Increased antibiotic susceptibility in *Neisseria gonorrhoeae* through adaptation to the cervical environment

Kevin C Ma[1¶], Tatum D Mortimer[1¶], Allison L Hicks[1], Nicole E Wheeler[2], Leonor Sánchez-Busó[2], Daniel Golparian[3], George Taiaroa[4], Daniel HF Rubin[1], Yi Wang[1], Deborah A Williamson[4], Magnus Unemo[3], Simon R Harris[5], Yonatan H Grad[1,6*]

[1] Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, USA

[2] Centre for Genomic Pathogen Surveillance, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom

[3] WHO Collaborating Centre for Gonorrhoea and other STIs, Swedish Reference Laboratory for STIs, Faculty of Medicine and Health, Örebro University, Örebro, Sweden

[4] Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology and Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

[5] Microbiotica Ltd, Biodata Innovation Centre, Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom

[6] Division of Infectious Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, USA

* Corresponding author

E-mail: ygrad@hsph.harvard.edu (YHG)

¶ These authors contributed equally to this work.

Abstract

Neisseria gonorrhoeae is an urgent public health threat due to rapidly increasing incidence and antibiotic resistance. In contrast with the trend of increasing resistance, clinical isolates that have reverted to susceptibility regularly appear, prompting questions about which pressures compete with antibiotics to shape gonococcal evolution. Here, we used genome-wide association on the largest collection of *N. gonorrhoeae* isolates to date (n=4852) to identify loss-of-function (LOF) mutations in the efflux pump *mtrCDE* operon as a mechanism of increased antibiotic susceptibility and demonstrate that these mutations are overrepresented in cervical isolates relative to urethral isolates (odds ratio (OR) = 3.74, 95% CI [1.98-6.70]). In support of a model in which pump expression incurs a fitness cost in this niche, cervical isolates were also enriched relative to urethral isolates in LOF mutations in the *mtrCDE* activator *mtrA* (OR = 8.60, 95% CI [4.96-14.57]) and in *farA*, a subunit of the FarAB efflux pump (OR = 6.25, 95% CI [3.90-9.83]). In total, approximately 2 in 5 cervical isolates (42.6%) contained a LOF mutation in either the efflux pump components *mtrC* or *farA* or the activator *mtrA.* Our findings extend beyond *N. gonorrhoeae* to other *Neisseria*: *mtrC* LOF mutations are rare (<1%) in the primarily nasopharyngeal-colonizing *N. meningitidis* in a collection of 14,798 genomes but enriched in a heterosexual urethritis-associated lineage (8.6%, $p = 9.90\times10^{-5}$), indicating that efflux pump downregulation contributes broadly to the adaptation of pathogenic *Neisseria* to the female urogenital tract. Overall, our findings highlight the impact of integrating microbial population genomics with host metadata and demonstrate how host environmental pressures can lead to increased antibiotic susceptibility.

Introduction

Neisseria gonorrhoeae is the causative agent of the sexually transmitted disease gonorrhea. Antibiotics have played a key role in shaping gonococcal evolution[1-3], with *N. gonorrhoeae* gaining resistance to each of the first line antibiotics used to treat it[4-6]. As *N. gonorrhoeae* is an obligate human pathogen, the mucosal niches it infects—most commonly including the urethra, cervix, pharynx, and rectum—must also influence its evolution[7]. The gonococcal phylogeny suggests the interaction of these factors, with an ancestral split between a drug-susceptible lineage circulating primarily in heterosexuals and a drug-resistant lineage circulating primarily in men who have sex with men[3].

Despite the deeply concerning increase in antibiotic resistance reported in gonococcal populations globally[8], some clinical isolates of *N. gonorrhoeae* have become more susceptible to antibiotics[9,10]. This unexpected phenomenon prompts questions about which environmental pressures could be drivers of increased susceptibility and the mechanisms by which suppression or reversion of resistance may occur. To address these questions, we analyzed the genomes of a global collection of clinical isolates together with patient demographic and clinical data to identify mutations associated with increased susceptibility and define the environments in which they appear.

Antibiotic susceptibility is influenced by unknown genetic determinants in *N. gonorrhoeae*

We first assessed how well variation in antibiotic resistance phenotype was captured by the presence and absence of known resistance markers. To do so, we assembled and examined a global dataset comprising the genomes and minimum inhibitory concentrations (MICs) of 4852 isolates collected across 65 countries and 38 years (Figure 1, Supplementary Table 1)[3,9-20]. We modeled log-transformed MICs using multiple regression on a panel of experimentally characterized resistance markers for the three most clinically relevant antibiotics[21-23] (Supplementary Table 2). This enabled us to make quantitative predictions of MIC based on known genotypic markers and to assess how well these markers predicted true MIC values. For the macrolide azithromycin, we observed that 434/4505 (9.63%) isolates had predicted MICs that deviated by two dilutions or more from their reported phenotypic values. The majority (59.4%) of these isolates had MICs that were lower than expected, indicative of increased susceptibility unexplained by the genetic determinants in our model. Overall MIC variance explained by known resistance mutations was relatively low (adjusted R^2 = 0.667), in agreement

with prior studies that employed whole-genome supervised learning algorithms to predict azithromycin resistance[24]. MIC variance explained by known resistance mutations was also low for ceftriaxone (adjusted R^2 = 0.674) but higher for ciprofloxacin (adjusted R^2 = 0.937), with 2.02% and 2.90% of strains, respectively, exhibiting two dilutions or lower reported MICs compared to predictions, similarly indicating unexplained susceptibility. The predictive modeling results therefore suggested unknown modifiers that promote susceptibility for multiple classes of antibiotics in *N. gonorrhoeae*.

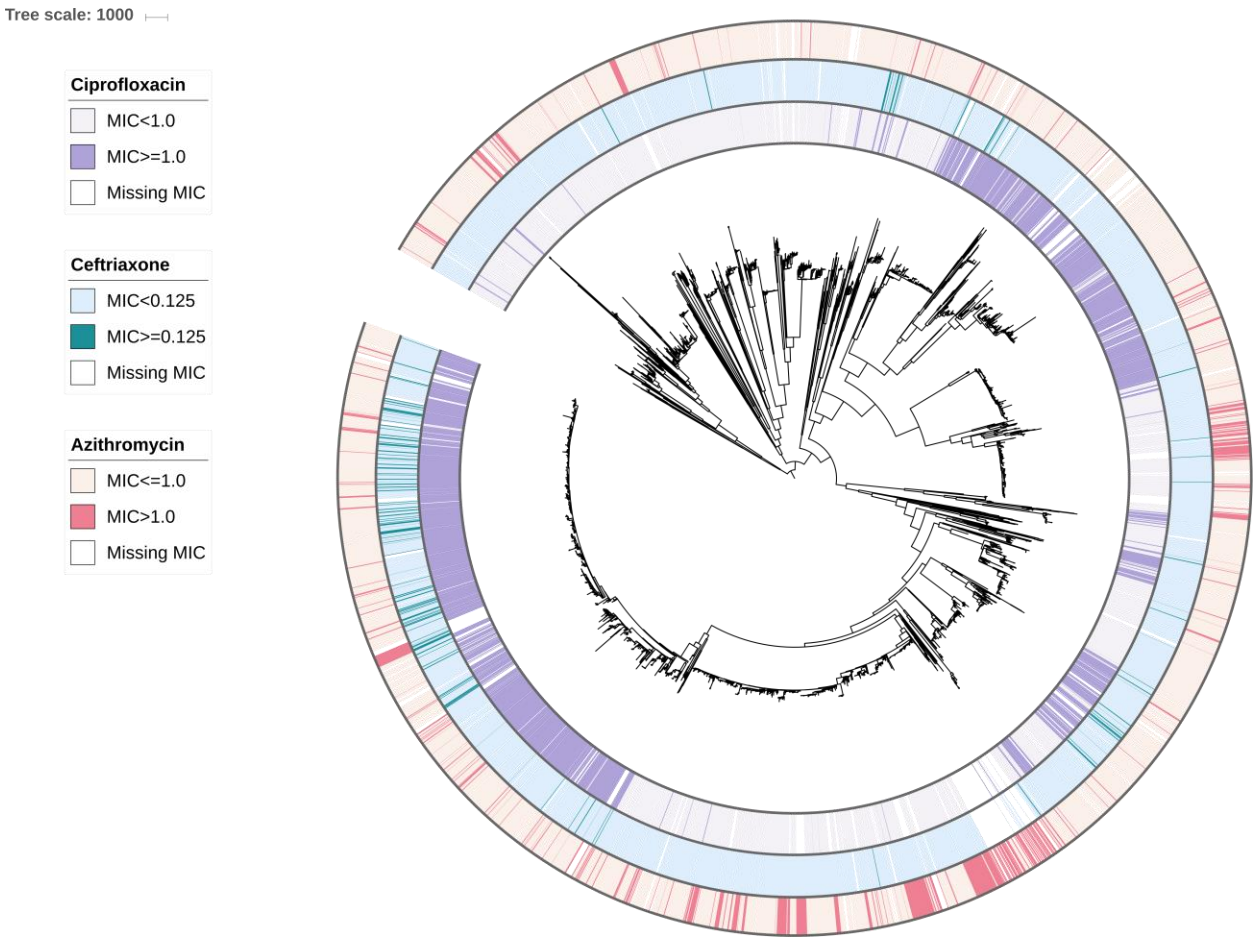


Figure 1 – Population structure and susceptibility profile of *N. gonorrhoeae* global meta-analysis collection. A midpoint rooted recombination-corrected maximum likelihood phylogeny of 4852 genomes based on 68697 SNPs (Supplementary Table 1) was annotated with binarized resistance (ciprofloxacin) or decreased susceptibility (azithromycin, ceftriaxone) values. Annotation rings are listed in order of ciprofloxacin (purple), ceftriaxone (blue), and azithromycin (pink) from innermost to outermost. Branch length represents total number of substitutions after removal of predicted recombination.

Microbial genome-wide association identifies a variant that contributes to susceptibility in the operon encoding the MtrCDE efflux pump

To identify novel antibiotic susceptibility loci in an unbiased manner, we conducted a bacterial genome-wide association study (GWAS). We used a linear mixed model framework to control for population structure, and we used unitigs constructed from genome assemblies to capture SNPs, indels, and accessory genome elements (see methods)[25-27]. Unitigs are a flexible representation of the genetic variation across a dataset that are constructed using compacted de Bruijn graphs, and have been previously applied as markers for microbial GWAS[28]. We performed a GWAS on the sequences of 4505 isolates with associated azithromycin MICs using a Bonferroni-corrected significance threshold of 3.38×10^{-7}. The linear mixed model adequately controlled for population structure (Supplementary Figure 1), and the proportion of phenotypic MIC variance attributable to genotype (i.e., narrow-sense heritability) estimated by the linear mixed model was high ($h^2 = 0.97$). In line with this, we observed highly significant unitigs with positive effect sizes corresponding to the known resistance substitutions C2611T and A2059G (*E. coli* numbering) in the 23S ribosomal RNA gene (Figure 2)[29]. The next most significant variant was a unitig associated with increased susceptibility that mapped to *mtrC* (β, or effect size on the $\log_2$-transformed MIC scale, = -2.82, 95% CI [-3.06, -2.57]; *p*-value = 2.81×10^{-108}). Overexpression of the *mtrCDE* efflux pump operon has been shown to decrease gonococcal susceptibility to a range of hydrophobic acids and antimicrobial agents[4,30], and conversely, knockout of the pump results in multi-drug hypersusceptibility[31]. To assess whether this *mtrC* variant was associated with increased susceptibility to other antibiotics, we performed GWAS for ceftriaxone (for which MICs were available from 4497 isolates) and for ciprofloxacin (4135 isolates). We recovered known ceftriaxone resistance mutations including recombination in the *penA* gene and ciprofloxacin resistance substitutions in DNA gyrase (*gyrA*). In agreement with the known pleiotropic effect of the MtrCDE efflux pump[31], we observed the same *mtrC* unitig at genome-wide significance associated with increased susceptibility to both ceftriaxone (β = -1.18, 95% CI [-1.34, -1.02]; *p*-value = 2.00×10^{-44}) and ciprofloxacin (β = -1.29, 95% CI [-1.54, -1.04]; *p*-value = 1.87×10^{-23}) (Figure 2). Across all three drugs, heritability estimates for this *mtrC* variant were comparable to that of prevalent major resistance determinants (azithromycin h^2: 0.323; ceftriaxone h^2: 0.208; ciprofloxacin h^2: 0.155), indicating that unexplained susceptibility in our model could be partially addressed by inclusion of this mutation.

Annotation of the *mtrC* unitig revealed that it represented a two base pair deletion in a 'GC' dinucleotide hexarepeat, leading to early termination of *mtrC* translation and probable loss of MtrCDE activity[32] (Figure 2 inset). We also checked whether the two base pair deletion would affect recognition by any of the gonococcal methylases[33], but no methylase target motif sites mapped to the hexarepeat or its direct surrounding sequences. A laboratory-generated gonococcal mutant with a four base pair deletion in this same *mtrC* dinucleotide hexarepeat exhibited multi-drug susceptibility[32], and clinical gonococcal isolates hypersensitive to erythromycin were shown to have mutations mapping to this locus[34]. To directly test the hypothesis that the two base pair deletion also contributed to increased susceptibility for the panel of antibiotics we examined, we complemented the mutation in a clinical isolate belonging to the multidrug-resistant lineage ST-1901[35] and observed significant increases in MICs for all three antibiotics, as predicted by the GWAS (Supplementary Table 4).

