST1901 and is the second most represented ST in the dataset
565 following NG-MAST ST2992, which mainly harbours resistance to tetracycline, benzylpenicillin
566 and sulfonamides (Figure 5).

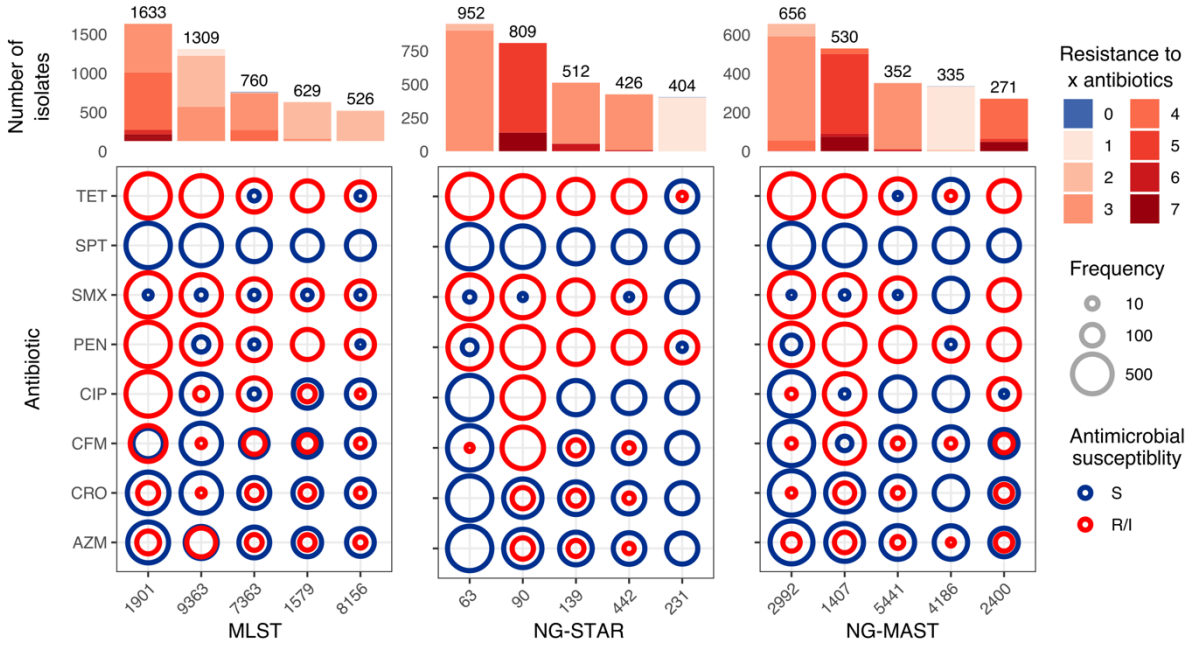


Figure 5. Predicted antimicrobial resistance (AMR) profiles of the top five Multi-Locus Sequence Typing (MLST), *N. gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) and *N. gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) types in the *N. gonorrhoeae* public data in Pathogenwatch. Coloured circles in the grids show the proportion of genomes from each ST which are predicted to have an intermediate (susceptible but increased exposure) or resistant phenotype, 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.

567 ***Case study: global expansion of an mtr mosaic-carrying clone***

568 The genetic mechanisms that have commonly been associated with an increased MIC of
569 azithromycin in *N. gonorrhoeae* are two mutations in the 23S rRNA gene (A2045G and C2597T
570 substitutions, in *N. gonorrhoeae* nomenclature) as well as mutations in *mtrR* and its promoter
571 (132, 133). As described above, other mechanisms have also been recently discovered that
572 increase the MIC of azithromycin (Table 2), such as mosaicism affecting the efflux pump-encoding
573 *mtrCDE* genes and its repressor *mtrR,* mainly when the mosaic spans the *mtrR* promoter region
574 and *mtrD* gene (23, 24). Some studies have recently reported the local expansion of azithromycin-
575 resistant *N. gonorrhoeae* lineages carrying an *mtr* mosaic in the USA (122, 123, 134) and
576 Australia (118). However, the extent of the dispersion of this mechanism to other parts of the
577 world has not been studied yet. Here, using the public genomes of *N. gonorrhoeae* in
578 Pathogenwatch, we have been able to explore this question.

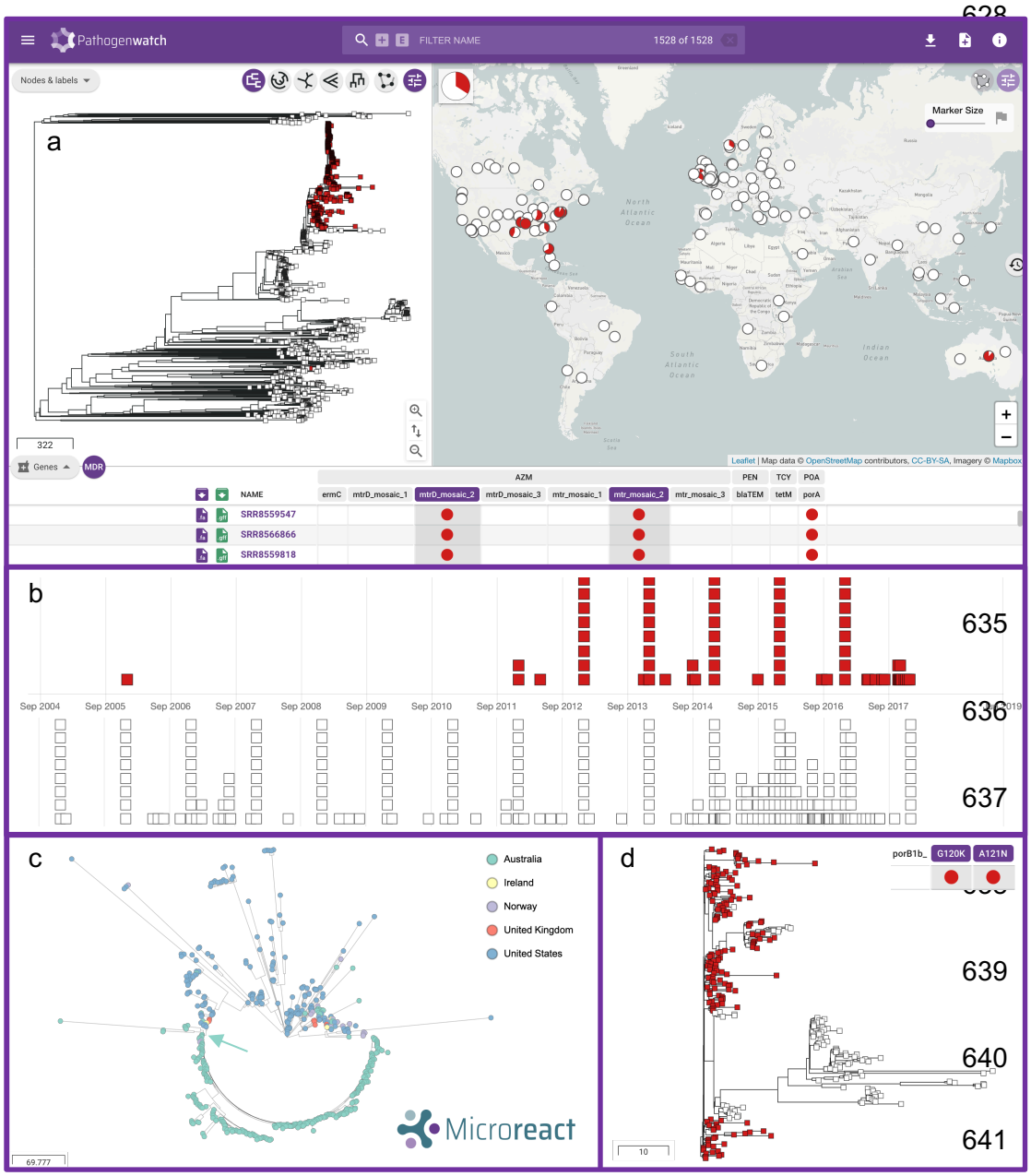
579 A total of 1,142 strains with genetic determinants of azithromycin resistance were selected in
580 Pathogenwatch and combined with 395 genomes from a global collection (64) for background
581 contextualization (see Pathogenwatch project in (135)) (Figure 6a). 571 of the strains predicted
582 to be resistant to azithromycin had some form of mosaic in the *mtrR* promoter and/or *mtrD* gene
583 of one of the three types described in Wadsworth *et al.* (2018) (23) and included in the
584 Pathogenwatch AMR library (Table 2). These mosaics have been experimentally proven to
585 increase MIC of azithromycin above 1 mg/L, which is the EUCAST ECOFF value as well as the
586 Clinical Laboratory and Standards Institute (CLSI) non-susceptibility breakpoint (23, 24). One of
587 the *N. lactamica*-like mosaics, termed here '*mtr*_mosaic.2', was by far the most extended, as it
588 was found in 545 genomes spanning the *mtrR* promoter and/or the *mtrD* gene, with 521 (95.6%)
589 of them spanning both regions. Twenty-five genomes contained a *N. meningitidis*-like mosaic
590 *mtrR* promoter and/or *mtrD* gene ('*mtr*_mosaic.1') and in only 9 (36%) of them the mosaic
591 spanned both loci. The *N. lactamica*-like '*mtr*_mosaic.3' was only found in isolate ERR855360
592 (GCGS834) from Los Angeles (USA, 2012), which is where the reference sequence for this

