Miguel Cascais Ferreira Pinto

Mestre em Biologia Humana e Ambiente

Genome-scale approaches to strengthen *Neisseria gonorrhoeae* epidemiological and antimicrobial resistance surveillance

Dissertação para obtenção do Grau de Doutor em Biologia

Orientador: Doutor João Paulo dos Santos Gomes, Investigador Auxiliar com Habilitação, Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P.

Co-orientador: Doutora Maria José Gonçalves Gaspar Borrego, Investigadora Auxiliar, Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P.

Co-orientador: Doutor Luís Jaime Gomes Ferreira da Silva Mota, Professor Auxiliar, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Júri:

Presidente:	Professora Doutora Isabel Maria Godinho de Sá Nogueira
Arguentes:	Doutor Mário Nuno Ramos d'Almeida Ramirez
	Doutora Raquel Sá Leão Domingues da Silva
Vogais:	Doutor João Paulo dos Santos Gomes
	Doutora Marta Aires-de-Sousa

Novembro, 2020

Miguel Cascais Ferreira Pinto

Mestre em Biologia Humana e Ambiente

Genome-scale approaches to strengthen *Neisseria gonorrhoeae* epidemiological and antimicrobial resistance surveillance

Copyright © **Miguel Pinto**

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

As secções desta dissertação já publicadas por editores para os quais foram transferidos direitos de cópia pelos autores, encontram-se devidamente identificadas ao longo da dissertação e são reproduzidas sob permissão dos editores originais e sujeitas às restrições de cópia impostas pelos mesmos.

Para a Andreia e para o Tiago

"I wish it need not have happened in my time," said Frodo.
"So do I," said Gandalf, "and so do all who live to see such times.
But that is not for them to decide.
All we have to decide is what to do with the time that is given us."
J.R.R. Tolkien
The Lord of the Rings – The Fellowship of the Ring

Acknowledgements

Firstly, I would like to thank Doctor João Paulo Gomes, for inviting me to collaborate and be a part of his research group, for all his teachings and help, for his friendship and availability, for defending my best interests, and finally for the freedom and trust he gave and continuously gives me in the pursuit of science and scientific research. Particularly, I would like to acknowledge his ability to recognize the human side of science (and its inherent struggles) and for making me feel part of a team. Since day one you have deposited a large amount of faith and trust in my abilities and I hope that in the end I have lived up to the expectations. Also, thank you for allowing me to take care of my family during this last year.

I would like to thank Doctor Maria José Borrego, for all her help and teachings in the field of sexually transmitted bacterial infections, for her friendship and availability, for always defending my best interests and for continuously worrying about my future and well-being. Some gestures will never be forgotten and I will always be grateful to you for pushing me.

Thanks to Doctor Jaime Mota and Doctor João Carriço for their support during the elaboration of this thesis.

Secondly, I would like to thank the director of the National Institute of Health (Instituto Nacional de Saúde, Doutor Ricardo Jorge, I.P.) and the coordinator of the Department of Infectious Diseases Doctor Jorge Machado, for allowing me the opportunity to perform this work within the premises of the institute.

I would also like to acknowledge the Portuguese Foundation for Science and Technology for supporting this study (Grant no. SFRH/BD/109264/2015).

Thirdly, I would like to thank Vítor Borges for his invaluable help and support, for his friendship, for all the great scientific and personal conversations, for his insights and constant availability (especially when it was clear that you were not available). Since your PhD thesis defense, you have raised the bar and have been a continuous source of aspiration and inspiration. Working with you in *Treponema pallidum* genomics has been to this day the best scientific experience I have ever had, and not because of the output, but because of the undertaking of a true collaborative process.

I would like to thank Joana Isidro for all her help and support, for her friendship, for the personal and scientific conversations, for the cooperative (although brief) drawing sessions, for her effort in sequencing most samples in the present work and for trusting me with her early bioinformatics training. Your ability to quickly learn and progress beyond my own abilities has

been (and still is) a source of inspiration for me to keep challenging myself and improve technically and scientifically.

I would like to thank Alexandra Nunes for her help and support throughout these years as well as Rita Ferreira and Mínia Antelo for their support during the initial stages of this endeavour.

I would like to thank Doctor Luís Vieira and his team of the Technology and Innovation Unit of the Department of Human Genetics of INSA, namely Catarina Cordeiro, Joana Mendonça, Dina Carpinteiro, Daniel Sampaio and Miguel Machado, with a special thanks to Sílvia Duarte.

I would like to thank everyone that works/worked in, or in direct collaboration with, the NRL-STI lab, namely Dora Cordeiro, João Carlos Rodrigues, Lúcia Reis, Cristina Correia, Ivone Água-Doce, Rui Matias and Ana Isabel Salas.

I would like to thank all the Master students that worked in the lab throughout these years: Zohra Lodhia, Ana Rita Caldeira, Aleksandra Azevedo, Marta Ferreira and Nádia Albuquerque.

I would like to thank everyone that entrusted me with their work, through fruitful collaborations that enriched my expertise, and those who made working at INSA a more cheerful environment, namely: Rita Macedo, Leonor Silveira, Andrea Santos, Raquel Rocha, Paula Bajanca-Lavado, Elizabeth Padua, Raquel Sabino, Aida González-Díaz, Maria João Simões, Célia Betencourt, Ana Pelerito and Rita Cordeiro.

A special thank you to Paulo Costa for the insightful scientific conversations, but mostly for his friendship and support during complicated times. In addition, I would like to thank Mariana Lopes Correia, André Grilo, Joana Paredes and Filipa Perreira for their support and all the unwinding board game nights.

Finally, I would like to thank my entire family for their support. To my beautiful wife Andreia for her constant support and encouragement, for being my rock, for believing in me and making me believe in myself (especially in the most desperate times). Throughout this endeavour, you got "stuck" with the ugliest side of things, so thank you for not wavering and putting up with my own frustrations. To my son Tiago, thank you for being such an incredible little human being and for all the amazing things you do. The best part of this last year was being able to spend so much time with you, to watch you grow, and to hear your amazing laughter which makes everything else seem smaller.

As a last note, I would like to acknowledge Doctor Maria João Silva, who many years ago, when times seemed darkest, provided me with the opportunity to rediscover my passion for science.

Resumo

Neisseria gonorrhoeae é o agente patogénico humano que causa a doença sexualmente transmissível gonorreia, constituindo um problema de saúde publica. Esta bactéria tem capacidade de adquirir resistência antimicrobiana (RAM) a múltiplas classes de antibióticos. Sem vacina, o controlo da gonorreia exige medidas preventivas e tratamentos antimicrobianos eficazes, assim como vigilância epidemiológica. Programas nacionais e internacionais estão a promover progressivamente a aplicação da sequenciação total do genoma (WGS) para monitorizar a circulação de *N. gonorrhoeae* e a emergência e propagação de RAM.

O principal objetivo desta dissertação foi o de contribuir para o fortalecimento da vigilância epidemiológica e de RAM em *N. gonorrhoeae* recorrendo a WGS. As principais tendências de RAM em Portugal ao longo de 16 anos foram identificadas, beneficiando dos dados colhidos pela Rede Nacional de Vigilância Laboratorial de Estirpes de *Neisseria gonorrhoeae* (PTGonoNET), sediada no Instituto Nacional de Saúde, Doutor Ricardo Jorge, I.P. (INSA). Através de dados de WGS Europeus, reportamos uma classificação de *N. gonorrhoeae* em genogrupos, os quais representam as suas principais linhagens em circulação, correlacionando-se com outras técnicas de tipagem e com perfis específicos de RAM. Aplicando uma abordagem gene-a-gene, efetuou-se o primeiro estudo à escala do genoma para *N. gonorrhoeae* em Portugal, revelando-se a diversidade genética das estirpes em circulação, assim como possíveis cadeias de transmissão, de forma a promover e/ou suportar a investigação epidemiológica. Finalmente, avaliou-se um método independente da cultura para obter dados de WGS diretamente de amostras clínicas e a sua aplicação para a vigilância epidemiológica e na deteção de RAM.

Os resultados apresentados nesta tese constituem um ponto de viragem para a consolidação da epidemiologia genómica dos gonococos em Portugal com a implementação de uma metodologia de vigilância baseada em WGS no Laboratório de Referência no INSA. Por ultimo, este trabalho contribui para o fortalecimento da vigilância de *N. gonorrhoeae* promovendo a monotorização da frequência e propagação geográfica de genogrupos e contribuindo para ações de Saúde Pública mais orientadas para controlar a propagação de RAM de *N. gonorrhoeae*.

Palavras chave: *Neisseria gonorrhoeae*; Resistência antimicrobiana; Vigilância; Sequenciação total do genoma

Abstract

Neisseria gonorrhoeae is the human pathogen responsible for the sexually transmitted disease gonorrhoea, whose burden remains a major public health concern. This bacterium has shown an extraordinary ability to develop antimicrobial resistance (AMR) to multiple classes of antimicrobials, with the advent of reaching a "superbug" status. With no available vaccine, managing gonorrhoea infections demands effective preventive measures, antibiotic treatments and epidemiological surveillance. National and international surveillance programmes are increasingly promoting the application of whole-genome sequencing (WGS) data to track *N. gonorrhoeae* circulation and the emergence and spread of AMR.

The major goal of the PhD dissertation was to strengthen *N. gonorrhoeae* epidemiological and AMR surveillance using WGS. Particularly, we disclose the major AMR trends observed in Portugal throughout 16 years, by reporting data from the National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNet), hosted at the Portuguese National Institute of Health (NIH). Using WGS data from across Europe, we report a comprehensive WGS-based genogroup assignment for *N. gonorrhoeae*. These genogroups represent main circulating lineages and were correlated with other typing techniques and linked to specific AMR signatures. Using a dynamic gene-by-gene approach, we performed the first genome-scale study of *N. gonorrhoeae* in Portugal, highlighting the genetic diversity of circulating strains, as well as potential transmission chains, which is essential to support epidemiological investigation. Finally, we evaluated a culture-independent strategy to obtain WGS data directly from clinical samples and its suitability for epidemiological surveillance and AMR detection.

The findings presented in this dissertation constituted a turning point to consolidate the genomic epidemiology of gonococci in Portugal through the implementation of a WGS-based surveillance methodology in the Portuguese NIH. Ultimately, this work enhances *N. gonorrhoeae* surveillance by promoting the prospective monitoring of genogroup frequency and geographic spread, towards more oriented Public Health actions to control the spread of *N. gonorrhoeae* AMR.

Keywords: *Neisseria gonorrhoeae*; Antimicrobial resistance; Surveillance; Whole-genome sequencing

Table of contents

Acknowledgements iii
Resumo v
Abstract vii
Table of contents ix
Figure Index xiii
Table Index xv
List of Abbreviations xvii
Author's notes: thesis organization, format and outline........ xxi

1. Chapter I – General introduction 1
1.1. The genus *Neisseria* 3
1.2. Gonorrhoea origin and history 3
1.3. *Neisseria gonorrhoeae* pathogenesis and impact in human health 4
1.3.1. Disease, pathogenesis and adaptation 4
1.3.2. Laboratory diagnosis and antimicrobial susceptibility testing 7
1.3.3. Antibiotic treatment therapy for gonorrhoea 8
1.3.4. Global epidemiology and the Portuguese panorama 10
1.3.5. *Neisseria gonorrhoeae* antimicrobial resistance burden 11
1.4. *Neisseria gonorrhoeae* genomics 13
1.4.1. General genomic features 13
1.4.2. Phase variation and major virulence factors 14
1.4.3. Genomic mechanisms of antimicrobial resistance 15
1.4.4. Traditional molecular typing 17
1.4.5. *Neisseria gonorrhoeae* in the whole-genome sequencing era 20
1.5. Aims and general research plan 22

2. Chapter II – Fifteen years of a nationwide culture collection of *Neisseria gonorrhoeae* antimicrobial resistance in Portugal 23
2.1. Abstract 25
2.2. Introduction 25
2.3. Methods 26
2.3.1. PTGonoNET isolate collection 26
2.3.2. Antimicrobial susceptibility testing 27
2.3.3. Isolate typing 27
2.3.4. Data availability and interactive visualization 28
2.3.5. Limitations of this study 28
2.4. Results 28
2.4.1. Demographic distribution of *Neisseria gonorrhoeae* isolates 28
2.4.2. Antibiotic resistance within the culture collection 30
2.4.3. Genetic diversity of *Neisseria gonorrhoeae* isolates 33
2.5. Discussion 34

3. Chapter III – *Neisseria gonorrhoeae* clustering to reveal major European WGS-based genogroups in association with antimicrobial resistance ………… 39

3.1. Abstract ……………………………………………………………………….. 41

3.2. Introduction …………………………………………………………………… 42

3.3. Methods ……………………………………………………………………….. 43

3.3.1. Selection and characterization of Portuguese isolates …………………... 43

3.3.2. Additional dataset for the European context and genome assembly ……. 44

3.3.3. *In silico* typing and antimicrobial resistance prediction ………………… 44

3.3.4. Gene-by-gene analysis …………………………………………………... 45

3.3.5. WGS-based genogroup classification …………………………………… 45

3.3.6. *Neisseria gonorrhoeae* genogroups and genetic clustering analysis ……. 46

3.3.7. Data availability ………………………………………………………… 47

3.4. Results ………………………………………………………………………… 47

3.4.1. *Neisseria gonorrhoeae* WGS-based genogroups ……………………….. 47

3.4.2. From traditional to WGS-based typing ………………………………….. 51

3.4.3. Distribution of *Neisseria gonorrhoeae* genogroups by European countries 54

3.4.4. WGS-based genogroups carrying antibiotic resistance determinants …... 57

3.4.5. Association between *Neisseria gonorrhoeae* genetic clustering and epidemiological data ……………………………………………………….. 61

3.5. Discussion …………………………………………………………………….. 63

4. Chapter IV – Insights into the *Neisseria gonorrhoeae* pan-genome: a preliminary assessment of genogroup-specific genetic signatures …………….. 69

4.1. Introduction …………………………………………………………………… 71

4.2. Methods ……………………………………………………………………….. 71

4.2.1. Construction of a draft *Neisseria gonorrhoeae* pan-genome ……………. 71

4.2.2. Scoring the presence and absence of accessory genes …………………… 72

4.3. Results ………………………………………………………………………… 72

4.3.1. Overview of the *Neisseria gonorrhoeae* pan-genome …………………… 72

4.3.2. Association between accessory genes and WGS-based genogroups …….. 74

4.4. Discussion …………………………………………………………………….. 76

5. Chapter V – First case of a cephalosporin resistant *Neisseria gonorrhoeae* isolate detected in Portugal ……………………………………………………… 79

5.1. Abstract ……………………………………………………………………….. 81

5.2. Introduction …………………………………………………………………… 81

5.3. Case description ……………………………………………………………… 81

5.4. Characterization of *Neisseria gonorrhoeae* isolate NGPT19535 …………….. 82

5.5. Discussion …………………………………………………………………….. 84

6. Chapter VI – *Neisseria gonorrhoeae* laboratory surveillance in portugal in the whole-genome sequencing era …………………………………………………… 87

6.1. Introduction …………………………………………………………………… 89

6.2. Methods ……………………………………………………………………….. 90

6.2.1. Portuguese *Neisseria gonorrhoeae* genomic dataset ……………………. 90

6.2.2. Whole-genome sequencing, genome assembly and typing …………….. 91
6.2.3. cgMLST and cluster analysis ……………………………………………. 91
6.2.4. Data availability …………………………………………………………. 92
6.3. Results …………………………………………………………………………. 92
6.3.1. Description of the PTGonoNet 2019 collection ……………………….. 92
6.3.2. Description of the Portuguese genomic dataset ………………………… 93
6.3.3. Overview of *Neisseria gonorrhoeae* genetic diversity circulating in Portugal from 2013 to 2019 ……………………………………………… 94
6.3.4. Antimicrobial resistance within *Neisseria gonorrhoeae* genogroups predominant in Portugal ………………………………………………… 97
6.3.5. Identification of potential epidemiologically linked isolates …………… 98
6.3.6. Highlights of *Neisseria gonorrhoeae* isolates needing close monitoring in Portugal ……………………………………………………………….. 105
6.4. Discussion ……………………………………………………………………….. 106

7. Chapter VII – Novel approach to capture the genome of *Neisseria gonorrhoeae* directly from clinical samples ………………………………………………. 109
7.1. Introduction ……………………………………………………………………. 111
7.2. Material and methods ………………………………………………………… 113
7.2.1. *Neisseria gonorrhoeae*-positive clinical samples ……………………….. 113
7.2.2. Generation of standard curves for real-time quantitative PCR ………... 114
7.2.3. qPCR for quantification of *Neisseria gonorrhoeae* and human cells …… 114
7.2.4. RNA “bait” design and synthesis for DNA capture …………………….. 115
7.2.5. SureSelect$^{XT\ HS}$ *Neisseria gonorrhoeae* enrichment and WGS directly from clinical samples ………………………………………………. 115
7.2.6. *Neisseria gonorrhoeae* read filtering and genome assembly …………… 115
7.2.7. Isolate typing, antimicrobial resistance prediction and phylogenetic analysis ……………………………………………………………….. 117
7.3. Results 117
7.3.1. WGS of *Neisseria gonorrhoeae* directly from clinical samples ………….. 117
7.3.2. Global strain typing and antimicrobial resistance profile ……………….. 120
7.3.3. Phylogenetic analysis ……………………………………………………. 123
7.4. Discussion ……………………………………………………………………….. 125

8. Chapter VIII – Final overview, concluding remarks and future directions …... 129
8.1. Final discussion ……………………………………………………………….. 131
8.2. Future perspectives …………………………………………………………..... 134

