

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Williams, E;Low, SJ;Pollock, GL;Taouk, ML;Prestedge, J;Howden, BP;Chow, EPF;Fairley, CK;Seib, KL;McCarthy, JS;Pasricha, S;Williamson, DA

Title:

Selecting candidate *Neisseria gonorrhoeae* strains for oropharyngeal gonorrhoea human challenge: a genomics-based analysis of clinical isolates

Date:

2025-09-01

Citation:

Williams, E., Low, S. J., Pollock, G. L., Taouk, M. L., Prestedge, J., Howden, B. P., Chow, E. P. F., Fairley, C. K., Seib, K. L., McCarthy, J. S., Pasricha, S. & Williamson, D. A. (2025). Selecting candidate *Neisseria gonorrhoeae* strains for oropharyngeal gonorrhoea human challenge: a genomics-based analysis of clinical isolates. *Lancet Microbe*, 6 (9), pp.101105-. <https://doi.org/10.1016/j.lanmic.2025.101105>.

Persistent Link:

<https://hdl.handle.net/11343/363463>

License:

CC BY-NC-ND

# Selecting candidate *Neisseria gonorrhoeae* strains for oropharyngeal gonorrhoea human challenge: a genomics-based analysis of clinical isolates

Eloise Williams\*, Soo Jen Low\*, Georgina L Pollock, Mona L Taouk, Jacqueline Prestedge, Benjamin P Howden, Eric P F Chow, Christopher K Fairley, Kate L Seib, James S McCarthy, Shivani Pasricha, Deborah A Williamson



## Summary

**Background** *Neisseria gonorrhoeae* is a human pathogen of major public health importance due to its increasing global prevalence and antimicrobial resistance (AMR). Evidence suggests that oropharyngeal infection plays a key role in *N gonorrhoeae* transmission and AMR; however, our understanding of oropharyngeal gonorrhoea pathogenesis is poor. A controlled human infection model (CHIM) for oropharyngeal gonorrhoea will improve understanding of infection and accelerate urgently needed novel gonorrhoea prevention and therapeutic strategies. As the first step in the development of this CHIM, we describe a systematic approach to CHIM strain selection that leverages genomics and clinical data.

**Methods** In this genomics-based analysis, we applied a systematic *N gonorrhoeae* challenge strain selection strategy incorporating genomic and clinical data to a primary dataset of clinical isolates of *N gonorrhoeae* collected from adult patients in Victoria, Australia, between Jan 1 and Dec 31, 2017, and July 1, 2019, and June 30, 2021. This selection strategy used clinical, phenotypic, and genomic characteristics to define a set of eight criteria that aimed to ensure the contemporary global clinical relevance of the candidate strains; select strains that would be applicable for the assessment of current and future gonorrhoea vaccines; and maximise participant safety by reducing the risk of disseminated gonococcal infection and clinically significant AMR. We applied these criteria to our primary dataset to generate a panel of potential challenge strains. From this final dataset of potential challenge strains, we predetermined that we would select up to ten isolates to proceed to the next stage of detailed phenotypic characterisation for final *N gonorrhoeae* CHIM strain selection.

**Findings** 5881 isolates comprised the primary dataset. After application of the selection criteria, most of the isolates (5795 [98.6%] of 5881) were excluded, mostly due to having clinically significant AMR and poor contemporary global clinical relevance. The remaining 86 *N gonorrhoeae* challenge strain candidates comprised five multilocus sequence types and six *N gonorrhoeae* multiantigen sequence types, many of which were represented by a single isolate. Of these 86 strains, five isolates were selected to maximise coverage of the phylogenetically distinct groups within the 86 candidate challenge strains and ensure representation of strains collected from various anatomical sites.

**Interpretation** We transparently describe a novel, systematic, and rational genomics-based strategy for oropharyngeal gonorrhoea CHIM strain selection that improves the efficiency and transparency of CHIM strain selection and enables identification of contemporary and clinically relevant potential challenge strains. A final *N gonorrhoeae* challenge strain will be selected from the subset of five shortlisted candidates after detailed phenotypic assessment.

**Funding** Medical Research Future Fund, Australian National Health and Medical Research Council and Australian Government Research Training Program.

**Copyright** Crown Copyright © 2025 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

*Neisseria gonorrhoeae*, the causative pathogen of gonorrhoea, is a major cause of morbidity in humans, with an estimated annual global prevalence of 87 million cases.<sup>1</sup> *N gonorrhoeae* infection causes a variety of clinical manifestations, ranging from asymptomatic mucosal infection (eg, oropharyngeal, rectal, and urogenital infection), to localised symptomatic disease (eg, urethritis, cervicitis, pelvic inflammatory disease, and conjunctivitis), and rarely invasive (eg, meningitis, endocarditis, and osteomyelitis) or disseminated infection.<sup>2</sup>

Urogenital gonorrhoea can cause long-term reproductive sequelae,<sup>2</sup> and infection in pregnancy is associated with adverse outcomes, including neonatal conjunctivitis.<sup>3</sup> Importantly, *N gonorrhoeae* has a propensity to acquire antimicrobial resistance (AMR), and has developed resistance to all available first-line antimicrobials.<sup>4</sup>

An effective *N gonorrhoeae* vaccine would have the potential to substantially reduce both the global increase in rates of *N gonorrhoeae* infection and AMR.<sup>5</sup> There is increased momentum in the development and deployment

*Lancet Microbe* 2025; 6: 101105

Published Online July 10, 2025  
<https://doi.org/10.1016/j.lanmic.2025.101105>

\*Contributed equally

**Department of Infectious Diseases** (E Williams MBBS, S J Low PhD, G L Pollock PhD, M L Taouk BSc, J Prestedge BSc, J S McCarthy MD, S Pasricha PhD, D A Williamson PhD) and **Microbiological Diagnostic Unit Public Health Laboratory** (B P Howden PhD), **University of Melbourne** at the Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia; **Victorian Infectious Diseases Reference Laboratory** (E Williams, J Prestedge) and **Victorian Infectious Diseases Service** (J S McCarthy), **Royal Melbourne Hospital** at the Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia; **Melbourne Sexual Health Centre**, **Alfred Health**, **Melbourne**, **VIC**, **Australia** (E P F Chow PhD, C K Fairley PhD); **School of Translational Medicine**, **Faculty of Medicine, Nursing and Health Sciences**, **Monash University**, **Melbourne**, **VIC**, **Australia** (E P F Chow, C K Fairley); **Melbourne School of Population and Global Health**, **University of Melbourne**, **Melbourne**, **VIC**, **Australia** (E P F Chow); **Institute for Biomedicine and Glycomics**, **Griffith University**, **Gold Coast**, **QLD**, **Australia** (K L Seib PhD); **Walter and Eliza Hall Institute of Medical Research**, **Melbourne**, **VIC**, **Australia** (J S McCarthy, G L Pollock); **School of Medicine**, **University of St Andrews**, **Fife**, **Scotland** (D A Williamson)

Correspondence to:  
Dr Eloise Williams, Department of Infectious Diseases, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, VIC 3000, Australia  
[eloise.williams3@unimelb.edu.au](mailto:eloise.williams3@unimelb.edu.au)

## Research in context

### Evidence before this study

Before undertaking this study, we considered the strategies used for selection of human challenge strains for bacterial controlled human infection models (CHIMs). We searched PubMed and EMBASE for studies published in English, between Jan 1, 2000, and Oct 1, 2023, with the search terms “human challenge model”, “human volunteer infection model”, “human experimental infection model”, “controlled human infection model”, “CHIM”, “strain selection”, and “strain”. Primary research studies that incorporated a bacterial CHIM, bacterial CHIM study protocols, and bacterial challenge strain selection or challenge strain characterisation studies were included. Relevant references of studies identified through this search were also reviewed. Although several contemporary CHIMs have been developed in the past decade, no studies have described a method for selecting a strain *a priori* from a defined database of isolates. Our search results also confirmed the absence of oropharyngeal gonorrhoea CHIMs in the literature. An oropharyngeal gonorrhoea CHIM incorporating a contemporary strain and modern manufacture techniques might aid the development of novel *N gonorrhoeae* vaccines and improve our understanding of the pathogenesis of contemporary oropharyngeal *N gonorrhoeae* infection.

### Added value of this study

This study is the first to describe the use of genomics and clinical data to select a panel of contemporary bacterial human challenge strains from one of the largest and most diverse clinical *Neisseria gonorrhoeae* reference collections globally. In addition, this study shows, to our knowledge for the first time, how a comprehensive bioinformatic analysis can be used *a priori* to select a panel of potential challenge strains from a large dataset. Previously described challenge strain selection studies have described the genotypic and phenotypic characteristics of several strains (either historical strains, strains available from research collections, or

strains repurposed from vaccine manufacture) with little rationale for why these strains were selected for further characterisation. The rational and systematic approach to strain selection described in this study improves the transparency of challenge strain selection, reduces the time and resources required to select a new challenge strain, and enables the identification of contemporary and clinically relevant potential challenge strains.