593 mosaic was extracted from. Of the studies where these *mtr* mosaic-carrying genomes were
594 obtained from, only those from the USA and Australia specifically targeted and found this genetic
595 determinant of resistance. The rest did not target this mosaic and some of them found strains with
596 unexplained increased MICs of azithromycin (69, 121, 129), which could partly be explained by
597 the presence of these *mtr* mosaics.

598 We observed one main lineage carrying mosaic 2 in *mtrR* promoter and *mtrD* gene (Figure 6a)
599 with 520 genomes. Of those, only 3 and 8 isolates carried the 23S rDNA A2045G and C2597T
600 mutations, respectively. Interestingly, the first strain in the database with this type of mosaic dates
601 from 2006 (18), however, it was not until the end of 2011-2012 when this lineage started to expand
602 (Figure 6b). Despite the genomic data contained in Pathogenwatch being biased to the amount
603 of data sequenced and published from each country and year, we can easily infer that this lineage
604 has spread across the world as we detect cases in Australia (n=293) (118), the USA (n=195) (18,
605 120, 122, 123), Norway (n=19) (121), the United Kingdom (n=11) (68, 119), and Ireland (n=3)
606 (129). A strong association was found to the country of isolation (Figure 6c), with a broad diversity
607 of sublineages having spread across the USA (strains mostly isolated between 2012 and 2016).
608 In contrast, an expansion of a particular clone, likely from a single main introduction, was observed
609 to have occurred in Australia (strains isolated in 2017), followed by a further divergence of a
610 subclone within the country which correlates with the loss of the *porB1b* G120K and A121N
611 mutations (Figure 6d), likely through a recombination event. Despite epidemiological data not
612 being available for the Australian study (118), from their work we know that the clusters carrying
613 an *mtr* mosaic were mostly linked to transmission between men, although bridging among MSM
614 and heterosexual populations was also observed.

615 The results from our case study show that there is an emerging lineage of *N. gonorrhoeae* that
616 has spread across the world and that is carrying a mosaic *mtr* that has been associated with low-
617 to-medium resistance to azithromycin. This global lineage, as well as others that may emerge
618 carrying this or other genetic AMR mechanisms, has to be closely monitored. For this purpose,
619 an up-to-date genomic epidemiology tool such as Pathogenwatch, which includes a list of genetic

AMR mechanisms approved by an expert group is a great resource for the scientific community. At the moment, Pathogenwatch includes references for three types of mosaics in the *mtrR* promoter and *mtrD* genes that have been experimentally proven to increase MIC of azithromycin (23, 24), and the detection of these mosaics on new genomes respond to a set of similarity rules (see Data availability section). However, we will keep the database updated with new experimentally-confirmed reference sequences that may arise from further studies as it is still unclear whether all mosaics affecting the *mtrCDE* efflux pump will cause a decreased susceptibility to azithromycin.



642 Figure 6. *N. gonorrhoeae* genomes carrying genetic AMR mechanisms associated to azithromycin resistance were
643 selected in Pathogenwatch (n=1,142) and combined with genomes from a global collection (64, 88) (total n=1528) for
644 background contextualization. (a) Main layout of the combined collection, with the emerging lineage carrying *mtr* mosaic
645 2 spanning the *mtrR* promoter and *mtrD* marked in red in the tree and the map. (b) Timeline of the genomes carrying
646 *mtr* mosaic 2 (in red) and other public genomes in the database without this genetic AMR mechanism. (c) Visualization
647 of the *mtr* mosaic 2-carrying lineage (n=520) spreading in the USA and Australia (see legend) using Microreact. The
648 arrow in turquoise colour marks the divergence of the Australian lineage, shown in more detail in (d) coloured by the
649 presence (in red) or absence (in white) of the *porB1b* G120K and A121N mutations. The Pathogenwatch project of this
650 case study can be explored in (135).

651

652 **Discussion**

653 We present a public health focussed *N. gonorrhoeae* framework at Pathogenwatch, an open
654 access platform for genomic surveillance supported by an expert group that can be adapted to
655 any public health or microbiology laboratory. Little bioinformatics expertise is required, and users
656 can choose to either upload raw short read data or assembled genomes. In both cases, the upload
657 of high-quality data is encouraged in the form of quality-checked reads and/or quality-checked
658 assemblies. Recent benchmark analyses show particular recommendations for long-read or
659 hybrid data (136) as well as short read-only data (40, 137). On upload, several analyses are run
660 on the genomes, and results for the three main typing schemes (MLST, NG-MAST and NG-STAR)
661 as well as the detection of genetic determinants of AMR and a prediction of phenotypic resistance
662 using these mechanisms can be obtained simultaneously. The library of AMR determinants
663 contained in Pathogenwatch for *N. gonorrhoeae* has been revised and extended to include the
664 latest mechanisms and epistatic interactions with experimental evidence of decreasing
665 susceptibility or increasing resistance to at least one of eight antibiotics (Tables 2). A test and
666 validation benchmark analyses revealed sensitivity and/or specificity values >90% for most of the
667 tested antibiotics (Additional file 1: Table S3). Sensitivity values for the antimicrobials in the
668 current dual treatment, azithromycin (80%) and ceftriaxone (50%), reflect the complexity of the
669 resistance mechanisms for these antibiotics, for which we can only explain part of the observed
670 phenotypic resistance. However, their specificity values were above 99% (Additional file 1: Table
671 S3), further strengthening the associations of the included AMR determinants in increasing MICs

672 of these antibiotics. It remains essential to perform phenotypic susceptibility testing so we can
673 detect inconsistencies between phenotypic and genotypic data that can lead to the identification
674 and subsequent verification of novel or unknown resistance mechanisms. This will allow to
675 continuously expand the list of genetic AMR mechanisms, and the AMR prediction from genomic
676 data will further improve.

677 The continuous increase in reporting of *N. gonorrhoeae* AMR isolates worldwide led to a call for
678 international collaborative action in 2017 to join efforts towards a global surveillance scheme. This
679 was part of the WHO global health sector strategy on STIs (2016-2021), which set the goal of
680 ending STI epidemics as a public health concern by year 2030 (7, 8). Several programmes are
681 currently in place at different global, regional or national levels to monitor gonococcal AMR trends,
682 emerging resistances and refine treatment guidelines and public health policies. This is the case
683 of, for example, the WHO Global Gonococcal Antimicrobial Surveillance Programme (WHO
684 GASP) (7, 8), the Euro-GASP in Europe (6, 15, 138), the Gonococcal Isolate Surveillance Project
685 (GISP) in the United States (139), the Canadian Gonococcal Antimicrobial Surveillance
686 Programme (140), the Gonococcal Surveillance Programme (AGSP) in Australia (141) or the
687 Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP) in England and
688 Wales (142). The WHO in collaboration with CDC has recently started an enhanced GASP
689 (EGASP) (143) in some sentinel countries such as the Philippines and Thailand (144), aimed at
690 collecting standardized and quality-assured epidemiological, clinical, microbiological and AMR
691 data. On top of these programs, WHO launched the Global AMR Surveillance System (GLASS)
692 in 2015 to foster national surveillance systems and enable standardized, comparable and
693 validated AMR data on priority human bacterial pathogens (145). Efforts are now underway to link
694 WHO GASP to GLASS. However, gonococcal AMR surveillance is still suboptimal or even lacking
695 in many locations, especially in LMICs, such as several parts of Asia, Central and Latin America,
696 Eastern Europe and Africa, which worryingly have the greatest incidence of gonorrhoea (3).
697 LMICs often have access to antimicrobials without prescription, have limited access to an optimal
698 treatment, lack the capacity needed to perform a laboratory diagnosis due to limited or non-

699 existent quality-assured laboratories, microbiological and bioinformatics expertise or training,
700 insufficient availability and exorbitant prices of some reagents on top of a lack of funding, which
701 altogether compromises infection control.