References ……………………………………………………………………………….. 135
Supplementary Material ……………………………………………………………….. 163

Figure index

Figure 1.1. | Depiction of typical and atypical anatomical sites infected by *Neisseria gonorrhoeae* and associated clinical manifestations …………………….. 5

Figure 1.2. | Overview of the different steps involved in *Neisseria gonorrhoeae* infection ……………………………………………………………….. 6

Figure 1.3. | The history of *Neisseria gonorrhoeae* antimicrobial resistance ………….. 9

Figure 1.4. | Number of notified cases of gonorrhoea per year in Portugal, from 1950 to 2018 ……………………………………………………………………… 11

Figure 1.5. | Current *Neisseria gonorrhoeae*-targeting antimicrobials and the corresponding resistance genes ………………………………………….. 15

Figure 2.1. | Map of Portugal exhibiting the geographic origin of the *Neisseria gonorrhoeae* isolates of the Portuguese National Institute of Health culture collection from 2003-2018 ………………………………………. 29

Figure 2.2. | Distribution of the *Neisseria gonorrhoeae* isolates by minimum inhibitory concentrations for all tested antimicrobials ……………………………… 31

Figure 2.3. | MIC_{50} and MIC_{90} value progression from 2010 up to 2018, for ceftriaxone, cefixime and azithromycin ……………………………………………….. 32

Figure 2.4. | *Neisseria gonorrhoeae* antibiotic susceptibility trends from 2003 up to 2018 …………………………………………………………………… 33

Figure 2.5. | Minimum spanning tree of 1042 concatenated and aligned sequences of *Neisseria gonorrhoeae* Multi-Antigen Sequence Types ………………… 34

Figure 3.1. | Relationship between cluster stability (assessed by Neighbourhood Adjusted Wallace Coefficient) and goeBURST threshold based on the MScgMLST scheme for all 3791 *Neisseria gonorrhoeae* isolates analysed 48

Figure 3.2. | Distribution of the percentage of pairwise allelic differences between isolates of all low-level WGS-based genogroups represented by at least 10 isolates ………………………………………………………………... 51

Figure 3.3. | Comparison of WGS-based genogroups, defined at two levels, and both traditional typing methods for *Neisseria gonorrhoeae* …………………… 52

Figure 3.4. | Cluster congruence (Adjusted Rand Index) and agreement (Adjusted Wallace Coefficient) between traditional MLST and NG-MAST typing and goeBURST clustering at all thresholds (based on the MScgMLST scheme) ………………………………………………………………... 53

Figure 3.5. | Overlap between *Neisseria gonorrhoeae* WGS-based genogroups defined in this study and PubMLST core-genome groups ………………………... 54

Figure 3.6. | Phylogeny of 3791 *Neisseria gonorrhoeae* isolates from Europe, based on a gene-by-gene approach using the MScgMLST scheme …………..... 55

Figure 3.7. | Distribution of low-level WGS-based genogroups by country ………….. 56

Figure 3.8. | Heatmap distribution and occurrence of the genetic determinants involved in antimicrobial resistance by high- and low-level WGS-based genogroups ……………………………………………………………… 58

Figure 3.9. | Major predicted antimicrobial resistance profiles observed within each low-level genogroup …………………………………………………….. 59

Figure 3.10. | Analysis of *Neisseria gonorrhoeae* WGS-based genetic clusters at low resolution level potentially concordant with epidemiological link ………. 62

Figure 4.1. | *Neisseria gonorrhoeae* pan- and core-genomes ………………………….. 73

Figure 4.2. | Functional diversity of core- and accessory-genomes of *Neisseria gonorrhoeae* ……………………………………………………………... 73

Figure 4.3. | Distribution of accessory-genome genes presenting more than 80% specificity and sensitivity across WGS-based genogroups ……………… 75

Figure 6.1. | Antimicrobial susceptibility trends of the 2019 PTGonoNet isolate collection ……………………………………………………………… 93

Figure 6.2. | Comparison of antimicrobial susceptibility data of the Portuguese genomic dataset and the PTGonoNet Collection from 2003 to 2019 …….. 94

Figure 6.3. | Phylogenetic distribution of novel *Neisseria gonorrhoeae* genomes from Portugal ……………………………………………………………….. 96

Figure 6.4. | Relative frequency by year of all top ten predominant genogroups observed in our PT genomic dataset ……………………………………… 97

Figure 6.5. | Phylogenetic distribution of *Neisseria gonorrhoeae* genomic clusters that include isolates from Portugal ……………………………………………. 99

Figure 6.6. | Distribution of genetic clusters enrolling Portuguese isolates by geographic region and time interval ……………………………………… 100

Figure 6.7. | *Neisseria gonorrhoeae* clusters composed by European isolates and more than five Portuguese isolates ……………………………………………….. 101

Figure 6.8. | *Neisseria gonorrhoeae* clusters exclusively composed by at least five Portuguese isolates ……………………………………………………… 103

Figure 6.9. | *Neisseria gonorrhoeae* clusters exclusively composed by three or four Portuguese isolates ……………………………………………………… 104

Figure 7.1. | Enrichment success versus the *Neisseria gonorrhoeae* and human genome copies in the input DNA ………………………………………………….. 118

Figure 7.2. | Taxonomic classification within *Neisseria* species of raw and filtered reads generated for all clinical samples …………………………………... 119

Figure 7.3. | Phylogenetic distribution of novel validated *Neisseria gonorrhoeae* genomes captured directly from clinical samples ………………………… 120

Figure 7.4. | *Neisseria gonorrhoeae* clusters enrolling genomes captured directly from clinical samples ……………………………………………………………... 124

Table index

Table 1.1. | Different treatment guidelines for gonorrhoea worldwide 10

Table 1.2. | Main antimicrobial resistance determinants described in *Neisseria gonorrhoeae* for previous and currently recommended antimicrobials for treatment of gonorrhoea 16

Table 1.3. | Most prevalent *Neisseria gonorrhoeae* Multi-antigen Sequence Typing (NG-MAST) sequence types worldwide by study 19

Table 2.1. | Distribution of *Neisseria gonorrhoeae* isolates for each collection year according to patient gender, sexual orientation, geographic region and site of infection 30

Table 3.1. | Summary description of the high-level WGS-based genogroups enrolling at least ten isolates 49

Table 3.2. | Summary description of the low-level WGS-based genogroups enrolling at least ten isolates 50

Table 5.1. | Antimicrobial susceptibility and genetic characterization of *Neisseria gonorrhoeae* strain NGPT19535, isolated in Portugal in 2019 83

Table 7.1. | Distribution of *Neisseria gonorrhoeae*-positive clinical samples available at INSA 112

Table 7.2. | Description of all clinical samples enrolled in the present study 113

Table 7.3. | Assembly statistics for the first and final draft assemblies 116

Table 7.4. | Strain typing statistics for all *Neisseria gonorrhoeae* genomes captured from clinical samples 121

Table 7.5. | Summary of all in silico detected antimicrobial resistance determinants ... 122

Supplementary Table S2.1. | Associated patient data, molecular typing and antibiotic susceptibility data of all viable *Neisseria gonorrhoeae* isolates collected from 2003 up to 2018 in Portugal on behalf of the PTGonoNET.......... 163

Supplementary Table S3.1. | Metadata, genome assembly statistics and antibiotic resistance phenotype and genotype data of the Portuguese *Neisseria gonorrhoeae* isolates enrolled in the present study.......................... 163

Supplementary Table S3.2. | Metadata, genome assembly statistics and antibiotic resistance genotype data of the publicly available European *Neisseria gonorrhoeae* isolates enrolled in the present study 163

Supplementary Table S3.3. | List of loci removed from the pubMLST *Neisseria gonorrhoeae* cgMLST v1 scheme, after preparation for chewBBACA..... 163

Supplementary Table S3.4. | List of loci enrolled in the Maximum Shared cgMLST scheme used for clustering analysis.. 163

Supplementary Table S4.1. | *Neisseria gonorrhoeae* pan-genome composition based on 3971 genomes 163

Supplementary Table S4.2. | Accessory genes associated with high- and low-level WGS-based genogroups, presenting simultaneously more than 80% sensitivity and specificity 163

Supplementary Table S4.3. | Distribution of accessory genes presenting simultaneously more than 80% sensitivity and specificity by high- and low-level WGS-genogroups ………………………………………….. 164

Supplementary Table S6.1. | Metadata, genome assembly statistics and antibiotic resistance phenotype and genotype data of the novel sequenced Portuguese *Neisseria gonorrhoeae* isolates enrolled in the present study .. 164

Supplementary Table S6.2. | Associated patient data and antibiotic susceptibility data of all viable *Neisseria gonorrhoeae* isolates collected in 2019 in Portugal on behalf of the PTGonoNET …………………………………………. 164

Supplementary Table S6.3. | Composition of genetic clusters enrolling Portugusese isolates ………………………………………………………………… 164

Supplementary Table S7.1. | Summary of the taxonomic classification of all raw reads for all clinical samples ………………………………………….. 164

Supplementary Table S7.2. | Summary of the taxonomic classification of all recovered reads (after mapping) for all clinical samples ………………… 164

List of abbreviations

ABC – ATP-binding cassette

AD – Allelic differences

AGSP – Australien Gonococcal Surveillance Program

AMR – Antimicrobial resistance

AR – Autonomous Region

ARI – Adjusted Rand Index

ATP – Adenosine triphosphate

AWC – Adjusted Wallace Coefficient

AZM – Azithromycin

BLAST – Basic Local Alignment Search Tool

bp – Base pairs

BURST – Based upon related sequence types

CDC – Centers for Disease Control and Prevention

CDS – Coding sequence

CEF – cephalosporin

CFM – Cefixime

cgMLST – Core-genome Multi-locus Sequence Typing

CIP – Ciprofloxacin

CLSI – Clinical and Laboratory Standards Institute

COG – Clusters of Orthologous Groups

CRO – Ceftriaxone

CTX – Cefotaxime

DC – Dendritic cell

DGI – Disseminated gonococcal infection

DGS – Portuguese General Health Directorate (Direção Geral da Saúde)

DHPS – Dihydropteroate synthase

DNA – Deoxyribonucleic acid

DOI – Digital object identifier

dRS3 – Duplicated repeat sequence 3

DUS – DNA uptake sequence

dXXX – Decreased susceptibility to

ECDC – European Centre for Disease Prevention and Control

ECOFF – Epidemiologic cut-off

EEA – European Economic Area

ENA – European Nucleotide Archive

EQA – External quality assessment

ESC – Extended-spectrum cephalosporin

EU – European Union

EUCAST – European Committee for Antimicrobial Susceptibility Testing

EURO-GASP – European Gonococcal Antimicrobial Surveillance Programme

FHA – Forkhead-associated

GEN – Gentamicin

GGI – Gonococcal genetic island

GISP – Gonococcal Isolate Surveillance Project

goeBURST – Global optimal eBURST

GRASP – Gonococcal Resistance to Antimicrobials Surveillance Programme

HBP – Heptose-1,7-bisphosphate

HGT – Horizontal gene transfer

HIV – Human immunodeficiency virus

IM – Inner membrane

INSA – Portuguese National Institute of Health (Instituto Nacional de Saúde, Doutor Ricardo Jorge, I.P.)

LOS – Lipooligosaccharide

MA – Metropolitan Area

Mbp – Mega base pairs

MIC – Minimum inhibitory concentration

MLST – Multi-locus Sequence Typing

MScgMLST – Maximum shared core-genome Multi-Locus Sequence Typing

MSM – Men who have sex with men

MSMW – Men who have sex with men and women

MST – Minimum spanning tree

NA – Not available

NAAT – Nucleic acid amplification technologies

nAWC – Neighbourhood Adjusted Wallace Coefficient

ND – Not determined

NEQAS – National External Quality Assessment Service

NG-MAST – *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing

NGS – Next-Generation Sequencing

NG-STAR – *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance

NIH – National Institute of Health

NLR – NOD-like receptor

NOD – Nucleotide-binding oligomerization domain-containing protein

NRL-STI – National Reference Laboratory for Sexually Transmitted Infections

NUT – Nomenclature of Territorial Units for Statistics

OM – Outer membrane

OMV – Outer membrane vesicles

Opa – Opacity

PBP – Penicillin-binding protein

PCR – Polymerase chain reaction

PEN – Penicillin

PHAC – Public Health Agency of Canada

PHE – Public Health England

PID – Pelvic inflammatory disease

PMN – Polymorphonuclear neutrophils

PT – Portuguese

PTGonoNET – National Laboratory Network for *Neisseria gonorrhoeae* Collection

PubMLST – Public databases for molecular typing and microbial genome diversity

QA – Quality assurance

QC – Quality control

qPCR – Real-time quantitative PCR

RAM – Resistência antimicrobiana

RI – Rifampicin

RNA – Ribonucleic acid

RPSBLAST – Reverse Position-Specific BLAST

rRNA – Ribosomal ribonucleic acid

rXXX – Resistant to

SI – Shannon Index

SINAVE – Sistema Nacional de Vigilância Epidemiológica

SNP – Single nucleotide polymorphism

SPT – Spectinomycin

ST – Sequence type

STD – Sexually transmitted disease

STI – Sexually transmitted infection

SUL – Sulfonamide

T4SS – Type IV secretion system

TET – Tetracycline

TIFA – TRAF-interacting protein with FHA domain-containing protein A

TLR – Toll-like receptor

TNF – Tumor necrosis factor

TRAF – TNF Receptor associated factors

UK – United Kingdom

USA – United States of America

wgMLST – Whole-genome Multi-locus Sequence Typing

WGS – Whole-genome sequencing

WGSCL – Whole-genome sequencing cluster

WHO – World Health Organization

XDR – Extensive drug resistant

Author's Notes: thesis organization, format and outline

The main body of this PhD dissertation is composed of eight chapters, including a general introduction of the subject, six research studies (either published or ongoing) and a final discussion. Its core is based on three published manuscripts (listed below) that are presented as individual chapters and three additional chapters that include proofs of concept of methodologies or exploratory/ongoing studies. As three chapters correspond to published manuscripts in peer reviewed international journals, they faithfully reproduce what was published. In this context, the chapter presentation order neither perfectly reflects the chronological order of the manuscript publication nor the time in which all studies were performed, as several of these were developed simultaneously and the time elapsing between the manuscript submission and publication largely depends on the journal and on the inherent revision requirements. As such, chapters were organized following a rational order taking into account the objectives delineated for this PhD work, and are preceded by a title page describing the reference of the publication (when applicable) and the personal contributions of the author of the present PhD thesis. In brief, each chapter includes the following contents:

Chapter I – General introduction

This chapter consists of a general introduction that intends to provide the reader with the state of the art in the subjects addressed in this doctoral dissertation focusing on the sexually transmitted pathogen *Neisseria gonorrhoeae*. It includes a historical context, a global overview of the major aspects of *N. gonorrhoeae* taxonomy, biology, pathogenesis, impact on human health and molecular epidemiology, followed by insights into the genetic features and antimicrobial resistance mechanisms of this bacterium and the contribution of the whole-genome sequencing era in this field. It ends with the description of the main objectives of this PhD project, which include the specific research questions that drove the investigations carried out on behalf of each chapter.

Chapter II – Fifteen years of a nationwide culture collection of *Neisseria gonorrhoeae* antimicrobial resistance in Portugal

This chapter corresponds to the following published manuscript: "*Pinto M, Rodrigues JC, Matias R, Água-Doce I, Cordeiro D, Correia C, Gomes JP, Borrego MJ and the PTGonoNet. (2020) Fifteen years of a nationwide culture collection of Neisseria gonorrhoeae antimicrobial resistance in Portugal. European Journal of Clinical Microbiology and Infectious Diseases; 39:1761–1770*". It reports all antimicrobial resistance data generated up to 2018 on behalf of the National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNet), hosted and supported by the National Reference Laboratory for Sexually

Transmitted Infections of the Portuguese National Institute of Health. The ultimate goal of this study was to disclose the overall *Neisseria gonorrhoeae* antimicrobial resistance trends in Portugal throughout fifteen years of surveillance.

Chapter III – *Neisseria gonorrhoeae* clustering to reveal major European WGS-based genogroups in association with antimicrobial resistance

This chapter corresponds to the following published manuscript: "*Pinto M, Borges V, Isidro J, Rodrigues JC, Vieira L, Borrego MJ, Gomes JP. (2020) Neisseria gonorrhoeae clustering to reveal major European WGS-based genogroups in association with antimicrobial resistance. Microbial Genomics (in press)*". The aim of this study was to perform *Neisseria gonorrhoeae* clustering into genogroups based on whole-genome sequencing data, for enhanced prospective laboratory surveillance. Particularly, it identifies the major circulating whole-genome sequencing-based genogroups in Europe and establishes a relationship between these and antimicrobial resistance.

Chapter IV – Insights into the *Neisseria gonorrhoeae* pan-genome: a preliminary assessment of genogroup-specific genetic signatures

This chapter corresponds to an exploratory study that takes advantage of the data generated in the previous chapter and complements it by analysing the pan-genome of *Neisseria gonorrhoeae*. Particularly, it explores *Neisseria gonorrhoeae*'s accessory-genome in order to associate the presence of specific genetic markers with the previously established whole-genome sequencing-based genogroups.