### Implications of all the available evidence

The selection of a human challenge strain is a pivotal step in the development of a CHIM. Judicious strain selection optimises the safety, generalisability, and clinical impact of CHIM studies. General criteria for an ideal challenge strain are well described, requiring that the strain is representative of contemporary circulating strains, is capable of causing infection at the intended anatomical site, possesses representative vaccine antigens and virulence factors, is susceptible to clinically relevant antimicrobials, is well characterised, and has a well defined provenance. Most challenge strains used in contemporary studies, however, are well defined historical strains, strains repurposed from vaccine manufacture, or strains available from research collections. The selection of an appropriate strain will vary according to organism-specific factors (such as strain diversity, clinical disease associations, virulence determinants, and vaccine antigens) and project-specific factors (such as the availability of well provenanced isolates, the resources available for detailed characterisation, and access to good-manufacturing-practice-grade production). In this study, we describe a rational, systematic, and transparent method for selecting a panel of challenge strains with genomics and clinical metadata for further phenotypic characterisation. The application of this method to challenge strain selection for other pathogens might contribute to an increase in the number of CHIMs that use contemporary challenge strains.

of *N gonorrhoeae* vaccines, with meningococcal B outer membrane vesicle (OMV) vaccines showing vaccine effectiveness of 31–46% against *N gonorrhoeae* infection in observational studies.<sup>6</sup> The development of vaccines designed specifically for *N gonorrhoeae* might further improve vaccine effectiveness.<sup>7</sup> Novel therapeutics are also required, with a number of first-in-class antimicrobials being explored for treatment of *N gonorrhoeae*, including zoliflodacin<sup>8</sup> and gepotidicin.<sup>9</sup> There is increasing evidence that the oropharynx plays a prominent role as both a reservoir for transmission of *N gonorrhoeae*<sup>10</sup> and an important environment for the horizontal transmission of resistance genes.<sup>11</sup> Modelling suggests that a vaccine without efficacy at the oropharynx would have little effect on reducing the prevalence of gonorrhoea.<sup>12</sup>

Controlled human infection models (CHIMs) are an important platform for assessing vaccines, therapeutics, and disease pathogenesis.<sup>13</sup> A CHIM for *N gonorrhoeae* urethritis was developed in the 1980s by investigators in the

USA with strains FA1090 and MS11mkC.<sup>14</sup> This urethritis CHIM has been used for studies of pathogenesis, particularly the role of potential virulence factors,<sup>14</sup> and is currently being used to investigate the efficacy of the meningococcal B OMV vaccine 4CMenB (Bexsero; NCT05294588). However, there have been no CHIMs for oropharyngeal gonorrhoea. An oropharyngeal gonorrhoea CHIM with a contemporary strain (ie, a strain isolated from clinical infection in the past decade) coupled with modern challenge strain manufacture techniques might accelerate the development of novel *N gonorrhoeae* vaccines and therapeutics and improve our understanding of the pathogenesis of oropharyngeal *N gonorrhoeae* infection. Modern challenge strain manufacture techniques include cell bank manufacture and release testing to produce quality controlled single-use vials for thaw from ultra-low temperature frozen storage on the day of challenge.

An effective and appropriate *N gonorrhoeae* challenge strain must fulfil several criteria to maximise the safety of

participants and optimise the generalisability of CHIM studies. Although several CHIMs have been developed in the past decade, few have systematically documented a process for strain selection<sup>15</sup> and none have described a method for selecting a strain a priori from a defined database of isolates. Bacterial CHIMs have generally used well characterised historical strains, such as the *Salmonella enterica* serovar Typhi Quail strain (isolated from the gallbladder of an individual with chronic carriage in 1958),<sup>16</sup> or contemporary strains after historical strains were found to have unfavourable characteristics, such as the presence of the *cag* pathogenicity island in *Helicobacter pylori* (associated with increased risk of peptic ulcer and gastric cancer)<sup>17</sup> or the presence of a lipooligosaccharide with molecular mimicry with human gangliosides in *Campylobacter jejuni* (which might induce Guillain-Barré syndrome).<sup>18</sup> This study bridges an important gap by providing a logical and transparent approach to optimising safety and generalisability in the initial stage of challenge strain selection, with novel application of genomic analysis for selection of candidate challenge strains.

The *N gonorrhoeae* laboratory-adapted challenge strains used in the urethritis CHIMs FA1090 and MS11mkC were originally isolated in the 1970s.<sup>14</sup> Given the evolution of *N gonorrhoeae* over time,<sup>19</sup> a contemporary *N gonorrhoeae* challenge strain might facilitate study of clinically relevant circulating *N gonorrhoeae*. In this study, we describe an approach to selecting CHIM challenge strains that combines available genomic and clinical data.

## Methods

### Data sources

*N gonorrhoeae* infection is a notifiable infection in Australia, with diagnostic microbiology laboratories requested to refer *N gonorrhoeae* isolates to state-based reference laboratories for antimicrobial susceptibility testing. Between Jan 1 and Dec 31, 2017, and July 1, 2019, and June 30, 2021, all referred *N gonorrhoeae* isolates from adult patients (ie, individuals aged 16 years or older) in the state of Victoria, Australia, underwent phenotypic antimicrobial susceptibility testing, DNA extraction, and whole-genome sequencing (appendix 1 pp 2–3). The two timeframes were due to impacts of the COVID-19 pandemic on genomic sequencing capacity during this period. Relevant clinical data accompanying the isolates were obtained from the referring pathology service, including the anatomical site of sample collection (appendix 2 p 1). The combined genomic data and metadata comprised the so-called primary dataset. No individual patient consent was required as the data were collected in accordance with the Victorian Public Health and Wellbeing Act, 2008.<sup>20</sup> No identifiable patient data were used in this study; therefore, ethical approval was not required.

### *N gonorrhoeae* CHIM strain selection criteria

We developed a set of eight criteria for *N gonorrhoeae* CHIM strain selection based on generalisability and safety

parameters (table 1).<sup>21</sup> We applied these criteria to our primary dataset to generate a panel of potential *N gonorrhoeae* challenge strains. A stepwise elimination of strains was performed for each key criterion in the order described below, with only the strains that fulfilled the requirements of each criterion retained at each stage. From the final dataset of potential challenge strains, we predetermined that we would select up to ten isolates to proceed to the next stage of detailed phenotypic characterisation with the key challenge agent quality metrics of identity, purity, potency, and stability<sup>22</sup> (appendix 1 p 12) to facilitate final *N gonorrhoeae* CHIM strain selection, which will be described in detail elsewhere.

### Generalisability criteria

To ensure global clinical relevance, we downloaded genome sequences and associated metadata for *N gonorrhoeae* isolates from Jan 1, 2015, to Oct 31, 2021, from Pathogenwatch and supplemented these with genomes from Asia and Africa, regions which were under-represented in the Pathogenwatch dataset compared with other regions (appendix 1 p 2). These genome sequences were identified through a literature review of clinical studies involving *N gonorrhoeae* isolates undertaken during the study period (appendix p 2). In-silico sequence typing with the multilocus sequence type (MLST), *N gonorrhoeae* multiantigen sequence type (NG-MAST), and *N gonorrhoeae* sequence typing for antimicrobial resistance (NG-STAR) typing schemes was performed on the genomic data from the primary and global datasets (appendix 1 pp 2–4). In addition, the primary dataset was classified into two major lineages (lineage A and lineage B) described in global phylogeographic analyses.<sup>23</sup> Geographical diversity within each sequence type was calculated with the Shannon diversity index from the combined primary and global datasets. An *N gonorrhoeae* isolate met our global clinical relevance criteria if it belonged to an MLST or NG-MAST sequence type that was present on at least three continents and had a Shannon index of  $\geq 1$  or a diversity ranking score (ie, Shannon index  $\times$  sequence type cluster size) greater than or equal to the 50th percentile.

To verify an association with contemporary oropharyngeal and urogenital *N gonorrhoeae* infection, we grouped anatomical sites of infection into six categories: urogenital, rectal, oropharyngeal, sterile site, ocular, or other or unknown (appendix 1 pp 4, 13). *N gonorrhoeae* isolates met our anatomical site criterion if they belonged to an MLST or NG-MAST that had shown oropharyngeal and urogenital infection in the combined primary and global datasets.