702 High throughput sequencing approaches have proved invaluable over traditional molecular
703 methods to identify AMR clones of bacterial pathogens, outbreaks, transmission networks and
704 national and international spread among others (28, 29). Genomic surveillance efforts to capture
705 the local and international spread of *N. gonorrhoeae* have resulted in several publications within
706 the last decade involving high throughput sequence data of thousands of isolates from many
707 locations across the world. The analysis of this data requires expertise, not always completely
708 available, in bioinformatics, genomics, genetics, AMR, phylogenetics, epidemiology, etc. For
709 lower-resourced settings, initiatives such as the NIHR Global Health Research Unit, Genomic
710 Surveillance of Antimicrobial Resistance (146) are essential to build genomic surveillance
711 capacity and provide the necessary microbiology and bioinformatics training for quality-assured
712 genomic surveillance of AMR.

713 One of the strengths of genomic epidemiology is being able to compare new genomes with
714 existing data from a broader geographical level, which provides additional information on, e.g. if
715 new cases are part of a single clonal expansion or multiple introductions from outside a specific
716 location. To support this, Pathogenwatch calculates phylogenetic trees from a set of genomes
717 selected as collections. Currently, over 12,000 isolates of *N. gonorrhoeae* have been sequenced
718 using high throughput approaches and publicly deposited on the ENA linked to a scientific
719 publication. We have quality-checked and assembled these data using a common pipeline and
720 we made it available through Pathogenwatch, with the aim of representing as much genomic
721 diversity of this pathogen as possible to serve as background for new analyses. These public
722 genomes are associated with at least 27 different scientific publications, and have been organized
723 in Pathogenwatch as individual collections (Additional file 1: Table S6). The clustering of strains
724 on the resulting reconstructions was found consistent with those in the original publications (some

725 examples in Figure S11), while differences in branch lengths may be attributed to the usage of
726 different reconstruction methods.

727 The power of Pathogenwatch to investigate questions of public health concern is reflected in a
728 case study (Figure 6). By selecting 1,142 azithromycin resistant strains from the public data in
729 Pathogenwatch in the context of a global collection (64), we observed one clone carrying *N.*
730 *lactamica*-like *mtr* mosaic ('mosaic_2') in both the *mtrR* promoter and *mtrD* genes, likely resulting
731 from the same recombination event. Strong geographical structure was found in these
732 azithromycin resistant strains, with isolates from the USA (mostly from 2012-2016) clearly
733 differentiated from those from Australia (from 2017), which show a more clonal dispersion, likely
734 from a single main introduction to the country followed by a rapid spread. Interestingly, a
735 sublineage of this Australian *mtr* mosaic-carrying clone seems to have also diverged after losing
736 the *porB1b* G120K and D121N mutations. It is important to note that the data from which these
737 inferences were derived was gathered from surveillance-based studies and outbreak
738 investigations, which may bias the observed global diversity of strains carrying this mosaic.
739 Phenotypic susceptibility data for azithromycin or epidemiological information were not available
740 for over half of these strains, thus impeding making further inferences. This reflects the need of
741 improving the submission of anonymized epidemiological and antimicrobial susceptibility data for
742 individual isolates rather than aggregated data to public repositories and/or as supplementary
743 information of the corresponding publications, as this is where the public data in Pathogenwatch
744 is coming from.

745 In this study, we have additionally gathered an advisory group of *N. gonorrhoeae* experts in
746 different fields such as AMR, microbiology, genetics, genomics, epidemiology and public health
747 who will consult and discuss current and future analytics to be included to address the global
748 public health needs of the community. We suggest this strategy as a role model for other
749 pathogens in this and other genomic surveillance platforms, so the end user, who may not have
750 full computational experience in some cases, can be confident that the analytics and databases
751 underlying this tool are appropriate, and can have access to all the results provided by

752 Pathogenwatch through uploading the data via a web browser. We are aware that this is a
753 constantly moving field and analytics will be expanded and updated in the future. These updates
754 will be discussed within an advisory group to make sure they are useful in the field and the way
755 results are reported is of use to different profiles (microbiologists, epidemiologists, public health
756 professionals, etc.). Future analytics that are under discussion include the automatic submission
757 of new MLST, NG-STAR and NG-MAST STs and alleles to the corresponding servers, e.g.
758 PubMLST (48) and the automatic submission of data to public archives such as the ENA. Inter-
759 connectivity and comparability of results with PubMLST is of particular interest, as this database
760 has traditionally been the reference for *Neisseria* sequence typing and genomics and it is widely
761 used by the *N. gonorrhoeae* community. Plasmid and *tetM/blaTEM* subtyping as recently
762 described (147) will also be considered within the development roadmap of Pathogenwatch.
763 Including a separate library to automatically screen targets of potential interest for vaccine design
764 (148-150) as well as targets of new antibiotics currently in phase III clinical trials (i.e. zoliflodacin
765 (151) or gepotidacin (152)) can also be an interesting addition to the scheme. Regarding AMR,
766 new methods for phenotypic prediction using genetic data are continuously being reported (62,
767 153, 154), especially those based on machine learning algorithms (155), and will be considered
768 for future versions of the platform. The prediction of MIC values or ranges instead of SIR
769 categories will allow users to decide whether to use EUCAST (156) or CLSI (157) guidelines for
770 categorization.

771

772 **Conclusions**

773 In summary, we present a genomic surveillance platform adapted to *N. gonorrhoeae*, one of the
774 main public health priorities compromising the control of AMR infections, where decisions on
775 existing and updated databases and analytics as well as how results are reported will be
776 discussed with an advisory board of experts in different public health areas. This will allow
777 scientists from both higher or lower resourced settings with different capacities regarding high

778 throughput sequencing, bioinformatics and data interpretation, to be able to use a reproducible
779 and quality-assured platform where analyse and contextualise genomic data resulting from the
780 investigation of treatment failures, outbreaks, transmission chains and networks at different
781 regional scales. This open access and reproducible platform constitutes one step further into an
782 international collaborative effort where countries can keep ownership of their data in line with
783 national STI and AMR surveillance and control programs while aligning with global strategies for
784 a joint action towards battling AMR *N. gonorrhoeae*.
785

786 **List of abbreviations**

787 AGSP: Australian Gonococcal Surveillance Programme
788 AMR: Antimicrobial Resistance
789 AZM: Azithromycin
790 CDC: Centers for Disease Control and Prevention
791 CFM: Cefixime
792 cgMLST: Core Genome Multi-Locus Sequence Typing
793 CIP: Ciprofloxacin
794 CLSI: Clinical Laboratory and Standards Institute
795 CRO: Ceftriaxone
796 ECOFF: Epidemiological Cut-Off
797 EGASP: Enhanced Gonococcal Antimicrobial Surveillance Programme
798 ENA: European Nucleotide Archive
799 ESCs: Extended Spectrum Cephalosporins
800 EUCAST: European Committee on Antimicrobial Susceptibility Testing
801 Euro-GASP: European Gonococcal Antimicrobial Surveillance Programme
802 FN: False Negative
803 FP: False Positive
804 GASP: 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/-/tree/master/ngonorrhoeae.