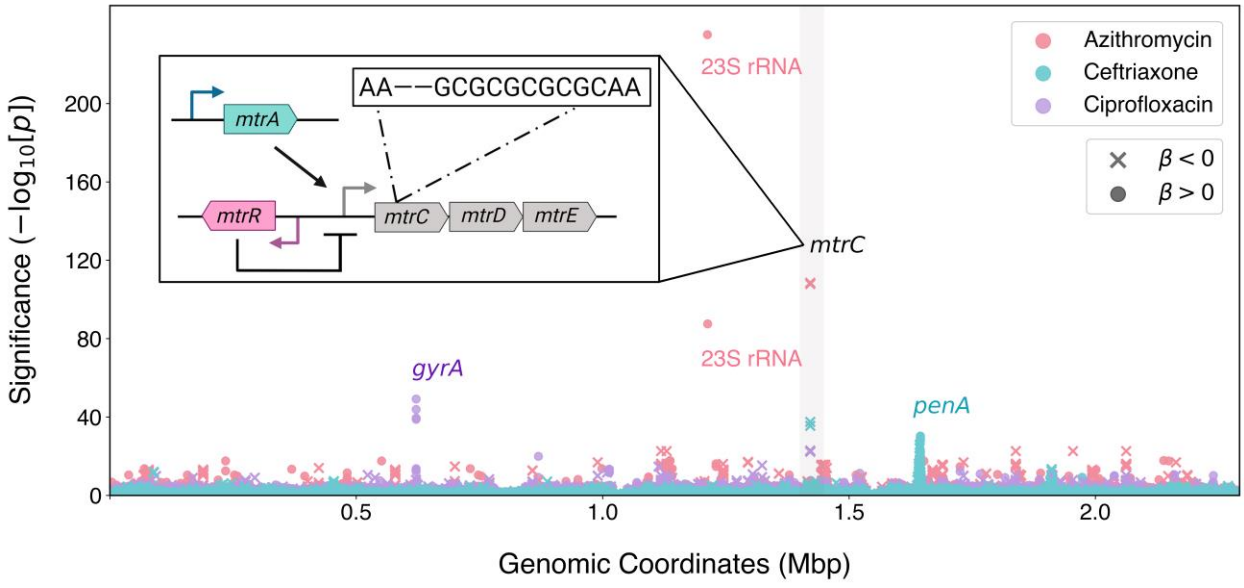


Figure 2 – GWAS identifies a variant mapping to *mtrC* associated with increased susceptibility to azithromycin, ceftriaxone, and ciprofloxacin. Negative $\log_{10}$-transformed *p*-values for unitigs tested in GWAS on MICs to azithromycin (pink, n=4505), ceftriaxone (blue, n=4497), and ciprofloxacin (purple, n=4135) are shown in the Manhattan plot. The sign of the GWAS regression coefficient β (with positive indicating an association with increased resistance and negative indicating an association with increased susceptibility) is indicated by symbol shape, as depicted in the legend. Labels indicate known influential resistance determinants, and the *mtrC* variant associated with increased susceptibility was highlighted in gray. A full list of the annotated significant unitigs for each antibiotic can be found in Supplementary Table 3. Inset: schematic of the *mtr* genetic regulon including structural genes *mtrCDE*, the activator *mtrA*, and the repressor *mtrR*. The approximate genomic location within *mtrC* and specific nucleotide

change of the *mtrC* GWAS variant relative to the gonococcal NCCP11945 reference genome (i.e., a two base pair deletion in a 'GC' dinucleotide repeat) is shown.

We searched for additional *mtrC* loss-of-function (LOF) mutations and found six clinical isolates with genomes encoding indels outside of the dinucleotide hexarepeat that also were associated with increased susceptibility (Supplementary Figure 2A). Ten isolates that had acquired the two base pair deletion also have a two base pair insertion elsewhere in *mtrC* that restores the original coding frame, suggesting that loss of MtrC function may be reverted by further mutation or recombination (Supplementary Figure 2A). In line with this, *mtrC* LOF mutations have emerged numerous times throughout the phylogeny (Supplementary Figure 3), indicative of possible repeated losses of a dinucleotide in the hexarepeat region due to DNA polymerase slippage, which may occur at a higher rate than single nucleotide nonsense mutations[36]. In total, including all strains with *mtrC* frameshift mutations and excluding revertants, we identified 185 isolates (3.82%) that encoded a LOF allele of *mtrC*. Presence of the *mtrC* LOF mutation in isolates with known resistance markers was correlated with significantly reduced MICs (Supplementary Figure 4), and inclusion of *mtrC* LOF mutations in our linear model increased adjusted R^2 values (azithromycin: 0.667 to 0.704; ceftriaxone: 0.674 to 0.690; ciprofloxacin: 0.937 to 0.939), decreased the proportion of strains with unexplained susceptibility (azithromycin: 5.73% to 3.88%; ceftriaxone: 2.02% to 1.73%; ciprofloxacin: 2.90% to 2.42%), and significantly improved model fit (p-value < 2.2×10^{-16} for all three antibiotics; Likelihood-ratio χ^2 test for nested models). *mtrC* LOF strains were identified in 28 of the 66 countries surveyed and ranged in isolation date from 2002 to 2017. Because most strains in this dataset were collected within the last two decades, we also examined a dataset of strains collected in Denmark from 1928 to 2013 to understand the historical prevalence of *mtrC* LOF mutations[37]. We observed an additional 10 strains with the 'GC' two base pair deletion ranging in isolation date from 1951-2000, indicating that *mtrC* LOF strains have either repeatedly arisen or persistently circulated for decades. Our results demonstrate that a relatively common mechanism of gonococcal acquired antibiotic susceptibility is a two base pair deletion in *mtrC* and that such mutations are globally and temporally widespread.

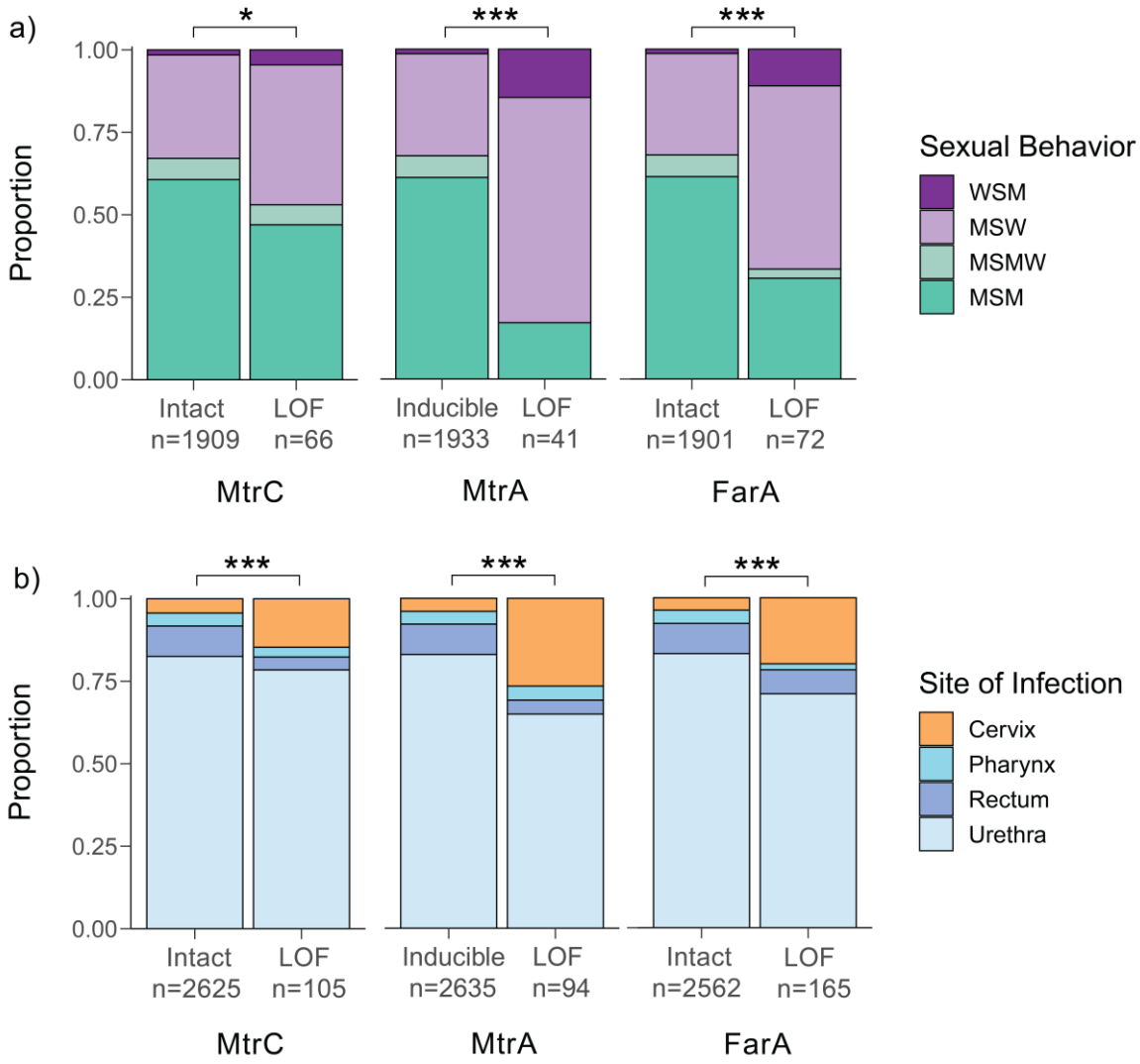
Loss of the MtrCDE efflux pump is associated with cervical infection

The MtrCDE pump has been demonstrated to play a critical role in gonococcal survival in the presence of human neutrophils and in the female murine genital tract model of gonococcal infection, and overexpression of *mtrCDE* results in substantial fitness benefits for dealing with

both antimicrobial and environmental pressures[38-41]. The relative frequency of the *mtrC* LOF mutations we observe (occurring in approximately 1 in every 25 isolates) thus seems unusual for a mutation predicted to be deleterious for human infection. *mtrC* LOF strains do not grow more or less quickly *in vitro* than *mtrC* wild-type strains, indicating that this mutation does not confer a simple fitness benefit due to reduced energetic cost[34,38,42]. Instead, we hypothesized that there are unique environments that select for non-functional efflux pump.

We aggregated patient-level metadata across included studies on sex partner preferences and anatomical site of infection. Sexual behavior and *mtrC* genotypic information was available for 1975 isolates from individual patients. There was a significant association between *mtrC* LOF and sexual behavior (p-value = 0.04021; Fisher's exact test) (Figure 3a), and *mtrC* LOF occurred more often in isolates from men who have sex with women (MSW) (28/626, 4.47%) compared to isolates from men who have sex with men (MSM) (31/1189, 2.61%) (OR = 1.75, 95% CI [1.00-3.04], p-value = 0.037; Fisher's exact test). To understand whether anatomical selective pressures contributed to this enrichment, we analyzed the site of infection and *mtrC* genotypic information available for 2730 isolates. *mtrC* LOF mutations were significantly associated with site of infection (p-value = 6.49×10^{-5}; Fisher's exact test) and were overrepresented particularly in cervical isolates: 16 out of 129 (12.4%) cervical isolates contained an *mtrC* LOF mutation compared to 82 out of 2249 urethral isolates (3.65%; OR = 3.74, 95% CI [1.98-6.70], p-value = 4.71×10^{-5}; Fisher's exact test), 3 out of 106 pharyngeal isolates (2.83%; OR = 4.83, 95% CI [1.33-26.63], p-value = 0.00769; Fisher's exact test), and 4 out of 246 rectal isolates (1.63%; OR = 8.52, 95% CI [2.67-35.787], p-value = 2.39×10^{-5}; Fisher's exact test) (Figure 3b). Because our meta-analysis collection comprises datasets potentially biased by preferential sampling for drug-resistant strains, we validated our epidemiological associations on a set of 2186 sequenced isolates, corresponding to all cultured isolates of *N. gonorrhoeae* in the state of Victoria, Australia in 2017[43]. We again observed significant associations between *mtrC* LOF and sexual behavior (p-value = 0.0180; Fisher's exact test) as well as anatomical site of infection (p-value = 0.0256; Fisher's exact test) (Supplementary Figure 5, Supplementary Text). *mtrC* LOF mutations were again overrepresented in cervical isolates: 9 out of 227 (3.96%) cervical isolates contained an *mtrC* LOF mutation compared to 15 out of 882 urethral isolates (1.70%; OR = 2.38, 95% CI [0.91-5.91], p-value = 0.0679; Fisher's exact test), 3 out of 386 pharyngeal isolates (0.78%; OR = 5.26, 95% CI [1.29-30.51], p-value = 0.0117; Fisher's exact test), and 7 out of 632 rectal isolates (1.11%; OR = 3.68, 95% CI [1.20-11.78], p-value = 0.0173; Fisher's exact test). These

231 results indicate that environmental pressures unique to female urogenital infection may select
232 for loss of the primary gonococcal efflux pump resulting in broadly increased susceptibility to
233 antibiotics and host-derived antimicrobial peptides.
234