To check for the presence of a broad range of putative vaccine antigens, we identified 50 putative vaccine antigens in the primary dataset (appendix 1 pp 4–5, 14–16). We defined the presence of a broad range of vaccine antigens as a key selection criterion for an *N gonorrhoeae* challenge strain. Given the large number of putative *N gonorrhoeae* vaccine antigens, *N gonorrhoeae* isolates in the primary dataset required the presence of a smaller subset of 11 key

For Pathogenwatch see <https://pathogen.watch/>

See Online for appendix 1

See Online for appendix 2

| Challenge strain selection criteria                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Rationale                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Generalisability</b>                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Global clinical relevance                                 | Strain belongs to an MLST or NG-MAST sequence type shown to cause contemporary clinical infection in a wide range of geographical settings, with the following criteria: presence of <i>Neisseria gonorrhoeae</i> of the sequence type on $\geq 3$ continents between 2015 and 2021; and the sequence type meets a predefined global diversity threshold based on the number of countries it is present in and the number of isolates per country (ie, Shannon diversity index $\geq 1$ or diversity ranking score $\geq$ the 50th percentile threshold) | Due to the genomic plasticity of <i>N gonorrhoeae</i> and the global diversity of <i>N gonorrhoeae</i> isolates from different geographical regions, a challenge strain should be prevalent across various geographical regions to ensure the global clinical relevance and generalisability of the CHIM                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| Causes infection at anatomical site of challenge          | Definite contemporary cause of oropharyngeal gonorrhoea                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | The oropharynx is the third most common site of <i>N gonorrhoeae</i> infection and increasing evidence suggests it is a major contributor of <i>N gonorrhoeae</i> transmission; the oropharynx represents a site with a high risk for the development of AMR due to horizontal transfer of AMR determinants from commensal <i>Neisseria</i> species; <i>N gonorrhoeae</i> vaccine modelling studies suggest that a vaccine with poor efficacy at the oropharynx would have little effect on <i>N gonorrhoeae</i> prevalence                                                                                                                                                                                                                                                                                                                                                                   |
| Causes infection at anatomical site of challenge          | Definite contemporary cause of urogenital gonorrhoea                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Urogenital gonorrhoea has the greatest short-term and long-term clinical impact on sexual and reproductive health and is an important clinical endpoint for vaccine and therapeutics trials; the challenge strain should be suitable for use in future gonorrhoea urethritis studies                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Applicable to putative vaccines                           | Presence of a broad range of vaccine antigens                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | A key function of an oropharyngeal CHIM is early assessment of novel vaccines; as such, the challenge strain should contain a broad range of putative vaccine antigens                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <b>Safety</b>                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Reduced risk of severe infection                          | Absence of genomic determinants associated with disseminated gonococcal infection                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | <i>N gonorrhoeae</i> rarely causes severe manifestations (eg, meningitis, endocarditis, and osteomyelitis) and disseminated gonococcal infection (eg, arthritis-tenosynovitis-rash syndrome); although no specific virulence factors completely dictate clinical disease manifestations, specific genomic determinants have consistently been associated with disseminated gonococcal infection and should be avoided in a challenge strain to maximise participant safety                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Reduced risk of severe infection                          | Exclusion of strains isolated from sterile sites                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | As a marker of potential increased virulence, <i>N gonorrhoeae</i> isolates clinically associated with severe or disseminated infection should be avoided to maximise participant safety; the selection of a strain that has not been associated with disseminated disease is likely to improve participant acceptability of a proposed oropharyngeal gonorrhoea CHIM                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Susceptible to multiple antimicrobials used for treatment | Absence of clinically significant phenotypic AMR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | The participant should be cured and the challenge strain eradicated from the oropharynx with well tolerated antimicrobial therapy; <i>N gonorrhoeae</i> has developed AMR to all first-line anti-gonococcal therapies; ensuring antimicrobial susceptibility to multiple lines of clinically relevant antimicrobials is paramount to the participant and public health safety of an oropharyngeal CHIM; selection of a strain susceptible to all clinically relevant antimicrobials will probably increase the participant acceptability of a proposed oropharyngeal gonorrhoea CHIM; and selection of a strain that is susceptible to all clinically relevant antimicrobials will reduce the public health risk associated with potential transmission of the challenge strain to non-challenge participants, a risk mitigation step that must be actively managed as part of any CHIM study |
| Susceptible to multiple antimicrobials used for treatment | Absence of clinically significant genomic AMR determinants                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Numerous AMR mechanisms have been described for <i>N gonorrhoeae</i> ; the exclusion of strains with plasmid-associated AMR genes or mutations within chromosomal genes that strongly correlate with phenotypic AMR in clinically relevant antimicrobials is key to participant safety, as this approach increases the likelihood that the participant can be cured and the challenge strain can be eradicated from the oropharynx with well tolerated antimicrobial therapy; this approach reduces the risk of transmission of AMR determinants to other oropharyngeal organisms, including those with the potential to cause invasive disease, such as <i>Neisseria meningitidis</i>                                                                                                                                                                                                        |

AMR=antimicrobial resistance. CHIM=controlled human infection model. MLST=multilocus sequence type. NG-MAST=*Neisseria gonorrhoeae* multiantigen sequence type.

**Table 1: Key criteria for *Neisseria gonorrhoeae* challenge strain selection for an oropharyngeal gonorrhoea CHIM**

vaccine antigen genes (appendix 1 pp 17–18) to meet our vaccine applicability challenge strain criterion. These antigens (including NHBA, PorB, and MetQ) were considered key vaccine targets due to their presence in vaccines that have progressed to human clinical trials<sup>24</sup> or their ability to induce immune responses or enhance gonococcal clearance in contemporary murine models.<sup>6</sup>

#### Safety criteria

Genes with putative associations to disseminated gonococcal infection are described in the supplementary materials (appendix 1 p 4). To reduce the risk of disseminated gonococcal infection, we applied a strain selection criterion that required the absence of the *porB* allele that encodes porB1A,

a genomic determinant that is strongly associated with disseminated gonococcal infection.

To reduce the risk of disseminated gonococcal infection or severe disease, we excluded all isolates in the primary dataset that were obtained from sterile, ocular, or unknown anatomical sites.

Genomic determinants of AMR were investigated by examining the alleles of the seven NG-STAR resistance genes and the presence of *blaTEM* and *tetM* genes for each isolate in the primary dataset. On the basis of a literature review, we defined and applied a set of genomic susceptibility criteria, which excluded isolates encoding genomic determinants highly associated with AMR, including mosaic or semi-mosaic *penA* genes (associated with

ceftriaxone resistance), *gyrA* or *parC* mutations (associated with ciprofloxacin resistance), 23S rRNA mutations (associated with azithromycin resistance), *mtrR* mutations or promoter disruption (associated with macrolide, tetracycline, and  $\beta$ -lactam resistance), or the presence of *blaTEM* (a mediator of plasmid-mediated high-level penicillin resistance) and *tetM* (a mediator of plasmid-mediated high-level tetracycline resistance) genes.<sup>25</sup> Additional genomic determinants of interest (*mtrC*, *mtrD*, *rpsJ*, *ermB*, *rplD*, and *rplV*) were also examined in the final dataset (appendix 2 p 7).

We applied phenotypic antimicrobial susceptibility criteria that included *N gonorrhoeae* strains with phenotypic susceptibility to ceftriaxone, ciprofloxacin, azithromycin, and tetracycline (appendix 1 pp 2, 19). Criteria for penicillin susceptibility were less stringent (including both strains susceptible to penicillin and those less susceptible) as penicillin is rarely used for contemporary treatment of gonorrhoea.

### Statistical analysis

Analyses were conducted with R (version 4.4.0). Sums, proportions, and percentiles were calculated with base R and the tidyverse package (v2.0.0) and data visualisations were generated with the following packages: data.table (v1.15.4); tidyverse (v2.0.0); rworldmap (v1.3–8); and ggvenn (v0.1.10). The Shannon index for sequence types was calculated with the vegan package (v2.6–6.1), and 1000 non-parametric bootstrap resampling was performed with the boot package (v1.3–30) and a custom function.<sup>26</sup> A phylogeny of the strains in the primary dataset was inferred with IQ-TREE from a core genome single nucleotide polymorphism alignment.

### Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

### Results

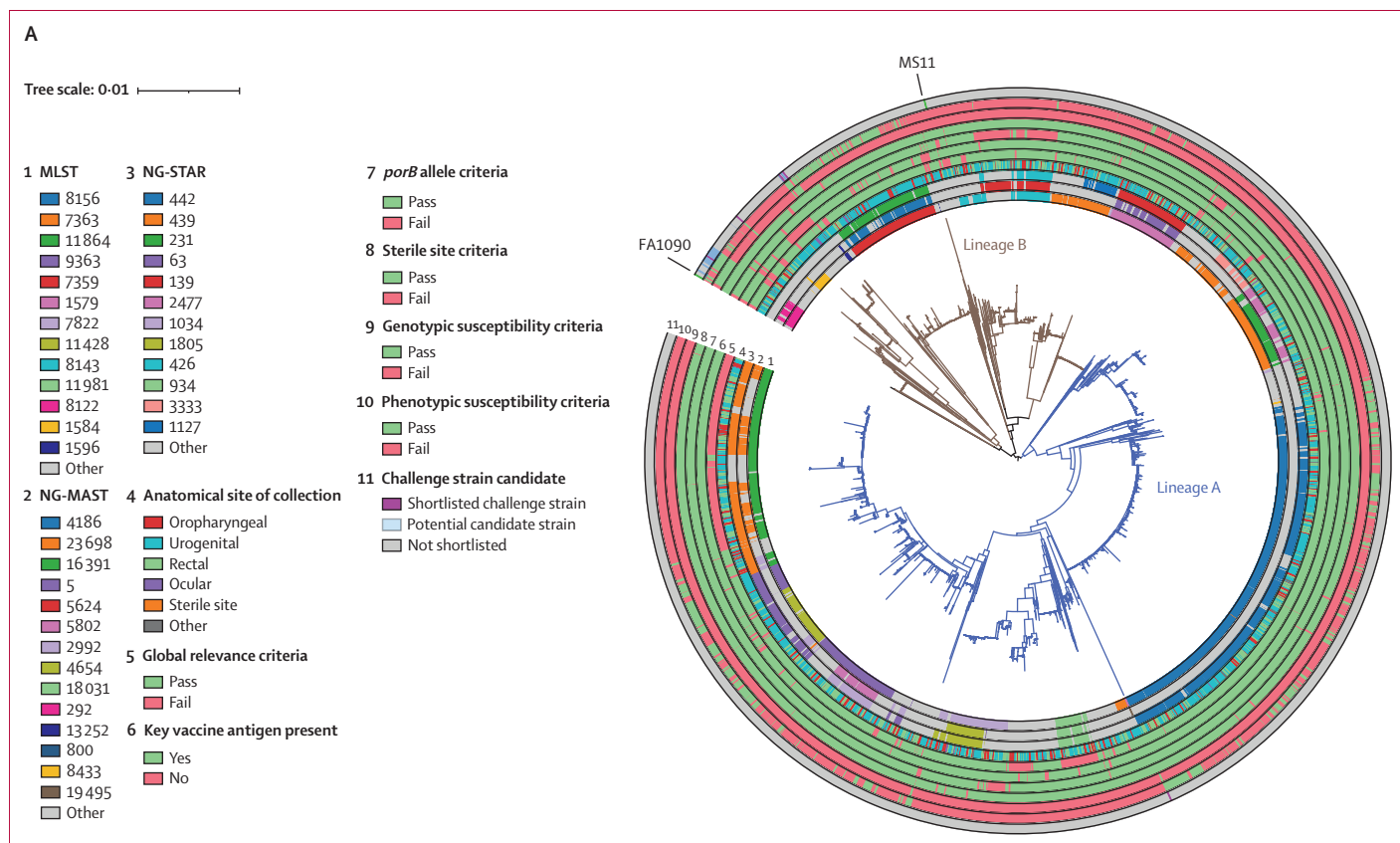
Overall, 5881 *N gonorrhoeae* isolates were included in the primary dataset, which formed two distinct lineages in the phylogenetic tree (4275 isolates in lineage A and 1606 isolates in lineage B; figure 1A; table 2; appendix 2 p 1). The primary dataset comprised 287 NG-MASTs, 269 NG-STARs, and 138 MLSTs. Eight genomes could not be assigned to known sequence types. 36 (0.6%) of 5881 isolates could not be assigned an MLST (appendix 2 p 3). The most common MLSTs were ST8156 (1279 [21.7%] of 5881 isolates), ST7363 (738 [12.5%] of 5881), and ST11864 (668 [11.4%] of 5881). 56 (40.6%) of 138 MLSTs were singletons. Almost half of isolates (2594 [44.1%] of 5881) could not be assigned to an existing NG-MAST. Of the remaining isolates, the most common NG-MASTs were ST4186 (270 [4.6%] of 5881 isolates), ST23698 (265 [4.5%] of 5881), and ST16391 (247 [4.2%] of 5881; figure 1A). 144 (50.2%) of 287 NG-MASTs were singletons (appendix 2 p 3). For the NG-STAR typing scheme, 546 (9.3%) of 5881 isolates could not be assigned to an existing sequence type. The

most common NG-STARs were ST442 (1186 [20.2%] of 5881 isolates), ST439 (598 [10.2%] of 5881), and ST231 (327 [5.6%] of 5881; figure 1A). 119 (44.2%) of 269 sequence types were singletons (appendix 2 p 3).

In total, 15 528 genome sequences were downloaded from Pathogenwatch spanning the period from Jan 1, 2015, to Oct 31, 2021, supplemented with 1290 sequences from Asia (577 sequences) and Africa (713 sequences) obtained from five studies identified through a literature review (appendix 1 p 2). These sequences comprised the global dataset (appendix 1 p 6), which covered all continents excluding Antarctica and 27 countries (appendix 2 p 2). Overall, 22 699 genomes were included in the combined primary and global datasets, with most genomes and resulting sequence type diversity originating from the USA (9118 [40.2%] of 22 699), Australia (5889 [25.9%] of 22 699), and Norway (2211 [9.7%] of 22 699; appendix 1 p 6). All other countries contributed fewer than 1000 isolates to the dataset. Few overlaps of sequence types were observed between the primary and global datasets for all sequence typing schemes (appendix 1 p 6), highlighting the extensive genomic diversity of *N gonorrhoeae* across different geographical regions.

In total, 112 (17%) of 648 MLSTs in the global dataset were present in the primary dataset. In addition, 26 (4%) of 648 MLSTs in the global dataset were unique to the primary dataset and not present in any other country within the sampled period (appendix 1 p 6). Of the 112 shared sequence types, 82 (73%) were found on at least three continents (appendix 1 pp 7–8; appendix 2 p 4). Overall, 4698 isolates assigned to 37 MLSTs met our predefined geographical diversity criteria (appendix 1 p 7). A total of 189 (9.2%) of 2389 NG-MASTs in the global dataset were also present in the primary dataset. In addition, 98 NG-MASTs were unique to the primary dataset (appendix 1 p 6). Of the 189 shared NG-MASTs, 97 (51%) were found on at least three continents (appendix 1 p 7; appendix 2 p 5). Overall, a total of 1056 isolates assigned to 35 NG-MASTs passed the geographical diversity criteria (appendix 1 p 7). In summary, 4768 (81.1%) of 5881 isolates from the primary dataset met the contemporary global relevance challenge strain criteria of belonging to an MLST or NG-MAST that was present on at least three continents and meeting the predefined geographical diversity threshold (ie, Shannon index  $\geq 1$  or diversity ranking score  $\geq$  the 50th percentile).

In the primary dataset, *N gonorrhoeae* was isolated from the following sites: urogenital (3136 [53.3%] of 5881 isolates); rectal (1702 [28.9%] of 5881); oropharyngeal (975 [16.6%] of 5881); ocular (28 [0.5%] of 5881); sterile sites (9 [0.2%] of 5881); and other or unknown (31 [0.5%] of 5881; table 2). Across the global dataset, the site of gonococcal isolation was unknown for a substantial proportion of isolates (12 957 [77.0%] of 16 818; appendix 1 p 8). For the remaining 3861 global isolates with a known anatomical site of isolation, 2333 (60.4%) were urogenital, 963 (24.9%) were rectal, 543 (14.1%) were oropharyngeal, 14 (0.4%) were ocular, and eight (0.2%) were sterile site isolates



(Figure 1 continues on next page)

(appendix 1 p 9). MLSTs and NG-MASTs without both oropharyngeal and urogenital representation were generally limited to small clusters, potentially due to genome under-representation (appendix 1 p 9). After exclusion of 120 isolates that did not have both oropharyngeal and urogenital representation by either MLST or NG-MAST (appendix 1 p 9), 5761 (98.0%) isolates from the primary dataset from at least 111 MLSTs (110 assigned sequence types plus unclassified sequence types) and 243 NG-MASTs (242 assigned sequence types plus unclassified sequence types) met the anatomical site criterion.

Most candidate vaccine antigens (44 [88%] of 50) were present in  $\geq 90\%$  of the genomes, including the 11 key vaccine antigens that have progressed to human clinical trials or murine vaccine studies (appendix 1 p 10). Of the 44 prevalent antigens, the least conserved in terms of nucleotide sequence variants were *porB* (with 996 unique alleles), *tbpB* (with 715 unique alleles), and *pilC* (with 644 unique alleles). The most conserved antigens in terms of nucleotide sequence variants were NGO0778 (with six alleles), NGO0416 (with eight alleles), *bamE* (with eight alleles), and *slcC* (with nine alleles; appendix 1 p 10; appendix 2 p 7). Overall, 5552 (94.4%) of 5881 isolates from the primary dataset met the defined criterion of possessing all 11 putative key vaccine antigens.

Across the primary dataset, 426 (7.2%) of 5881 isolates contained a *porB1A* allele (appendix 1 p 11) and were excluded on the basis of the presence of a genomic determinant associated with disseminated gonococcal infection. Three isolates with unassigned *porB* alleles were also excluded. As such, 5452 (92.7%) of 5881 isolates from the primary dataset were retained.

Overall, nine (0.2%) of 5881 isolates in the primary dataset were from sterile sites (three [33.3%] from blood and six [66.7%] from joint fluid). 28 (0.5%) of 5881 isolates were from ocular sites and 31 (0.5%) were from unknown sites (table 2), all of which met the criteria for exclusion from the final dataset. Overall, 5813 (98.8%) of 5881 isolates from the primary dataset were retained after application of this criterion.