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). At the time of preparation of this manuscript, LSB was supported by the Li Ka Shing Foundation (Big Data Institute, University of Oxford) and the Centre for Genomic Pathogen Surveillance (CGPS, http://pathogensurveillance.net). At the time of review and publication of this manuscript, LSB is funded by Plan GenT (CDEI-06/20-B), Conselleria de Sanidad Universal y Salud Pública, Generalitat Valenciana (Valencia, Spain). DMA is supported by the Li Ka Shing Foundation (Big Data Institute, University of Oxford) and the

859 Centre for Genomic Pathogen Surveillance (CGPS). DMA and SA are supported by the National
860 Institute for Health Research (UK) Global Health Research Unit on Genomic Surveillance of AMR
861 (16_136_111). The department of MJC receives funding from the European Centre for Disease
862 Prevention and Control and the National Institute for Health Research (Health Protection
863 Research Unit) for gonococcal whole-genome sequencing. YHG is supported by the NIH/NIAID
864 grants R01 AI132606 and R01 AI153521. KCM is supported by the NSF GRFP grant
865 number DGE1745303. TDM is supported by the National Institute of Allergy and Infectious
866 Diseases at the National Institutes of Health [1 F32 AI145157-01]. WMS is a recipient of a Senior
867 Research Career Scientist Award from the Biomedical Laboratory Research and Development
868 Service of the Department of Veterans. Work on antibiotic resistance in his laboratory is supported
869 by NIH grants R37 AI-021150 and R01 AI-147609. The content of this article is solely the
870 responsibility of the authors and does not necessarily represent the official views of the
871 Department of Veterans Affairs, The National Institutes of Health or the United States
872 Government. The findings and conclusions in this article are those of the author(s) and do not
873 necessarily represent the official position of the Centers for Disease Control and Prevention. The
874 WHO Collaborating Centre for Gonorrhoea and other STIs represented by DG and MU receives
875 funding from the European Centre for Disease Prevention and Control and the World Health
876 Organization. This publication made use of the Neisseria Multi-Locus Sequence Typing website
877 (https://pubmlst.org/neisseria/) sited at the University of Oxford (48) and funded by Wellcome and
878 European Union.
879

880 ***Authors' contributions***

881 DMA conceived the Pathogenwatch application. CY, RG, KA, BT, AU and DMA developed the
882 Pathogenwatch application. LSB and DMA contributed to the conception and design of the work.
883 CY and LSB generated, updated and benchmarked the *N. gonorrhoeae* AMR library. BT, CY, AU
884 and LSB obtained, quality-checked and reassembled the raw data from the ENA. LSB revised the
885 assembled data, obtained all metadata available from the corresponding scientific publications
886 and created collections. LSB, CY and DMA analysed the data. LSB and DMA drafted the

887 manuscript. LSB, DMA, CY, SA, KCM, TDM, DG, MJC, YHG, IM, BHR, WMS, GS, KT, TW, SRH
888 and MU contributed to the acquisition, technical and scientific interpretation and discussion of the
889 data. LSB, DMA, MJC, YHG, IM, BHR, WMS, GS, KT, TW and MU agreed to participate in the *N.*
890 *gonorrhoeae* Pathogenwatch Scientific Steering Group before the preparation of this manuscript,
891 and participated in virtual discussions. All authors read and approved the final manuscript.
892

893 ***Acknowledgements***
894 We would like to thank MJC, YHG, IM, BHR, WMS, GS, KT, TW and MU for their support on the
895 development of the *N. gonorrhoeae* Pathogenwatch scheme and the creation of the *N.*
896 *gonorrhoeae* Pathogenwatch Scientific Steering Group.
897

898 ***References***

899 1. World Health Organization (WHO). Global priority list of antibiotic-resistant bacteria to
900 guide research, discovery, and development of new antibiotics.
901 http://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-
902 ET_NM_WHO.pdf. 2017.
903 2. Centers for Disease Prevention and Control. U.S. Department of Health and Human
904 Services, Atlanta, GA. Antibiotic Resistance Threats in the United States, 2019.
905 https://www.cdc.gov/drugresistance/pdf/threats-report/2019-ar-threats-report-508.pdf. 2019.
906 3. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al.
907 Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates,
908 2016. Bull World Health Organ. 2019;97(8):548-62P.
909 4. World Health Organization (WHO). Report on global sexually transmitted infection
910 surveillance 2018. http://apps.who.int/iris/bitstream/handle/10665/277258/9789241565691-
911 eng.pdf?ua=1. 2019.
912 5. Unemo M, Shafer WM. Antimicrobial resistance in *Neisseria gonorrhoeae* in the 21st
913 century: past, evolution, and future. Clin Microbiol Rev. 2014;27(3):587-613.
914 6. Cole MJ, Spiteri G, Town K, Unemo M, Hoffmann S, Chisholm SA, et al. Risk factors for
915 antimicrobial-resistant N*eisseria gonorrhoeae* in Europe. Sex Transm Dis. 2014;41(12):723-9.
916 7. Wi T, Lahra MM, Ndowa F, Bala M, Dillon JR, Ramon-Pardo P, et al. Antimicrobial
917 resistance in *Neisseria gonorrhoeae*: Global surveillance and a call for international
918 collaborative action. PLoS Med. 2017;14(7):e1002344.
919 8. Unemo M, Lahra MM, Cole M, Galarza P, Ndowa F, Martin I, et al. World Health
920 Organization Global Gonococcal Antimicrobial Surveillance Program (WHO GASP): review of
921 new data and evidence to inform international collaborative actions and research efforts. Sex
922 Health. 2019;16(5):412-25.
923 9. Fifer H, Natarajan U, Jones L, Alexander S, Hughes G, Golparian D, et al. Failure of
924 Dual Antimicrobial Therapy in Treatment of Gonorrhea. N Engl J Med. 2016;374(25):2504-6.
925 10. Jennison A, Whiley D, Lahra M, Graham R, Cole M, Hughes G, et al. Genetic
926 relatedness of ceftriaxone-resistant and high-level azithromycin resistant *Neisseria gonorrhoeae*
927 cases, United Kingdom and Australia, February to April 2018. . Euro Surveill.
928 2019;24(8):1900118.

929 11. Lahra MM, Martin I, Demczuk W, Jennison AV, Lee KI, Nakayama SI, et al. Cooperative
930 Recognition of Internationally Disseminated Ceftriaxone-Resistant *Neisseria gonorrhoeae*
931 Strain. Emerg Infect Dis. 2018;24(4).
932 12. Fifer H, Saunders J, Soni S, Sadiq ST, FitzGerald M. 2018 UK national guideline for the
933 management of infection with *Neisseria gonorrhoeae*. Int J STD AIDS. 2020;31(1):4-15.
934 13. Deguchi T, Yasuda M, Asano M, Tada K, Iwata H, Komeda H, et al. DNA gyrase
935 mutations in quinolone-resistant clinical isolates of *Neisseria gonorrhoeae.* Antimicrob Agents
936 Chemother. 1995;39(2):561-3.
937 14. Tanaka M, Takahashi K, Saika T, Kobayashi I, Ueno T, Kumazawa J. Development of
938 fluoroquinolone resistance and mutations involving GyrA and ParC proteins among *Neisseria*
939 *gonorrhoeae* isolates in Japan. J Urol. 1998;159(6):2215-9.
940 15. Harris SR, Cole MJ, Spiteri G, Sanchez-Buso L, Golparian D, Jacobsson S, et al. Public
941 health surveillance of multidrug-resistant clones of *Neisseria gonorrhoeae* in Europe: a genomic
942 survey. Lancet Infect Dis. 2018;18(7):758-68.
943 16. Spratt BG. Hybrid penicillin-binding proteins in penicillin-resistant strains of *Neisseria*
944 *gonorrhoeae*. Nature. 1988;332(6160):173-6.
945 17. Ohnishi M, Golparian D, Shimuta K, Saika T, Hoshina S, Iwasaku K, et al. Is *Neisseria*
946 *gonorrhoeae* initiating a future era of untreatable gonorrhea?: detailed characterization of the
947 first strain with high-level resistance to ceftriaxone. Antimicrob Agents Chemother.
948 2011;55(7):3538-45.
949 18. Grad YH, Harris SR, Kirkcaldy RD, Green AG, Marks DS, Bentley SD, et al. Genomic
950 epidemiology of gonococcal resistance to extended spectrum cephalosporins, macrolides, and
951 fluoroquinolones in the US, 2000-2013. J Infect Dis. 2016;214:1579-87.
952 19. Abrams AJ, Kirkcaldy RD, Pettus K, Fox JL, Kubin G, Trees DL. A Case of Decreased
953 Susceptibility to Ceftriaxone in *Neisseria gonorrhoeae* in the Absence of a Mosaic Penicillin-
954 Binding Protein 2 (penA) Allele. Sex Transm Dis. 2017;44(8):492-4.
955 20. Lindberg R, Fredlund H, Nicholas R, Unemo M. *Neisseria gonorrhoeae* isolates with
956 reduced susceptibility to cefixime and ceftriaxone: association with genetic polymorphisms in
957 penA, mtrR, porB1b, and ponA. Antimicrob Agents Chemother. 2007;51(6):2117-22.
958 21. Palace SG, Wang Y, Rubin DH, Welsh MA, Mortimer TD, Cole K, et al. RNA polymerase
959 mutations cause cephalosporin resistance in clinical *Neisseria gonorrhoeae* isolates. Elife.
960 2020;9.
961 22. Unemo M, Golparian D, Sanchez-Buso L, Grad Y, Jacobsson S, Ohnishi M, et al. The
962 novel 2016 WHO *Neisseria gonorrhoeae* reference strains for global quality assurance of
963 laboratory investigations: phenotypic, genetic and reference genome characterization. J
964 Antimicrob Chemother. 2016;71(11):3096-108.
965 23. Wadsworth CB, Arnold BJ, Sater MRA, Grad YH. Azithromycin Resistance through
966 Interspecific Acquisition of an Epistasis-Dependent Efflux Pump Component and Transcriptional
967 Regulator in *Neisseria gonorrhoeae*. mBio. 2018;9(4).
968 24. Rouquette-Loughlin CE, Reimche JL, Balthazar JT, Dhulipala V, Gernert KM, Kersh EN,
969 et al. Mechanistic Basis for Decreased Antimicrobial Susceptibility in a Clinical Isolate of
970 *Neisseria gonorrhoeae* Possessing a Mosaic-Like mtr Efflux Pump Locus. mBio. 2018;9(6).
971 25. Ma KC, Mortimer TD, Duckett MA, Hicks AL, Wheeler NE, Sanchez-Buso L, et al.
972 Increased power from conditional bacterial genome-wide association identifies macrolide
973 resistance mutations in *Neisseria gonorrhoeae.* Nat Commun. 2020;11(1):5374.
974 26. Ma KC, Mortimer TD, Hicks AL, Wheeler NE, Sanchez-Buso L, Golparian D, et al.
975 Adaptation to the cervical environment is associated with increased antibiotic susceptibility in
976 *Neisseria gonorrhoeae.* Nat Commun. 2020;11(1):4126.
977 27. Unemo M, Dillon JA. Review and international recommendation of methods for typing
978 *Neisseria gonorrhoeae* isolates and their implications for improved knowledge of gonococcal
979 epidemiology, treatment, and biology. Clin Microbiol Rev. 2011;24(3):447-58.
980 28. Loman NJ, Pallen MJ. Twenty years of bacterial genome sequencing. Nat Rev Microbiol.
981 2015;13(12):787-94.

