SUPPLEMENTARY METHODS

DNA extraction and whole genome sequencing

DNA extraction and WGS were performed at the Microbiological Diagnostic Unit Public Health Laboratory at the University of Melbourne. Genomic DNA was isolated from a single colony using a QIAsymphony SP platform (Qiagen). WGS was performed using the Illumina NextSeq platform with 150 base-pair paired-end reads. All sequencing reads are available on the National Center for Biotechnology's Sequence Read Archive under BioProject PRJNA449254 (Table S1).

Bioinformatic analysis

Sequences were analysed using the Nullarbor pipeline (Seeman T, available at: https://github.com/tseemann/nullarbor). In brief, reads were trimmed to remove adaptor sequences and low-quality bases with Trimmomatic,[1] and Kraken (v0.10.5-beta) was used to investigate for contamination.[2] Reads were aligned to the NCCP11945reference genome [3] using the Burrows-Wheeler Aligner MEM (v 0.7.15-r1140) algorithm.[4] Samples with at least 45x depth of coverage and 85% genome coverage were retained for analysis. SNPs were identified using Freebayes (v1.0.2) with a minimum depth of coverage of 10x and allelic frequency of 0.9 required to confidently call a single nucleotide polymorphism (SNP).[5] For all isolates, sequence reads covered >89% of the reference genome, A total of 12,202 core SNP sites were identified, with 1809 genes identified as core genes (present in >99% of isolates).

De novo assemblies were also performed using SPAdes (v.3.9.0),[6] as part of the Nullarbor pipeline, with genes annotated using Prokka (v.1.12-beta, Seemann T, available at: https://github.com/tseemann/prokka). NG-MAST and MLST results were inferred in silico with NGMASTER,[7] and mlst (https://github.com/tseemann/mlst) respectively.

Blocks of recombination were detected using Gubbins (v.2.2.0-1),[8] 10 with the core full alignment provided as input. The final recombination-masked fasta file was used to produce a maximum likelihood tree in RaXML (version 8).[9]

REFERENCES

1. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**: 2114–2120.
2. Wood D, Salzberg, SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014; **15**: R46.
3. Chung GT, Yoo JS, Oh HB, et al. Complete genome sequence of *Neisseria gonorrhoeae* NCCP11945. *J Bacteriol* 2008; **190**(17): 6035–6.
4. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWAMEM. *arXiv* 2013; 1303.3997v1302.
5. Garrison E, Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv* 2012; 1207.3907v1202.
6. Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–77.
7. Kwong JC, Goncalves da Silva A, Dyet K, et al. NGMASTER: in silico multi-antigen sequence typing for Neisseria gonorrhoeae. *Microbial Genomics* 2016; **2**: e000076.
8. Croucher NJ, Page AJ, Connor TR, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015; **43**: e15.
9. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014; **30**(9): 1312–3.