Within the primary dataset, 826 (14.0%) of 5881 isolates contained mosaic or semi-mosaic *penA* genes, 2203 (37.5%) contained a *gyrA* mutation, 1908 (32.4%) contained a *parC* mutation, 95 (1.6%) contained a 23S rRNA mutation, and 5611 (95.4%) contained an *mtrR* mutation or promoter disruption. *blaTEM* was present in 1000 (17.0%) of 5881 isolates and *tetM* was present in 791 (13.5%) of 5881 isolates (table 2; appendix 1 p 11). Only 29 (0.5%) isolates contained wild-type alleles for *penA*, *mtrR*, *porB*, *ponA*, *gyrA*, *parC*, and 23S rRNA genes and non-mosaic

*penA* genes, and did not contain the *tetM* and *bla*<sub>TEM</sub> genes. Overall, 128 (2.2%) of 5881 isolates met the predefined criterion for the absence of clinically significant genomic AMR determinants.

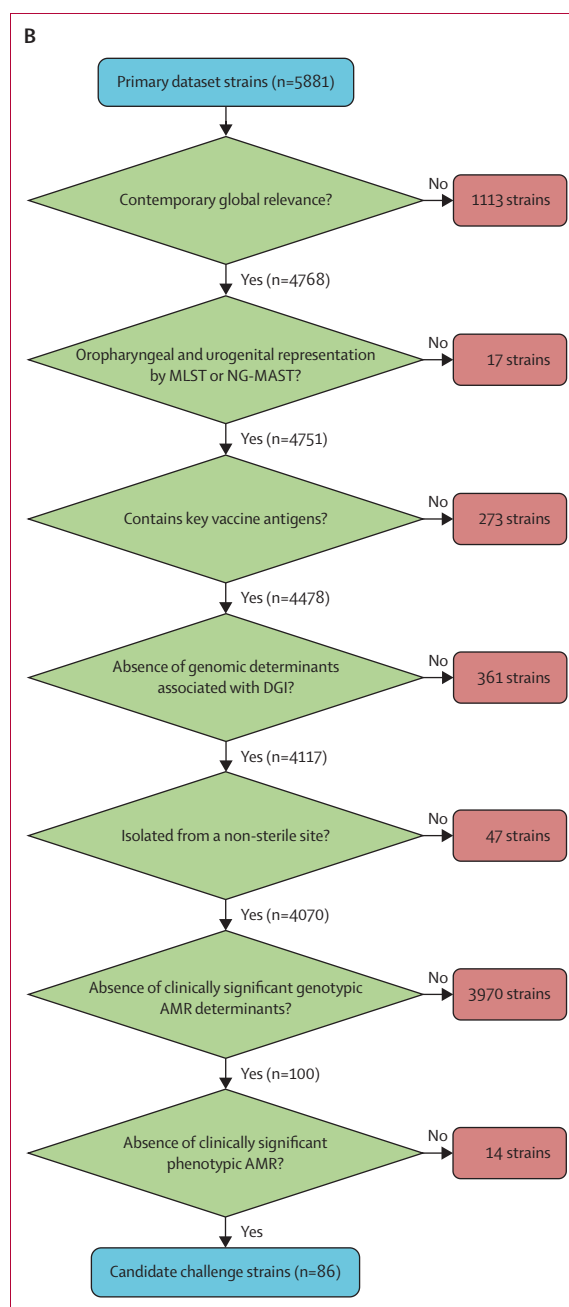
In the primary dataset, 5786 (98.4%) of 5881 isolates were phenotypically susceptible to ceftriaxone, 3676 (62.5%) were susceptible to ciprofloxacin, 5491 (93.4%) were susceptible to azithromycin, and 2449 (41.6%) were susceptible to tetracycline. 28 (0.5%) of 5881 isolates were susceptible to penicillin and 4192 (71.3%) of 5881 isolates were less susceptible to penicillin (table 2). In total, 1324 (22.5%) of 5881 isolates met the phenotypic AMR criterion, which required isolates to be susceptible to ceftriaxone, ciprofloxacin, azithromycin, and tetracycline and either susceptible or less susceptible to penicillin (appendix 1 p 11).

A stepwise elimination of strains was performed for each key criterion in the order discussed above, retaining only the strains fulfilling the requirements at each stage. This elimination resulted in a final dataset of 86 challenge strain candidates (figure 1B). Further characterisation of additional genomic determinants of resistance was performed for these isolates (appendix 2 p 7). Most strains were excluded due to clinically significant AMR or poor contemporary global relevance (figure 1A, B; appendix 1 p 11). The challenge strain candidate dataset comprised five MLSTs and six NG-MASTs (figure 2A, B) that mostly formed a single clade within the global *N gonorrhoeae* phylogeny (figure 2C).

A subset of five isolates was selected from the candidate challenge strain dataset of 86 strains (figure 3; appendix 2 p 7). These isolates were selected to maximise coverage of the phylogenetically distinct groups within the 86 candidate strains, including clusters B.1, B.2, and B.3, and a single lineage A isolate. One isolate each was selected from cluster B.1, cluster B.3, and lineage A. Two isolates were selected from cluster B.2, as this cluster included several isolates with a wild-type NG-STAR allele profile. The five shortlisted isolates represent each of the assigned MLSTs identified within the candidate challenge strain dataset, with four isolates from lineage B and one isolate from lineage A (including isolates from urogenital, rectal, and oropharyngeal sites). Among the three clusters observed in lineage B, the available isolate with the fewest genomic determinants of resistance or the lowest minimum inhibitory concentration to clinically relevant anti-gonococcal antimicrobials was selected. When there was no notable difference in susceptibility characteristics (appendix 2 p 1), consideration was given to ensure representation of strains collected from various anatomical sites among the five isolates.

## Discussion

In this study, we describe a systematic framework for the selection of a contemporary *N gonorrhoeae* challenge strain from a large genomic epidemiological dataset. *N gonorrhoeae* is a highly complex organism, characterised by extensive genomic diversity within and across geographical regions, numerous AMR mechanisms, and complex virulence mechanisms (such as phase variability,



**Figure 1: Overview of the genomes in the primary dataset and the strain elimination process followed in this study**

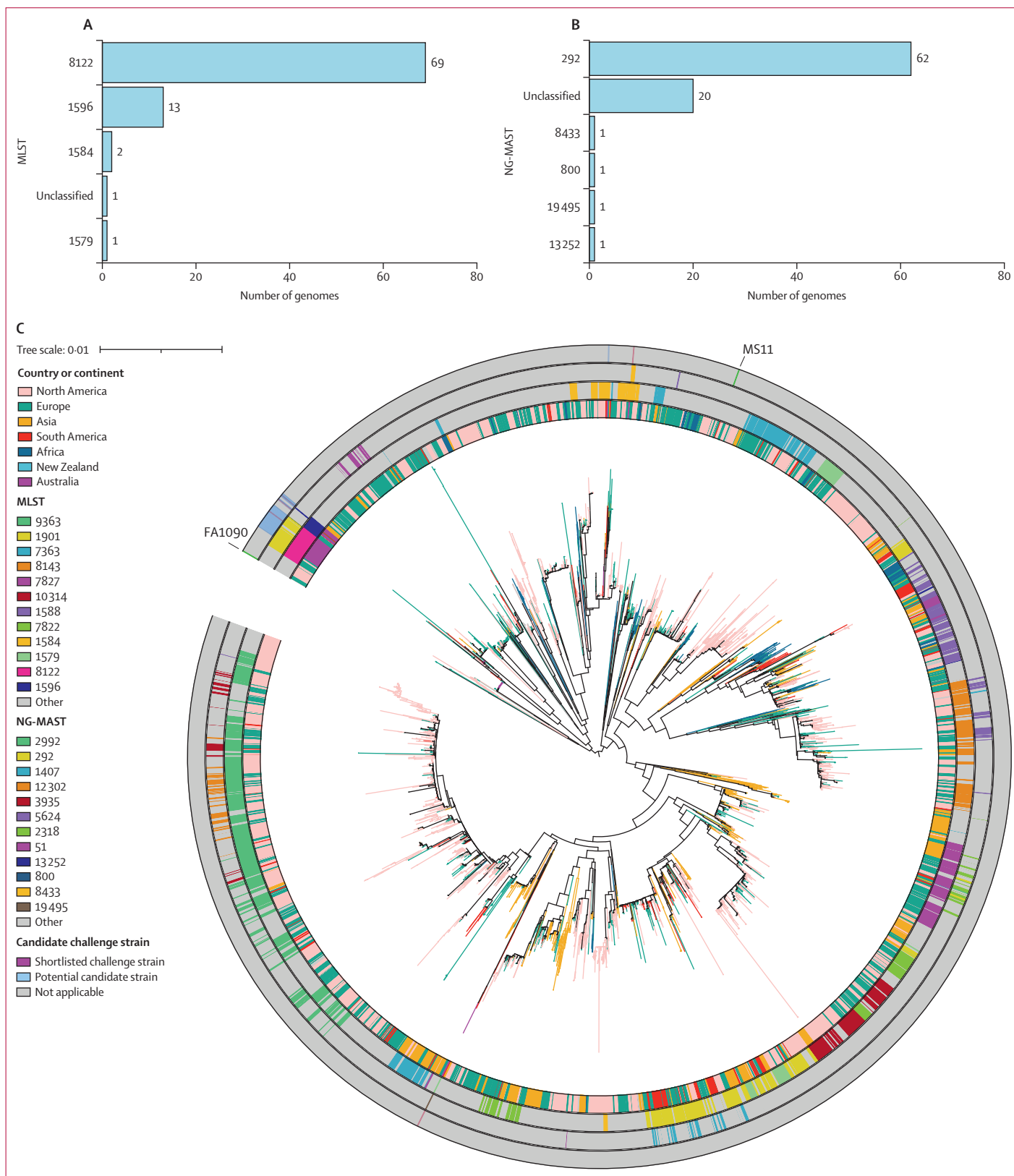
(A) Midpoint-rooted phylogeny of *Neisseria gonorrhoeae* genomes in the primary dataset and strains FA1090 and MS11 (which were included in a controlled human infection model for *N gonorrhoeae* urethritis developed in the 1980s).<sup>14</sup> This phylogeny was inferred with IQ-TREE from a core genome single nucleotide polymorphism alignment. Rings 1–3 are coloured by sequence types according to the MLST, NG-MAST, and NG-STAR typing schemes for sequence types containing at least 100 genome representatives. Rings 5–10 indicate whether or not a strain fulfils specific key selection criteria. Tips are coloured by final strain designation after applying all key selection criteria. The outermost ring (ring 11) indicates whether a strain passed all the defined key selection criteria. The positions of FA1090 and MS11 within the phylogeny are indicated on the outermost ring (green lines). The scale bar is in units of nucleotide substitutions per site. (B) Stepwise strain elimination according to seven defined key selection criteria. The numbers of strains passing and failing each criterion are shown up to the final set of 86 candidate challenge strains. AMR=antimicrobial resistance. DGI=disseminated gonococcal infection. MLST=multilocus sequence type. NG-MAST=*Neisseria gonorrhoeae* multiantigen sequence type. NG-STAR=*Neisseria gonorrhoeae* sequence typing for antimicrobial resistance.