982 29. Balloux F, Bronstad Brynildsrud O, van Dorp L, Shaw LP, Chen H, Harris KA, et al. From
983 Theory to Practice: Translating Whole-Genome Sequencing (WGS) into the Clinic. Trends
984 Microbiol. 2018;26(12):1035-48.
985 30. Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to
986 study antimicrobial resistance. Nat Rev Genet. 2019;20(6):356-70.
987 31. Hendriksen RS, Bortolaia V, Tate H, Tyson GH, Aarestrup FM, McDermott PF. Using
988 Genomics to Track Global Antimicrobial Resistance. Front Public Health. 2019;7:242.
989 32. React. https://reactjs.org/. Accessed 24 November 2020.
990 33. Material Design Lite. https://getmdl.io. Accessed 24 November 2020.
991 34. Phylocanvas. http://phylocanvas.org. Accessed 24 November 2020.
992 35. Leaflet. https://leafletjs.com/. Accessed 24 November 2020.
993 36. Sigma. http://sigmajs.org/. Accessed 24 November 2020.
994 37. Argimon S, Abudahab K, Goater RJ, Fedosejev A, Bhai J, Glasner C, et al. Microreact:
995 visualizing and sharing data for genomic epidemiology and phylogeography. Microb Genom.
996 2016;2(11):e000093.
997 38. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid
998 large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691-3.
999 39. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch technical
1000 descriptions. Core assignment. https://cgps.gitbook.io/pathogenwatch/technical-
1001 descriptions/core-genome-tree/core-assignment. Accessed 24 November 2020.
1002 40. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a
1003 new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol.
1004 2012;19(5):455-77.
1005 41. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch technical
1006 descriptions. Short read assembly. https://cgps.gitbook.io/pathogenwatch/technical-
1007 descriptions/short-read-assembly. Accessed 24 November 2020.
1008 42. CGPS. Pathogenwatch technical descriptions. Speciator.
1009 https://cgps.gitbook.io/pathogenwatch/technical-descriptions/species-assignment/speciator.
1010 Accessed 24 November 2020.
1011 43. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch technical
1012 descriptions. Core filter. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/core-
1013 genome-tree/core-filter. Accessed 24 November 2020.
1014 44. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch technical
1015 descriptions. Tree construction. https://cgps.gitbook.io/pathogenwatch/technical-
1016 descriptions/core-genome-tree/tree-construction. Accessed 24 November 2020.
1017 45. Bennett JS, Jolley KA, Sparling PF, Saunders NJ, Hart CA, Feavers IM, et al. Species
1018 status of *Neisseria gonorrhoeae*: evolutionary and epidemiological inferences from multilocus
1019 sequence typing. BMC Biol. 2007;5:35.
1020 46. Harrison OB, Cehovin A, Skett J, Jolley KA, Massari P, Genco CA, et al. *Neisseria*
1021 *gonorrhoeae* Population Genomics: Use of the Gonococcal Core Genome to Improve
1022 Surveillance of Antimicrobial Resistance. J Infect Dis. 2020;222(11):1816-25.
1023 47. PubMLST. *Neisseria* Multi-Locus Sequence Typing website.
1024 https://pubmlst.org/neisseria/. Accessed 24 November 2020.
1025 48. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb
1026 software, the PubMLST.org website and their applications. Wellcome Open Res. 2018;3:124.
1027 49. Martin IM, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Rapid sequence-based
1028 identification of gonococcal transmission clusters in a large metropolitan area. J Infect Dis.
1029 2004;189(8):1497-505.
1030 50. NG-MAST. *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing database.
1031 http://www.ng-mast.net/. Accessed 24 November 2020.
1032 51. Demczuk W, Sidhu S, Unemo M, Whiley DM, Allen VG, Dillon JR, et al. *Neisseria*
1033 *gonorrhoeae* Sequence Typing for Antimicrobial Resistance, a Novel Antimicrobial Resistance
1034 Multilocus Typing Scheme for Tracking Global Dissemination of N. gonorrhoeae Strains. J Clin
1035 Microbiol. 2017;55(5):1454-68.