235

236 **Figure 3 – Gonococcal *mtrC*, *mtrA*, and *farA* LOF mutations are associated with patient**
237 **sexual behavior and site of infection in the global meta-analysis dataset.** a) Sexual
238 behavior of patients infected with isolates with either intact or LOF alleles of *mtrC* (left), *mtrA*
239 (middle), or *farA* (right). b) Site of infection in patients infected with isolates with either intact or
240 LOF alleles of *mtrC* (left), *mtrA* (middle), or *farA* (right). *mtrA* alleles were predicted as LOF only
241 in the absence of other epistatic Mtr overexpression mutations. Statistical significance between
242 genotype and patient metadata was assessed by Fisher's exact test: * $p<0.05$, ** $p<0.01$, and
243 *** $p<0.001$. WSM = women who have sex with men, MSW = men who have sex with women,
244 MSMW = men who have sex with men and women, MSM = men who have sex with men.
245 **Activator loss-of-function offers an additional level of adaptive regulatory control**

The association of *mtrC* LOF mutations with cervical specimens suggests that other mutations that downregulate expression of the *mtrCDE* operon should also promote adaptation to the cervical niche. The MtrCDE efflux pump regulon comprises the MtrR repressor and the MtrA activator (Figure 2 inset), the latter of which exists in two allelic forms: a wild-type functional gene capable of inducing *mtrCDE* expression and a variant rendered non-functional by an 11-bp deletion near the 5' end of the gene[44] (Supplementary Figure 2B). Knocking out *mtrA* has a detrimental effect on fitness in the gonococcal mouse model, and epistatic *mtrR* mutations resulting in overexpression of *mtrCDE* compensate for this fitness defect by masking the effect of the *mtrA* knockout[39]. Prior work assessing the genomic diversity of *mtrA* in a set of 922 primarily male urethral specimens found only four isolates with the 11-bp deletion (0.43%)[45], seemingly in agreement with the *in vivo* importance of *mtrA*. However, in our global meta-analysis dataset, 362/4842 isolates (7.48%) were predicted to be *mtrA* LOF, of which the majority (357/362, 98.6%) were due to the 11-bp deletion. Of the 4842 isolates, 268 (5.53%) had *mtrA* LOF mutations in non-*mtrCDE* overexpression backgrounds (as defined by the absence of known *mtrR* promoter or coding sequence mutations or *mtrCDE* mosaic alleles) and therefore not epistatically masked. We repeated our epidemiological associations on these *mtrA* LOF strains without concurrent overexpression mutations and observed highly significant associations with reported patient sexual behavior (p-value = 1.81×10^{-11}; Fisher's exact test) and site of infection (p-value = 1.64×10^{-12}; Fisher's exact test) (Figure 3). As with *mtrC* LOF mutations, *mtrA* LOF mutations were significantly overrepresented in cervical isolates: 25 out of 129 (19.4%) cervical isolates contained an *mtrA* LOF mutation compared to 61 out of 2248 urethral isolates (2.71%; OR = 8.60, 95% CI [4.96-14.57], p-value = 4.60×10^{-13}; Fisher's exact test), 4 out of 106 pharyngeal isolates (3.78%; OR = 6.09, 95% CI [2.00-24.93], p-value = 0.000240; Fisher's exact test), and 4 out of 246 rectal isolates (1.63%; OR = 14.43, 95% CI [4.81-58.52], p-value = 3.00×10^{-9}; Fisher's exact test). In the Australian validation cohort[43], the majority (81/85, 95.3%) of *mtrA* LOF strains had concurrent *mtrCDE* overexpression mutations, so it was not possible to test for these associations. In such genetic backgrounds where overexpression mutations mask the effect of *mtrA* LOF, *mtrC* LOF is the preferred method of efflux pump downregulation: the majority of *mtrC* LOF mutations in both the global dataset (174/180, 96.7%) and the Australian cohort (33/35, 94.3%) occurred in backgrounds with known *mtr* overexpression mutations. Phylogenetic analysis showed that the distribution of *mtrA* LOF differed from that of *mtrC* LOF with fewer introductions but more sustained transmission and that the two mutations were largely non-overlapping (Supplementary Figure 3). Our results

indicate that multiple adaptive paths for MtrCDE efflux pump downregulation exist depending on genetic interactions with other concurrent mutations in the *mtrCDE* regulon.

A second class of proton-dependent efflux pumps appears to contribute to adaptation via loss of function mutations

The associations we observed in the *mtrCDE* regulon raised the question of the mechanism by which the cervical environment could select for pump downregulation. Recent work on *Pseudomonas* suggested one possible model: overexpression of homologous *P. aeruginosa* efflux pumps belonging to the same resistance/nodulation/cell division (RND) proton/substrate antiporter family as MtrCDE results in a fitness cost due to increased cytoplasmic acidification[46]. This fitness cost was only observed in anaerobic conditions, where aerobic respiration cannot be used to dissipate excess protons efficiently[46]. Analogous conditions in the female urogenital tract, potentially augmented by environmental acidity, could create a similar selective pressure during human infection that leads to pump downregulation or loss.

This model predicts that adaptation to these conditions would similarly result in downregulation of FarAB, the other proton-substrate antiporter efflux pump in *N. gonorrhoeae*. FarAB is a member of the major facilitator superfamily (MFS) of efflux pumps and effluxes long-chain fatty acids[47,48]. In our global dataset, 332/4838 (6.86%) of isolates were predicted to have *farA* LOF mutations, of which the majority (316/332; 95.2%) were due to indels in a homopolymeric stretch of eight 'T' nucleotides near the 5' end of the gene (Supplementary Figure 2C). *farA* LOF mutations were associated with patient sexual behavior (p-value = 5.06×10^{-10}; Fisher's exact test) and site of infection (p-value = 1.78×10^{-12}; Fisher's exact test) and overrepresented in cervical isolates: 33 out of 129 (25.6%) cervical isolates contained a *farA* LOF mutation compared to 117 out of 2246 urethral isolates (5.21%; OR = 6.25, 95% CI [3.90-9.83], p-value = 3.24×10^{-13}; Fisher's exact test), 3 out of 106 pharyngeal isolates (2.83%; OR = 11.70, 95% CI [3.50-61.61], p-value = 3.80×10^{-7}; Fisher's exact test), and 12 out of 246 rectal isolates (4.88%; OR = 6.66, 95% CI [3.19-14.80], p-value = 1.57×10^{-8}; Fisher's exact test) (Figure 3). *farA* LOF mutations were prevalent also in our Australian validation dataset[43] (225/2180; 10.32%) and again associated with sexual behavior (p-value < 2.20×10^{-16}; Fisher's exact test) and site of infection (p-value < 2.20×10^{-16}; Fisher's exact test) (Supplementary Figure 5). The phylogenetic distribution of *farA* LOF indicated sustained transmission (Supplementary Figure 3) and overlapped with that of *mtrA* LOF mutations, potentially indicating additive contributions to cervical adaptation. Furthermore, MtrR activates *farAB* expression by repressing the *farR*

repressor[49]. This cross-talk between the two efflux pump operons indicates that in *mtrCDE* overexpression strains where MtrR activity is impaired, the effect of *farA* LOF – like *mtrA* LOF – may be masked. We did not observe frequent LOF mutations in the sodium gradient-dependent MATE family efflux pump NorM[50] or in the ATP-dependent ABC family pump MacAB[51] (Supplementary Table 5). The prevalence and cervical enrichment of *farA* LOF mutations and the relative rarity of LOF mutations in other non-proton motive force-driven pumps suggests that cytoplasmic acidification may be a mechanism by which the female urogenital tract selects for efflux pump loss.

Meningococcal evolution of efflux pump loss-of-function is driven by urogenital adaptation

N. meningitidis, a species closely related to *N. gonorrhoeae,* colonizes the oropharyngeal tract and can cause invasive disease, including meningitis and septicemia[52]. We characterized *mtrC* diversity in a collection of 14,798 *N. meningitides* genomes, reasoning that the cervical environmental pressures that select for efflux pump LOF in the gonococcus will be rarely encountered by the meningococcus. In agreement with this, the 'GC' hexarepeat associated with most gonococcal *mtrC* LOF mutations was less conserved in *N. meningitidis*; only 9684/14798 (65.4%) isolates contained an intact hexarepeat compared to 4644/4847 (95.8%) of *N. gonorrhoeae* isolates (*p*-value < 2.2e-16; Fisher's exact test). In this same collection, we observed *mtrC* LOF due to deletions in the hexarepeat region in only 82 meningococcal isolates (0.55%), with a similar frequency of 25/4059 (0.62%) in a curated dataset comprising all invasive meningococcal disease isolates collected in the UK from 2009-2013[53]. The observed interruption of 'GC' dinucleotide repeats, predicted to result in a lower mutation rate[54], and the relative rarity of *mtrC* LOF mutations suggests that efflux pump loss is not generally adaptive in *N. meningitidis.* However, a urogenitally-adapted meningococcal lineage has recently emerged in the US associated with outbreaks of non-gonococcal urethritis in heterosexual patients[55-57]. In isolates from this lineage, the prevalence of *mtrC* LOF mutations was 18/207 (8.70%), substantially higher than typical *N. meningitidis* and comparable to the prevalence of gonococcal *mtrC* LOF mutations in MSW in our global dataset (4.47%). We compared the frequency of *mtrC* LOF mutations in the urogenital lineage to geographically and genetically matched isolates (i.e., all publicly available n=456 PubMLST ST-11 North American isolates) and observed a significant difference in prevalence (18 out of 207 or 8.70% versus 2 out of 249 or 0.80%; OR = 11.71, 95% CI [2.75-105.37], *p*-value = 3.31×10^{-5}; Fisher's exact test). Most *mtrC* LOF mutations occurred due to the same hexarepeat two base pair deletion that we

previously observed for *N. gonorrhoeae*, and in line with this, *mtrC* LOF in this urogenital lineage arose multiple times independently similarly to gonococcal *mtrC* LOF mutations (Figure 4, Supplementary Figure 3). *farA* LOF mutations were not observed in this meningococcal lineage. We conclude that MtrCDE efflux pump LOF is rare in typical meningococcal strains that inhabit the oropharynx but elevated in frequency in a unique urogenitally-adapted lineage circulating in heterosexuals, indicative of potential ongoing adaptation to the cervical niche. Our results suggest that efflux pump loss is broadly adaptive for cervical colonization across pathogenic *Neisseria*.

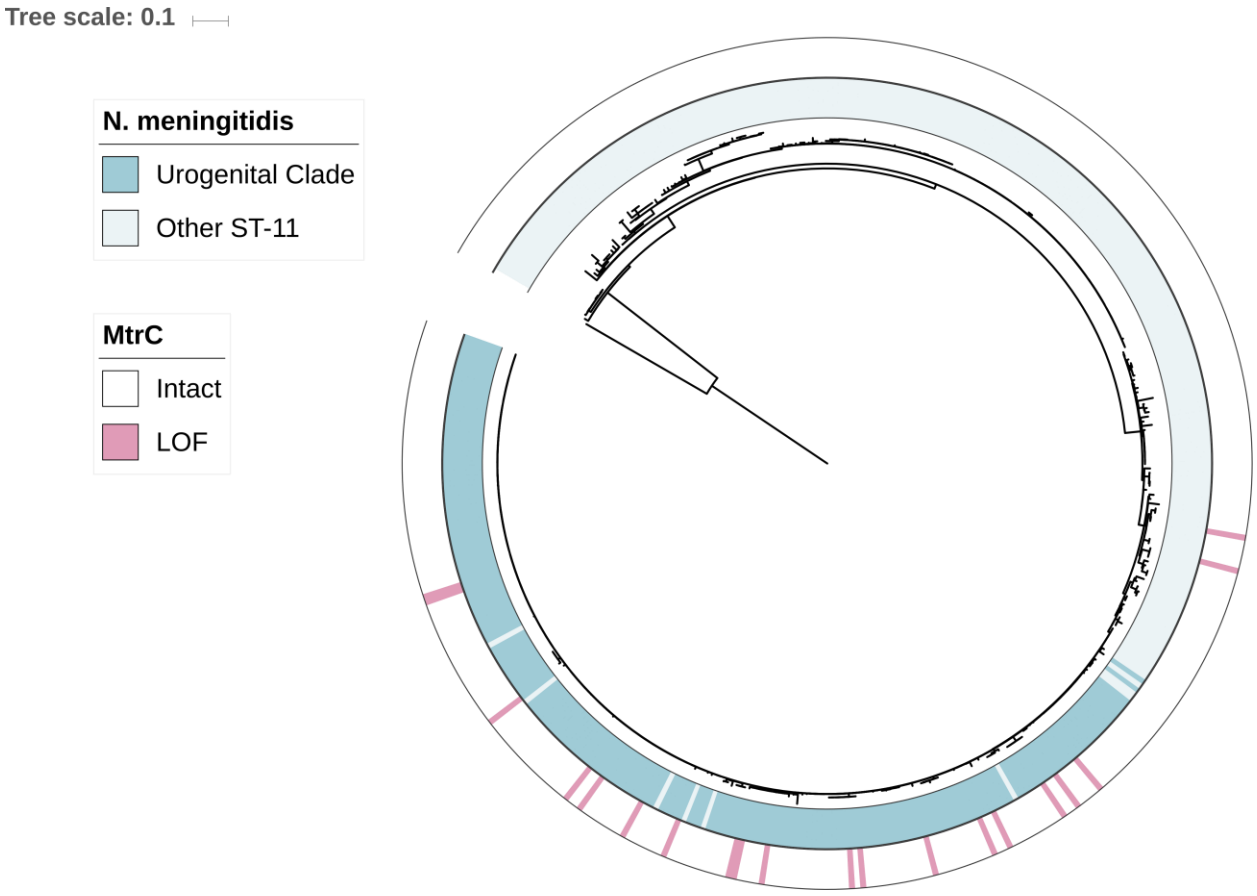


Figure 4 – *mtrC* LOF mutations are enriched in a lineage of ST-11 urogenitally-adapted *N. meningitidis*. A core-genome maximum likelihood phylogeny based on 25045 SNPs was estimated of all North American ST-11 *N. meningitidis* strains from PubMLST (n=456; accessed 2019-09-03) rooted with meningococcal reference genome MC58. Membership in ST-11 urogenital clade (blue) defined as in Retchless et al., 2018[56]. Genomes with *mtrC* LOF mutations are indicated in pink. Branch length represents substitutions per site.