|                                                                               | n or n (%; N=5881) |
|-------------------------------------------------------------------------------|--------------------|
| <b>Unique sequence types</b>                                                  |                    |
| Multilocus sequence type                                                      | 1385               |
| NG-MAST                                                                       | 28 793             |
| NG-STAR                                                                       | 269                |
| <b>Year of isolation</b>                                                      |                    |
| 2017                                                                          | 2167 (36.8%)       |
| 2019                                                                          | 1225 (20.8%)       |
| 2020                                                                          | 1557 (26.5%)       |
| 2021                                                                          | 931 (15.8%)        |
| <b>Anatomical site of isolation</b>                                           |                    |
| Urogenital                                                                    | 3136 (53.3%)       |
| Rectal                                                                        | 1702 (28.9%)       |
| Oropharyngeal                                                                 | 975 (16.6%)        |
| Other                                                                         | 31 (0.5%)          |
| Ocular                                                                        | 28 (0.5%)          |
| Sterile                                                                       | 9 (0.2%)           |
| <b>Phenotypic antimicrobial susceptibility†</b>                               |                    |
| Penicillin                                                                    |                    |
| Susceptible                                                                   | 28 (0.5%)          |
| Less susceptible                                                              | 4192 (71.3%)       |
| Resistant                                                                     | 1661 (28.2%)       |
| Ceftriaxone                                                                   |                    |
| Susceptible                                                                   | 5786 (98.4%)       |
| Decreased susceptibility                                                      | 94 (1.6%)          |
| Non-susceptible                                                               | 1 (<0.1%)          |
| Ciprofloxacin                                                                 |                    |
| Susceptible                                                                   | 3676 (62.5%)       |
| Less susceptible                                                              | 69 (1.2%)          |
| Resistant                                                                     | 2136 (36.3%)       |
| Azithromycin                                                                  |                    |
| Susceptible                                                                   | 5491 (93.4%)       |
| Resistant                                                                     | 390 (6.6%)         |
| Tetracycline                                                                  |                    |
| Susceptible                                                                   | 2449 (41.6%)       |
| Intermediate                                                                  | 1923 (32.7%)       |
| Resistant                                                                     | 1509 (25.7%)       |
| Spectinomycin                                                                 |                    |
| Susceptible                                                                   | 2631 (44.7%)       |
| Resistant                                                                     | 0                  |
| Not tested                                                                    | 3250 (55.3%)       |
| <b>Genomic determinants associated with disseminated gonococcal infection</b> |                    |
| PorB structure                                                                |                    |
| PorB1A                                                                        | 426 (7.2%)         |
| PorB1B                                                                        | 5455 (92.8%)       |
| <b>Genomic determinants of antimicrobial resistance</b>                       |                    |
| penA                                                                          |                    |
| Mosaic                                                                        | 812 (13.8%)        |
| Semi-mosaic                                                                   | 14 (0.2%)          |
| Non-mosaic                                                                    | 4933 (83.9%)       |
| Ala311Val                                                                     | 3 (0.1%)           |
| Ala501Val/Thr                                                                 | 248 (4.2%)         |
| <b>(Continued from previous column)</b>                                       |                    |
| Asn513Tyr                                                                     | 811 (13.8%)        |
| Ala517Gly                                                                     | 4897 (83.3%)       |
| Gly543Ser                                                                     | 521 (8.9%)         |
| Other‡                                                                        | 122 (2.1%)         |
| Non-mosaic or wild-type                                                       | 30 (0.5%)          |
| mtrR                                                                          |                    |
| -35Ala del                                                                    | 2062 (35.1%)       |
| Ala39Thr                                                                      | 3141 (53.4%)       |
| Gly45Asp                                                                      | 106 (1.8%)         |
| 161 base-pair deletion                                                        | 2 (<0.1%)          |
| Meningitidis-like promoter                                                    | 377 (6.4%)         |
| Disrupted WHO-P-like promoter                                                 | 1 (<0.1%)          |
| Other‡                                                                        | 59 (1.0%)          |
| Wild-type                                                                     | 270 (4.6%)         |
| porB                                                                          |                    |
| Gly120Asp/Lys/Asn/Arg                                                         | 1371 (23.3%)       |
| Ala121Asp/Gly/Asn/Ser/Val                                                     | 1706 (29.0%)       |
| porB1A                                                                        | 426 (7.2%)         |
| Other‡                                                                        | 3 (0.1%)           |
| Wild-type                                                                     | 3742 (63.6%)       |
| ponA                                                                          |                    |
| Leu421Pro                                                                     | 2579 (43.9%)       |
| Wild-type                                                                     | 3302 (56.1%)       |
| gyrA                                                                          |                    |
| Ser91Phe/Tyr                                                                  | 2198 (37.4%)       |
| Asp95Ala/Gly/Asn/Tyr                                                          | 2198 (37.4%)       |
| Wild-type                                                                     | 3678 (62.5%)       |
| parC                                                                          |                    |
| Asp86Asn                                                                      | 916 (15.6%)        |
| Ser87Asn/Arg                                                                  | 995 (16.9%)        |
| Ser88Pro                                                                      | 46 (0.8%)          |
| Glu91Lys                                                                      | 1 (<0.1%)          |
| Wild-type                                                                     | 3973 (67.6%)       |
| 23S rRNA                                                                      |                    |
| Ala2059Gly                                                                    | 2 (<0.1%)          |
| Cys2611Thr                                                                    | 78 (1.3%)          |
| Other‡                                                                        | 15 (0.3%)          |
| Wild-type                                                                     | 5786 (98.4%)       |
| blaTEM                                                                        |                    |
| Present                                                                       | 1000 (17.0%)       |
| Absent                                                                        | 4881 (83.0%)       |
| tetM                                                                          |                    |
| Present                                                                       | 791 (13.5%)        |
| Absent                                                                        | 5090 (86.5%)       |

(Table 2 continues in next column)

NG-MAST=N gonorrhoeae multiantigen sequence type. NG-STAR=N gonorrhoeae Sequence Typing for Antimicrobial Resistance. WHO-P=WHO-P reference strain. \*N gonorrhoeae isolates submitted to a state public health reference laboratory in Melbourne, VIC, Australia, from Jan 1 to Dec 31, 2017, and from July 1, 2019, to June 30, 2021. †Phenotypic antimicrobial susceptibility testing performed with the Australian Gonococcal Surveillance Programme agar dilution method. ‡Non-wild-type with no exact allele match.

**Table 2: Characteristics of the primary dataset\* of *Neisseria gonorrhoeae* isolates**



serum resistance, and host immune manipulation).<sup>27</sup> The challenge strain selection criteria were therefore judiciously chosen to ensure a sufficient number of isolates would be available to achieve diversity among the final challenge strain panel, while also ensuring the clinically relevant safety and generalisability criteria were met.