1036 52. NG-STAR. *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance.
1037 https://ngstar.canada.ca/. Accessed 24 November 2020.
1038 53. Kwong JC, Gonçalves da Silva A, Dyet K, Williamson DA, Stinear TP, Howden BP, et al.
1039 NGMASTER: in silico Multi-Antigen Sequence Typing for *Neisseria gonorrhoeae*. Microb
1040 Genom. 2016;2(8):e000076.
1041 54. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch technical
1042 descriptions. MLST. https://cgps.gitbook.io/pathogenwatch/technical-descriptions/typing-
1043 methods/mlst. Accessed 24 November 2020.
1044 55. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch technical
1045 descriptions. cgMLST clustering. https://cgps.gitbook.io/pathogenwatch/technical-
1046 descriptions/cgmlst-clusters. Accessed 24 November 2020.
1047 56. Sibson R. SLINK: An optimally efficient algorithm for the single-link cluster method. The
1048 Computer Journal. 1973;16(1):30-4.
1049 57. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch technical
1050 descriptions. Pathogenwatch AMR. https://cgps.gitbook.io/pathogenwatch/technical-
1051 descriptions/antimicrobial-resistance-prediction/pw-amr. Accessed 24 November 2020.
1052 58. Chisholm SA, Wilson J, Alexander S, Tripodo F, Al-Shahib A, Schaefer U, et al. An
1053 outbreak of high-level azithromycin resistant *Neisseria gonorrhoeae* in England. Sex Transm
1054 Infect. 2015;0:1-3.
1055 59. Golparian D, Harris SR, Sanchez-Buso L, Hoffmann S, Shafer WM, Bentley SD, et al.
1056 Genomic evolution of *Neisseria gonorrhoeae* since the preantibiotic era (1928-2013):
1057 antimicrobial use/misuse selects for resistance and drives evolution. BMC Genomics.
1058 2020;21(1):116.
1059 60. Demczuk W, Lynch T, Martin I, Van Domselaar G, Graham M, Bharat A, et al. Whole-
1060 genome phylogenomic heterogeneity of *Neisseria gonorrhoeae* isolates with decreased
1061 cephalosporin susceptibility collected in Canada between 1989 and 2013. J Clin Microbiol.
1062 2015;53(1):191-200.
1063 61. Demczuk W, Martin I, Peterson S, Bharat A, Van Domselaar G, Graham M, et al.
1064 Genomic Epidemiology and Molecular Resistance Mechanisms of Azithromycin-Resistant
1065 *Neisseria gonorrhoeae* in Canada from 1997 to 2014. J Clin Microbiol. 2016;54(5):1304-13.
1066 62. Eyre DW, De Silva D, Cole K, Peters J, Cole MJ, Grad YH, et al. WGS to predict
1067 antibiotic MICs for *Neisseria gonorrhoeae*. J Antimicrob Chemother. 2017;72(7):1937-47.
1068 63. Fifer H, Cole M, Hughes G, Padfield S, Smolarchuk C, Woodford N, et al. Sustained
1069 transmission of high-level azithromycin-resistant *Neisseria gonorrhoeae* in England: an
1070 observational study. Lancet Infect Dis. 2018;18(5):573-81.
1071 64. Sanchez-Buso L, Golparian D, Corander J, Grad YH, Ohnishi M, Flemming R, et al. The
1072 impact of antimicrobials on gonococcal evolution. Nat Microbiol. 2019.
1073 65. Grad YH, Kirkcaldy RD, Trees D, Dordel J, Harris SR, Goldstein E, et al. Genomic
1074 epidemiology of *Neisseria gonorrhoeae* with reduced susceptibility to cefixime in the USA: a
1075 retrospective observational study. Lancet Infect Dis. 2014;14(3):220-6.
1076 66. Jacobsson S, Golparian D, Cole M, Spiteri G, Martin I, Bergheim T, et al. WGS analysis
1077 and molecular resistance mechanisms of azithromycin-resistant (MIC >2 mg/L) *Neisseria*
1078 *gonorrhoeae* isolates in Europe from 2009 to 2014. J Antimicrob Chemother. 2016.
1079 67. Lee RS, Seemann T, Heffernan H, Kwong JC, Goncalves da Silva A, Carter GP, et al.
1080 Genomic epidemiology and antimicrobial resistance of *Neisseria gonorrhoeae* in New Zealand.
1081 J Antimicrob Chemother. 2018;73(2):353-64.
1082 68. Town K, Harris S, Sanchez-Buso L, Cole MJ, Pitt R, Fifer H, et al. Genomic and
1083 Phenotypic Variability in *Neisseria gonorrhoeae* Antimicrobial Susceptibility, England. Emerg
1084 Infect Dis. 2020;26(3):505-15.
1085 69. Yahara K, Nakayama SI, Shimuta K, Lee KI, Morita M, Kawahata T, et al. Genomic
1086 surveillance of *Neisseria gonorrhoeae* to investigate the distribution and evolution of
1087 antimicrobial-resistance determinants and lineages. Microb Genom. 2018;4(8).
1088 70. Kwong JC, Chow EPF, Stevens K, Stinear TP, Seemann T, Fairley CK, et al. Whole-
1089 genome sequencing reveals transmission of gonococcal antibiotic resistance among men who
1090 have sex with men: an observational study. Sex Transm Infect. 2018;94(2):151-7.

1091 71. European Committee on Antimicrobial Susceptibility Testing (EUCAST). Breakpoint
1092 tables for interpretation of MICs and zone diameters. Version 9.0.
1093 https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_9.0_Brea
1094 kpoint_Tables.pdf. Accessed 24 November 2020.
1095 72. Stevenson M, Nunes T, Heuer C, Marshall J, Sanchez J, Thornton R, et al. epiR: Tools
1096 for the Analysis of Epidemiological Data. R package version 1.0-14. https://CRAN.R-
1097 project.org/package=epiR. 2020.
1098 73. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated
1099 platform for visualization and analysis of high-throughput sequence-based experimental data.
1100 Bioinformatics. 2012;28(4):464-9.
1101 74. Chisholm SA, Dave J, Ison CA. High-level azithromycin resistance occurs in *Neisseria*
1102 *gonorrhoeae* as a result of a single point mutation in the 23S rRNA genes. Antimicrob Agents
1103 Chemother. 2010;54(9):3812-6.
1104 75. Underwood A. Gobal Health Research Unit (GHRU) assembly pipeline.
1105 https://gitlab.com/cgps/ghru/pipelines/assembly. Accessed 24 November 2020.
1106 76. Bioinformatics B. FastQC. https://bioinformatics.babraham.ac.uk/projects/fastqc/.
1107 Accessed 24 November 2020.
1108 77. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
1109 data. Bioinformatics. 2014;30(15):2114-20.
1110 78. Song L, Florea L, Langmead B. Lighter: fast and memory-efficient sequencing error
1111 correction without counting. Genome Biol. 2014;15(11):509.
1112 79. Low AJ, Koziol AG, Manninger PA, Blais B, Carrillo CD. ConFindr: rapid detection of
1113 intraspecies and cross-species contamination in bacterial whole-genome sequence data. PeerJ.
1114 2019;7:e6995.
1115 80. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash:
1116 fast genome and metagenome distance estimation using MinHash. Genome Biol.
1117 2016;17(1):132.
1118 81. Li H. Seqtk. https://github.com/lh3/seqtk. Accessed 24 November 2020.
1119 82. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome
1120 assemblies. Bioinformatics. 2011;27(21):2957-63.
1121 83. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for
1122 genome assemblies. Bioinformatics. 2013;29(8):1072-5.
1123 84. Underwood A. Bactinspector. https://gitlab.com/antunderwood/bactinspector. Accessed
1124 24 November 2020.
1125 85. Underwood A. Qualifyr. https://gitlab.com/cgps/qualifyr. Accessed 24 November 2020.
1126 86. European Nucleotide Archive API portal. https://www.ebi.ac.uk/ena/portal/api/. Accessed
1127 24 November 2020.
1128 87. Sánchez-Busó L. *N. gonorrhoeae* Pathogenwatch video demo.
1129 https://vimeo.com/375402189. 2020
1130 88. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch collection.
1131 Sánchez-Busó et al. (2019). https://pathogen.watch/collection/9stz0m94hin7-sanchez-buso-et-
1132 al-2019.
1133 89. Page A, Taylor B, Keane J. Multilocus sequence typing by blast from de novo
1134 assemblies against PubMLST. Journal of Open Source Software. 2016;1(8).
1135 90. Ng LK, Martin I, Liu G, Bryden L. Mutation in 23S rRNA associated with macrolide
1136 resistance in *Neisseria gonorrhoeae*. Antimicrob Agents Chemother. 2002;46(9):3020-5.
1137 91. Roberts MC, Chung WO, Roe D, Xia M, Marquez C, Borthagaray G, et al. Erythromycin-
1138 resistant *Neisseria gonorrhoeae* and oral commensal *Neisseria* spp. carry known rRNA
1139 methylase genes. Antimicrob Agents Chemother. 1999;43(6):1367-72.
1140 92. Cousin S, Jr., Whittington WL, Roberts MC. Acquired macrolide resistance genes in
1141 pathogenic *Neisseria* spp. isolated between 1940 and 1987. Antimicrob Agents Chemother.
1142 2003;47(12):3877-80.
1143 93. Luna VA, Cousin S, Jr., Whittington WL, Roberts MC. Identification of the conjugative
1144 mef gene in clinical *Acinetobacter junii* and *Neisseria gonorrhoeae* isolates. Antimicrob Agents
1145 Chemother. 2000;44(9):2503-6.