Discussion

In an era in which widespread antimicrobial pressure has led to the emergence of extensively drug-resistant *N. gonorrhoeae*[58], isolates that appear to have reverted to susceptibility still arise[9,10], demonstrating that antibiotic and host environmental pressures interact to shape the evolution of *N. gonorrhoeae*. Here, we showed that frameshift-mediated truncations in the *mtrC* component of the MtrCDE efflux pump are the primary mechanism for epistatic increases in antibiotic susceptibility across a global collection of clinical gonococcal isolates, as suggested by prior work[32,34]. *mtrC* LOF mutations are enriched in cervical isolates and a frameshifted form of the pump activator MtrA exhibits similar trends, supporting a model in which reduced or eliminated *mtrCDE* efflux pump expression contributes to adaptation to the female genital tract. We hypothesized that the mechanism by which this occurs is through increased cytoplasmic acidification in anaerobic conditions[46] and demonstrated that LOF mutations in *farA*, encoding a subunit of the other proton motive force-driven pump FarAB, were likewise enriched in cervical isolates. The LOF mutations we observed in *mtrC* and *farA* primarily occurred in short homopolymeric sequences (though with low numbers of repeated units) and thus may occur at a frequency higher than baseline mutation rate, similar to other resistance suppressor mutations[59]. In total, 42.6% of cervical isolates in the global dataset and 32.6% in the validation dataset contained a LOF mutation in either *mtrC, farA,* or *mtrA*, indicating that efflux pump downregulation via multiple genetic mechanisms is prevalent in cervical infection. These results complement prior studies suggesting that *mtrR* LOF resulting in increased resistance to fecal lipids plays a critical role in gonococcal adaptation to the rectal environment[60,61], and taken together suggest a model in which the fitness benefit of efflux pump expression is highly context dependent.

Other selective forces could also have contributed to the observed enrichment of LOF mutations in cervical isolates. For instance, iron levels modulate *mtrCDE* expression through Fur (the ferric uptake regulator) and MpeR[62]. Iron limitation results in increased expression of *mtrCDE*, and conversely, iron enrichment result in decreased expression, suggesting a fitness cost for *mtrCDE* expression during high iron conditions. Variation in environmental iron levels, such as in the menstrual cycle, may provide another selective pressure for LOF mutations to arise particularly when MtrR function is impaired through active site or promoter mutations. Differing rates of antibiotic use for gonorrhea in men and women due to increased asymptomatic infection in women might also select for *mtrC* LOF mutations, but this would not

399 explain the associations we observed for the non-antibiotic substrate efflux pump *farAB* or the
400 increased frequency of *mtrC* LOF mutations in urogenitally-adapted meningococci. RNA
401 sequencing from men and women infected with gonorrhea demonstrated a 4-fold lower
402 expression of *mtrCDE* in women, re-affirming the idea that efflux pump expression in the
403 female genital tract incurs a fitness cost[63].

404

405 Despite significant associations, only a proportion of cervical isolates exhibited these LOF
406 genotypes, suggesting variation in cervix-associated pressures or indicating that cervical
407 culture specimens were obtained before niche pressures could select for pump
408 downregulation. This variation could also lead to mixed populations of efflux pump WT and
409 LOF strains; however, because only one clonal isolate per site per patient is typically
410 sequenced in clinical surveillance studies, we would be unable to detect this intra-host patient
411 diversity. Targeted amplicon sequencing of LOF loci directly from patients in future studies
412 would help to assess whether this intra-host diversity plays a role in infection and transmission.
413 Additionally, while the cervix is the primary site of infection and source for culture in women,
414 the selective pressures at play may include other sites more broadly in the female genital tract
415 and may be influenced by the presence of other microbial species both pathogenic and
416 commensal.

417

418 Our model extended to the other pathogenic *Neisseria* species, *N. meningitidis*, in that a
419 urogenital clade transmitting in primarily heterosexual populations appeared to be undergoing
420 further urogenital adaptation via the same *mtrC* frameshift mutation that was most commonly
421 observed for *N. gonorrhoeae*. In the absence of data on cases of cervicitis, we hypothesized
422 that for this meningococcal lineage, efflux pump LOF emerged in the female urogenital tract
423 and was transmitted to heterosexual men resulting in the enrichment we observed. Efflux
424 pumps are common across Gram-negative bacteria[64], and their loss may be a general
425 adaptive strategy for species that face similar pressures as *N. gonorrhoeae* and urogenitally-
426 adapted *N. meningitidis*. In support of this, clinical isolates of *Pseudomonas aeruginosa* with
427 truncations in genes homologous to *mtrC*[65,66] and exhibiting antibiotic hypersensitivity have
428 been obtained from cystic fibrosis patients, in whom the thick mucus in airway environments
429 can likewise exhibit increased acidity and decreased oxygen availability[67,68].

430

Our results also suggest potential therapeutic avenues for addressing the emergence of multidrug-resistant gonococcal strains. Selective knockdown of MtrCDE homologs in other bacteria via antisense RNA[69] and bacteriophages[70] has successfully re-sensitized resistant strains and enhanced antibiotic efficacy, and ectopic expression in *N. gonorrhoeae* of the *mtrR* repressor in a cephalosporin-resistant strain enhances gonococcal killing by β-lactam antibiotics in the mouse model[71]. Our population-wide estimated effect sizes for *mtrC* LOF mutations provide a prediction for the re-sensitization effect of MtrCDE knockdown across multiple genetic backgrounds and suggest particularly strong effects for the macrolide azithromycin (Supplementary Figure 4). Because the correlation between MIC differences and clinical efficacy is still not well understood[72,73], follow up studies to assess treatment efficacy differences in patients with and without *mtrC* LOF strains can help to quantify the expected effect of MtrCDE knockdown in the clinical context.

In summary, by analysis of population genomics and patient clinical data, we have shown that pathogenic *Neisseria* can use multiple avenues of efflux pump perturbation as an adaptive strategy to respond to host environmental pressures and illustrate how these host pressures may result in increased antibiotic susceptibility in *N. gonorrhoeae*.

Methods

Genomics pipeline: Reads for isolates with either associated azithromycin, ciprofloxacin, or ceftriaxone MIC metadata were downloaded from datasets listed in Supplementary Table 1. Reads were inspected using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and removed if GC content diverged notably from expected values (approximately 52-54%) or if base quality was systematically poor. We mapped read data to the NCCP11945 reference genome (RefSeq accession: NC_011035.1) using BWA-MEM (version 0.7.17-r1188)[74,75] and deduplicated reads using Picard (version 2.8.0) (https://github.com/broadinstitute/picard). BamQC in Qualimap (version 2.2.1)[76] was run to identify samples with less than 70% of reads aligned or samples with less than 40X coverage, which were discarded. We used Pilon (version 1.16)[77] to call variants with mindepth set to 10 and minmq set to 20 and generated pseudogenomes from Pilon VCFs by including all PASS sites and alternate alleles with AF > 0.9; all other sites were assigned as 'N'. Samples with greater than 15% of sites across the genome missing were also excluded. We created *de novo a*ssemblies using SPAdes (version 3.12.0 run using 8 threads, paired end reads where available, and the --careful flag set)[78] and quality filtered contigs to ensure coverage greater than 10X, length greater than 500 base pairs, and total genome size approximately equal to the FA1090 genome size (2.0 to 2.3 Mbp). We annotated assemblies with Prokka (version 1.13)[79], and clustered core genes using Roary (version 3.12)[80] (flags -z -e -n -v -s -i 92) and core intergenic regions using piggy (version 1.2)[81]. A recombination-corrected phylogeny of all isolates was constructed by running Gubbins (version 2.3.4) on the aligned pseudogenomes and visualized in iTOL (version 4.4.2)[82-84]. All isolates with associated metadata and accession numbers are listed in Supplementary Tables 6 and 7.

Resistance allele calling: Known resistance determinants in single-copy genes were called by identifying expected SNPs in the pseudogenomes. For categorizing mosaic alleles of *mtr*, we ran BLASTn (version 2.6.0)[85] on the *de novo* assemblies using a query sequence from FA1090 (Genbank accession: NC_002946.2) comprising the *mtr* intergenic promoter region and *mtrCDE*. BLAST results were aligned using MAFFT (version 7.450)[86] and clustered into distinct allelic families using FastBAPS (version 1.0.0)[87]. We confirmed that horizontally-transferred *mtr* alleles associated with resistance from prior studies[5] corresponded to distinct clusters in FastBAPS. A similar approach was used to cluster *penA* alleles after running

BLASTn with a *penA* reference sequence from FA1090. Variant calling in the multi-copy 23S rRNA locus was done by mapping to a modified NCCP11945 reference genome containing only one copy of the 23S rRNA and analyzing variant allele frequencies[88]. We identified truncated MtrR proteins using Prokka annotations, and mutations in the *mtr* promoter region associated with upregulation of *mtrCDE* (A deletion and TT insertion in inverted repeat, *mtr* 120) using an alignment of the *mtr* promoter from piggy output.

Phenotype processing and linear models: We doubled GISP azithromycin MICs before 2005 to account for the GISP MIC protocol testing change[89]. Samples with binary resistance phenotypes (i.e., “SUS” and “RES”) were discarded. For samples with MICs listed as above or below a threshold (indicated by greater than or less than symbols), the MIC was set to equal the provided threshold. MICs were $\log_2$-transformed for use as continuous outcome variables in linear modeling and GWAS. We modeled transformed MICs using a panel of known resistance markers[22,23] and included the recently characterized mosaic *mtrCDE* alleles[5] and *rplD* G70D substitution[90] conferring azithromycin resistance, as well as isolate country of origin. Formulas called by the lm function in R (version 3.5.1) for each drug were (with codon or nucleotide site indicated after each gene or rRNA, respectively):

Azithromycin: Log_AZI ~ Country + MtrR 39 + MtrR 45 + MtrR LOF + *mtrR* promoter + *mtrRCDE* BAPS + RplD G70D + 23S rRNA 2059 + 23S rRNA 2611

Ceftriaxone: Log_CRO ~ Country + MtrR 39 + MtrR 45 + MtrR LOF + *mtrR* promoter + *penA* BAPS + PonA 421 + PenA 501 + PenA 542 + PenA 551 + PorB 120 + PorB 121

Ciprofloxacin: Log_CIP ~ Country + MtrR 39 + MtrR 45 + MtrR LOF + *mtrR* promoter + GyrA 91 + GyrA 95 + ParC 86 + ParC 87 + ParC 91 + PorB 120 + PorB 121

To visualize the continuous MICs using thresholds as on Figure 1, we binarized MICs using the CLSI resistance breakpoint for ciprofloxacin, the CLSI non-susceptibility breakpoint for azithromycin, and the CDC GISP surveillance breakpoint for ceftriaxone.

GWAS and unitig annotation: We used a regression-based GWAS approach to identify novel susceptibility mutations. In particular, we employed a linear mixed model with a random effect to control for the confounding influence of population structure and a fixed effect to control for isolate country of origin. Though the outcome variable ($\log_2$-transformed MICs) is the same, in

512 contrast to the linear modeling approach described above, which models the linear, additive
513 effect of multiple, known resistance mutations, regression in a GWAS is usually run
514 independently and univariately on each variant for all identified variants in the genome,
515 providing a systematic way to identify novel contributors to the outcome variable. Linear mixed
516 model GWAS was run using Pyseer (version 1.2.0 with default allele frequency filters) on the
517 480,902 unitigs generated from GATB (version 1.3.0); the recombination-corrected phylogeny
518 from Gubbins was used to parameterize the Pyseer population structure random effects term
519 and isolate country of origin was included as a fixed effect covariate. To create the Manhattan
520 plot, we mapped all unitigs from the GWAS using BWA-MEM (modified parameters: -B 2 and -
521 O 3) to the pan-susceptible WHO F strain reference genome (Genbank accession:
522 GCA_900087635.2) edited to contain only one locus of the 23S rRNA. Significant unitigs were
523 annotated using Pyseer's annotation pipeline. Unitigs mapping to multiple sites in the genome
524 and in or near the highly variable *pilE* (encoding pilus subunit) or *piiC* (encoding opacity protein
525 family) genes were excluded, as were unitigs less than twenty base pairs in length. Due to
526 redundancy and linkage, variants will be spanned by multiple overlapping unitigs with similar
527 frequencies and *p*-values. For ease of interpretation, we grouped unitigs within 50 base pairs
528 of each other and represented each cluster by the most significant unitig. Unitigs with allele
529 frequency greater than 50% were also excluded as they represented the majority allele. Unitig
530 clusters were then annotated by gene or adjacent genes for unitigs mapping to intergenic
531 regions and further analyzed for predicted functional effect relative to the WHO F reference
532 genome in Geneious Prime (version 2019.2.1, https://www.geneious.com).