Chapter V – First case of a cephalosporin resistant *Neisseria gonorrhoeae* isolate detected in Portugal

This chapter corresponds to the following published manuscript: "*Pinto M, Matias R, Rodrigues JC, Duarte S, Vieira L, Gonçalves I, Gonçalves MJ, Ramos MH, Gomes JP, Borrego MJ. (2020) Cephalosporin-Resistant Neisseria gonorrhoeae Isolated in Portugal, 2019. Sexually Transmitted Diseases; 47(11):e54–e56*". It reports the first case detected in Portugal of a cephalosporin resistant *Neisseria gonorrhoeae* isolate, in 2019. This chapter corresponds to a timely analysis performed in the context of Public Health needs, as it was important to rapidly report the existence of an isolate with such resistance profile in circulation in our country. As such, in terms of organization, this work was placed between Chapter III and Chapter VI, as it took advantage of the dataset generated in Chapter III to perform the analyses, with this isolate being integrated in the global Portuguese genomic study (Chapter VI) only after the manuscript publication.

Chapter VI – *Neisseria gonorrhoeae* laboratory surveillance in Portugal in the whole-genome sequencing era

This chapter corresponds to the first genome-scale study of *Neisseria gonorrhoeae* in Portugal, enrolling isolates collected from 2003 up to 2019. As such, it includes an in-depth analysis of all isolates sequenced in Chapter III and V, as well as novel isolates that were sequenced *a posteriori* to constitute a genomic sub-set representing the year 2019. Additionally, as Chapter II is presented as it was published at the time, the description of the all PTGonoNET isolates collected during 2019 were included in the present chapter as an update. In this context, this chapter applies the described methodology in Chapter III to all sequenced Portuguese *Neisseria gonorrhoeae* isolates to date, in order to disclose potential transmission chains occurring in Portugal. This work constituted a turning point to consolidate the genomic epidemiology of gonococci in Portugal that will be carried forward at the Portuguese National Reference Laboratory for Sexually Transmitted Infections of the Portuguese National Institute of Health.

Chapter VII – Novel approach to capture the genome of *Neisseria gonorrhoeae* directly from clinical samples

This chapter corresponds to an ongoing study that evaluates the application of a culture-independent strategy to obtain whole-genome sequencing data directly from clinical samples collected from distinct anatomic sites. Particularly, it aimed to bypass the culture step prior to whole-genome sequencing and assess its suitability for epidemiological surveillance and antimicrobial resistance detection.

Chapter VIII – Final overview, concluding remarks and future directions

This chapter provides a global overview of the subjects addressed throughout the chapters, highlighting the main results and conclusions achieved in this PhD dissertation. As each chapter contains its own section discussing the obtained results, only the most relevant findings are addressed here in order to avoid unnecessary redundancy. New research questions that can be addressed in future follow-up studies, raised by this work, are also highlighted.

Considering the dissimilar layouts and in-text reference styles adopted by the different scientific journals where the manuscripts were published, all chapters were formatted in a unique style. In this regard, a single section of "References" is presented, with all references being cited by sequential numbers and listed according to the order in which they appear in the dissertation. Additionally, published supplementary figures were included in the main text to facilitate reading.

Data availability

Finally, annexes relative to each chapter are also compiled in a last section of this thesis (referred to as "Supplementary Material"). As most annexes corresponded to large Tables with extensive information that could not be legibly formatted in the text, for the sake of simplicity these data are made available digitally through an online publication (Figshare_Project), with hyperlinks to these provided in this Section.

All *Neisseria gonorrhoeae* raw read data generated in this study were published in the European Nucleotide Archive (ENA) under BioProject accession no. PRJEB36482 (www.ebi.ac.uk/ena/data/view/PRJEB36482), with all individual ENA sample accession numbers detailed in the Supplementary Material. All Next-Generation Sequencing runs were performed by the Technology and Innovation Unit of the Department of Human Genetics of the National Institute of Health, Dr. Ricardo Jorge, I.P. In order to potentiate visualization, interpretation and reproducibility of data, additional supplementary support materials for the study performed in Chapter III were published in ZENODO and are available at https://doi.org/10.5281/zenodo.3946223. Likewise, for Chapters II and VI, two Projects were created in the MicroReact web-platform, and made publicly available in order to analyse dynamically antimicrobial resistance and molecular data presented in these studies by year and geographic region:

- https://microreact.org/project/4U7YJ-70k/240ea921;
- https://microreact.org/project/q8vwip8qPLJAhRwz3RedFE/ec303d1f.

This work was supported the Portuguese Foundation for Science and Technology under the individual Grant no. **SFRH/BD/109264/2015**.

CHAPTER I

General introduction

1. General introduction

1.1. The genus *Neisseria*

The genus *Neisseria* belongs to the family *Neisseriaceae* of the order *Neisseriales* [1], which is placed in the class *Betaproteobacteria*. Currently the genus *Neisseria* consists of 35 Gram negative species (according to the List of Prokaryotic names with Standing in Nomenclature [2]) that in general grow optimally under aerobic conditions at temperatures from 35 to 37 °C. Most are cocci with a diameter of up to 2 µm, presenting as single bacteria or in pairs (diplococci) [3]. Nevertheless, novel *Neisseria* species are still being characterized [4]. The members of this genus have a high affinity for the mucous membranes of animals and humans [3], but can also be found as free-living organisms in the environment [5]. Several species can be isolated from humans, namely *Neisseria gonorrhoeae*, *N. meningitidis*, *N. cinerea*, *N. elongata*, *N. flavescens*, *N. lactamica*, *N. mucosa*, *N. polysaccharea*, *N. sicca*, and *N. subflava*, while in animals the most predominant species are *N. animalis*, *N. animaloris*, *N. zoodegmatis* (cats and dogs), *N. denitrificans* (guinea pigs), *N. dentiae* (domestic cows), *N. macacae* (rhesus monkeys), and *N. weaveri* (dogs) [3, 5]. While most human *Neisseria* species normally inhabit the upper respiratory tract, only causing disease in an opportunistic fashion, in immune-compromised hosts, or systemic infections [5], species of animal origin are not associated with human disease and have only rarely caused infections in human wounds after animal bites [6]. The species *N. gonorrhoeae* and *N. meningitidis* are considered to be exclusive human pathogens. While the latter is a commensal of the human oropharynx, capable of causing acute disease in healthy individuals, *N. gonorrhoeae* is always considered a transmitted pathogen. Still, *N. meningitidis* is the only species that expresses a polysaccharide capsule, of which thirteen different serogroups can be distinguishable [7]. Contrasting with many of the species in the *Neisseria* genus that are not nutritionally demanding, *N. gonorrhoeae* and *N. meningitidis* are fastidious bacteria, showing particular susceptibility to unfavourable environmental factors, such as extreme temperatures, desiccation, and alkaline or acidic conditions. Finally, all species in this genus are naturally competent for DNA uptake and display a high frequency of horizontal gene transfer (HGT) [8].

1.2. Gonorrhoea origin and history

Sexually transmitted diseases (STD) are known to have existed from antiquity and are recorded in many of humanity's earliest surviving texts, suggesting that sexually transmitted infections (STI) such as syphilis, gonorrhoea, chlamydiosis and herpes have been a constant bane of human health. Gonorrhoea is considered an old malady or, as George Luys famously put it, "gonorrhoea is as old as the world" [9]. Its history is long (reviewed in [10]), with its causative

agent, the obligate human pathogen *N. gonorrhoeae* (gonococcus), having coexisted with humanity and caused this STD for a long time. Evidence suggests that gonorrhoea existed in early human cultures, as seen referenced in ancient Egyptian texts believed to date back as far as 3000 BC [11], in the Bible [12], in descriptions from physicians from Ancient Greece to the Roman Empire which carried on to the Middle Ages and were further elaborated in the Renaissance [13]. Indeed, the first mention of the word “gonorrhoea” is attributed to Aelius Galen (129–216 A.D.), a Greek physician, surgeon and philosopher of the Roman Empire [13], while the discovery of the gonococcus came much later, credited to the German microbiologist Albert Ludwig Siegmund Neisser (1855–1916) [14], to whom the bacteria *Neisseria gonorrhoeae* owes its name [15, 16]. From then on, gonococcal biology would rapidly be unveiled, from its consistent isolation to its association with clinical manifestations and, in the twentieth century, gonorrhoea was well accepted as an STI caused by a singular human pathogen. Culture provided the grounds for microbiology and molecular biology based studies on *N. gonorrhoeae*, which contributed to disclose the structure of this pathogen and its virulence factors, while immunobiology allowed the understanding of the interactions with its human host, its innate and adaptive immune responses and to develop effective treatments and prophylactic measures. Still, sulfanilamide (belonging to the sulfonamide functional group), the first drug used to reliably cure gonorrhoea was only introduced in 1938 [17], with gonococcal resistance to this drug and to others rapidly emerging [18–20], signalling the beginning of an enduring arms race between antimicrobials and the gonococcus (discussed further in another section of this Chapter). Finally, in 2000, the genomic era brought the release of the first complete genome of a *N. gonorrhoeae* strain (FA 1090, Genbank accession number NC_002946), with thousands of genomes being sequenced since then (discussed further in another section of this Chapter). This new era has challenged our understanding of gonorrhoea, not only suggesting that the origin of modern gonococci populations could lie around the sixteenth century, but also that it has been continuously shaped by antimicrobial treatment [21]. Nevertheless, with all the knowledge gained until now, the development of an effective vaccine for this disease remains to be accomplished.

1.3. *Neisseria gonorrhoeae* pathogenesis and impact in human health

1.3.1. Disease, pathogenesis and adaptation

N. gonorrhoeae is generally transmitted from infected individuals during sexual activity, by direct contact between the mucosal membranes of the urogenital tract, anal canal, and oropharynx. Occasionally it also causes infections in the eye, as well as other anatomic sites (**Figure 1.1.**) [20, 22, 23]. Gonorrhoea is associated with significant morbidity with the most common results for urogenital gonorrhoea being urethritis in men and cervicitis in women. While

most rectal infections and almost all pharyngeal infections are asymptomatic, it has been considered that the vast majority of urethral infections in men (about 90%) present with discharge or dysuria. In contrast, in women urethral and cervical infections are more commonly asymptomatic, as less than 50% develop symptoms [20, 24]. Although in cases of "uncomplicated cervicitis" women do not report symptoms, upon direct examination of the cervix they often manifest physical signs of inflammation [25, 26].

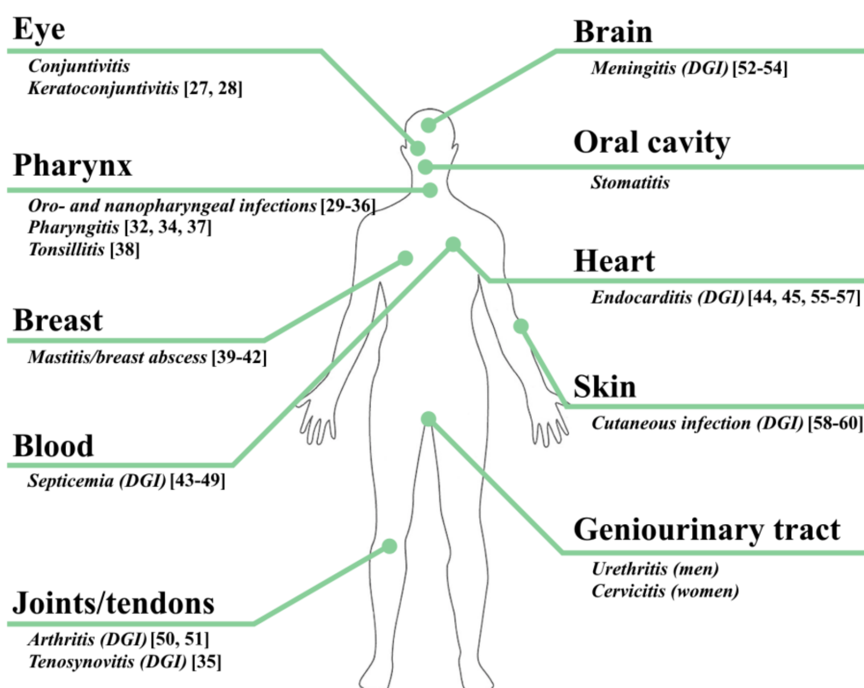

Figure 1.1. | Depiction of typical and atypical anatomical sites infected by *Neisseria gonorrhoeae* and associated clinical manifestations. Corresponding references refer to clinical case reports of unusual infections. Many of the unusual gonococcal infections are either associated with preceding Disseminated Gonococcal Infection (DGI) or considered the cause of subsequent gonococcal septicemia and/or other manifestations of DGI. (Adapted from [23])

Gonorrhoea infections are usually treatable by antibiotic therapy and most individuals who develop symptoms are no longer infectious after seeking medical attention. Nevertheless, it is estimated that about two-thirds of all infected men are asymptomatic (mainly due to oropharyngeal and anoretal infections) and the major source of spread [22]. Moreover, if left untreated, *N. gonorrhoeae* can cause ascending infections resulting in epididymo-orchitis, salpingitis and pelvic inflammatory disease (PID), by transluminal dissemination starting from the urogenital mucosae [61, 62]. These have a large impact on the female reproductive system as they may result in tubal factor infertility, ectopic pregnancy and chronic pelvic pain. Additionally, these can lead to adverse pregnancy outcomes, like low birth weight infants, small for gestational age infants, conjunctivitis (ophthalmia neonatorum) and oropharyngeal infections through transmission to newborns [49, 63]. In rare cases, in Disseminated Gonococcal Infection (DGI), *N. gonorrhoeae* can enter the bloodstream and spread systemically causing severe complications

(**Figure 1.1.**) like fever/septicemia [43–48], arthritis [50, 51, 64], tenosynovitis [35], endocarditis [44, 45, 55–57] or meningitis [52–54].

As an obligate human colonizer, *N. gonorrhoeae* has developed unique features that are crucial to its pathogenesis and adaptation to its host (reviewed in [65]). Its infection process is summarized in **Figure 1.2**. *N. gonorrhoeae*'s survival depends on a successful transmission between hosts, as this pathogen cannot survive *ex-vivo*, unlike *N. meningitidis* which can survive for short periods of time and spread via respiratory droplets [65]. After transmission, *N. gonorrhoeae* adheres, colonizes and invades the mucosal epithelium of the lower and upper genitourinary tract, which is mediated by bacterial surface structures such as type IV pili, opacity (Opa) proteins, lipooligosaccharide (LOS) and the major outer membrane (OM) protein porin PorB. Type IV pili and Opa proteins are believed to be essential for colonization and immune evasion, the latter through antigenic and phase variation mechanisms [66–69].

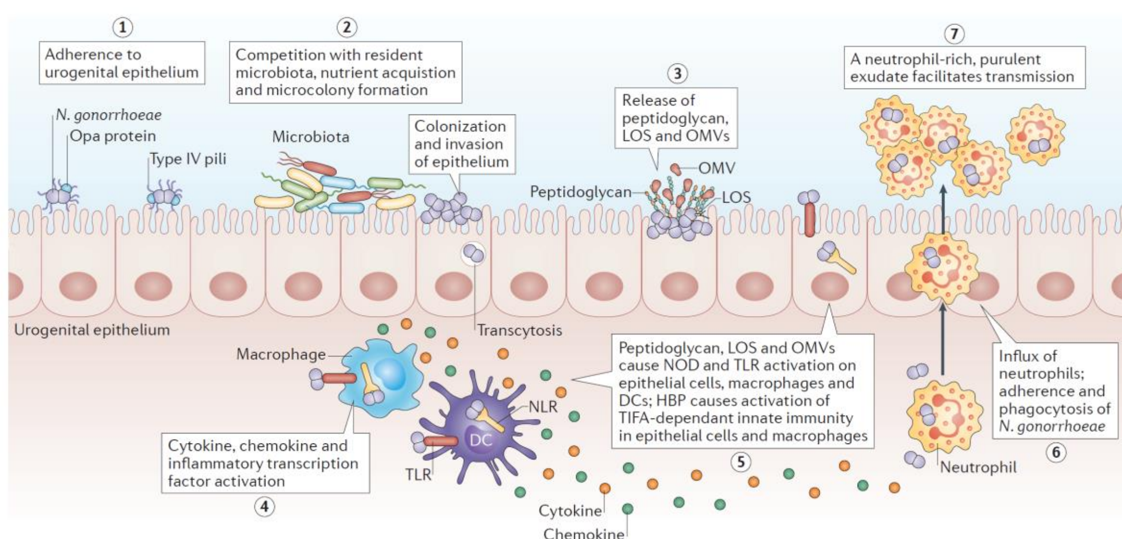

Figure 1.2. | Overview of the different steps involved in *Neisseria gonorrhoeae* infection. LOS- lipooligosaccharide; OMV – Outer membrane vesicles; NOD – Nucleotide-binding oligomerization domain-containing protein; TLR – Toll-like receptor; NLR – NOD-like receptor; DC – Dendritic cell; HBP – Heptose-1,7-bisphosphate; TIFA – TRAF-interacting protein with FHA domain-containing protein A. Opa – opacity. [65].

LOS play an important role in adherence and invasion of the host's cells, as variation in these affect the immune cell recognition, and its sialylation affects serum resistance through complement evasion [70–72]. Once it has adhered and colonized, *N. gonorrhoeae* is also capable of invasion and transcytosis of the epithelium which could lead to DGI [22, 65]. During the initial stages of infection, *N. gonorrhoeae* starts to release fragments of peptidoglycan, LOS and OM vesicles that trigger the innate immune response, which results in a localized influx of polymorphonuclear neutrophils (PMNs), macrophages and dendritic cells [73–76] that interact with and phagocytose the bacteria (**Figure 1.2.**). It is then the interaction and subsequent influx

of PMNs that makes up a purulent exudate that facilitates transmission [77]. Still, the bacterial and host factors that contribute to more asymptomatic infections are poorly understood. More sensitive surveillance, screening and diagnostics have been considered necessary to characterize bacterial strains and host cell factors that contribute to asymptomatic colonization [65].