Comparison of the isolates within our geographically restricted primary dataset and available genomes from around the world showed the extensive genomic diversity of *N gonorrhoeae* both within and across different geographical regions. This genomic diversity represents an important challenge to identification of an appropriate *N gonorrhoeae* challenge strain, for which generalisability is a key feature. To maximise generalisability, we did not focus exclusively on genomic diversity but also included parameters such as anatomical site of infection and geographical location. We observed a high prevalence of putative vaccine antigen targets in the *N gonorrhoeae* isolates in our primary dataset. This finding was expected, as high prevalence within globally circulating strains is a key feature of an ideal vaccine target.<sup>28</sup>

To maximise safety, we removed isolates with genotypic or phenotypic resistance profiles. However, applying a selection strategy that only included isolates with no genotypic or phenotypic resistance substantially restricted the number of candidate isolates, compromising the global relevance of the final strain panel. Accordingly, we prioritised clinically relevant genotypic or phenotypic resistance (eg, by removing isolates with phenotypic resistance to ceftriaxone, azithromycin, ciprofloxacin, and tetracycline, but keeping isolates that were less susceptible to penicillin). Despite this approach, the majority of our candidate challenge strains formed a single clade within the global *N gonorrhoeae* phylogeny, with most isolates belonging to lineage B (which is associated with heterosexual transmission networks and antimicrobial susceptibility), and only one isolate belonging to lineage A (which is associated with high-risk sexual networks and multidrug resistance).<sup>23</sup> This finding shows the difficulty of optimising safety but ensuring generalisability in challenge strain selection for an organism such as *N gonorrhoeae* with a high prevalence of AMR. Furthermore, this finding shows how our genomics-based strain selection approach can be rapidly adapted and modified to select bacterial strains that suit the evolving goals of CHIM strain selection over time.

We have selected a subset of potential challenge strains for phenotypic assessment, which will include detailed phenotypic characterisation, growth kinetics, infectivity, serum sensitivity, and amenability to manufacturing processes. Although our shortlisting process aimed to include a

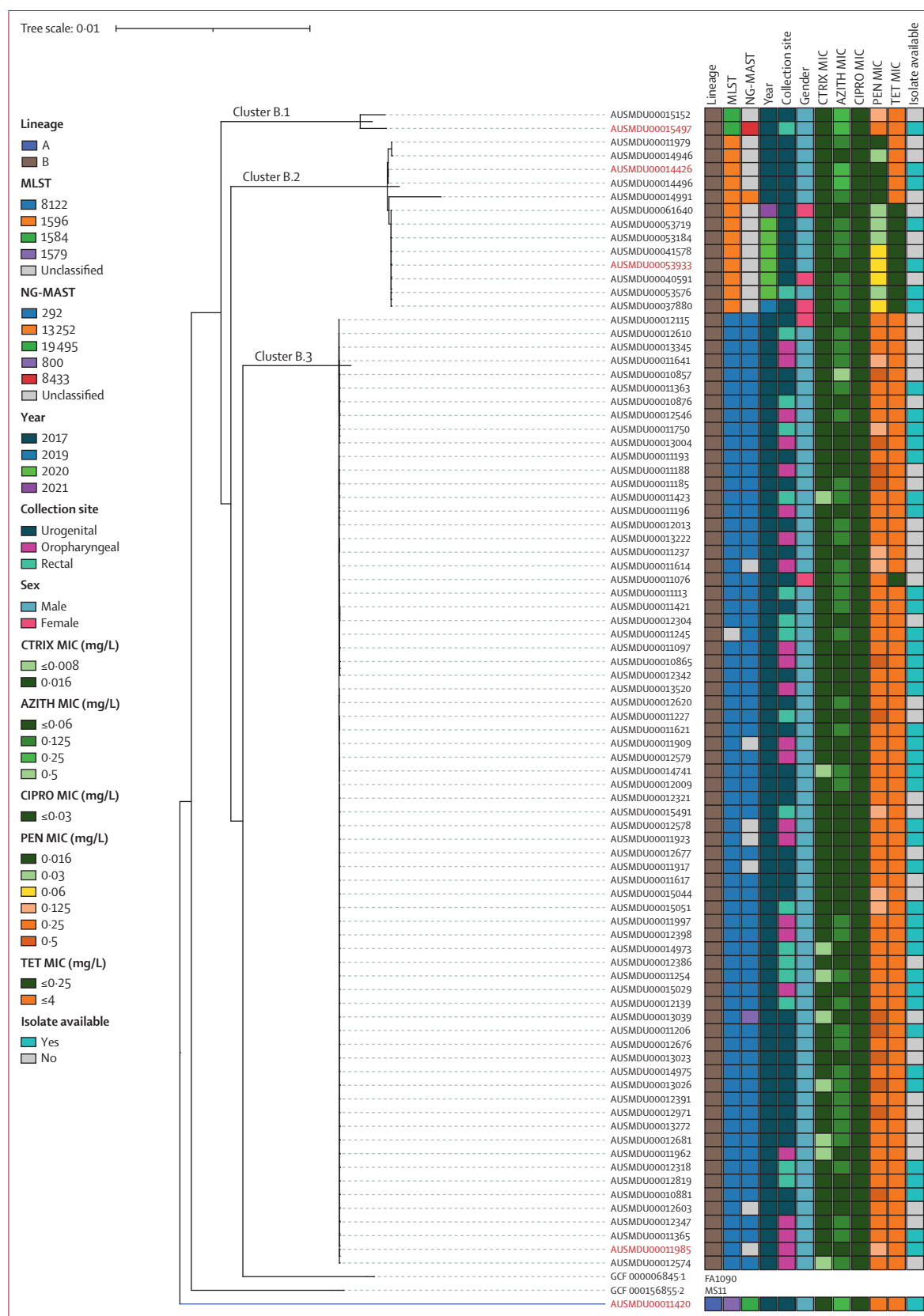
diversity of challenge strain candidates for further characterisation, an inherent limitation of CHIM studies is that they are restricted to the use of a single or very small number of challenge strains. The diversity of *N gonorrhoeae* precludes identification of a single representative challenge strain. However, through our systematic approach to challenge strain selection, we aim to identify an isolate that is amenable to oropharyngeal cell infectivity, does not have phenotypic characteristics associated with disseminated infection (eg, serum sensitivity), and remains stable in cell bank manufacturing conditions.

This genomic study has several limitations. First, the publicly available genomic data we used to contextualise our data are inherently biased. In many genomic studies, genomes are generated in the context of research performed for unrelated purposes and isolates might therefore be missing key metadata. For example, most genomic sequences did not have a description of the anatomical site of isolation. This missing information might have resulted in an underestimation of MLSTs and NG-MASTs associated with oropharyngeal and urogenital infection. Second, there was an over-representation of genomes from high-income countries. The generalisability of the challenge strains selected with our method might be reduced for geographical regions with poorer representation, such as the eastern Mediterranean. Finally, the data included in our primary dataset were only obtained from culture isolates. *N gonorrhoeae* infections diagnosed with only nucleic acid amplification testing were not included, which might represent a notable selection bias. In Australia in 2020, 76% of *N gonorrhoeae* infections were diagnosed with nucleic acid amplification testing alone.<sup>29</sup>

In summary, through the application of a genomics-based strain selection method, we have successfully identified a shortlist of potential *N gonorrhoeae* challenge strains. From this shortlist, the five isolates described in this study will proceed to detailed phenotypic assessment and assessment of stability in cell bank manufacturing conditions. Through this process, a contemporary *N gonorrhoeae* challenge strain will be developed for implementation in a novel oropharyngeal *N gonorrhoeae* CHIM. Leveraging a wealth of publicly available genomic data, our approach provides a template for investigators to rationally select an optimal panel of potential CHIM challenge strains for further phenotypic characterisation from similar datasets for other pathogens. This approach substantially reduces the time and resources required to select a contemporary challenge strain and improves the transparency of challenge strain selection. It also allows investigators to undertake an iterative approach to strain selection, balancing ideal strain characteristics with the

**Figure 2: Sequence types and global phylogenetic context of the 86 challenge strain candidates**

(A) Distribution of MLST sequence types for the 86 candidate challenge strains. (B) Distribution of NG-MAST sequence types for the 86 candidate challenge strains. (C) Phylogeny of 6686 *Neisseria gonorrhoeae* genomes, comprising 86 shortlisted genomes from the primary dataset and 6600 subsampled genomes from the global dataset. Annotation rings include country or continent, MLST, and NG-MAST sequence types, and the 86 isolates that were shortlisted from the primary dataset with the key selection criteria. Only MLST and NG-MAST sequence types with more than 150 and 50 representatives, respectively, are included in the legend. The positions of strains FA1090 and MS11, which were included in a controlled human infection model for *Neisseria gonorrhoeae* urethritis developed in the 1980s,<sup>34</sup> are indicated on the outermost ring (green lines). The scale bar is in units of nucleotide substitutions per site. MLST=multilocus sequence type. NG-MAST=Neisseria gonorrhoeae multiantigen sequence type.



**Figure 3: Annotated subtree of the 86 challenge strain candidates and strains FA1090 and MS11**

The 86 potential candidate strains were pruned from figure 1. Strains FA1090 and MS11 were included in a controlled human infection model for *Neisseria gonorrhoeae* urethritis developed in the 1980s.<sup>14</sup> The 5 strains shortlisted for phenotypic assessment are highlighted in red. Metadata are annotated. The last column indicates the availability of the isolate for phenotypic characterisation. The scale bar is in units of nucleotide substitutions per site. AZITH=azithromycin. CIPRO=ciprofloxacin. CTRIX=ceftriaxone. MIC=minimum inhibitory concentration. MLST=multilocus sequence type. NG-MAST=*Neisseria gonorrhoeae* multiantigen sequence type. PEN=penicillin. TET=tetracycline.

features of real-world isolates, and enables rapid adaptation to suit the evolving goals of CHIMs.