1146 94. Rouquette-Loughlin CE, Balthazar JT, Shafer WM. Characterization of the MacA-MacB
1147 efflux system in *Neisseria gonorrhoeae.* J Antimicrob Chemother. 2005;56(5):856-60.
1148 95. Veal WL, Nicholas RA, Shafer WM. Overexpression of the MtrC-MtrD-MtrE efflux pump
1149 due to an mtrR mutation is required for chromosomally mediated penicillin resistance in
1150 Neisseria gonorrhoeae. J Bacteriol. 2002;184(20):5619-24.
1151 96. Cousin SL, Jr., Whittington WL, Roberts MC. Acquired macrolide resistance genes and
1152 the 1 bp deletion in the mtrR promoter in *Neisseria gonorrhoeae*. J Antimicrob Chemother.
1153 2003;51(1):131-3.
1154 97. Warner DM, Shafer WM, Jerse AE. Clinically relevant mutations that cause derepression
1155 of the *Neisseria gonorrhoeae* MtrC-MtrD-MtrE Efflux pump system confer different levels of
1156 antimicrobial resistance and in vivo fitness. Mol Microbiol. 2008;70(2):462-78.
1157 98. Shafer WM, Balthazar JT, Hagman KE, Morse SA. Missense mutations that alter the
1158 DNA-binding domain of the MtrR protein occur frequently in rectal isolates of *Neisseria*
1159 *gonorrhoeae* that are resistant to faecal lipids. Microbiology. 1995;141 (Pt 4):907-11.
1160 99. Tomberg J, Unemo M, Davies C, Nicholas RA. Molecular and structural analysis of
1161 mosaic variants of penicillin-binding protein 2 conferring decreased susceptibility to expanded-
1162 spectrum cephalosporins in *Neisseria gonorrhoeae*: role of epistatic mutations. Biochemistry.
1163 2010;49(37):8062-70.
1164 100. Tomberg J, Unemo M, Ohnishi M, Davies C, Nicholas RA. Identification of amino acids
1165 conferring high-level resistance to expanded-spectrum cephalosporins in the penA gene from
1166 *Neisseria gonorrhoeae* strain H041. Antimicrob Agents Chemother. 2013;57(7):3029-36.
1167 101. Unemo M, Golparian D, Nicholas R, Ohnishi M, Gallay A, Sednaoui P. High-level
1168 cefixime- and ceftriaxone-resistant *Neisseria gonorrhoeae* in France: novel penA mosaic allele
1169 in a successful international clone causes treatment failure. Antimicrob Agents Chemother.
1170 2012;56(3):1273-80.
1171 102. Belland RJ, Morrison SG, Ison C, Huang WM. *Neisseria gonorrhoeae* acquires
1172 mutations in analogous regions of gyrA and parC in fluoroquinolone-resistant isolates. Mol
1173 Microbiol. 1994;14(2):371-80.
1174 103. Rouquette-Loughlin C, Dunham SA, Kuhn M, Balthazar JT, Shafer WM. The NorM efflux
1175 pump of *Neisseria gonorrhoeae* and *Neisseria meningitidis* recognizes antimicrobial cationic
1176 compounds. J Bacteriol. 2003;185(3):1101-6.
1177 104. Lindback E, Rahman M, Jalal S, Wretlind B. Mutations in gyrA, gyrB, parC, and parE in
1178 quinolone-resistant strains of *Neisseria gonorrhoeae*. APMIS. 2002;110(9):651-7.
1179 105. Hu M, Nandi S, Davies C, Nicholas RA. High-level chromosomally mediated tetracycline
1180 resistance in *Neisseria gonorrhoeae* results from a point mutation in the rpsJ gene encoding
1181 ribosomal protein S10 in combination with the mtrR and penB resistance determinants.
1182 Antimicrob Agents Chemother. 2005;49(10):4327-34.
1183 106. Morse SA, Johnson SR, Biddle JW, Roberts MC. High-level tetracycline resistance in
1184 Neisseria gonorrhoeae is result of acquisition of streptococcal tetM determinant. Antimicrob
1185 Agents Chemother. 1986;30(5):664-70.
1186 107. Ashford WA, Golash RG, Hemming VG. Penicillinase-producing *Neisseria gonorrhoeae*.
1187 Lancet. 1976;2(7987):657-8.
1188 108. Ropp PA, Hu M, Olesky M, Nicholas RA. Mutations in ponA, the gene encoding
1189 penicillin-binding protein 1, and a novel locus, penC, are required for high-level chromosomally
1190 mediated penicillin resistance in *Neisseria gonorrhoeae.* Antimicrob Agents Chemother.
1191 2002;46(3):769-77.
1192 109. Olesky M, Hobbs M, Nicholas RA. Identification and analysis of amino acid mutations in
1193 porin IB that mediate intermediate-level resistance to penicillin and tetracycline in *Neisseria*
1194 *gonorrhoeae.* Antimicrob Agents Chemother. 2002;46(9):2811-20.
1195 110. Bilgin N, Richter AA, Ehrenberg M, Dahlberg AE, Kurland CG. Ribosomal RNA and
1196 protein mutants resistant to spectinomycin. EMBO J. 1990;9(3):735-9.
1197 111. Unemo M, Golparian D, Skogen V, Olsen AO, Moi H, Syversen G, et al. *Neisseria*
1198 *gonorrhoeae* strain with high-level resistance to spectinomycin due to a novel resistance
1199 mechanism (mutated ribosomal protein S5) verified in Norway. Antimicrob Agents Chemother.
1200 2013;57(2):1057-61.

1201 112. Fiebelkorn KR, Crawford SA, Jorgensen JH. Mutations in folP associated with elevated
1202 sulfonamide MICs for *Neisseria meningitidis* clinical isolates from five continents. Antimicrob
1203 Agents Chemother. 2005;49(2):536-40.
1204 113. Zhao S, Duncan M, Tomberg J, Davies C, Unemo M, Nicholas RA. Genetics of
1205 chromosomally mediated intermediate resistance to ceftriaxone and cefixime in *Neisseria*
1206 *gonorrhoeae*. Antimicrob Agents Chemother. 2009;53(9):3744-51.
1207 114. Ison CA, Golparian D, Saunders P, Chisholm S, Unemo M. Evolution of *Neisseria*
1208 *gonorrhoeae* is a continuing challenge for molecular detection of gonorrhoea: false negative
1209 gonococcal porA mutants are spreading internationally. Sex Transm Infect. 2013;89(3):197-201.
1210 115. Kersh EN, Allen V, Ransom E, Schmerer M, Cyr S, Workowski K, et al. Rationale for a
1211 *Neisseria gonorrhoeae* Susceptible-only Interpretive Breakpoint for Azithromycin. Clin Infect Dis.
1212 2020;70(5):798-804.
1213 116. Cole MJ, Tan W, Fifer H, Brittain C, Duley L, Hepburn T, et al. Gentamicin, azithromycin
1214 and ceftriaxone in the treatment of gonorrhoea: the relationship between antibiotic MIC and
1215 clinical outcome. J Antimicrob Chemother. 2020;75(2):449-57.
1216 117. Town K, Field N, Harris SR, Sanchez-Buso L, Cole MJ, Pitt R, et al. Phylogenomic
1217 analysis of *Neisseria gonorrhoeae* transmission to assess sexual mixing and HIV transmission
1218 risk in England: a cross-sectional, observational, whole-genome sequencing study. Lancet
1219 Infect Dis. 2020;20(4):478-86.
1220 118. Williamson DA, Chow EPF, Gorrie CL, Seemann T, Ingle DJ, Higgins N, et al. Bridging
1221 of *Neisseria gonorrhoeae* lineages across sexual networks in the HIV pre-exposure prophylaxis
1222 era. Nat Commun. 2019;10(1):3988.
1223 119. De Silva D, Peters J, Cole K, Cole MJ, Cresswell F, Dean G, et al. Whole-genome
1224 sequencing to determine transmission of *Neisseria gonorrhoeae*: an observational study. Lancet
1225 Infect Dis. 2016;16(11):1295-303.
1226 120. Mortimer TD, Pathela P, Crawley A, Rakeman JL, Lin Y, Harris SR, et al. The
1227 distribution and spread of susceptible and resistant *Neisseria gonorrhoeae* across demographic
1228 groups in a major metropolitan center. Clin Infect Dis. 2020.
1229 121. Alfsnes K, Eldholm V, Olsen AO, Brynildsrud OB, Bohlin J, Steinbakk M, et al. Genomic
1230 epidemiology and population structure of *Neisseria gonorrhoeae* in Norway, 2016-2017. Microb
1231 Genom. 2020.
1232 122. Thomas JC, Seby S, Abrams AJ, Cartee J, Lucking S, Vidyaprakash E, et al. Evidence
1233 of Recent Genomic Evolution in Gonococcal Strains With Decreased Susceptibility to
1234 Cephalosporins or Azithromycin in the United States, 2014-2016. J Infect Dis. 2019;220(2):294-
1235 305.
1236 123. Schmerer MW, Abrams AJ, Seby S, Thomas JCt, Cartee J, Lucking S, et al. Genomic
1237 Characterization of *Neisseria gonorrhoeae* Strains from 2016 U.S. Sentinel Surveillance
1238 Displaying Reduced Susceptibility to Azithromycin. Antimicrob Agents Chemother. 2020;64(5).
1239 124. Lan PT, Golparian D, Ringlander J, Van Hung L, Van Thuong N, Unemo M. Genomic
1240 analysis and antimicrobial resistance of *Neisseria gonorrhoeae* isolates from Vietnam in 2011
1241 and 2015-16. J Antimicrob Chemother. 2020;75(6):1432-8.
1242 125. Didelot X, Dordel J, Whittles LK, Collins C, Bilek N, Bishop CJ, et al. Genomic Analysis
1243 and Comparison of Two Gonorrhea Outbreaks. MBio. 2016;7(3).
1244 126. Osnes MN, Didelot X, Korne-Elenbaas J, Alfsnes K, Brynildsrud OB, Syversen G, et al.
1245 Sudden emergence of a *Neisseria gonorrhoeae* clade with reduced susceptibility to extended-
1246 spectrum cephalosporins, Norway. Microb Genom. 2020.
1247 127. Cehovin A, Harrison OB, Lewis SB, Ward PN, Ngetsa C, Graham SM, et al.
1248 Identification of Novel *Neisseria gonorrhoeae* Lineages Harboring Resistance Plasmids in
1249 Coastal Kenya. J Infect Dis. 2018;218(5):801-8.
1250 128. Buckley C, Forde BM, Trembizki E, Lahra MM, Beatson SA, Whiley DM. Use of whole
1251 genome sequencing to investigate an increase in *Neisseria gonorrhoeae* infection among
1252 women in urban areas of Australia. Sci Rep. 2018;8(1):1503.
1253 129. Ryan L, Golparian D, Fennelly N, Rose L, Walsh P, Lawlor B, et al. Antimicrobial
1254 resistance and molecular epidemiology using whole-genome sequencing of *Neisseria*