1.3.2. Laboratory diagnosis and antimicrobial susceptibility testing

As it is considered that many gonorrhoea symptoms are not specific of the disease, an accurate laboratory diagnosis should be performed on symptomatic individuals including, if possible, antimicrobial susceptibility testing for *N. gonorrhoeae* positive cases. Routine screening of individuals that are part of risk groups is also recommended, such as sex workers or HIV-positive individuals. Routine laboratory diagnosis of *N. gonorrhoeae* can be performed in urogenital, anorectal, pharyngeal or conjunctival swab specimens, first-catch urine, blood, ascitic or synovial fluids [24]. Currently, there are several methods available for *N. gonorrhoeae* detection, the most commonly used being culture and nucleic acid amplification technologies (NAAT). The use of DNA probe assays, antigen tests and serology to detect antibodies are not recommended due to a lack of sensitivity and specificity [78]. Several studies have shown that NAAT offer higher sensitivity than culture [78–84], mainly due to the fact that bacteria do not need to be viable and that it can be performed with success on diverse specimen types, including non-invasive biological samples (e.g. urine) [78, 80, 85]. In addition, NAAT provide faster results than culture, with less hands-on time, have the possibility of being scalable and automated for high throughput testing [86–88], making them cheaper, and can be designed to test multiple pathogens simultaneously, which is the case currently for *N. gonorrhoeae, Chlamydia trachomatis, Mycoplasma genitalium* and *Trichomonas vaginalis* [78, 88]. In light of this, culture is being progressively replaced as a laboratory diagnosis method. So far, the major disadvantage of NAAT-based laboratory diagnosis is that it cannot provide information about antimicrobial susceptibility, which still requires bacterial culture. The latter is performed on rich culture medium such as GC base or Chocolate Agar plus PolyViteX, which provide the necessary growth factors required by this fastidious bacterium, with plates required to be incubated at 35-37 ºC, in a high humidity and 4-6% CO_2 enriched atmosphere [89]. After 18 to 24h of incubation, the gonococci appear as shiny grey colonies, which can be further identified by microscopy after Gram's staining, together with a positive reaction to cytochrome oxidase [24]. Although bacterial culture can be sensitive, confirmation of *Neisseria* species can be further performed through biochemical [90], immunological [90, 91], spectrometric [92, 93] or molecular tests [82, 83].

Furthermore, for *N. gonorrhoeae*, antimicrobial susceptibility testing is one of the most important procedures to perform, as it can be crucial for establishing an effective antimicrobial therapy, due to this pathogen's remarkable ability to acquire resistance. Antimicrobial

susceptibility testing indicates the minimum inhibitory concentration (MIC) (i.e., concentration that inhibits growth) of a given antimicrobial agent, and for *N. gonorrhoeae* this is accomplished via two main options: agar dilution method or the MIC gradient strip method (E-test) [24]. Interpretation of these assays should be based on breakpoint standards either following the guidelines from the Clinical and Laboratory Standards Institute (CLSI) (https://clsi.org/) or the European Committee for Antimicrobial Susceptibility Testing (EUCAST) [94], which divide MIC breakpoints into "susceptible", "intermediate", and "resistant" categories. Quality control assessments are required regardless of the test method, either inter- or intra-laboratory, using the World Health Organization (WHO) control strains [95, 96]. In fact, while the WHO recommends the agar dilution method, the mostly used method should be the MIC gradient strip test, as it is easier to perform and correlates well with the agar dilution methods [97, 98]. Additionally, as some *N. gonorrhoeae* strains exhibit high-level resistance to penicillin, associated with carriage of a plasmid with the β-lactamase encoding gene, the Nitrocefin test should also be performed in order to assess β-lactamase production [24].

1.3.3. Antibiotic treatment therapy for gonorrhoea

In the absence of a vaccine, gonorrhoea treatment has always relied on an effective antibiotic therapy. Several antibiotics have been historically used and since fallen out of use for the treatment of *N. gonorrhoeae* infections (**Figure 1.3.**). In fact, gonococci have shown the ability to acquire antimicrobial resistance (AMR) to multiple classes of antibiotics (the genetic mechanisms underlying resistance and its epidemiology will be discussed in a latter section of this Chapter), including β-lactams, tetracyclines, macrolides and quinolones, reaching dramatic levels. This is due to the combination of the remarkable genetic adaptation capability of *N. gonorrhoeae* and an indiscriminate use of a wide range of antibiotics to treat gonorrhoea [19, 65, 99–101]. Indeed, within ten years of the introduction of sulfonamides, *N. gonorrhoeae* became resistant to this class of antimicrobials and its use was no longer recommended. Then, penicillin became the drug of choice but only until the mid-1970s due to the reports of rapid decrease of susceptibility. Tetracyclines rapidly met the same fate as penicillins, and resistance to fluoroquinolones is currently so widespread globally that they can no longer be recommended as a first choice for the empiric treatment of gonorrhoea. Macrolides seemed to be the answer but resistance was shown to be rapidly selected [19, 100, 101]. Finally, third generation cephalosporins were introduced in the 1980s and 1990s to treat gonorrhoea, but resistance [102–104] and clinical failures have also been reported since 2010 [105–108].

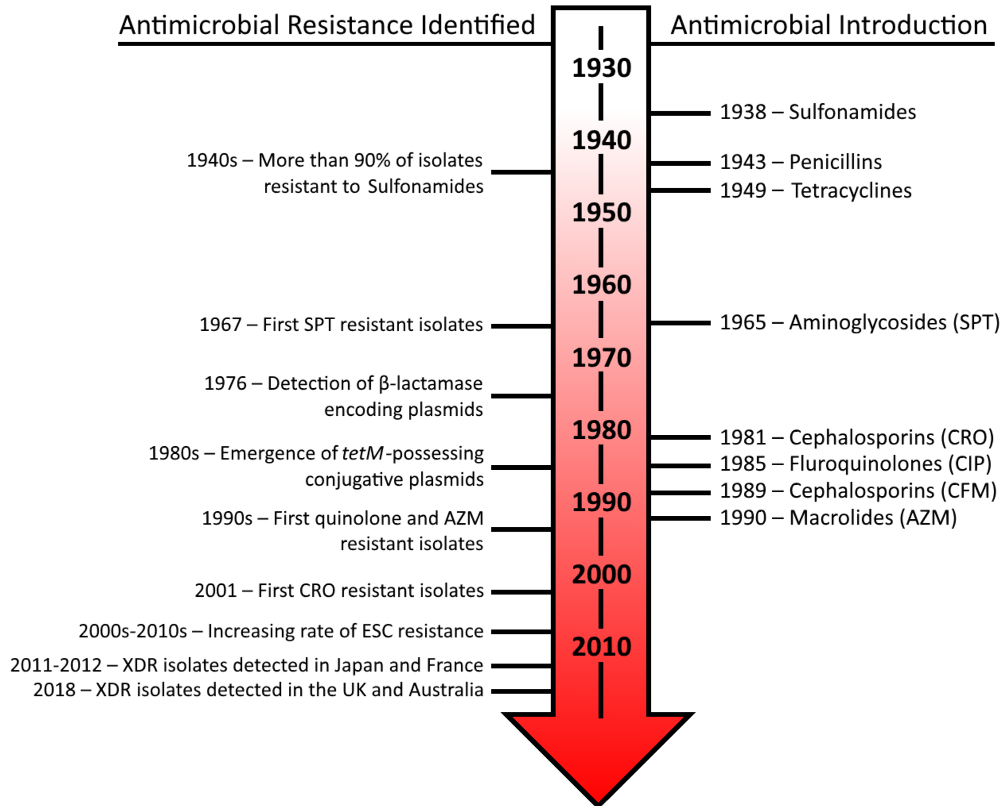

Figure 1.3. | The history of *Neisseria gonorrhoeae* antimicrobial resistance. The right panel shows the date of introduction of each antimicrobial into clinical use and the left panel shows the year in which resistance was first observed. CRO – Ceftriaxone; AZM – Azithromycin; CFM – Cefixime; SPT – Spectinomycin; ESC – Extended-spectrum cephalosporins; XDR – Extensive drug resistant. Adapted from [100, 109]

With treatment strategies in jeopardy and the fear that gonorrhoea will become untreatable, the WHO [110], the European Centre for Disease Prevention and Control (ECDC) [111] and the Centers for Disease Control and Prevention (CDC) [112] issued action plans since 2012, and raised the awareness of healthcare professionals to define and monitor gonorrhoea treatment failures worldwide. In fact, this lead to the CDC classifying this pathogen as a "superbug" and the WHO considering it as a "Priority 2" microorganism (i.e., Priority pathogen for research and development of new antibiotics). In general, the treatment for *N. gonorrhoeae* infections is given empirically at the first clinical visit, which evidently means that antimicrobial susceptibility is rarely performed prior to prescription following the WHO guidelines [113]. Worldwide health organizations have established empirical treatment guidelines for urogenital gonorrhoea cases, which consist in a dual therapy combining a single dose of a third generation cephalosporin (either ceftriaxone or cefixime) and a single dose of a macrolide (azithromycin), while for retreatment of cases after treatment failure other antibiotic regiments are recommended (Summarized in **Table 1.1.**) [113–119]. Finally, research on new treatment options for gonorrhoea is ongoing and follows either the repurposing of existing antimicrobials (or novel combinations), the development of new antimicrobials or the development of alternative therapies [109, 120].

Table 1.1. | Different treatment guidelines for gonorrhoea worldwide.

WHO [113]	**Australasia** [114]	**Canada** [115]	**USA** [116]	**UK** [117]	**EU** [118]	**New Zealand** [119]
CRO [250 mg IM] + AZM [1 g O] ***Or*** CFM [400 mg O] + AZM [1 g O]	CRO [500 mg IM] + AZM [1 g O]	CRO [250 mg IM] + AZM [1 g O]	CRO [250 mg IM] + AZM [1 g O]	CRO [500 mg IM] + AZM [1 g O]	CRO [500 mg IM] + AZM [1 g O]	CRO [250 mg IM] + AZM [1 g O]
CRO [500 mg IM] + AZM [2 g O] ***Or*** CFM [800 mg O] + AZM [2 g O] ***Or*** GEN [240 mg IM] + AZM [2 g O] ***Or*** SPT [2 g IM] + AZM [2 g O]	---	CFM [800 mg O] + AZM [1 g O] ***Or*** SPT [2 g IM] + AZM [1 g O]	CFM [400 mg O] + AZM [1 g O]	CFM [400 mg O] + AZM [1 g O] ***Or*** SPT [2 g IM] + AZM [1 g O] ***Or*** CTX [500 mg IM] + AZM [1 g O]	CFM [400 mg O] + AZM [2 g O] ***Or*** SPT [2 g IM] + AZM [2 g O]	SPT [2 g IM] + AZM [1 g O] ***Or*** GEN [240 mg IM] + AZM [2 g O]

All treatment guidelines are single dose. Top line are the recommended empirical treatment for uncomplicated gonorrhoea, while the bottom line corresponds to alternative treatment guidelines, mainly used for cases of retreatment after treatment failure. IM – Intramuscular; O – Oral; WHO – World Health Organization; USA – United States of America; UK – United Kingdom; EU – European Union; CRO – Ceftriaxone; AZM – Azithromycin; CFM – Cefixime; SPT – Spectinomycin; GEN – Gentamicin; CTX – Cefotaxime (Adapted from [109])

1.3.4. Global epidemiology and the Portuguese panorama

Gonorrhoea disease burden has increased in the last decade and remains a major public health concern worldwide. Gonorrhoea is still the second most common STI and from 2012 up to 2016, the WHO observed an increase from 78.3 million to 87.0 million new cases of *N. gonorrhoeae* worldwide in individuals aged from 15 to 49 years [113, 121]. Worldwide data also show that the African Region and the Region of the Americas are the most affected, presenting the highest estimated incidence rates, in both men and women [122]. In 2018, the CDC reported more than 580 000 cases in the USA, which reflects a 64% increase since 2014 and the highest number reported since 1991 [123]. Additionally, data suggest that men who have sex with men (MSM) are disproportionately affected by this disease [124]. In Europe, in 2017, the ECDC reported around 90 000 notified cases across 27 European Union/European Economic Area (EU/EEA) countries [125], an increase of 17% when compared with 2016 [126]. The countries with the highest notification rates per 100 000 population were the United Kingdom (UK), Ireland, Denmark, Iceland, Norway and Sweden, while the lowest rates were observed in Bulgaria, Croatia, Cyprus, Poland and Romania [125]. Moreover, notification was three fold higher for men than for women, with MSM accounting for almost half (47%) of all reported male cases, and the largest proportion of cases were reported among the age group of 25 to 35 years old [125].

In Portugal, data on gonorrhoea cases have been reported to the Portuguese General Health Directorate (DGS; Direção Geral da Saúde) at least since the 1950s with data showing a

rampant decrease in notified cases from 1970 to 1995 and very few cases reported during the early 2000s [127] (**Figure 1.4.**). Since the beginning of 2015, a nationwide integrated electronic platform was implemented in our country for public health surveillance (SINAVE; Sistema Nacional de Vigilância Epidemiológica), including the mandatory notification of all gonorrhoea cases, either confirmed by direct examination (restricted to urethral samples of symptomatic men), culture or by NAAT. Noteworthy, there is a separation between laboratory and clinical reporting in Portugal and, to date, DGS only validates gonorrhoea cases if these are reported by clinicians. Data from 2015-2018 evidence an 113% increase of notified cases of gonorrhoea (**Figure 1.4.**) (from 468 in 2015 to 976 cases in 2018; data retrieved from the Ministry of Health open-access platform - https://www.sns.gov.pt/transparencia/). Moreover, reports show that the highest notification rates hail from the Lisbon and Oporto Metropolitan Areas [128, 129], and that most cases hail from individuals aged from 15 to 34 years, with ~90% of cases being from men since 2015 [127]. Still, the electronic notification system cannot include AMR data due to lack of standardized cultural methods and to the unavailability of molecular AMR tests, which hampers an overview of the current state of *N. gonorrhoeae* AMR in our country.

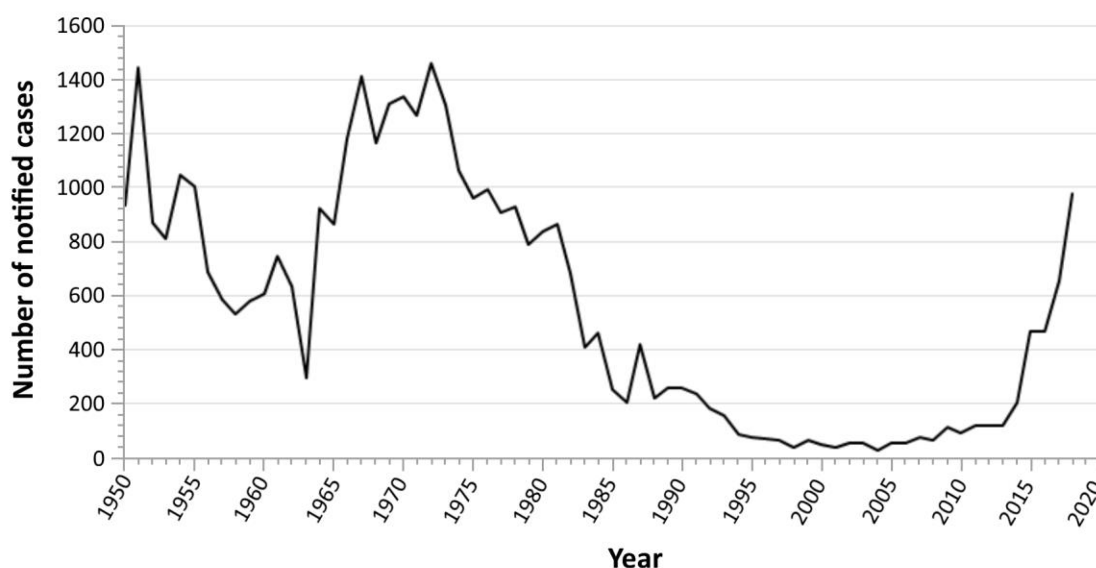

Figure 1.4. | Number of notified cases of gonorrhoea per year in Portugal, from 1950 to 2018 (Adapted from [127].