533

534 **Identifying LOF and upregulation alleles:** To identify predicted LOF alleles of efflux pump
535 proteins, we ran BLASTn on the *de novo* assemblies using a query sequence from FA1090
536 (reference genome FA19 was used for *mtrA*). Sequences that were full-length or
537 approximately full-length (+/- 5 nucleotides) beginning with expected start codons were
538 translated using Python (version 3.6.5) and Biopython (version 1.69)[91]. Peptides shorter than
539 90% of the expected full-length size of the protein were further analyzed using Geneious Prime
540 (version 2019.2.1, https://www.geneious.com) to identify the nucleotide mutations resulting in
541 predicted LOF by alignment of the nucleotide sequences. We called *mtrCDE* overexpression
542 status by identifying the presence of any of the known *mtrR* promoter mutations, MtrR coding
543 sequence mutations, and mosaic *mtrCDE* alleles.

544

Experimental validation: *N. gonorrhoeae* culture was conducted on GCB agar (Difco) plates supplemented with 1% Kellogg's supplements[92] at 37°C in a 5% CO2 atmosphere. Antimicrobial susceptibility testing was conducted on GCB agar supplemented with 1% IsoVitaleX (Becton Dickinson) using Etests (bioMérieux) at 37°C in a 5% CO2 atmosphere. We selected a clinical isolate (NY0195[93]) from the multidrug-resistant lineage ST-1901[35] that contained an *mtrC* LOF mutation mediated by a two base pair hexarepeat deletion and confirmed via Etests that its MIC matched, within one dilution, its reported MIC. Isolate NY0195 contained mosaic *penA* allele XXXIV conferring cephalosporin reduced susceptibility and the *gyrA* S91F substitution conferring ciprofloxacin resistance[9]. We complemented the *mtrC* LOF mutation in this strain by transforming it via electroporation[92] with a 2kb PCR product containing a *Neisserial* DNA uptake sequence and an in-frame *mtrC* allele, obtained by colony PCR from a neighboring isolate (GCGS0759). After obtaining transformants by selecting on an azithromycin 0.05 μg/mL GCB plate supplemented with Kellogg's supplement, we confirmed successful transformation by Sanger sequencing of the *mtrC* gene. No spontaneous mutants on azithromycin 0.05 μg/mL plates were observed after conducting control transformations in the absence of GCGS0759 *mtrC* PCR product. We conducted antimicrobial susceptibility testing in triplicate using Etests, assessing statistical significance between parental and transformant MICs by a two-sample t-test.

Metadata analysis: Patient metadata were collected from the following publications from Supplementary Table 1 that had information on site of infection: Demczuk et al., 2015, Demczuk et al., 2016, Ezewudo et al., 2015, Grad et al., 2014 and 2016, Kwong et al., 2017, Lee et al., 2018, and Mortimer et al., 2020. Sites of infection were standardized across datasets using a common ontology (i.e., specified as urethra, rectum, pharynx, cervix, or other). Fisher's two-sided exact test in R (version 3.5.1) was used to infer whether there was nonrandom association between *mtrC* LOF presence and either anatomical site of infection or sexual behavior. For sexual behavior analysis, isolates cultured from multiple sites on the same patient were counted as only one data point.

Meningococcal *mtrC* analysis: *mtrC* alleles from *N. meningitidis* assembled genomes were downloaded from PubMLST (n=14798; accessed 2019-09-03) by setting (species = "Neisseria meningitidis"), filtering by (Sequence bin size >= 2 Mbp), and exporting sequences for Locus "NEIS1634"[94]. *mtrC* LOF alleles were identified as described above. We generated a core-

genome maximum likelihood phylogeny of all North American ST-11 *N. meningitidis* strains from PubMLST (n=456; accessed 2019-09-03) rooted with meningococcal reference genome MC58 (Genbank accession: AE002098.2) using Roary (version 3.12) (flags -z -e -n -v -s -i 92) and annotated it using metadata from Retchless et al., 2018[56] (see Supplementary Table 8 for PubMLST IDs). Overrepresentation of *mtrC* LOF alleles in the US urogenital lineage compared to selected control datasets was assessed using Fisher's two-sided exact test in R (version 3.5.1).

Data and code availability

Accession numbers for all sequences used are listed in the supplementary tables. Code to reproduce the analyses and figures is available at https://github.com/gradlab/mtrC-GWAS or from the authors upon request. Source data underlying Figure 2 is available in the supplementary tables or at the above GitHub link.

Acknowledgements

This work was supported by the NIH/NIAID grant 1R01AI132606-01 and the Smith Family Foundation. TDM is additionally supported by the NIH/NIAID 1 F32 AI145157-01, and KCM is additionally supported by the NSF GRFP. DHFR was supported by award Number T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. DAW is supported by an Early Career Fellowship from the National Health and Medical Research Council of Australia (GNT1123854). Portions of this research were conducted on the O2 high-performance computing cluster, supported by the Research Computing Group at Harvard Medical School. This publication made use of the Meningitis Research Foundation Meningococcus Genome Library (http://www.meningitis.org/research/genome) developed by Public Health England, the Wellcome Trust Sanger Institute, and the University of Oxford as a collaboration and funded by the Meningitis Research Foundation. The authors additionally thank Crista Wadsworth, Samantha Palace, and other members of the Grad Lab for helpful comments during development of the project.

Publication	Study summary	Timespan	Included number
Mortimer et al., 2020[93]	Transmission and AMR surveillance in New York City, USA	2011-2015	888
Sánchez-Busó et al., 2019[3]	Worldwide phylogeography and evolution of gonococcus	1979-2012	408
Yahara et al., 2018[10]	AMR surveillance in Kyoto and Osaka, Japan	1996-2015	260
Ryan et al., 2018[11]	AMR surveillance in Ireland	2012-2016	39
Harris et al., 2018[12]	Genomic survey across European Euro-GASP participant countries (n=20)	2013	1048
Fifer et al., 2018[13]	High-level azithromycin resistance outbreak in UK	2004-2017	50
Lee et al., 2018[14]	Genomic epidemiology in New Zealand	2014-2015	397
Kwong et al., 2017[15]	Transmission among MSM in Melbourne, Australia	2005-2014	94
Eyre et al., 2017[22] and De Silva et al., 2016[20]	Transmission in Brighton, UK	2004-2011	231
Grad et al., 2016[9] and 2014[17]	AMR surveillance across CDC GISP clinics, USA	2000-2013	1100
Demczuk et al., 2016[19]	Azithromycin resistance surveillance in Canada	1991-2014	199
Demczuk et al., 2015[18]	Cephalosporin decreased susceptibility surveillance in Canada	1989-2013	114
Ezewudo et al., 2015[16]	Population structure and AMR surveillance	1982-2011	54

Supplementary Table 1 – Datasets included in global meta-analysis collection. All isolates that passed genomics quality control filters (see methods) with associated azithromycin, or ceftriaxone, or ciprofloxacin metadata were included (n=4852 in total). Euro-GASP = European Gonococcal Antimicrobial Surveillance Program; CDC GISP = Centers for Disease Control and Prevention Gonococcal Isolate Surveillance Project.

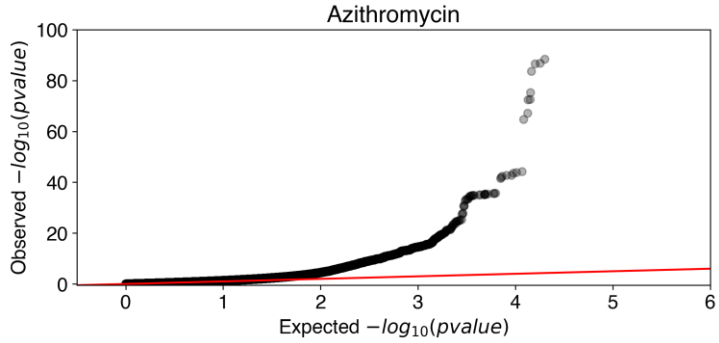
Strain	Azithromycin	Ceftriaxone	Ciprofloxacin
NY0195	0.064, 0.094, 0.094	0.016, 0.023, 0.023	6, 6, 4
NY0195 *mtrC*(in-frame)	0.5, 0.75, 0.75	0.064, 0.064, 0.064	≥ 32, ≥ 32, ≥ 32
	p = 0.0187	p = 0.00289	p = 0.000624

Supplementary Table 4 – Restoration of the *mtrC* coding frame in a clinical isolate by transformation increases MICs to azithromycin, ceftriaxone, and ciprofloxacin. MIC Etests were conducted in triplicate with all results reported. Statistical significance for MIC differences between parental strain and transformant strain was assessed by a two-sample t-test after setting the value of "≥ 32" to 32.

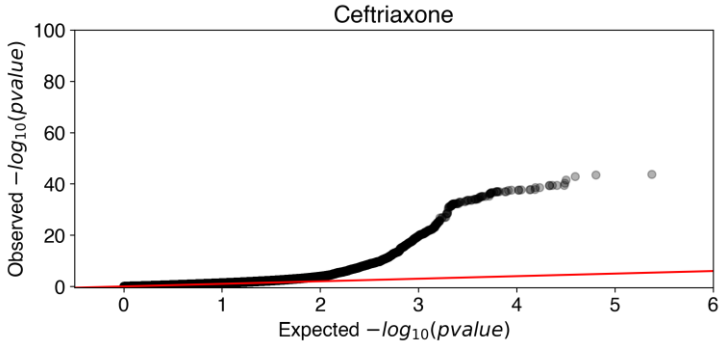
Dataset	Gene	N LOF	N Total	Percentage
Global	*farA*	332	4838	6.86%
Global	*farB*	2	4850	0.04%
Global	*norM*	2	4852	0.04%
Global	*macA*	1	4847	0.02%
Global	*macB*	13	4845	0.27%
Global	*mtrR*	386	4845	7.97%
Global	*mtrA*	268 (362)	4842	5.45% (7.48%)
Global	*mtrC*	185	4847	3.82%
Global	*mtrD*	10	4815	0.21%
Global	*mtrE*	0	4849	0.00%
Australia	*farA*	225	2180	10.32%
Australia	*farB*	1	2186	0.05%
Australia	*norM*	0	2186	0.00%
Australia	*macA*	0	2186	0.00%
Australia	*macB*	0	2186	0.00%
Australia	*mtrA*	4 (85)	2186	0.18% (3.89%)
Australia	*mtrR*	253	2183	11.59%
Australia	*mtrC*	35	2186	1.60%
Australia	*mtrD*	2	2185	0.09%
Australia	*mtrE*	0	2186	0.00%

Supplementary Table 5 – Prevalence of predicted LOF mutations in efflux pump genes for both global and Australian datasets. Counts for *mtrA* LOF mutations in the absence of other epistatic *mtrCDE* overexpression mutations are listed first, with counts for total number of *mtrA* LOF mutations regardless of *mtrCDE* overexpression status listed in parentheses.

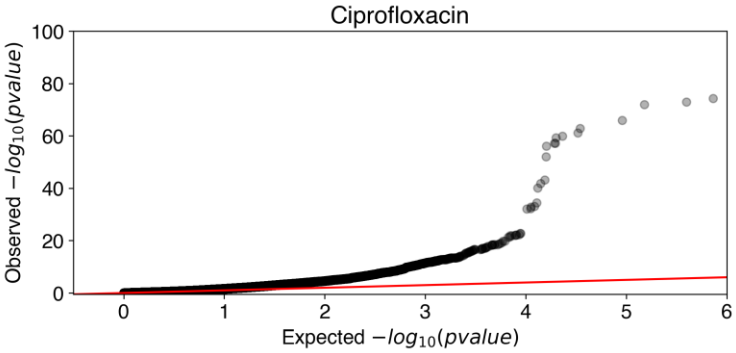
a)