1255 *gonorrhoeae* in Ireland, 2014-2016: focus on extended-spectrum cephalosporins and
1256 azithromycin. Eur J Clin Microbiol Infect Dis. 2018;37(9):1661-72.
1257 130. Wind CM, de Vries E, Schim van der Loeff MF, van Rooijen MS, van Dam AP, Demczuk
1258 WHB, et al. Decreased Azithromycin Susceptibility of *Neisseria gonorrhoeae* Isolates in Patients
1259 Recently Treated with Azithromycin. Clin Infect Dis. 2017;65(1):37-45.
1260 131. Ezewudo MN, Joseph SJ, Castillo-Ramirez S, Dean D, Del Rio C, Didelot X, et al.
1261 Population structure of *Neisseria gonorrhoeae* based on whole genome data and its relationship
1262 with antibiotic resistance. PeerJ. 2015;3:e806.
1263 132. Unemo M, Golparian D, Eyre DW. Antimicrobial Resistance in *Neisseria gonorrhoeae*
1264 and Treatment of Gonorrhea. Methods Mol Biol. 2019;1997:37-58.
1265 133. Mortimer TD, Grad YH. Applications of genomics to slow the spread of multidrug-
1266 resistant *Neisseria gonorrhoeae.* Ann N Y Acad Sci. 2019;1435(1):93-109.
1267 134. Gernert KM, Seby S, Schmerer MW, Thomas JCt, Pham CD, Cyr SS, et al. Azithromycin
1268 susceptibility of *Neisseria gonorrhoeae* in the USA in 2017: a genomic analysis of surveillance
1269 data. Lancet Microbe. 2020;1(4):e154-e64.
1270 135. Centre for Genomic Pathogen Surveillance (CGPS). Pathogenwatch collection. AZM
1271 resistant gonococcus in global context. https://pathogen.watch/collection/8mk1i3e55olu-azm-
1272 resistant-gonococcus-in-global-context. Accessed 24 November 2020.
1273 136. Wick RR, Holt KE. Benchmarking of long-read assemblers for prokaryote whole genome
1274 sequencing. F1000Res. 2019;8:2138.
1275 137. Souvorov A, Agarwala R, Lipman DJ. SKESA: strategic k-mer extension for scrupulous
1276 assemblies. Genome Biol. 2018;19(1):153.
1277 138. Spiteri G, Cole M, Unemo M, Hoffmann S, Ison C, van de Laar M. The European
1278 Gonococcal Antimicrobial Surveillance Programme (Euro-GASP)--a sentinel approach in the
1279 European Union (EU)/European Economic Area (EEA). Sex Transm Infect. 2013;89 Suppl
1280 4:iv16-8.
1281 139. Schwarcz SK, Zenilman JM, Schnell D, Knapp JS, Hook EW, 3rd, Thompson S, et al.
1282 National surveillance of antimicrobial resistance in *Neisseria gonorrhoeae*. The Gonococcal
1283 Isolate Surveillance Project. JAMA. 1990;264(11):1413-7.
1284 140. Public Health Agency of Canada National Microbiology Laboratory. National surveillance
1285 of antimicrobial susceptibilities of *Neisseria gonorrhoeae*. Annual summary 2017.
1286 https://www.canada.ca/en/public-health/services/publications/drugs-health-products/national-
1287 surveillance-antimicrobial-susceptibilities-neisseria-gonorrhoeae-annual-summary-2017.html.
1288 2017.
1289 141. Lahra MM, Enriquez R, George CRR. Australian Gonococcal Surveillance Programme
1290 Annual Report, 2017. Commun Dis Intell (2018). 2019;43.
1291 142. Paine TC, Fenton KA, Herring A, Turner A, Ison C, Martin I, et al. GRASP: a new
1292 national sentinel surveillance initiative for monitoring gonococcal antimicrobial resistance in
1293 England and Wales. Sex Transm Infect. 2001;77(6):398-401.
1294 143. Weston EJ, Wi T, Papp J. Strengthening Global Surveillance for Antimicrobial Drug-
1295 Resistant *Neisseria gonorrhoeae* through the Enhanced Gonococcal Antimicrobial Surveillance
1296 Program. Emerg Infect Dis. 2017;23(13).
1297 144. Sirivongrangson P, Girdthep N, Sukwicha W, Buasakul P, Tongtoyai J, Weston E, et al.
1298 The first year of the global Enhanced Gonococcal Antimicrobial Surveillance Programme
1299 (EGASP) in Bangkok, Thailand, 2015-2016. PLoS One. 2018;13(11):e0206419.
1300 145. Seale AC, Gordon NC, Islam J, Peacock SJ, Scott JAG. AMR Surveillance in low and
1301 middle-income settings - A roadmap for participation in the Global Antimicrobial Surveillance
1302 System (GLASS). Wellcome Open Res. 2017;2:92.
1303 146. Global Health Research Unit (GHRU), Genomic Surveillance of Antimicrobial
1304 Resistance. https://ghru.pathogensurveillance.net/. Accessed 24 November 2020.
1305 147. Cehovin A, Jolley KA, Maiden MCJ, Harrison OB, Tang CM. Association of *Neisseria*
1306 *gonorrhoeae* Plasmids With Distinct Lineages and The Economic Status of Their Country of
1307 Origin. J Infect Dis. 2020;222(11):1826-36.

148. 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.
149. Marjuki H, Topaz N, Joseph SJ, Gernert KM, Kersh EN, Antimicrobial-Resistant *Neisseria gonorrhoeae* Working Group, et al. Genetic Similarity of Gonococcal Homologs to Meningococcal Outer Membrane Proteins of Serogroup B Vaccine. mBio. 2019;10(5).
150. Russell MW, Jerse AE, Gray-Owen SD. Progress Toward a Gonococcal Vaccine: The Way Forward. Front Immunol. 2019;10:2417.
151. 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.
152. 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.
153. Institute of Medicine (US) 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.
154. 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).
155. 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.
156. European Committee on Antimicrobial Susceptibility Testing (EUCAST). EUCAST clinical breakpoints. https://eucast.org/clinical_breakpoints/. Accessed 2nd December 2020.
157. Clinical Laboratory and Standards Institute (CLSI) guidelines. https://clsi.org/. Accessed 2nd December 2020.