1.3.5. *Neisseria gonorrhoeae* antimicrobial resistance burden

Following the international calls for action to monitor and manage the spread of antimicrobial resistant *N. gonorrhoeae* [110–112], several national and international surveillance programmes have been implemented worldwide [130], namely the Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP) promoted by the Public Health England (PHE)

in the UK [131–133], the European Gonococcal Antimicrobial Surveillance Programme (EURO-GASP) in Europe [134, 135], the Gonococcal Isolate Surveillance Project (GISP) in the USA [136, 137], the Canadian National Programme in Canada [138, 139] and the Australian Gonococcal Surveillance Program (AGSP) [140, 141]. These surveillance programmes, along with other studies have allowed to observe that, in the last decade, there has been a steady rise in MICs and resistance for cephalosporins [103, 104, 133, 137, 141–145]. Moreover, since the first report of resistance [146, 147] and clinical failure to cephalosporins [102, 105], novel resistant strains and even clinical failures have been confirmed [106–108, 139, 148, 149]. Worldwide increase in azithromycin resistance is also being reported [150–154] with resistant strains associated with recent outbreaks [155–157], one of which in Australia leading to changes in treatment guidelines in this country [158]. Multidrug resistant and extensively drug resistant gonococcal strains are being reported in Europe [106, 150, 159, 160], Asia [152, 161, 162], the USA [137, 153], Canada [139] and Australia [141, 157]. The GISP's latest report refers to 2014 (sampling size of 5093), where resistance to tetracycline was at 25.3%, to ciprofloxacin at 19.2% and to penicillin at 16.2% [137]. In the USA, azithromycin breakpoint is defined at MIC ≥ 2 mg/L, contrasting with the EUCAST and CLSI guidelines that are based on an Epidemiologic cut-off (ECOFF) of 1 mg/L (i.e., MIC > 1 mg/L), and data showed an increase from 2013 to 2014 (from 0.6 to 2.5%), although no isolates simultaneously presented resistance to cephalosporins [137]. Additionally, the GISP also reports an overall increase in cefixime and ceftriaxone resistance in the USA in 2014 [137] when compared with previous years [136]. In Australia, decreased susceptibility to ceftriaxone was found to be at 1.7% (sampling size of 9006) in 2018, with two extensively drug-resistant isolates being detected [141]. Resistance to azithromycin decreased from 9.3 to 6.2% from 2018 to 2017, but several high-level resistant isolates were reported [140, 141]. Canada, through its National Surveillance Programme, confirmed its first ceftriaxone resistant *N. gonorrhoeae* in 2017, and reported (sampling size of 5290) an increase of multi-drug resistant isolates (i.e., isolates with decreased susceptibility to cephalosporins or resistance to azithromycin, along with resistance to at least two other antimicrobials) from 4.4% in 2013 to 12.2% in 2017 [139]. Although ciprofloxacin and tetracycline resistance remains high, at 50.1% and 45.9%, respectively, Canadian data revealed that the overall decreased susceptibility to ceftriaxone, cefixime and azithromycin have decreased since 2013 [139]. Still, in 2017, the EURO-GASP did not detect any isolate with resistance to ceftriaxone (sampling size of 3248) and results showed stable cefixime and azithromycin resistance (1.9% and 7.5%, respectively) [135] when compared with the previous year (2.1% and 7.5%, respectively) [134]. Nevertheless, 15 countries reported cefixime-resistant isolates, and 23 countries reported azithromycin-resistant isolates, revealing that resistance is becoming widespread. Ciprofloxacin resistance remained high at 46.5%, the same value as reported for 2016 [134, 135]. In contrast, the latest data from isolates collected from 2018 (sampling size of 1455) on behalf of the GRASP showed a moderate

increase in resistance, when compared to 2017, to azithromycin (from 9.3% to 9.7%), ciprofloxacin (from 36.4% to 39.8%), cefixime (from 1.7% to 2.1%), penicillin (from 10.8% to 12.4%) and tetracycline (from 48.5% to 52.8%), with no resistance to ceftriaxone being detected, albeit an increase in MICs was observed [132, 133]. Additionally, in 2018, the UK reported three cases of extensively drug-resistant *N. gonorrhoeae* with epidemiological links to sexual networks in Southeast Asia or Spain [133]. Noteworthy, most programmes report that overall AMR is more prevalent in isolates from MSM individuals [133, 135, 137]. Besides the limited data sent to and reported by the EURO-GASP, the overall global trends of *N. gonorrhoeae* AMR in Portugal are currently undisclosed. Nonetheless, up to 2018 no resistant isolate to cephalosporins was reported and data from EURO-GASP (110 isolates per year) for Portugal pointed it as one of the countries with highest resistance to azithromycin in Europe [134, 135].

1.4. *Neisseria gonorrhoeae* genomics

1.4.1. General genomic features

N. gonorrhoeae possesses a ~2.2 Mbp circular chromosome with ~2000 coding sequences. It can also carry plasmids, either the cryptic plasmid (it does not contain AMR genes or virulence factors) or plasmids associated with AMR [163] that can be transferred between pathogenic and commensal species [164, 165]. A repertoire of genetic elements enables increased genomic plasticity in *N. gonorrhoeae*, which likely contribute to host colonization, invasion and survival, such as DNA transformation [166] and variation of displayed antigens [167]. In fact, one of the most outstanding features of *Neisseria* spp. is the amount and variety of repetitive DNA sequences present in their genomes [168, 169]. The most abundant repeated sequences in the genome are the DNA uptake sequences (DUS), which are nonpalindromic 10-bp sequences present ~2000 times in the genome that function in DNA transformation involving same-species strains. Represented ~200 times in the chromosome are 20-bp sequences called duplicated repeat sequence 3 (dRS3) [168, 169] that act as sites for phage integration [170] and genome rearrangements [171]. Additionally, *N. gonorrhoeae* possess a gonococcal genetic island (GGI) encoding a type IV secretion system (T4SS) that secretes single-stranded DNA into the extracellular environment providing substrate for DNA transformation [172, 173]. Besides these genomic features used to facilitate HGT, *N. gonorrhoeae* also has several mechanisms of phase variation used to generate genetic diversity within its main virulence factors towards pathogenesis and adaptation.

1.4.2. Phase variation and major virulence factors

Phase variation has a vital role in *N. gonorrhoeae*'s ability to adapt to the various environmental niches it colonizes. It is defined as reversible changes in the expression state of genes, either by switching between an expressed (ON) and unexpressed (OFF) state, through transcription regulation or gene truncation, or by switching between two distinct forms of a gene product [168, 174]. This stochastic switching in virulence-associated genes or regulatory systems allows the infecting bacterial population to maintain diversity in order to adapt and survive in light of different selective pressures. In bacteria, phase variation can be mediated by invertible DNA segments, differential methylation, changes in length of homopolymeric tracts or short tandem repeats, causing shifts in the reading frame (when occurring within the coding sequence of a gene). It is estimated that more than 100 genes are phase-regulated in pathogenic *Neisseria*. Several genes identified as virulence factors in *N. gonorrhoeae* and *N. meningitidis* are phase variable, displaying either homopolymeric tracts or short tandem repeats [175–180]. The distinct products from phase-variable genes contribute to the continuous change in the expressed repertoire of outer surface components of antigens, and this antigenic variation is one of the more effective strategies used by pathogens to evade the immune system, persist and enable reinfection. These antigenic variation systems can generate functional variants that alter specific interactions between virulence factors and host molecules, with the variability of the Type IV pili, the LOS antigen and the OM Opa proteins being the major targets of phase variation contributing to a successful pathogenesis [168]. In *N. gonorrhoeae*, the *lgt* genes (*lgtABCDE*) encode a set of glycosyltransferases that transfer sugars to the LOS core [181]. A subset of these genes contain homopolymeric repeats in their 5' ends that allow different possible combinations of glycosyltransferases, thus producing different LOS structures putatively inducing different immunological responses [182–184]. Regarding the Opa proteins, *N. gonorrhoeae* can possess up to eleven different Opa-encoding genes, dispersed throughout the genome [185, 186], with each gene containing from two up to 20 adjacent CTCTT repeats in the coding sequence [69, 178] that can independently phase-vary using slipped-strand mispairing from ON and OFF states, representing a process of antigenic variation [168, 187]. Another mechanism of generating antigenic variation involves the pilin genes through gene conversion (reviewed in [168]), which is defined as nonreciprocal transfer of DNA, by recombination, from a donor homolog to a recipient locus without the donor locus being changed in the process. The Type IV pili of *Neisseria* are long filamentous surface structures involved in DNA transformation, twitching motility and adherence to epithelial cells, and its major component is encoded by the *pilE* locus [168, 188]. *N. gonorrhoeae* possesses several chromosomally scattered truncated pilin genes (*pilS*) that share sequence similarity to the *pilE* gene, acting as genetic reservoirs [168, 188–190].

It is the non-reciprocal transfer of DNA from *pilS* copies to *pilE* that results in variable regions, which encode the exposed regions of the pilus surface, recognized by antibodies [188, 190, 191].

1.4.3. Genomic mechanisms of antimicrobial resistance

Even though the drivers for AMR in *N. gonorrhoeae* are still not fully disclosed, it is suggested that AMR mainly arises due to spontaneous mutation and/or total or partial gene acquisition (by HGT), which are effectively selected due to antibiotic pressure and can be transferred to other gonococci by transformation [120, 166]. Whereas novel mutations are still being discovered, resistance mediated by HGT is generally well described, as are the examples of genes carried on conjugal plasmids in certain strains that confer resistance to tetracyclines and β-lactams. The ability of the gonococcal genome to undergo continuous mutation and internal recombination, resulting in rapidly evolving gonococcal populations, is well known and several mutations have already been described that are partially responsible for an acquired AMR (**Figure 1.5.** and **Table 1.2.**) [100, 120, 192].

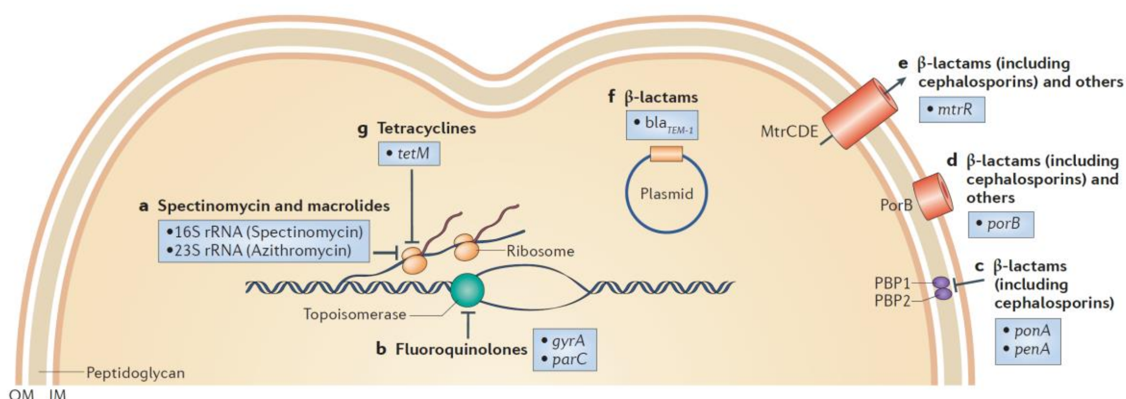

Figure 1.5. | Current *Neisseria gonorrhoeae*-targeting antimicrobials and the corresponding resistance genes. IM – Inner membrane; OM – outer membrane; PBP – Penicillin-binding protein. [100]

Sulfonamides target the bacterial dihydropteroate synthase (DHPS) enzymes, to inhibit the folic acid synthesis in the gonococci, with its resistance due to alterations in the *folP* gene encoding the drug target DHPS [96, 193, 194]. Tetracycline binds to the ribosome, and it is the presence of the plasmid-encoded tetracycline resistance protein TetM and its interaction with the ribosome that leads to high-level resistance to tetracyclines [195, 196]. Additionally, alterations in *rpsJ* (encoding for ribosomal protein S10) also contribute to tetracycline resistance, by reducing the affinity of 30S ribosome for tetracycline [197]. Although rifampicin is not used for gonorrhoea treatment, resistance to this drug in *Neisseria* has been described by point mutations in the *rpoB* gene (encoding the RNA polymerase beta-subunit) [198, 199] also present in *N. gonorrhoeae* [96]. Resistance to β-lactam antibiotics (such as penicillins, cephalosporins and carbapenems) occurs through mutations in *ponA* [200] and *penA* [102, 105, 201, 202], encoding the penicillin-

Table 1.2. | Main antimicrobial resistance determinants described in *Neisseria gonorrhoeae* for previous and currently recommended antimicrobials for the treatment of gonorrhoea

Genetic element	Product	Resistance determinant	Effect	Affected Antimicrobials	Reference
penA	Penicillin-binding protein 2	A311V, I312M, V316T, V316P, T483S, A501V, N512Y, G545S, A501P, A501V, A501T, G542S, P551S, P551L	Reduced β-lactam acetylation of PBP2	CRO, CFM, PEN	[105, 203]
mtrR promoter	---	Deletion of A in repeat (-35A), A → C in repeat (-38), 2 bp insertion	Overexpression of MtrCDE efflux pump	AZM, CRO, CFM, PEN, TET	[204–208]
mtrR	MtrR repressor protein	A39T, G45D, truncation	Overexpression of MtrCDE efflux pump	AZM, CRO, CFM, PEN, TET	[209]
penB (*porB1b*)	Major outer membrane Porin B	G120K, A121D, A121N	Reduced influx	CRO, CFM, PEN, TET	[210]
ponA	Penicilin-binding protein 1	L421P	Reduced β-lactam acetylation of PBP1	CRO, CFM, PEN	[200]
blaTEM	TEM-1 or TEM-135 penicillinase	Presence of *bla*TEM-1/*bla*TEM-135-encoding plasmids	Penicillinase production	PEN	[211, 212]
*23SrRNA**	Ribosome subunit 50S	C2611T, A2059G	Four copies of these genes present, increasing resistance with increased number of copies with SNPs via decreased binding to 50S ribosome	AZM	[213, 214]
gyrA	DNA-gyrase subunit A	S91F, D95N, D95G	Reduced quinolone binding to DNA gyrase	CIP	[96, 215]
parC	Topoisomerase IV subunit C	D86N, S87R, S87I, S87W, S88P, E91K	Reduced quinolone binding to topoisomerase IV	CIP	[96, 215]
rpsJ	Ribosomal protein S10	V57M	Reduced affinity of 30S ribosome for tetracycline	TET	[197]
tetM	TetM	Presence of plasmid	TetM resembles elongation factor G, binds 30S ribotype and prevents tetracycline binding	TET	[195, 196]
*pilQ***	Pore-forming secretin PilQ of the type IV pili	E666K	Reduced influx via pore-forming secretin PilQ	CRO, CFM, PEN	[216]
ermBCF	rRNA methylases	Presence	Methylate 23S RNA to block binding	AZM	[217]
macAB	MacAB efflux pump	Promoter mutation	Overexpression of efflux pump	AZM	[218]
mef	Mef efflux pump	Presence	Efflux pump	AZM	[219]
ereAB	Erythromycin esterase type I and II	Presence	Macrolide esterase	AZM	[207]
norM	NorM efflux pump	Promoter mutation	Overexpression of efflux pump	CIP	[220]
folP	Dihydropteroate synthase	R228S	Reduce target affinity	SUL	[20]
*16SrRNA**	Ribosome subunit 30S	C1192U	Reduces the affinity of the drug for the ribosomal target.	SPT	[20, 203]
rpsE	30S ribosomal protein S	T24P, V25del, K26E	Disrupts the binding to the ribosomal target	SPT	[20, 203]
rpoB	β-subunit of DNA-dependent RNA polymerase	H552	Reduce target affinity	RI	[96, 198, 199]

**Escherichia coli* numbering used. **only found in laboratory strains. PEN – penicillins; TET – tetracycline; CIP – ciprofloxacin; AZM – azithromycin; CRO – ceftriaxone; CFM – cefixime; SPT – spectinomycin, SUL – sulfonamides; RI – rifampicin. PBP – Penicillin-binding protein. Adapted from [20, 96, 192, 203]

-binding proteins 1 (PBP1) and 2 (PBP2) respectively, which are required for peptidoglycan synthesis. The presence of the conjugative plasmid containing the gene *blaTEM-1* (encoding the TEM-1 penicillinase) also confers resistance through inactivation, as the encoded protein hydrolyses the β-lactam ring of penicillin [211, 212]. Fluoroquinolones act by inhibiting the activity of topoisomerases that are involved in DNA replication, and resistance to these is mostly mediated by mutations in the topoisomerase-encoding genes *gyrA* (DNA-gyrase subunit A), *parC* (topoisomerase IV subunit C) and *parE* (toposiomerase IV subunit B) [96, 215, 221]. Spectinomycin and azithromycin bind to ribosomes to inhibit protein synthesis, thus mutations in the 16S and 23S rRNA-encoding genes confer resistance to spectinomycin [19, 95, 222] and azithromycin [213, 214], respectively. Studies have shown that the level of resistance to azithromycin depends on the number of copies of *23SrRNA* (*N. gonorrhoeae* possess four copies) that carry the mutation A2059G, with high-level resistance to this antimicrobial (i.e., MIC > 256 mg/L) being observed when all four copies carry this variant [213, 214]. Additionally, for spectinomycin, alterations in the *rpsE*-encoded 30S ribosomal protein S were also confirmed to result in high- or low-level resistance [223, 224]. *N. gonorrhoeae* can also display genetic alterations in specific genes that lead to a broader decreased susceptibility and resistance. For instance, alterations in PorB, the major OM porin (encoded by *penB/porB*) reduce the permeability of this protein, conferring resistance to various classes of antimicrobials, such as penicillins, cephalosporins or tetracyclines [210]. Resistance can also be mediated by alteration in efflux pump systems, as is the cases of the MtrC-MtrD-MtrE [209] or MacA-MacB [218] efflux systems that expel antibiotics of various classes from *N. gonorrhoeae* cells. Mutations in efflux pump-associated genes can lead to an increased activity of the efflux pump complex, conferring resistance to multiple classes of antimicrobials. This is the case of the promoter of *mtrR* (which encodes the MtrR repressor protein), where alterations lead to an overexpression of the efflux pump [204–208] that are mainly associated with azithromycin decreased susceptibility and resistance while affecting other antimicrobials.