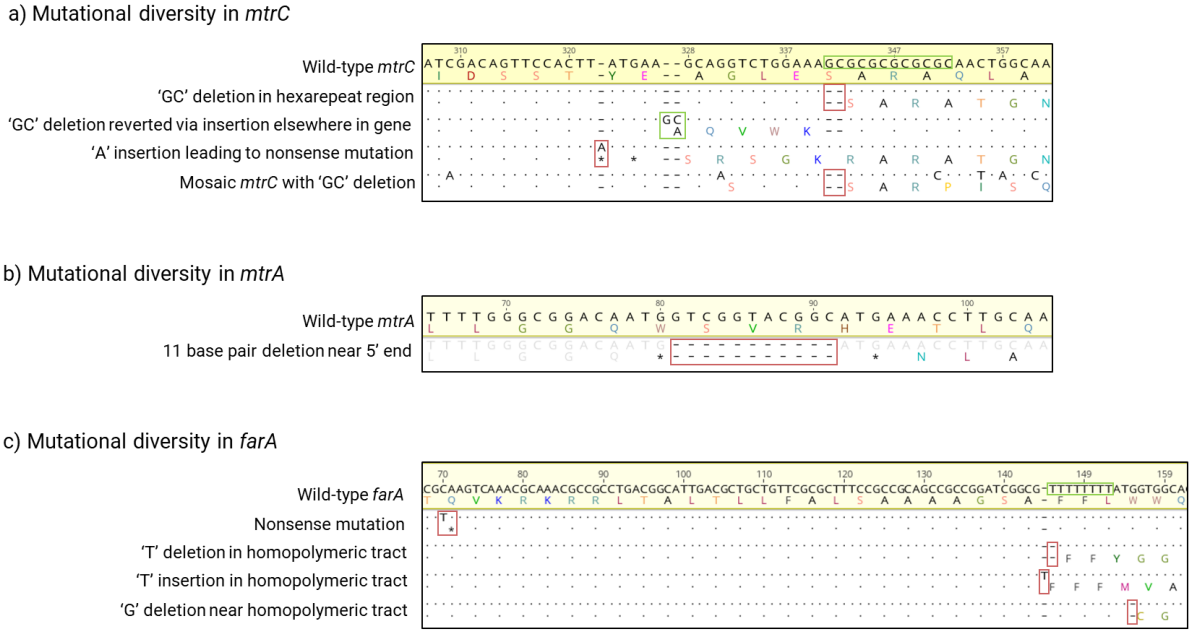
b)



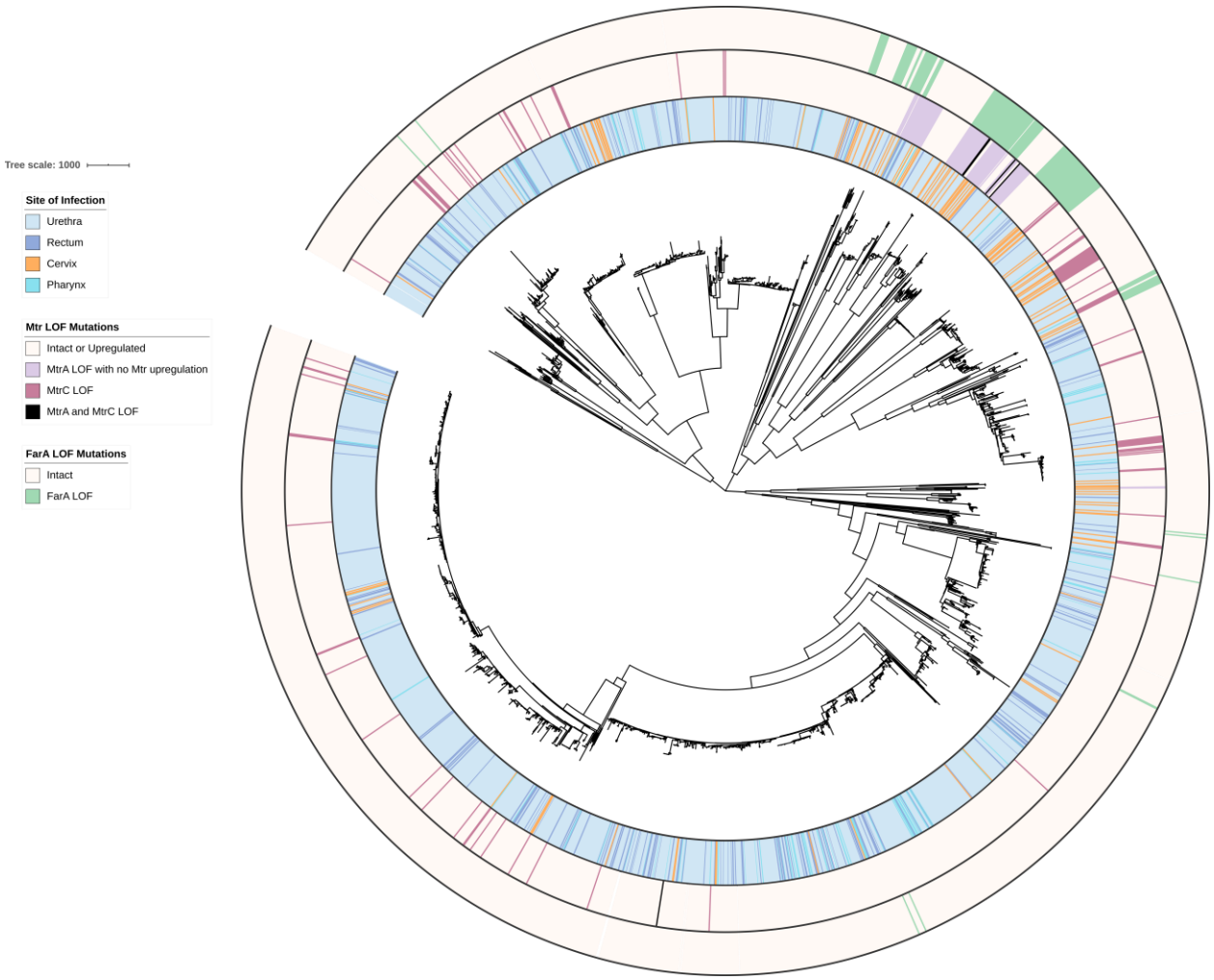
c)



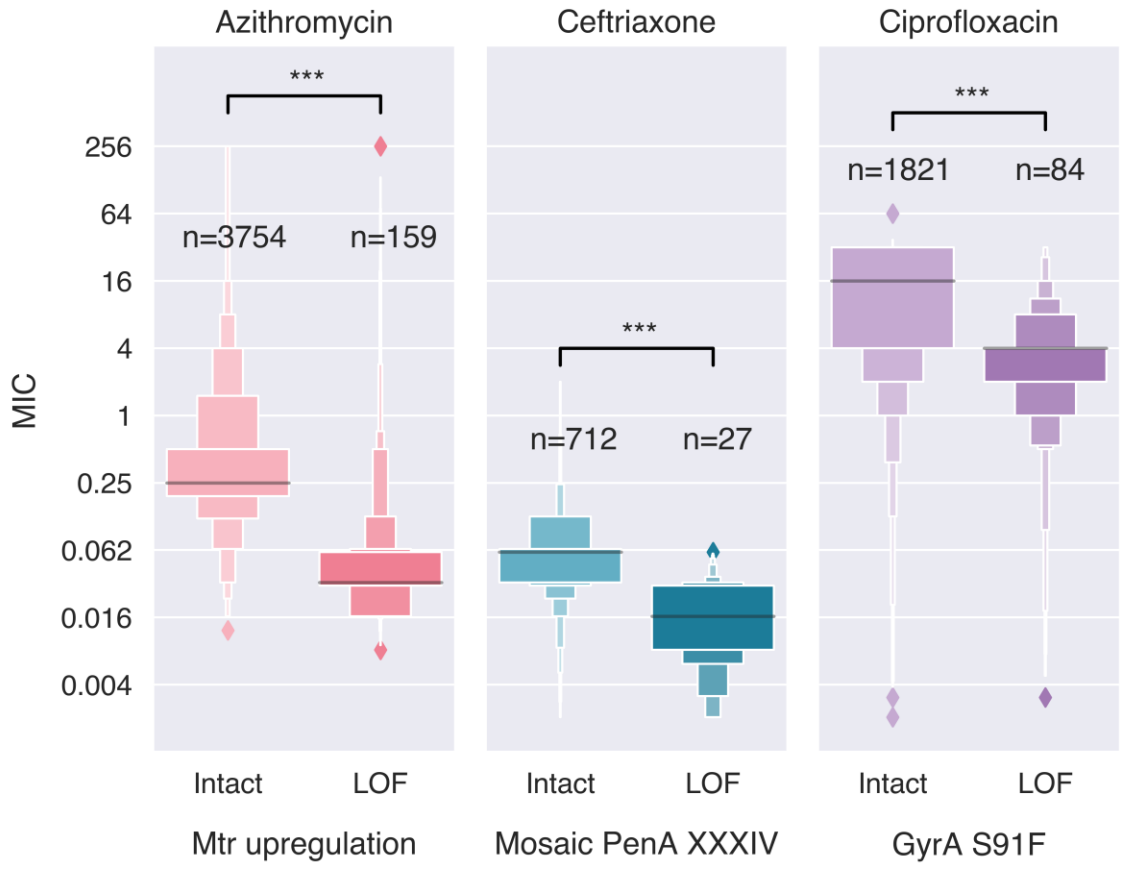
Supplementary Figure 1 – **Diagnostic Q-Q plots of expected versus observed *p*-values for GWAS on a) azithromycin, b) ceftriaxone, and c) ciprofloxacin.** In the absence of confounders such as population structure, *p*-values are distributed uniformly and would be expected to lie along the y=x line (in red) before diverging at higher -$\log_{10}$(*p*-values) due to true causal variants[95]. Q-Q plots for all three antibiotics appear to be well-behaved, indicating that the steps we have taken to control for population structure (i.e., using a linear mixed model parameterized by the recombination-corrected phylogeny) were adequate. Highly significant markers corresponding to diverging variants at higher -$\log_{10}$(*p*-values) were confirmed to map to known causal variants for all three antibiotics (see Supplementary Table 3).



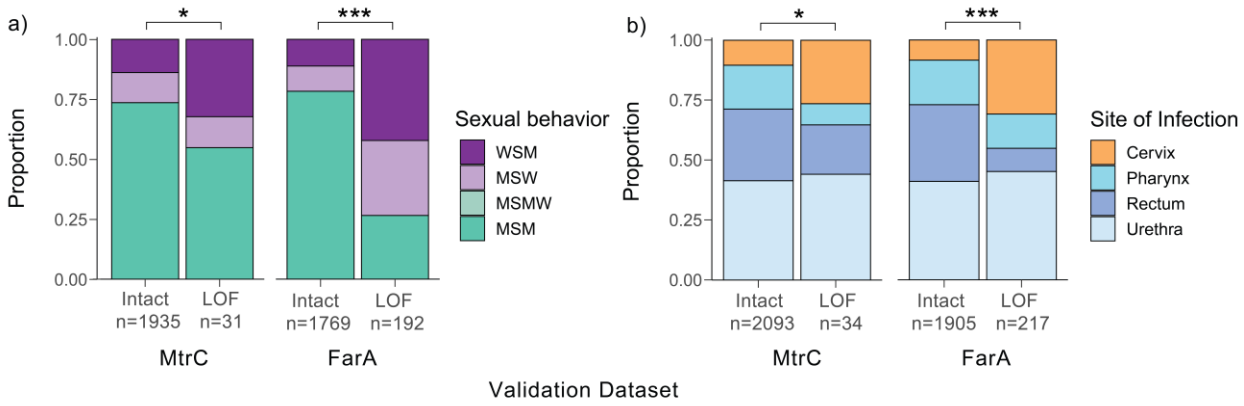
Supplementary Figure 2 – Alignment of nucleotide sequences for strains with representative LOF mutations observed in a) *mtrC*, b) *mtrA*, and c) *farA* in the global dataset. The wild-type reference sequences (FA1090 for *mtrC* and *farA*, FA19 for *mtrA*) are shown at the top of the alignment highlighted in yellow. Nucleotide sequences were depicted in black with the corresponding amino acid translations directly under. Dots in LOF sequences represent exact match to the wild-type reference sequence. For *mtrC*, the hexarepeat tract was boxed in the reference genome in green, and mutations leading to LOFs were boxed in red. For *mtrA*, the 11-bp deletion leading to *mtrA* LOF was boxed in red. For *farA*, the repeat tract of Ts was boxed in green, and mutations leading to LOFs were boxed in red. All alignments were visualized in Geneious Prime (see methods).



Supplementary Figure 3 – **Phylogenetic distribution of gonococcal *mtrC*, MtrA, and *farA* LOF alleles with patient site of infection (n=2742) in global dataset.** The recombination-corrected maximum likelihood phylogeny based on 36347 SNPs is shown annotated with rings (from innermost to outermost) for site of infection, *mtrCDE* regulon LOF mutations, and *farA* LOF mutations. Branch length represents total number of substitutions after removal of predicted recombination.