#### Contributors

EW was involved in the conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration and writing of the original draft. SJL was involved in the conceptualisation, data curation, formal analysis, investigation, methodology, data visualisation and writing of the original draft. GLP was involved in the investigation and methodology. MLT was involved in the data curation and investigation. JP was involved in the project administration and funding acquisition. BPH was involved in administration and resources. EPFC and CKF were involved in the investigation and resources. KLS was involved in the methodology. JSM was involved in the conceptualisation, funding acquisition, methodology, resources, and supervision. SP was involved in the conceptualisation, investigation, methodology, and supervision. DAW was involved in the conceptualisation, funding acquisition, investigation, methodology, resources, data visualisation, and supervision. All authors contributed to reviewing and editing the manuscript. EW, SJL, and DAW accessed and verified the data. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

#### Declaration of interests

We declare no competing interests.

#### Data sharing

Accessions for the *Neisseria gonorrhoeae* sequences in the primary dataset and associated metadata can be found in appendix 2 (p 1). Further data supporting the findings of this study are available in appendix 1 and appendix 2. Short-read sequence data for all isolates in the primary dataset analysed in this study are available from the National Centre for Biotechnology Information database under BioProject PRJNA520805.

#### Acknowledgments

We would like to thank the scientific staff of the Microbiological Diagnostics Unit Public Health Laboratory (University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia) who performed antimicrobial susceptibility testing and whole-genome sequencing for this study. This project was supported by a Medical Research Future Fund Clinician Researchers Applied Research in Health Grant (grant number MRFAR000354). EW is supported by a postgraduate scholarship from the Australian National Health and Medical Research Council (NHMRC; grant number GNT2005380). MLT is supported by an Australian Government Research Training Program scholarship. BPH is supported by an NHMRC Leadership Investigator Grant (grant number GNT1196103). EPFC is supported by an NHMRC Emerging Leadership Investigator Grant (grant number GNT1172873). CKF is supported by an NHMRC Leadership Investigator Grant (grant number GNT1172900). KLS is supported by an NHMRC Leadership Investigator Grant (grant number GNT2017383). JSM was supported by an NHMRC Leadership Investigator Grant (grant number GNT2016396). DAW is supported by an NHMRC Investigator Grant (grant number GNT1174555).

#### References

- Rowley J, Vander Hoorn S, Korenromp E, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bull World Health Organ* 2019; **97**: 548–62.
- Unemo M, Seifert HS, Hook EW 3rd, Hawkes S, Ndowa F, Dillon JR. Gonorrhoea. *Nat Rev Dis Primers* 2019; **5**: 79.
- Vallely LM, Egli-Gany D, Wand H, et al. Adverse pregnancy and neonatal outcomes associated with *Neisseria gonorrhoeae*: systematic review and meta-analysis. *Sex Transm Infect* 2021; **97**: 104–11.
- Unemo M, Shafer WM. Antimicrobial resistance in *Neisseria gonorrhoeae* in the 21st century: past, evolution, and future. *Clin Microbiol Rev* 2014; **27**: 587–613.
- WHO. Global health sector strategy on sexually transmitted infections 2016–2021: towards ending STIs. Geneva: World Health Organization, 2016.
- Williams E, Seib KL, Fairley CK, et al. *Neisseria gonorrhoeae* vaccines: a contemporary overview. *Clin Microbiol Rev* 2024; **37**: e0009423.
- WHO. Bacterial vaccines in clinical and preclinical development: an overview and analysis. Geneva: World Health Organization, 2022.
- Taylor SN, Marrazzo J, Batteiger BE, et al. Single-dose zoliflodacin (ETX0914) for treatment of urogenital gonorrhea. *N Engl J Med* 2018; **379**: 1835–45.
- Taylor SN, Morris DH, Avery AK, et al. Gepotidacin for the treatment of uncomplicated urogenital gonorrhea: a phase 2, randomized, dose-ranging, single-oral dose evaluation. *Clin Infect Dis* 2018; **67**: 504–12.
- Fairley CK, Cornelisse VJ, Hocking JS, Chow EPF. Models of gonorrhoea transmission from the mouth and saliva. *Lancet Infect Dis* 2019; **19**: e360–66.
- Lewis DA. Will targeting oropharyngeal gonorrhoea delay the further emergence of drug-resistant *Neisseria gonorrhoeae* strains? *Sex Transm Infect* 2015; **91**: 234–37.
- Hui BB, Padaniya TN, Rebuli N, et al. A gonococcal vaccine has the potential to rapidly reduce the incidence of *Neisseria gonorrhoeae* infection among urban men who have sex with men. *J Infect Dis* 2022; **225**: 983–93.
- Abo YN, Jamrozik E, McCarthy JS, Roestenberg M, Steer AC, Osowski J. Strategic and scientific contributions of human challenge trials for vaccine development: facts versus fantasy. *Lancet Infect Dis* 2023; **23**: e533–46.
- Hobbs MM, Sparling PF, Cohen MS, Shafer WM, Deal CD, Jerse AE. Experimental gonococcal infection in male volunteers: cumulative experience with *Neisseria gonorrhoeae* strains FA1090 and MS11mkC. *Front Microbiol* 2011; **2**: 123.
- Osowski J, Azzopardi KI, McIntyre L, et al. A controlled human infection model of group A *Streptococcus* pharyngitis: which strain and why? *mSphere* 2019; **4**: e00647-18.
- Waddington CS, Darton TC, Jones C, et al. An outpatient, ambulant-design, controlled human infection model using escalating doses of *Salmonella* Typhi challenge delivered in sodium bicarbonate solution. *Clin Infect Dis* 2014; **58**: 1230–40.
- Graham DY, Opekun AR, Osato MS, et al. Challenge model for *Helicobacter pylori* infection in human volunteers. *Gut* 2004; **53**: 1235–43.
- Poly F, Read TD, Chen YH, et al. Characterization of two *Campylobacter jejuni* strains for use in volunteer experimental-infection studies. *Infect Immun* 2008; **76**: 5655–67.
- Golparian D, Harris SR, Sánchez-Busó L, et al. Genomic evolution of *Neisseria gonorrhoeae* since the preantibiotic era (1928–2013): antimicrobial use/misuse selects for resistance and drives evolution. *BMC Genomics* 2020; **21**: 116.
- Parliament of Victoria. Public Health and Wellbeing Act 2008. Parliament of Victoria. <https://www.legislation.vic.gov.au/in-force/acts/public-health-and-wellbeing-act-2008/067> (accessed Jan 11, 2024).
- Darton TC, Blohmke CJ, Moorthy VS, et al. Design, recruitment, and microbiological considerations in human challenge studies. *Lancet Infect Dis* 2015; **15**: 840–51.
- La C, Bell A, Trouvin JH, et al. Consideration on the principles of development and manufacturing qualities of challenge agents for use in human infection models. London: Wellcome Trust, 2022.
- Sánchez-Busó L, Golparian D, Corander J, et al. The impact of antimicrobials on gonococcal evolution. *Nat Microbiol* 2019; **4**: 1941–50.
- Molina JM, Bercot B, Assoumou L, et al. Doxycycline prophylaxis and meningococcal group B vaccine to prevent bacterial sexually transmitted infections in France (ANRS 174 DOXYVAC): a multicentre, open-label, randomised trial with a 2 × 2 factorial design. *Lancet Infect Dis* 2024; **24**: 1093–104.
- Eyre DW, De Silva D, Cole K, et al. WGS to predict antibiotic MICs for *Neisseria gonorrhoeae*. *J Antimicrob Chemother* 2017; **72**: 1937–47.
- Gardener M. Community ecology. London: Pelagic Publishing, 2014.
- Russell MW, Jerse AE, Gray-Owen SD. Progress toward a gonococcal vaccine: the way forward. *Front Immunol* 2019; **10**: 2417.
- Edwards JL, Jennings MP, Seib KL. *Neisseria gonorrhoeae* vaccine development: hope on the horizon? *Curr Opin Infect Dis* 2018; **31**: 246–50.
- Australian Government Department of Health and Aged Care. Communicable Diseases Intelligence: Australian gonococcal surveillance programme annual report. Australian Government Department of Health and Aged Care, 2022. [https://www1.health.gov.au/internet/main/publishing.nsf/Content/2A15CD097063EF40CA2587CE008354F1/\\$File/australian\\_gonococcal\\_surveillance\\_programme\\_annual\\_report\\_2021.pdf](https://www1.health.gov.au/internet/main/publishing.nsf/Content/2A15CD097063EF40CA2587CE008354F1/$File/australian_gonococcal_surveillance_programme_annual_report_2021.pdf) (accessed Jan 11, 2024).