1.4.4. Traditional molecular typing

N. gonorrhoeae typing is traditionally performed either using the seven loci Multi-Locus Sequence typing (MLST) scheme for *Neisseria* [225, 226] or the specific two loci *Neisseria gonorrhoeae* Multi-Antigen Sequence typing (NG-MAST) scheme [227]. Both typing methods can identify clusters of circulating isolates and can be used to create a general epidemiological picture [228]. The MLST scheme for *Neisseria* was first developed for *N. meningitidis* [225] and then further extended to include *N. lactamica* and *N. gonorrhoeae* [226]. Currently, it has more than 15 000 defined profiles (data from http://pubmlst.org/neisseria/), enrolling seven loci, namely: *abcZ* (ABC transporter ATP-binding protein), *adk* (adenylate kinase), *aroE* (shikimate

dehydrogenase), *fumC* (fumarate hydratase class II), *gdh* (glucose-6-phosphate 1-dehydrogenase), *pdhC* (pyruvate dehydrogenase) and *pgm* (phosphoglucomutase). This scheme has allowed to perform epidemiological surveillance, post-vaccine studies and population and evolutionary analyses for *N. meningitidis* [226, 228]. Still, to monitor *N. gonorrhoeae*, researchers rely on NG-MAST, a specific scheme based on internal fragments of two highly polymorphic loci encoding for gonococcal antigens, namely *porB* (490 bp fragment), which encodes the gonococcal OM porin B and *tbpB* (390 bp fragment), which encodes the beta-subunit of the transferrin-binding protein [227]. Due to the high variability of the loci, this scheme currently has more than 19 000 defined profiles (data from http://pubmlst.org/neisseria/). This lead to several studies now agglomerating sequence types (ST) based on sequence identity, in order to cluster isolates into reportable genogroups [207, 208, 229–232]. However, this approach is not without flaws as the agglomeration of isolates is dependent on both the method used (e.g. using clustering algorithms or simple similar sequence identity) and the dataset under analysis, and its nomenclature is based on the most prevalent ST found within a dataset, meaning that it cannot be directly transposed to novel studies [208, 231, 232]. Still, the use of NG-MAST is still performed in order to maintain a certain historical continuity in the current whole-genome sequencing (WGS) era, and provides the possibility to compare "old" isolates with isolates characterized in recent years.

The most predominant NG-MAST STs found in several countries are described in **Table 1.3.** Briefly, from 2009 up to 2013, the most common ST detected in Europe was ST1407 (belonging to NG-MAST genogroup G1407, which includes 330 different STs [230]), as it predominated in several countries, namely Austria, Belgium, Italy, the Netherlands, Portugal, Romania, Slovenia, Spain, Switzerland and the UK [232–234], followed by ST2992 (G2992), ST225 (G225) and ST2400 (G2400). More importantly, ST1407 has been associated with multi-drug resistance, as most isolates belonging to this ST have presented resistance or decreased susceptibility (or both) to cefixime, ceftriaxone, azithromycin and ciprofloxacin [232, 233]. Additionally, between 2010 and 2012 this was also the most common ST found in Canada [231, 235, 236], the USA [237, 238], Japan [239] and Argentina [240], while also being described for the first time in China in 2013 [241]. Nevertheless, in Canada, from 2013 up to 2017, the most prevalent STs seem to have shifted towards ST5985, ST10451, ST12302 and ST14698 [139], while in China ST1407 did not predominate up to 2017 [242–244]. Since 2014, besides G1407 other NG-MAST genogroups that warrant surveillance have emerged worldwide, as these presented resistance to ceftriaxone, with MICs above 0.5 mg/L, namely G564, G5267, G11110, G11018, G21, and more importantly G1866, which also present high-level azithromycin resistance (i.e., MIC > 256 mg/L) [230].

Table 1.3. | Most prevalent *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) sequence types worldwide by study

Continent	Country	Study period	# of isolates tested	Most prevalent ST (%)	2nd most prevalent ST (%)	Reference
Africa	Ghana	2012-2015	44	ST8948 (20%)	ST10251 (7%)	[245]
Asia	China	2007 and 2012	278	ST568 (6%)	ST270 (4%)	[246]
		2014-2015	126	ST7469 (5%)	ST1866 (4%)	[247]
		2014-2017	55	ST14781 (7%)	ST1766 (5%)	[242]
		2013-2017	380	ST5308 (3%)	ST5061 (2%)	[243]
		2015-2017	379	ST4539 (4%)	ST3741 (3%)	[244]
		2012-2013	920	ST2318 (5%)	ST1866 (5%)	[241]
	India	2010-2012	60	ST6058 (22%)	ST621 (3%)	[248]
		2010-2013	100	ST6058 (21%)	ST9774 (4%)	[249]
	Indonesia	2014	78	**ST1407 (51%)**	ST2992 (17%)	[250]
	Japan	1996-2016	400	ST2958 (5%)	ST4016 (5%)	[251]
		2010-2012	193	**ST1407 (16%)**	ST4186 (5%)	[239]
	Pakistan	2012-2014	94	ST338 (3%)	ST3328 (3%)	[252]
	South Korea	2012-2017	593	ST10668 (13%)	ST15016 (10%)	[253]
	Vietnam	2011	108	ST4787 (10%)	ST7720 (5%)	[254]
Australia	Australia	2011-2013	128	ST758 (20%)	ST4186 (2%)	[255]
Europe	Austria	2010	50	**ST1407 (32%)**	ST387 (10%)	[233]
		2013	55	ST3785 (16%)	ST11575 (11%)	[232]
	Belarus	2010-2013	193	ST1993 (27%)	ST807 (7%)	[256]
	Belgium	2010	50	**ST1407 (14%)**	ST387 (12%)	[233]
		2013	55	**ST1407 (16%)**	ST2992 (15%)	[232]
	Denmark	2010	50	ST225 (12%)	ST3158 (12%)	[233]
		2013	55	ST1993 (20%)	**ST1407 (15%)**	[232]
	France	2010	50	ST2 (14%)	ST2992 (14%)	[233]
		2013	57	ST645 (9%)	ST4995 (5%)	[232]
		2010-2012	176	ST8922 (6%)	ST1285 (5%)	[257]
	Germany	2010	50	ST25 (32%)	**ST1407 (28%)**	[233]
		2013	47	ST4995 (9%)	ST5441 (6%)	[232]
	Greece	2010	50	ST5405 (16%)	ST5505 (16%)	[233]
		2013	48	ST3128 (19%)	ST225 (10%)	[232]
	Hungary	2010	17	ST5332 (29%)	ST356 (12%)	[233]
		2013	48	**ST1407 (21%)**	ST995 (13%)	[232]
		2014	50	**ST1407 (10%)**	ST4995 (10%)	[258]
	Ireland	2010	50	ST2992 (32%)	ST384 (12%)	[233]
		2008-2016	58	**ST1407 (9%)**	ST2992 (7%)	[259, 260]
	Italy	2010	50	**ST1407 (34%)**	ST2992 (10%)	[233]
		2013	49	ST2992 (18%)	ST6360 (14%)	[232]
	Latvia	2009-2010	29	ST3227 (41%)	ST5185 (10%)	[233]
		2013	38	ST5 (37%)	ST10828 (13%)	[232]
	Malta	2009-2010	50	ST225 (46%)	ST5463 (12%)	[233]
		2013	20	ST2992 (35%)	**ST1407 (10%)**	[232]
	Netherlands	2010	100	**ST1407 (16%)**	ST2992 (7%)	[233]
		2013	66	ST2992 (14%)	ST2400 (11%)	[232]
		2009-2017	344	**ST1407 (15%)**	ST2400 (12%)	[261]
	Norway	2009-2010	49	ST2992 (23%)	**ST1407 (14%)**	[233]
		2013	55	**ST1407 (9%)**	ST4275 (7%)	[232]
	Poland	2010-2012	228	**ST1407 (25%)**	ST2992 (7%)	[262]
	Portugal	2009-2010	50	**ST1407 (14%)**	ST1034 (8%)	[233]
		2013	109	**ST1407 (16%)**	ST7445 (11%)	[232]
	Romania	2010	9	**ST1407 (22%)**	ST4120 (22%)	[233]
	Russia	2011-2012	521	ST807 (8%)	ST5714 (6%)	[263]
		2013-2018	804	ST807 (8%)	ST228 (3%)	[264]
		2015	123	ST9476 (11%)	ST807 (8%)	[265]
	Slovakia	2010	50	ST437 (26%)	ST5595 (20%)	[233]
		2013	56	**ST1407 (16%)**	ST2992 (9%)	[232]

NG-MAST ST1407 is highlighted in bold. ST – Sequence Type.

Table 1.3. | Most prevalent *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) sequence types worldwide by study (cont.)

Continent	Country	Study period	# of isolates tested	Most prevalent ST (%)	2nd most prevalent ST (%)	Reference
Europe	Slovenia	2009-2010	50	**ST1407 (28%)**	ST225 (16%)	[233]
		2006-2012	194	**ST1407 (9%)**	ST21 (6%)	[266]
		2013	54	ST21 (13%)	ST10800 (11%)	[232]
	Spain	2009-2010	100	**ST1407 (28%)**	ST2992 (9%)	[233]
		2016	104	ST5441 (8%)	ST2318 (7%)	[267]
		2012-2016	134	**ST1407 (6%)**	ST14958 (6%)	[268]
		2013	116	**ST1407 (9%)**	ST2400 (9%)	[232]
	Sweden	2010	50	ST225 (8%)	ST2992 (6%)	[233]
		2013	50	ST5445 (10%)	ST2400 (6%)	[232]
	Switzerland	1998-2001	26	ST8672 (12%)	ST189 (8%)	[269]
		2009-2012	34	**ST1407 (18%)**	ST2992 (9%)	[269]
	United Kingdom	2010	100	**ST1407 (14%)**	ST2 (6%)	[233]
		2013	106	ST2992 (11%)	ST51 (10%)	[232]
		1995-2004	237	ST225 (44%)	ST12 (35%)	[270]
		2004-2017	101	ST9768 (51%)	ST649 (19%)	[271]
		2004-2015	1841	ST2992 (13%)	**ST1407 (8%)**	[272]
North America	Canada	1989-2013	180	**ST1407 (19%)**	ST225 (7%)	[231]
		2010	1233	**ST1407 (13%)**	ST3150 (11%)	[236]
		2001-2010	155	ST3158 (26%)	ST225 (19%)	[273]
		2017	5290	ST12302 (24%)	ST5985 (8%)	[139]
	Greenland	2012-2015	102	ST210 (55%)	ST9816 (30%)	[274]
	United States of America	2011	59	**ST1407 (41%)**	ST7268 (17%)	[237]
		2009-2010	236	**ST1407 (23%)**	ST5895 (6%)	[238]
South America	Argentina	2011-2016	158	**ST1407 (45%)**	ST3378 (4%)	[240]
		2009-2013	42	**ST1407 (52%)**	ST3431 (10%)	[275]

NG-MAST ST1407 is highlighted in bold. ST – Sequence Type.

1.4.5. *Neisseria gonorrhoeae* in the whole-genome sequencing era

As stated by Blank and Daskalakis [276] in order to control gonorrhoea in a population we need "access to screening, routine assessment of patients' sexual practices to guide the identification of anatomical sites requiring specimen collection, laboratory capacity to perform testing, diagnostic technology that can characterize the organism and its antibiotic susceptibility, systems for gathering that information to guide treatment recommendations, and above all, effective and simple antibiotic therapy". In light of this, WGS, coupled with bioinformatics, could fulfil several of these requirements, as increased genomic data has already allowed to widen our knowledge on *N. gonorrhoeae*'s origin, evolution, population structure, pathogenesis, AMR mechanisms (including the identification of novel loci and tracking of the spread of AMR determinants) and epidemiology, providing a detailed epidemiological analysis of *N. gonorrhoeae* transmission and contact tracing [277]. WGS also complements other laboratory techniques in order to improve diagnostics and treatment in the clinic and inform public health policies to limit the impact of antibiotic resistance [277]. In fact, at the time of writing the present PhD dissertation, there are more than 15 000 *N. gonorrhoeae* genomes publicly available from across the world, namely from Europe [21, 96, 106, 155, 232, 259, 260, 270-272, 278–282], Australia

[21, 96, 283–289], North America [21, 208, 231, 238, 290–294], South America [21, 240, 281, 292], Asia [21, 96, 295–298] and Africa [21, 299]. WGS data are now being generated by most national or international gonococcal AMR surveillance programs in order to monitor the spread of resistant *N. gonorrhoeae* strains. These have allowed to unveil the current AMR genetic trends of circulating *N. gonorrhoeae* strains in several countries and how the genetic determinants for AMR are transmitted [21, 232, 282, 291, 300], as well as to disclose local transmission chains and outbreaks [155, 270, 272, 287, 288]. The ever evolving bioinformatics tools under development are making it possible to rapidly perform *in silico* isolate typing from genome assemblies either by NG-MAST [286] or MLST [301, 302], as well as to identify the presence/absence of AMR genetic determinants directly from WGS reads [302] or genome assemblies [303]. Additionally, WGS data coupled with bioinformatics pipelines have also been used to predict antimicrobial MICs directly from genetic data, either using Illumina technologies for large-scale analysis [192, 304, 305] or Oxford Nanopore technologies for a faster point-of-care response [306]. Nonetheless, one remaining issue of these studies is the continuous reliance on traditional typing methods (i.e., NG-MAST and MLST) to report epidemiological data, which is due to the fact that we are in an ongoing transition phase. Thus, a WGS-based typing strategy has not yet been implemented and traditional typing is still required for backward compatibility in results communication. Recently, a novel typing methodology has been proposed by the Public Health Agency of Canada, the *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR; https://ngstar.canada.ca/), which is based on seven well-characterized genes associated with AMR (namely, *penA*, *mtrR*, *porB*, *ponA*, *gyrA*, *parC* and *23SrRNA*) [303]. However, although useful for tracking the dissemination of AMR determinants, this method does not take advantage of the whole genome data. Still, in an era where culture is being progressively replaced by NAAT for the diagnosis of gonorrhoea, it is crucial to develop culture-independent methodologies that can trustworthily identify the genetic drivers of *N. gonorrhoeae* AMR.

Currently, two main methods are being applied in order to take advantage of WGS data for isolate comparison, namely single-nucleotide polymorphism (SNP)-based or gene-by-gene approaches [307]. The use of gene-by-gene approaches [308] enables the comparison of genomes (either complete or draft) using a predefined set of loci, collected in a scheme. These can enrol either core-loci (i.e., loci present in all or the vast majority of isolates) thus being referred to as core-genome MLST (cgMLST) schemes, or a combination of core and accessory loci (i.e., loci present in only a fraction of isolates), being referred to as whole-genome MLST (wgMLST) schemes. Contrary to SNP-based methodologies, gene-by-gene approaches are independent of a reference genome/strain, and have gained appeal for sub-typing due to the portability of allele nomenclature, being, for instance, the recently adopted approach for WGS-typing of food-borne pathogens [309, 310]. Still, for *N. gonorrhoeae*, most studies are have relied on SNP-based approaches [21, 232, 270, 272] and only recently has a cgMLST scheme being published [301,

311]. Thus, the transition to a WGS-based laboratory surveillance of *N. gonorrhoeae* remains to be defined. As such, with the advent of the new large-scale genomics era and the transition to a WGS-based routine laboratory surveillance, novel approaches are required in order to facilitate data communication and comparability, namely a WGS-based typing methodology, along with harmonized and easily communicable nomenclature that contemplates AMR data. In this context, in Portugal, as of the time of writing, no genome-scale study has ever been performed, and thus the genetic diversity of circulating strains remains to be disclosed. Additionally, acting as the National Reference Laboratory for Sexually Transmitted Infections (NRL-STI) of the Portuguese National Institute of Health (INSA), whose mission is to epidemiologically survey STI pathogens such as *N. gonorrhoeae* and report AMR data, the *modus operandi* for its transition to WGS-based surveillance also needs to be defined.

1.5. Aims and general research plan

The global goal of this PhD dissertation was to strengthen *N. gonorrhoeae* epidemiological and AMR surveillance using WGS. This work was driven by the progressive need of transition to a WGS-based laboratory surveillance for *N. gonorrhoeae*, and the lack of information surrounding AMR trends and the genetic diversity of circulating strains in Portugal. To achieve this overall objective, several specific aims were pursued in the course of this PhD work, namely:

i) to disclose the major *N. gonorrhoeae* AMR trends observed in Portugal throughout sixteen years of laboratory surveillance (Chapter II and Chapter VI);

ii) to achieve a comprehensive WGS-based genogroup assignment for *N. gonorrhoeae*, representing its major lineages, to be used in the transition to WGS-based surveillance (Chapter III);

iii) to link the identified genogroups with specific genetic and AMR signatures (Chapter III and Chapter IV)

iv) to establish a WGS-based molecular surveillance methodology in order to disclose the genetic diversity and AMR profiles of circulating strains, as well as potential transmissions chains, in Portugal (Chapter V and Chapter VI)

v) to evaluate the effectiveness of a culture-independent strategy to obtain WGS data directly from clinical samples and its suitability for epidemiological surveillance and AMR detection (Chapter VII).

CHAPTER II

Fifteen years of a nationwide culture collection of *Neisseria gonorrhoeae* antimicrobial resistance in Portugal

---

This Chapter was published in

Pinto M, Rodrigues JC, Matias R, Água-Doce I, Cordeiro D, Correia C, Gomes JP, Borrego MJ and the PTGonoNet. (2020) **Fifteen years of a nationwide culture collection of *Neisseria gonorrhoeae* antimicrobial resistance in Portugal**. *European Journal of Clinical Microbiology and Infectious Diseas*es; 39:1761–1770. **DOI**: 10.1007/s10096-020-03907-7.

Personal contributions

MP contributed to the study design, performed wet-lab experiments and research, analyzed the data and wrote the manuscript.