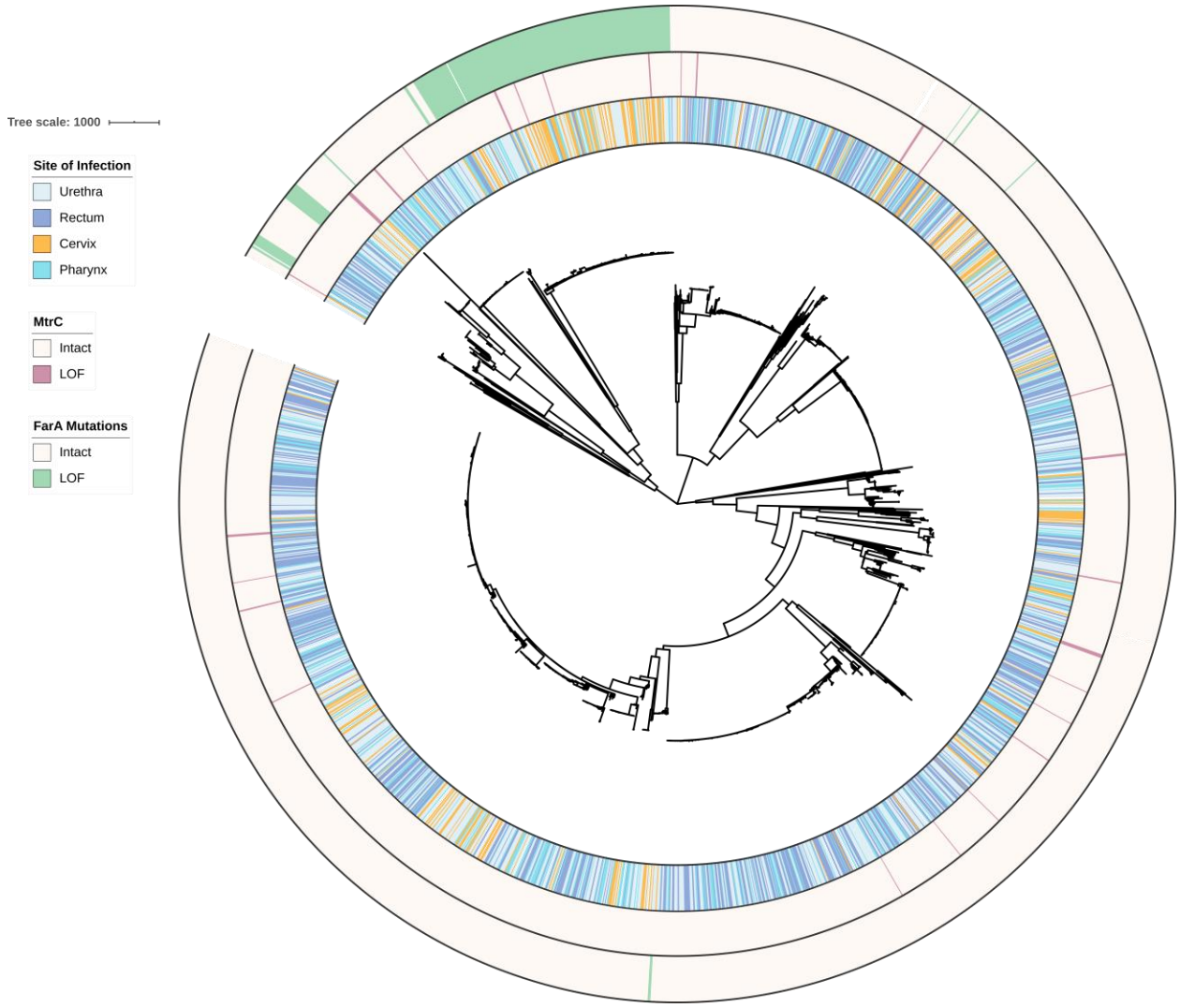
Supplementary Figure 4 – **MIC distributions for isolates with indicated resistance determinants for azithromycin, ceftriaxone, and ciprofloxacin respectively, stratified by *mtrC* genotypic status.** Statistical significance between *mtrC* intact and LOF MIC distributions was assessed by Mann-Whitney U Test: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. Enhanced box plots were drawn using default settings in the Python Seaborn visualization package (version 0.9.0).



682

683 **Supplementary Figure 5 – Gonococcal *mtrC* and *farA* LOF mutations are associated with**
684 **sexual behavior and site of infection in the validation dataset.** a) Sex partner information in
685 patients infected with isolates with either intact or LOF alleles of *mtrC* (left) or *farA* (right).
686 MSMW patients were labelled as MSM in the validation dataset[43]. b) Site of infection in patients
687 infected with isolates with either intact or LOF alleles of *mtrC* (left) or *farA* (right) datasets.
688 Statistical significance between intact versus LOF patient metadata distributions was assessed
689 by Fisher's exact test: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. WSM = women who have sex with
690 men, MSW = men who have sex with women, MSMW = men who have sex with men and
691 women, MSM = men who have sex with men.
692
693

694



Supplementary Figure 6 – **Phylogenetic distribution of gonococcal *mtrC* LOF alleles with patient site of infection (n=2186) in the validation dataset.** The recombination-corrected maximum likelihood phylogeny based on 26669 SNPs is shown annotated with rings (from innermost to outermost) for site of infection, *mtrC* LOF mutations, and *farA* LOF mutations. Branch length represents total number of substitutions after removal of predicted recombination.

References

1 Olesen, S. W. *et al.* Azithromycin susceptibility in Neisseria gonorrhoeae and seasonal macrolide use. *J Infect Dis*, doi:10.1093/infdis/jiy551 (2018).

2 Olesen, S. W. & Grad, Y. H. Deciphering the impact of bystander selection for antibiotic resistance in Neisseria gonorrhoeae. *J Infect Dis*, doi:10.1093/infdis/jiz156 (2019).

3 Sanchez-Buso, L. *et al.* The impact of antimicrobials on gonococcal evolution. *Nat Microbiol*, doi:10.1038/s41564-019-0501-y (2019).

4 Unemo, M. & Shafer, W. M. Antimicrobial resistance in Neisseria gonorrhoeae in the 21st century: past, evolution, and future. *Clin Microbiol Rev* **27**, 587-613, doi:10.1128/CMR.00010-14 (2014).

5 Wadsworth, C. B., Arnold, B. J., Sater, M. R. A. & Grad, Y. H. Azithromycin Resistance through Interspecific Acquisition of an Epistasis-Dependent Efflux Pump Component and Transcriptional Regulator in Neisseria gonorrhoeae. *MBio* **9**, doi:10.1128/mBio.01419-18 (2018).

6 Rouquette-Loughlin, C. E. *et al.* Mechanistic Basis for Decreased Antimicrobial Susceptibility in a Clinical Isolate of Neisseria gonorrhoeae Possessing a Mosaic-Like mtr Efflux Pump Locus. *MBio* **9**, doi:10.1128/mBio.02281-18 (2018).

7 Quillin, S. J. & Seifert, H. S. Neisseria gonorrhoeae host adaptation and pathogenesis. *Nat Rev Microbiol* **16**, 226-240, doi:10.1038/nrmicro.2017.169 (2018).

8 Wi, T. *et al.* Antimicrobial resistance in Neisseria gonorrhoeae: Global surveillance and a call for international collaborative action. *PLoS Med* **14**, e1002344, doi:10.1371/journal.pmed.1002344 (2017).

9 Grad, Y. H. *et al.* Genomic Epidemiology of Gonococcal Resistance to Extended-Spectrum Cephalosporins, Macrolides, and Fluoroquinolones in the United States, 2000-2013. *J Infect Dis* **214**, 1579-1587, doi:10.1093/infdis/jiw420 (2016).

10 Yahara, K. *et al.* Genomic surveillance of Neisseria gonorrhoeae to investigate the distribution and evolution of antimicrobial-resistance determinants and lineages. *Microb Genom* **4**, doi:10.1099/mgen.0.000205 (2018).

11 Ryan, L. *et al.* Antimicrobial resistance and molecular epidemiology using whole-genome sequencing of Neisseria gonorrhoeae in Ireland, 2014-2016: focus on extended-spectrum cephalosporins and azithromycin. *Eur J Clin Microbiol Infect Dis*, doi:10.1007/s10096-018-3296-5 (2018).

12 Harris, S. R. *et al.* Public health surveillance of multidrug-resistant clones of Neisseria gonorrhoeae in Europe: a genomic survey. *Lancet Infect Dis* **18**, 758-768, doi:10.1016/S1473-3099(18)30225-1 (2018).

13 Fifer, H. *et al.* Sustained transmission of high-level azithromycin-resistant Neisseria gonorrhoeae in England: an observational study. *Lancet Infect Dis* **18**, 573-581, doi:10.1016/S1473-3099(18)30122-1 (2018).

14 Lee, R. S. *et al.* Genomic epidemiology and antimicrobial resistance of Neisseria gonorrhoeae in New Zealand. *J Antimicrob Chemother* **73**, 353-364, doi:10.1093/jac/dkx405 (2018).

15 Kwong, J. C. *et al.* Whole-genome sequencing reveals transmission of gonococcal antibiotic resistance among men who have sex with men: an observational study. *Sex Transm Infect* **94**, 151-157, doi:10.1136/sextrans-2017-053287 (2018).

16 Ezewudo, M. N. *et al.* Population structure of Neisseria gonorrhoeae based on whole genome data and its relationship with antibiotic resistance. *PeerJ* **3**, e806, doi:10.7717/peerj.806 (2015).

17 Grad, Y. H. *et al.* Genomic epidemiology of Neisseria gonorrhoeae with reduced susceptibility to cefixime in the USA: a retrospective observational study. *Lancet Infect Dis* **14**, 220-226, doi:10.1016/S1473-3099(13)70693-5 (2014).

18 Demczuk, W. *et al.* Whole-genome phylogenomic heterogeneity of Neisseria gonorrhoeae isolates with decreased cephalosporin susceptibility collected in Canada between 1989 and 2013. *J Clin Microbiol* **53**, 191-200, doi:10.1128/JCM.02589-14 (2015).

19 Demczuk, W. *et al.* Genomic Epidemiology and Molecular Resistance Mechanisms of Azithromycin-Resistant Neisseria gonorrhoeae in Canada from 1997 to 2014. *J Clin Microbiol* **54**, 1304-1313, doi:10.1128/JCM.03195-15 (2016).

20 De Silva, D. *et al.* Whole-genome sequencing to determine transmission of Neisseria gonorrhoeae: an observational study. *Lancet Infect Dis* **16**, 1295-1303, doi:10.1016/S1473-3099(16)30157-8 (2016).

21 Wadsworth, C. B., Arnold, B. J., Abdul Sater, M. R. & Grad, Y. H. Azithromycin resistance through interspecific acquisition of an epistasis dependent efflux pump component and transcriptional regulator in Neisseria gonorrhoeae. . *BioRxiv* (2018).

22 Eyre, D. W. *et al.* WGS to predict antibiotic MICs for Neisseria gonorrhoeae. *J Antimicrob Chemother* **72**, 1937-1947, doi:10.1093/jac/dkx067 (2017).

23 Demczuk, W. *et al.* Equations To Predict Antimicrobial MICs in Neisseria gonorrhoeae Using Molecular Antimicrobial Resistance Determinants. *Antimicrob Agents Chemother* **64**, doi:10.1128/AAC.02005-19 (2020).

24 Hicks, A. L. *et al.* Evaluation of parameters affecting performance and reliability of machine learning-based antibiotic susceptibility testing from whole genome sequencing data. *PLoS Comput Biol* **15**, e1007349, doi:10.1371/journal.pcbi.1007349 (2019).

25 Earle, S. G. *et al.* Identifying lineage effects when controlling for population structure improves power in bacterial association studies. *Nat Microbiol* **1**, 16041, doi:10.1038/nmicrobiol.2016.41 (2016).

26 Lees, J. A., Galardini, M., Bentley, S. D., Weiser, J. N. & Corander, J. pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics*, bty539-bty539, doi:10.1093/bioinformatics/bty539 (2018).

27 Jaillard, M. *et al.* A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. *PLoS Genet* **14**, e1007758, doi:10.1371/journal.pgen.1007758 (2018).

28 Jaillard, M. *et al.* A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. *PLoS Genet.* **14**, e1007758, doi:10.1371/journal.pgen.1007758 (2018).

29 Mortimer, T. D. & Grad, Y. H. Applications of genomics to slow the spread of multidrug-resistant Neisseria gonorrhoeae. *Ann N Y Acad Sci*, doi:10.1111/nyas.13871 (2018).

30 Veal, W. L., Nicholas, R. A. & Shafer, W. M. Overexpression of the MtrC-MtrD-MtrE efflux pump due to an mtrR mutation is required for chromosomally mediated penicillin resistance in Neisseria gonorrhoeae. *J Bacteriol* **184**, 5619-5624, doi:10.1128/jb.184.20.5619-5624.2002 (2002).

31 Golparian, D., Shafer, W. M., Ohnishi, M. & Unemo, M. Importance of multidrug efflux pumps in the antimicrobial resistance property of clinical multidrug-resistant isolates of Neisseria gonorrhoeae. *Antimicrob Agents Chemother* **58**, 3556-3559, doi:10.1128/AAC.00038-14 (2014).

32 Veal, W. L. *et al.* Loss-of-function mutations in the mtr efflux system of Neisseria gonorrhoeae. *Microbiology* **144 (Pt 3)**, 621-627, doi:10.1099/00221287-144-3-621 (1998).

33 Sanchez-Buso, L., Golparian, D., Parkhill, J., Unemo, M. & Harris, S. R. Genetic variation regulates the activation and specificity of Restriction-Modification systems in Neisseria gonorrhoeae. *Sci Rep* **9**, 14685, doi:10.1038/s41598-019-51102-2 (2019).

34 Eisenstein, B. I. & Sparling, P. F. Mutations to increased antibiotic sensitivity in naturally-occurring gonococci. *Nature* **271**, 242-244, doi:10.1038/271242a0 (1978).

35 Shimuta, K. *et al.* Emergence and evolution of internationally disseminated cephalosporin-resistant Neisseria gonorrhoeae clones from 1995 to 2005 in Japan. *BMC Infect Dis* **15**, 378, doi:10.1186/s12879-015-1110-x (2015).

36 Bichara, M., Pinet, I., Schumacher, S. & Fuchs, R. P. Mechanisms of dinucleotide repeat instability in Escherichia coli. *Genetics* **154**, 533-542 (2000).

37 Golparian, D. *et al.* Genomic evolution of Neisseria gonorrhoeae since the preantibiotic era (1928-2013): antimicrobial use/misuse selects for resistance and drives evolution. *BMC Genomics* **21**, 116, doi:10.1186/s12864-020-6511-6 (2020).