2. Fifteen years of a nationwide culture collection of *Neisseria gonorrhoeae* antimicrobial resistance in Portugal

2.1. Abstract

Neisseria gonorrhoeae antimicrobial resistance (AMR) and gonorrhoea disease burden remain major public health concerns worldwide. To contribute to the supranational demands to monitor and manage the spread of antimicrobial resistant *N. gonorrhoeae*, the Portuguese NIH promoted the creation of the National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNet). The present study reports the *N. gonorrhoeae* major AMR trends observed from 2003 up to 2018. All isolates described in the present study constitute the opportunistic ongoing *N. gonorrhoeae* isolate collection supported by the National Reference Laboratory for Sexually Transmitted Infections of the Portuguese NIH, enrolling strains isolated in 35 different public and private laboratories. Minimum inhibitory concentrations were determined using E-tests for azithromycin, benzylpenicillin, cefixime, ceftriaxone, ciprofloxacin, gentamicin, spectinomycin and tetracycline. Molecular typing was determined using NG-MAST. AMR data of 2596 country-spread isolates show that 87.67% of all *N. gonorrhoeae* isolates presented decreased susceptibility to at least one antimicrobial. A continuous decreased susceptibility and resistance to penicillin, tetracycline and ciprofloxacin can be observed along the years. However, no decreased susceptibility for cephalosporins was observed until 2018, while for azithromycin this was always low. The most common observed NG-MAST genogroups were G1407, G7445, G225, G2 and G1034. This study evidences the advantages of a nationwide collection of isolates and of centralized AMR testing to respond to supranational (EURO-GASP) requirements, while providing unprecedented data on AMR in the context of 15 years of surveillance.

Keywords: *Neisseria gonorrhoeae*; antimicrobial resistance; surveillance; sexually transmitted infection

2.2. Introduction

Neisseria gonorrhoeae is the obligate human pathogen that causes the sexually transmitted disease gonorrhoea associated with significant morbidity [22]. In 2016, the World Health Organization (WHO) estimated 87 million new cases of gonorrhoea worldwide [121] and the European Centre for Disease Prevention and Control (ECDC) reported in Europe around 90 thousand cases in 27 countries in 2017 [125]. In the absence of a vaccine, managing *N. gonorrhoeae* mainly relies on preventive measures, epidemiological surveillance and an effective

antibiotic treatment. Nevertheless, this bacterium has shown the ability to acquire antimicrobial resistance (AMR) to multiple classes of antibiotics [120]. Currently, the WHO recommends empirical treatment of gonorrhoea with a dual antimicrobial therapy comprising ceftriaxone and azithromycin [113]. However, during the last decade there has been a steady rise in minimum inhibitory concentrations (MICs) for cephalosporins, with reports of resistant strains and even clinical failures [103, 104, 106, 107], although none in Portugal up to 2018; alongside, there was an worldwide increase in azithromycin resistance [150, 152–154, 158]. With treatment strategies in jeopardy and the fear that gonorrhoea becomes untreatable, the WHO, the ECDC and the Centers for Disease Control and Prevention (CDC) issued action plans since 2012 and raised the awareness of healthcare professionals to define and monitor gonorrhoea treatment failures worldwide.

In Portugal, a nationwide integrated electronic platform was implemented in 2015 for public health surveillance [127, 129], including the mandatory notification of all gonorrhoea cases to the Portuguese General Health Directorate (DGS), either confirmed by culture or by nucleic acids amplification methods. Data from 2015-2018 evidence an 113% increase of notified cases of gonorrhoea (from 468 in 2015 to 976 cases in 2018; data retrieved from the Ministry of Health open-access platform - https://www.sns.gov.pt/transparencia/). However, the electronic notification system cannot include AMR data due to lack of standardized culture methods and to the unavailability of molecular AMR tests.

The National Reference Laboratory for Sexually Transmitted Infections (NRL-STI) of the Portuguese National Institute of Health (INSA) started a modest process of collecting *N. gonorrhoeae* strains in 2003. In 2010, INSA formalized the creation of the Portuguese National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNet) to widen the sampling of isolates and contribute to supra-national demands. PTGonoNET aims to: collect and characterize circulating *N. gonorrhoeae* strains; monitor AMR emergence; and identify potential outbreak and transmission chains. This Network has allowed to contribute for the European Gonococcal Antimicrobial Surveillance Programme (EURO-GASP) [150, 154, 232], which requires the reporting on AMR and, since 2013, whole-genome sequencing [232]. In this context, the present study aims to report the *N. gonorrhoeae* antibiotic resistance trends in Portugal for eight antimicrobials, from 2003 up to 2018, while displaying the advantages of the creation of an opportunistic centralized nationwide isolate collection.

2.3. Methods

2.3.1. PTGonoNET isolate collection

All isolates were obtained from the ongoing *N. gonorrhoeae* isolate collection supported by the NRL-STI of INSA. INSA receives strains from 35 countrywide spread public and private

laboratories that voluntarily participate in the PTGonoNet. INSA receives anonymized countrywide isolates with minimal available epidemiological information, namely site of infection, gender, age, sexual orientation (rarely declared) and region of residence of the patient. Each viable isolate is tested for antimicrobial susceptibility before storage at -80ºC. For this study, when the patient's region of residence data was unavailable, the location of the reporting laboratory was attributed to the isolate. For isolates obtained from the rectum of male patients, the sexual orientation was assumed as 'MSM' (Men who have sex with men), even when that information was missing.

2.3.2. Antimicrobial susceptibility testing

Upon reception at INSA, isolates were sub-cultured in Chocolate agar (BioMérieux, France) before susceptibility testing. Suspensions of 24h-cultures were prepared equivalent to McFarland's standard 0.5 and MIC for azithromycin, benzylpenicillin, cefixime, ceftriaxone, ciprofloxacin, gentamicin, spectinomycin and tetracycline were determined using E-tests (BioMérieux, France) in Chocolate agar PolyViteX or GC agar 1% Isovitalex, 24h at 36ºC in 5% CO_2. For 234 isolates from 2006-2010, no MIC values were available for the present study, as antibiotic susceptibility was determined by agar dilution breakpoint technique at Public Health England, which provided solely qualitative results [312] (**Supplementary Table S2.1.**). All but 85 isolates were tested for β-lactamase production using the chromogenic reagent Nitrocefin (Oxoid, UK) according to the manufacturer's instructions. INSA participates in the National External Quality Assessment Service (NEQAS) quality assurance for antimicrobial susceptibility testing promoted by EURO-GASP. Isolate antibiotic susceptibility results were interpreted according to 2019 European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoint definitions [94], with interpretation of azithromycin updated for all isolates, accordingly [i.e., according to the epidemiological cut-off (ECOFF) at 1 mg/L]. As gentamicin susceptibility lacks an established breakpoint to date, isolates were classified based on interpretive results from previous studies [313, 314] as susceptible when MIC $\leq$ 4 and resistant when > 16 mg/L, for analysis purposes.

2.3.3. Isolate typing

For 1042 *N. gonorrhoeae* isolates, molecular typing was performed by using the *Neisseria gonorrhoeae* Multi Antigen Sequence Typing (NG-MAST) method, as previously described [315]. Briefly, after DNA extraction, the two highly polymorphic loci, *porB* and *tbpB*, were subjected to PCR amplification and sequencing. Alleles and sequence types (STs) were assigned from the international database of the NG-MAST website (http://www.ng-mast.net/).

Concatenated and aligned NG-MAST *porB* and *tbpB* sequences were used to construct minimum spanning tree using goeBURST, implemented in the PHYLOViZ platform [316]. All sequences with 99% identity were grouped in a single genogroup for analysis purposes, where the Genogroup number refers to the most prevalent ST (genogroup was only attributed to isolates cluster with more than five isolates). All determined alleles, ST and genogroups are reported in **Supplementary Table S2.1**.

2.3.4. Data availability and interactive visualization

To potentiate visualization and interpretation, data is made available through the MicroReact web-platform [317], in order to analyse dynamically AMR and molecular data presented in this study by year and geographic region (https://microreact.org/project/4U7YJ-70k/240ea921). The full characterization of all isolates, including metadata, molecular data and antibiotic susceptibility data, are detailed in **Supplementary Table S2.1**.

2.3.5. Limitations of this study

There is a separation between laboratory and clinical reporting in Portugal and, to date, DGS only validates gonorrhoea cases if these are reported by clinicians. On this regard, and considering that all isolates received at INSA are anonymized, it is not possible to assess which of them were notified to DGS (either by laboratories or by clinicians). In addition, although the described dataset enrols the participation of 35 countrywide spread public and private laboratories (including the major hospitals and private laboratory networks), it cannot be assured that the epidemic in Portugal is 100% reflected by the described dataset.

2.4. Results

2.4.1. Demographic distribution of *Neisseria gonorrhoeae* isolates

Up to 2018, INSA has received 2596 viable isolates from 201 different cities/towns across the country (**Figure 2.1.**). Total isolates distribution from 2003 up to 2018, according to the second tier of Nomenclature of Territorial Units for Statistics (NUTII) classifier is as follows: 72.61% from the Lisbon Metropolitan Area, 17.10% from the North, 5.12% from the Center, 3.08% from the Algarve, 1.89% from the Alentejo, 0.12% from the Azores Autonomous Region and 0.04% from the Madeira Autonomous Region (**Table 2.1.**), with only one isolate with untraceable data. The vast majority of isolates were collected from men (90.37%), of which 17.65% were MSM, although sexual orientation was mostly undeclared (76.19%). Most isolates

were retrieved from urogenital tract swabs (97.65%), with 86.71% from the urethra, 8.67% from the cervix (or the vagina) and 2.27% from the rectum (**Table 2.1.**). Still, 2.35% of isolates were collected from blood, the conjunctiva, pharynx, articular and ascitic fluids, sperm and penile abscess pus (**Supplementary Table S2.1.**). Regarding patient age distribution (**Table 2.1.**), more than half of the culture collection hails from young adults, particularly from 20 to 29 (42.60%) and 30 to 39 (26.35%) years of age, followed by adults aged from 40 to 59 (16.95%).

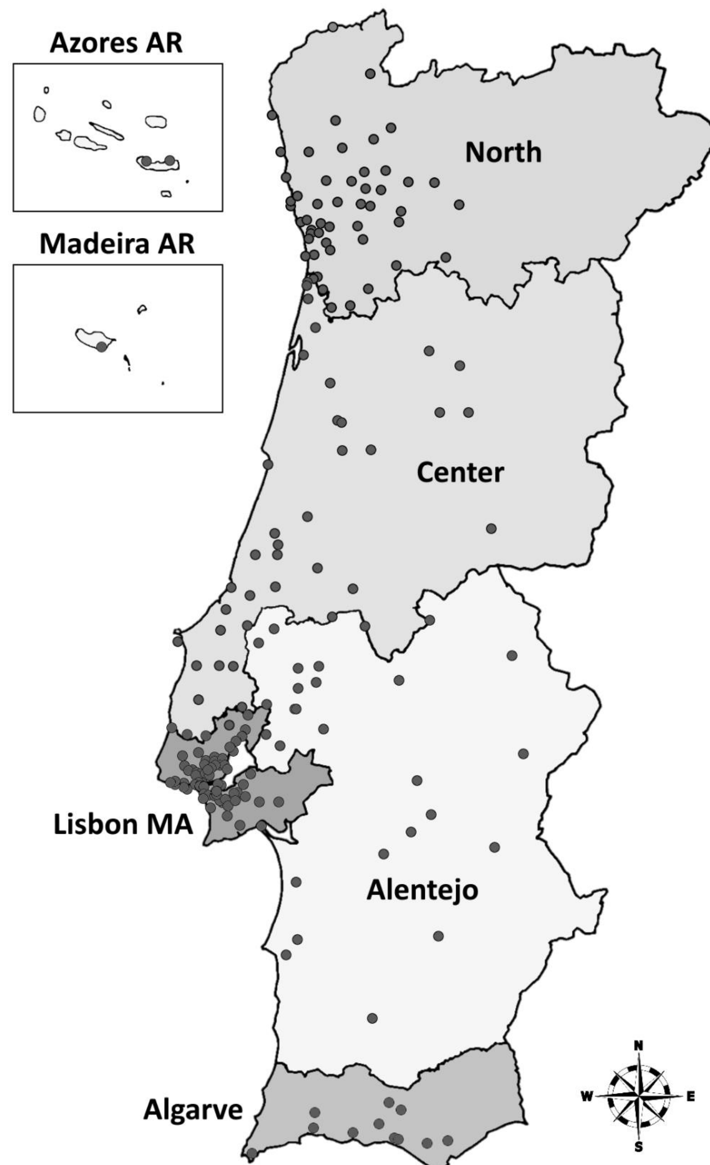

Figure 2.1. | Map of Portugal exhibiting the geographic origin of the *Neisseria gonorrhoeae* isolates of the Portuguese National Institute of Health (INSA) culture collection from 2003-2018. Map is divided by geographic regions according to the second tier of the Nomenclature of Territorial Units for Statistics (NUTII), and each dot displays a city/town representative of a specific NUTII. AR – Autonomous Region; MA – Metropolitan Area.

Table 2.1. | Distribution of *Neisseria gonorrhoeae* isolates for each collection year according to patient gender, sexual orientation, geographic region and site of infection.

	Year																
	2003	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**Total**
Gender																	
Male	15	13	18	26	73	48	70	65	100	96	140	155	227	314	483	503	2346
Female	2	6	2	3	7	10	10	8	8	13	21	16	29	26	35	44	240
Transsexual	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1
Undisclosed	.	.	.	.	.	.	2	.	.	.	2	2	.	1	1	1	9
Sexual Orientation																	
Heterosexual	7	7	4	9	18	9	12	14	6	12	35	9	11	5	9	24	191
MSM	7	10	13	18	28	36	64	31	7	13	15	16	24	25	36	71	414
MSMW	1	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	4
Not applicable	.	.	.	.	.	.	2	.	.	.	2	2	1	.	1	1	9
Undeclared	2	2	3	2	34	13	4	28	95	84	111	146	220	311	470	453	1978
Age (years)																	
< 10	.	.	.	.	.	.	2	1	1	.	2	3	2	3	3	3	20
10-19	2	1	1	.	3	7	15	7	10	9	28	16	37	46	49	54	285
20-29	7	6	9	16	38	32	41	24	50	51	59	56	111	140	221	245	1106
30-39	6	8	7	12	25	10	17	25	28	32	36	49	52	86	141	150	684
40-59	1	4	3	1	14	9	7	15	18	15	32	42	42	60	91	86	440
> 60	.	.	.	.	.	.	.	1	1	2	6	7	12	6	14	6	55
Unknown	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5	6
Geographic Region																	
Lisbon MA	15	19	18	29	64	53	78	55	84	84	112	108	160	212	384	410	1885
North	.	.	.	.	12	4	2	15	21	14	31	48	59	81	78	79	444
Algarve	.	.	.	.	3	.	2	1	3	5	6	6	7	7	20	20	80
Center	2	.	1	.	1	1	.	2	.	2	11	9	16	30	25	33	133
Alentejo	.	.	.	.	.	.	.	.	.	4	3	2	14	9	11	6	49
Azores AR	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	1	3
Madeira AR	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	1
Unknown	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	1
Site of infection																	
Urethra	15	13	15	23	68	45	66	64	99	94	136	152	225	304	457	475	2251
Cervix	2	6	2	3	5	10	10	6	5	12	21	11	24	19	32	34	202
Rectum	.	.	2	3	5	3	2	2	1	.	1	.	.	8	14	18	59
Other	.	.	1	0	2	0	4	1	3	3	5	10	7	10	16	19	81
No data	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	3
Antibiotic susceptibility testing																	
Benzylpenicilin	.	.	.	29	78	58	82	12	.	.	.	.	2	340	519	549	1669
Ciprofloxacin	16	10	20	29	78	58	82	73	104	109	163	173	251	341	519	549	2575
Tetracycline	.	.	.	29	70	52	.	.	.	.	.	.	2	340	519	549	1561
Azithromycin	16	10	20	29	70	52	78	73	108	109	163	173	251	341	519	549	2561
Ceftriaxone	16	10	20	29	77	58	82	73	108	109	163	173	251	341	519	549	2578
Cefixime	16	10	20	.	4	.	82	73	108	109	163	173	251	341	519	549	2418
Spectinomycin	16	10	20	28	53	57	82	73	108	109	163	173	251	340	519	549	2551
Gentamicin	10	.	.	.	.	.	78	73	107	109	163	173	251	340	519	549	2372
β-lactamase	.	.	.	28	78	58	82	52	108	109	163	173	251	341	519	549	2511
NG-MAST	16	17	20	28	76	51	76	72	108	109	141	144	33	30	121	.	1042
Total number of isolates	17	19	20	29	80	58	82	73	108	109	163	173	256	341	519	549	2596

AR – Autonomous Region; MA – Metropolitan Area; MSM – Men who have sex with men; MSMW – Men who have sex with men and women. NG-MAST – *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing. Geographic regions are presented according to the second tier of the Nomenclature of Territorial Units for Statistics (NUTII).

2.4.2. Antibiotic resistance within the culture collection

In the INSA collection, 87.67% of all isolates presented decreased susceptibility to at least one antimicrobial, with only 66 isolates presenting full susceptibility among the 1410 isolates for which all antimicrobials were tested (**Table 2.1.**). For benzylpenicillin, 16.42% of isolates were susceptible, 13.24% were resistant and 70.34% presented intermediate resistance. β-lactamase production was positive in 9.80% of isolates. Decreased susceptibility to tetracycline was observed in 83.15% of isolates, with 65.48% of these being fully resistant and 34.52%

presenting intermediate resistance**.** Regarding azithromycin, 5.04% of tested isolates presented resistance. Thirty-two isolates had MIC above 2, with two isolates reaching values of 256 mg/L (i.e., high level of resistance) (**Figure 2.2.**). Noteworthy, while for the other antimicrobials the EUCAST breakpoints have remained unchanged up to 2019, from 2010-2019 the EUCAST breakpoint for azithromycin resistance was set at 0.5 mg/L [94, 318], which meant that at least 738 isolates had been previously characterized as intermediately resistant, corresponding to 33.86% of the total.