38 Handing, J. W., Ragland, S. A., Bharathan, U. V. & Criss, A. K. The MtrCDE Efflux Pump Contributes to Survival of Neisseria gonorrhoeae From Human Neutrophils and Their Antimicrobial Components. *Front Microbiol* **9**, 2688, doi:10.3389/fmicb.2018.02688 (2018).

39 Warner, D. M., Folster, J. P., Shafer, W. M. & Jerse, A. E. Regulation of the MtrC-MtrD-MtrE efflux-pump system modulates the in vivo fitness of Neisseria gonorrhoeae. *J Infect Dis* **196**, 1804-1812, doi:10.1086/522964 (2007).

40 Warner, D. M., Shafer, W. M. & Jerse, A. E. Clinically relevant mutations that cause derepression of the Neisseria gonorrhoeae MtrC-MtrD-MtrE Efflux pump system confer different levels of antimicrobial resistance and in vivo fitness. *Mol Microbiol* **70**, 462-478, doi:10.1111/j.1365-2958.2008.06424.x (2008).

41 Folster, J. P. *et al.* MtrR modulates rpoH expression and levels of antimicrobial resistance in Neisseria gonorrhoeae. *J Bacteriol* **191**, 287-297, doi:10.1128/JB.01165-08 (2009).

42 Wang, S. *et al.* Gonococcal MtrE and its surface-expressed Loop 2 are immunogenic and elicit bactericidal antibodies. *J Infect* **77**, 191-204, doi:10.1016/j.jinf.2018.06.001 (2018).

43 Williamson, D. A. *et al.* Bridging of Neisseria gonorrhoeae lineages across sexual networks in the HIV pre-exposure prophylaxis era. *Nat Commun* **10**, 3988, doi:10.1038/s41467-019-12053-4 (2019).

44 Rouquette, C., Harmon, J. B. & Shafer, W. M. Induction of the mtrCDE-encoded efflux pump system of Neisseria gonorrhoeae requires MtrA, an AraC-like protein. *Mol Microbiol* **33**, 651-658, doi:10.1046/j.1365-2958.1999.01517.x (1999).

45 Vidyaprakash, E., Abrams, A. J., Shafer, W. M. & Trees, D. L. Whole-Genome Sequencing of a Large Panel of Contemporary Neisseria gonorrhoeae Clinical Isolates Indicates that a Wild-Type mtrA Gene Is Common: Implications for Inducible Antimicrobial Resistance. *Antimicrob Agents Chemother* **61**, doi:10.1128/AAC.00262-17 (2017).

46 Olivares Pacheco, J., Alvarez-Ortega, C., Alcalde Rico, M. & Martinez, J. L. Metabolic Compensation of Fitness Costs Is a General Outcome for Antibiotic-Resistant Pseudomonas aeruginosa Mutants Overexpressing Efflux Pumps. *MBio* **8**, doi:10.1128/mBio.00500-17 (2017).

47 Zgurskaya, H. I. Multicomponent drug efflux complexes: architecture and mechanism of assembly. *Future Microbiol* **4**, 919-932, doi:10.2217/fmb.09.62 (2009).

48 Lee, E. H. & Shafer, W. M. The farAB-encoded efflux pump mediates resistance of gonococci to long-chained antibacterial fatty acids. *Mol Microbiol* **33**, 839-845, doi:10.1046/j.1365-2958.1999.01530.x (1999).

49 Lee, E. H., Rouquette-Loughlin, C., Folster, J. P. & Shafer, W. M. FarR regulates the farAB-encoded efflux pump of Neisseria gonorrhoeae via an MtrR regulatory mechanism. *J Bacteriol* **185**, 7145-7152, doi:10.1128/jb.185.24.7145-7152.2003 (2003).

50 Rouquette-Loughlin, C., Dunham, S. A., Kuhn, M., Balthazar, J. T. & Shafer, W. M. The NorM efflux pump of Neisseria gonorrhoeae and Neisseria meningitidis recognizes

antimicrobial cationic compounds. *J Bacteriol* **185**, 1101-1106, doi:10.1128/jb.185.3.1101-1106.2003 (2003).

51 Rouquette-Loughlin, C. E., Balthazar, J. T. & Shafer, W. M. Characterization of the MacA-MacB efflux system in Neisseria gonorrhoeae. *J Antimicrob Chemother* **56**, 856-860, doi:10.1093/jac/dki333 (2005).

52 Stephens, D. S., Greenwood, B. & Brandtzaeg, P. Epidemic meningitis, meningococcaemia, and Neisseria meningitidis. *Lancet* **369**, 2196-2210, doi:10.1016/S0140-6736(07)61016-2 (2007).

53 Hill, D. M. *et al.* Genomic epidemiology of age-associated meningococcal lineages in national surveillance: an observational cohort study. *Lancet Infect Dis* **15**, 1420-1428, doi:10.1016/S1473-3099(15)00267-4 (2015).

54 Bichara, M., Schumacher, S. & Fuchs, R. P. Genetic instability within monotonous runs of CpG sequences in Escherichia coli. *Genetics* **140**, 897-907 (1995).

55 Bazan, J. A. *et al.* Notes from the Field: Increase in Neisseria meningitidis-Associated Urethritis Among Men at Two Sentinel Clinics - Columbus, Ohio, and Oakland County, Michigan, 2015. *MMWR Morb Mortal Wkly Rep* **65**, 550-552, doi:10.15585/mmwr.mm6521a5 (2016).

56 Retchless, A. C. *et al.* Expansion of a urethritis-associated Neisseria meningitidis clade in the United States with concurrent acquisition of N. gonorrhoeae alleles. *BMC Genomics* **19**, 176, doi:10.1186/s12864-018-4560-x (2018).

57 Tzeng, Y. L. *et al.* Emergence of a new Neisseria meningitidis clonal complex 11 lineage 11.2 clade as an effective urogenital pathogen. *Proc Natl Acad Sci U S A* **114**, 4237-4242, doi:10.1073/pnas.1620971114 (2017).

58 Whiley, D. M., Jennison, A., Pearson, J. & Lahra, M. M. Genetic characterisation of Neisseria gonorrhoeae resistant to both ceftriaxone and azithromycin. *Lancet Infect Dis* **18**, 717-718, doi:10.1016/S1473-3099(18)30340-2 (2018).

59 Kime, L. *et al.* Transient Silencing of Antibiotic Resistance by Mutation Represents a Significant Potential Source of Unanticipated Therapeutic Failure. *MBio* **10**, doi:10.1128/mBio.01755-19 (2019).

60 Morse, S. A. *et al.* Gonococcal strains from homosexual men have outer membranes with reduced permeability to hydrophobic molecules. *Infect Immun* **37**, 432-438 (1982).

61 Shafer, W. M., Balthazar, J. T., Hagman, K. E. & Morse, S. A. Missense mutations that alter the DNA-binding domain of the MtrR protein occur frequently in rectal isolates of Neisseria gonorrhoeae that are resistant to faecal lipids. *Microbiology* **141 (Pt 4)**, 907-911, doi:10.1099/13500872-141-4-907 (1995).

62 Mercante, A. D. *et al.* MpeR regulates the mtr efflux locus in Neisseria gonorrhoeae and modulates antimicrobial resistance by an iron-responsive mechanism. *Antimicrob Agents Chemother* **56**, 1491-1501, doi:10.1128/AAC.06112-11 (2012).

63 Nudel, K. *et al.* Transcriptome Analysis of Neisseria gonorrhoeae during Natural Infection Reveals Differential Expression of Antibiotic Resistance Determinants between Men and Women. *mSphere* **3**, doi:10.1128/mSphereDirect.00312-18 (2018).

64 Du, D. *et al.* Multidrug efflux pumps: structure, function and regulation. *Nat Rev Microbiol* **16**, 523-539, doi:10.1038/s41579-018-0048-6 (2018).

65 Vettoretti, L. *et al.* Efflux unbalance in Pseudomonas aeruginosa isolates from cystic fibrosis patients. *Antimicrob Agents Chemother* **53**, 1987-1997, doi:10.1128/AAC.01024-08 (2009).

66 Chalhoub, H. *et al.* Mechanisms of intrinsic resistance and acquired susceptibility of Pseudomonas aeruginosa isolated from cystic fibrosis patients to temocillin, a revived antibiotic. *Sci Rep* **7**, 40208, doi:10.1038/srep40208 (2017).

67 Garnett, J. P. *et al.* Hyperglycaemia and Pseudomonas aeruginosa acidify cystic fibrosis airway surface liquid by elevating epithelial monocarboxylate transporter 2 dependent lactate-H(+) secretion. *Sci Rep* **6**, 37955, doi:10.1038/srep37955 (2016).

68 Worlitzsch, D. *et al.* Effects of reduced mucus oxygen concentration in airway Pseudomonas infections of cystic fibrosis patients. *J Clin Invest* **109**, 317-325, doi:10.1172/JCI13870 (2002).

69 Ayhan, D. H. *et al.* Sequence-Specific Targeting of Bacterial Resistance Genes Increases Antibiotic Efficacy. *PLoS Biol* **14**, e1002552, doi:10.1371/journal.pbio.1002552 (2016).

70 Chan, B. K. *et al.* Phage selection restores antibiotic sensitivity in MDR Pseudomonas aeruginosa. *Sci Rep* **6**, 26717, doi:10.1038/srep26717 (2016).

71 Chen, S. *et al.* Could Dampening Expression of the Neisseria gonorrhoeae mtrCDE-Encoded Efflux Pump Be a Strategy To Preserve Currently or Resurrect Formerly Used Antibiotics To Treat Gonorrhea? *MBio* **10**, doi:10.1128/mBio.01576-19 (2019).

72 Cole, M. J. *et al.* Gentamicin, azithromycin and ceftriaxone in the treatment of gonorrhoea: the relationship between antibiotic MIC and clinical outcome. *J Antimicrob Chemother* **75**, 449-457, doi:10.1093/jac/dkz436 (2020).

73 Tapsall, J. W. *et al.* Failure of azithromycin therapy in gonorrhea and discorrelation with laboratory test parameters. *Sex Transm Dis* **25**, 505-508, doi:10.1097/00007435-199811000-00002 (1998).

74 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

75 Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint* **1303.3997** (2013).

76 Garcia-Alcalde, F. *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**, 2678-2679, doi:10.1093/bioinformatics/bts503 (2012).

77 Walker, B. J. *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**, e112963, doi:10.1371/journal.pone.0112963 (2014).

78 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455-477, doi:10.1089/cmb.2012.0021 (2012).

79 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069, doi:10.1093/bioinformatics/btu153 (2014).

80 Page, A. J. *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691-3693, doi:10.1093/bioinformatics/btv421 (2015).

81 Thorpe, H. A., Bayliss, S. C., Sheppard, S. K. & Feil, E. J. Piggy: a rapid, large-scale pan-genome analysis tool for intergenic regions in bacteria. *Gigascience* **7**, 1-11, doi:10.1093/gigascience/giy015 (2018).

82 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256-W259, doi:10.1093/nar/gkz239 (2019).

83 Croucher, N. J. *et al.* Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* **43**, e15, doi:10.1093/nar/gku1196 (2015).

84 Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313, doi:10.1093/bioinformatics/btu033 (2014).

85 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410, doi:10.1016/S0022-2836(05)80360-2 (1990).

86 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059-3066, doi:10.1093/nar/gkf436 (2002).

87 Tonkin-Hill, G., Lees, J. A., Bentley, S. D., Frost, S. D. W. & Corander, J. Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Res* **47**, 5539-5549, doi:10.1093/nar/gkz361 (2019).

88 Johnson, S. R., Grad, Y., Abrams, A. J., Pettus, K. & Trees, D. L. Use of whole-genome sequencing data to analyze 23S rRNA-mediated azithromycin resistance. *Int J Antimicrob Agents* **49**, 252-254, doi:10.1016/j.ijantimicag.2016.10.023 (2017).

89 Kersh, E. N. *et al.* Rationale for a Neisseria gonorrhoeae Susceptible Only Interpretive Breakpoint for Azithromycin. *Clin Infect Dis*, doi:10.1093/cid/ciz292 (2019).

90 Ma, K. C. *et al.* Increased power from bacterial genome-wide association conditional on known effects identifies Neisseria gonorrhoeae macrolide resistance mutations in the 50S ribosomal protein L4. *bioRxiv* (2020).

91 Cock, P. J. *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422-1423, doi:10.1093/bioinformatics/btp163 (2009).

92 Dillard, J. P. Genetic Manipulation of Neisseria gonorrhoeae. *Curr Protoc Microbiol* **Chapter 4**, Unit4A 2, doi:10.1002/9780471729259.mc04a02s23 (2011).

93 Mortimer, T. D. *et al.* The distribution and spread of susceptible and resistant Neisseria gonorrhoeae across demographic groups in a major metropolitan center. *medRxiv*, 2020.2004.2030.20086413, doi:10.1101/2020.04.30.20086413 (2020).

94 Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* **3**, 124, doi:10.12688/wellcomeopenres.14826.1 (2018).

95 Power, R. A., Parkhill, J. & de Oliveira, T. Microbial genome-wide association studies: lessons from human GWAS. *Nat Rev Genet* **18**, 41-50, doi:10.1038/nrg.2016.132 (2017).