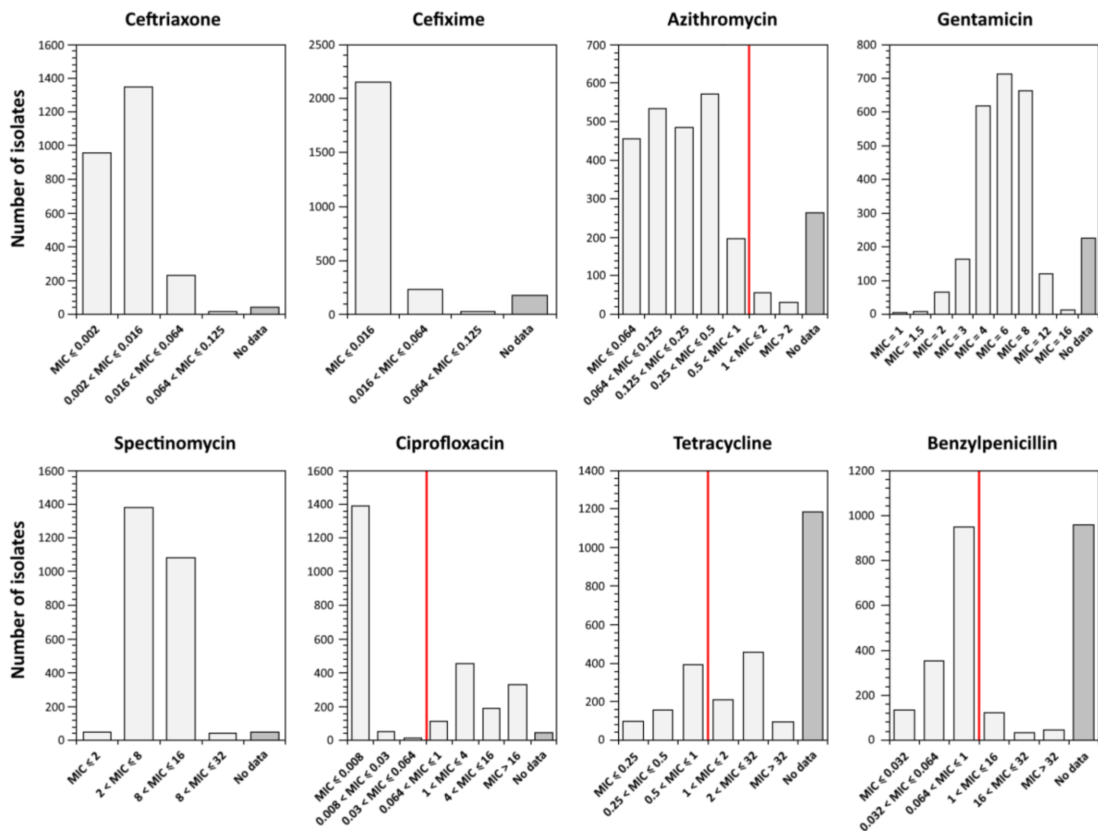

Figure 2.2. | Distribution of the *Neisseria gonorrhoeae* isolates by minimum inhibitory concentrations (MIC) for all tested antimicrobials. Red line refer to the EUCAST breakpoint for antimicrobial resistance.

Data from the last eight years show that MIC_{50} and MIC_{90} values for this antimicrobial are below the ECOFF value of 1 mg/L (**Figure 2.3.**) and have remained relatively constant. For ciprofloxacin, 43.07% were fully resistant and only ten presented intermediate resistance with the majority of resistant isolates exhibiting MICs either between 1 and 4 mg/L (n=457) or above 16 mg/L (n=335). All isolates were susceptible to spectinomycin, with 2507 isolates exhibiting MIC between 2 and 16 mg/L and none > 32 mg/L. All tested isolates for ceftriaxone and cefixime were susceptible, with only 17 and 24 isolates, respectively, having MIC close to the currently defined limit of 0.125 mg/L for resistance classification. Additionally, MIC_{50} values for both antimicrobials has remained low and MIC_{90} values show a decrease since 2013 (**Figure 2.3.**).

Although no breakpoint has been established for gentamicin, we observed that 63.62% of isolates presented intermediate resistance [313, 314], with the vast majority of isolates presenting MICs of 4, 6 or 8 mg/L, with the maximum 16 mg/L detected for 13 isolate (**Figure 2.2.**).

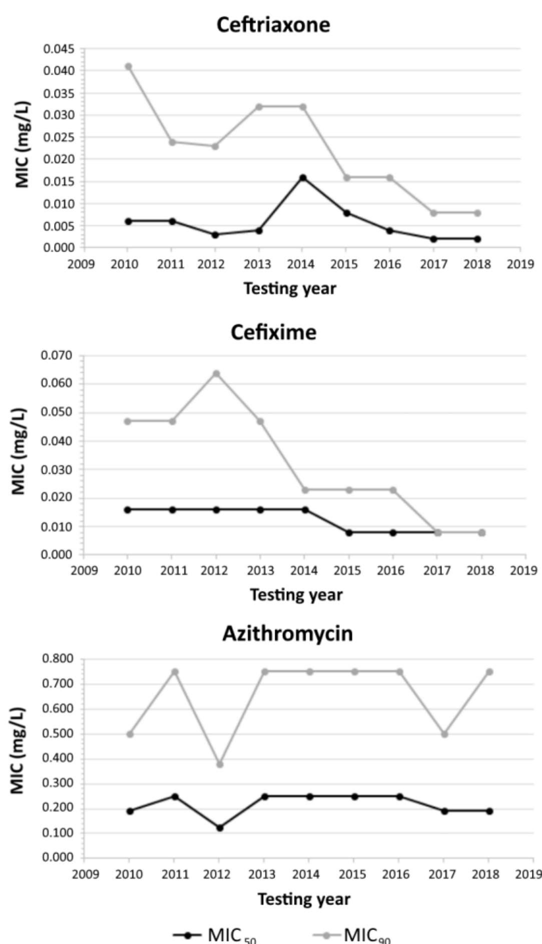

Figure 2.3. | MIC_{50} and MIC_{90} value progression from 2010 up to 2018, for ceftriaxone, cefixime and azithromycin.

A chronological analysis revealed consistent trends for some antimicrobials. In fact, throughout the years, a mean of ~40% of isolates presented resistance to ciprofloxacin (**Figure 2.4.**) and ~80% presented decreased susceptibility to tetracycline and penicillin (although there was a period lacking testing for these two antimicrobials). Data on azithromycin susceptibility shows that the percentage of resistance has also remained relatively stable with values always below 10%, and the highest value observed in 2007 (~8%). This contrasts with the previous classification that showed azithromycin resistance fluctuating throughout the year but with consistent high values since 2010, reaching almost 50% in 2013 (**Figure 2.4.**). Results show an increasing trend in potential intermediate resistance to gentamicin from 2009 up to 2018, reaching ~70% of all isolates collected in 2018 (**Figure 2.4.**). No overall trend between decreased antimicrobial susceptibility and geographic location of origin was observed (see Methods section).

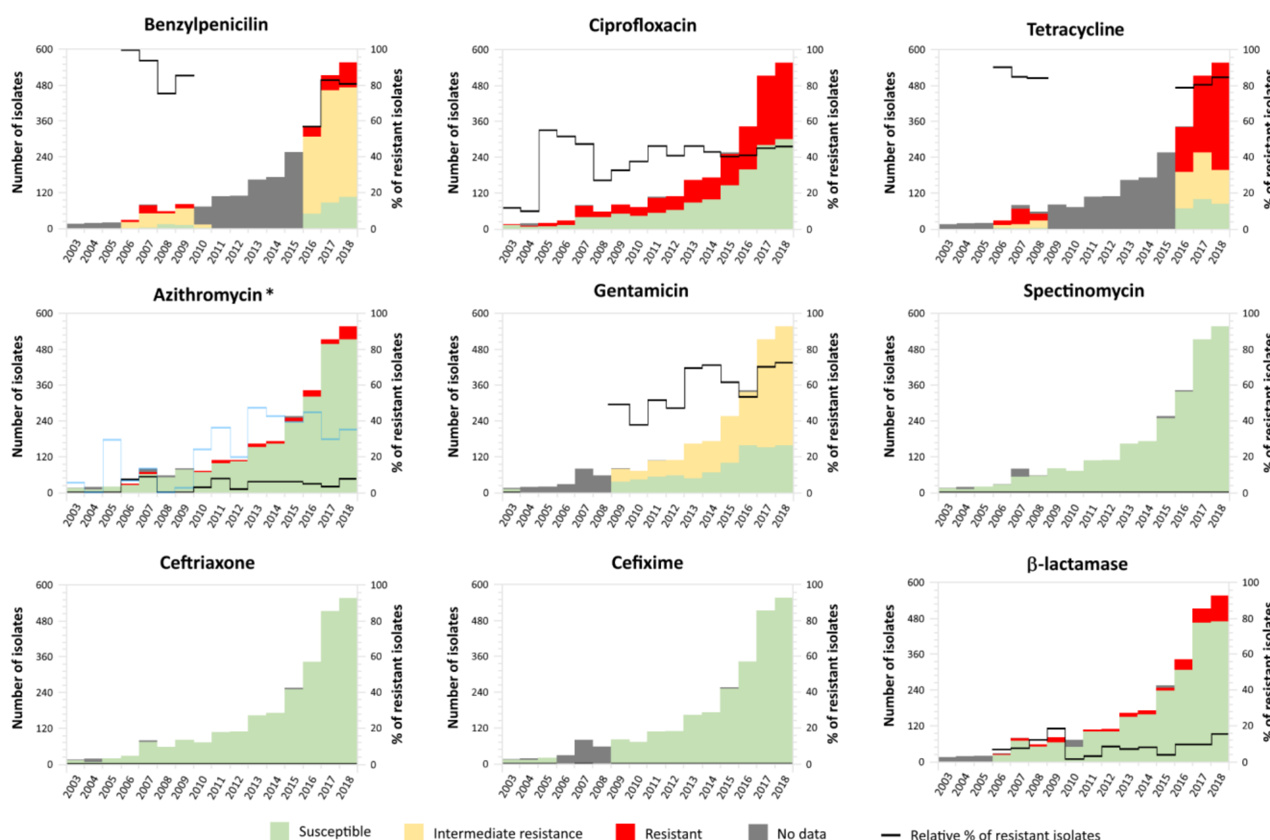

Figure 2.4. | *Neisseria gonorrhoeae* antibiotic susceptibility trends from 2003 up to 2018. Absolute number of isolates by year are relative to the primary *y*-axis. Black lines refer to the relative percentage of resistant isolates (including intermediate resistant) with values' correspondence in the secondary *y*-axis. *blue line represents the percentage of resistant isolates obtained using the previous EUCAST classification for azithromycin. Noteworthy, although the percentage of resistant isolates to benzylpenicillin and tetracycline is high from 2006 to 2010, data correspond to a lower number of isolates when compared with the latter years. Antibiotic resistance results were interpreted according to 2019 EUCAST breakpoint definitions [94] as follows: azithromycin resistant when E-test MIC ≥ 1 mg/L; benzylpenicillin susceptible when MIC ≤ 0.06 and resistant > 1 mg/L; ceftriaxone and cefixime resistant when MIC > 0.125 mg/L; ciprofloxacin susceptible when MIC ≤ 0.03 and resistant > 0.06 mg/L; spectinomycin resistant when MIC > 64 mg/L; and tetracycline susceptible when MIC ≤ 0.5 and resistant > 1 mg/L. Gentamicin isolates were classified based on interpretive results from previous studies [313, 314] as susceptible when MIC ≤ 4 and resistant when > 16 mg/L, for analysis purposes.

2.4.3. Genetic diversity of *Neisseria gonorrhoeae* isolates

Molecular typing results performed on 1042 isolates, from 2003 up to 2017, revealed the existence of 391 distinct ST (**Supplementary Table S2.1.**), that could be classified into 41 distinct NG-MAST genogroups (composed of more than five isolates) representing 76.2% of all typed isolates. This reflects the high genetic diversity of *N. gonorrhoeae* and data show that throughout the years the five most common genogroups found in this dataset were G1407 (n=135), G7445 (67), G225 (66), G2 (58) and G1034 (42) (**Figure 2.5.**), which correspond to one-third of all STs. Moreover, each of these genogroups are composed by isolates collected from at least six different years, with data suggesting that G1407 is the more persistent genogroup in

Portugal, being detected from 2007 up to 2017. In contrast, the second most prevalent genogroup, G7445, seems to have been introduced later (2011) and was frequently detected in 2014 (**Figure 2.5.**).

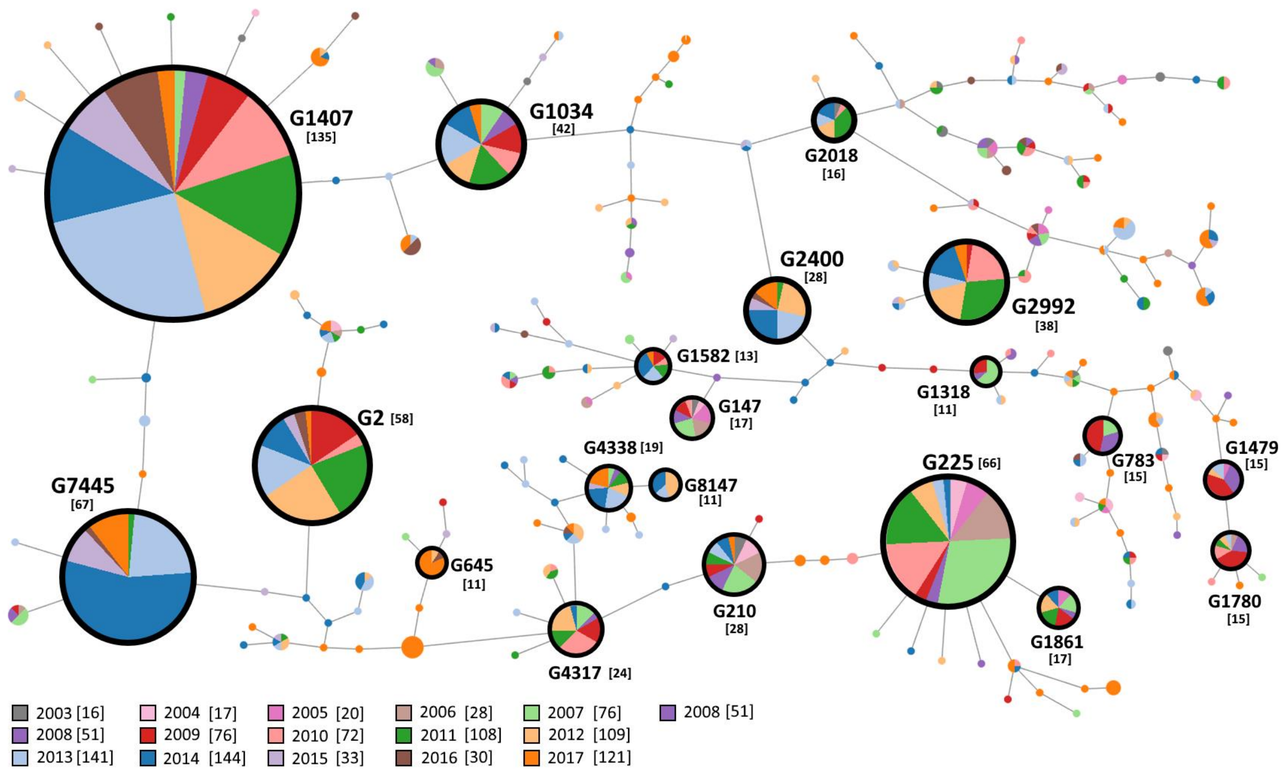

Figure 2.5. | Minimum spanning tree of 1042 concatenated and aligned sequences of *Neisseria gonorrhoeae* Multi-Antigen Sequence Types (NG-MAST). Nodes with sequences with more 99% identity have been collapsed in order to reflect NG-MAST genogroups. Each node is coloured by the collection year of each isolates and NG-MAST Genogroups represented by more than 10 isolates are highlighted by black circles. Number in brackets correspond to the number of isolates in each genogroup and year.

2.5. Discussion

Before 2010, INSA only received strains sporadically, with the vast majority hailing from the Lisbon Metropolitan Area, which hampered accurate data acquisition and reporting on antimicrobial susceptibility in Portugal. Since then, isolates collection has grown considerably, reaching now geographical coverage of all regions of the country. Although laboratory participation in the PTGonoNet remains voluntary, data evidence that healthcare professionals are increasingly aware of the need to monitor the AMR of *N. gonorrhoeae* alongside diagnosis. The participation provides them with the opportunity to become aware of the resistance profile of their isolates and to integrate their data at a national level through the collection. Results show that the vast majority of isolates are collected in the two major metropolitan areas of Portugal, located in Lisbon and Oporto, where the major sexual networks are likely sited. This follows the trends reported by the DGS [127, 129], as these two metropolises are the main entry points into

Portugal, present the highest population density and display prominent tourism. In terms of gender, age and sexual orientation, isolates collected from men aged between 20 and 30 years are still overrepresented, which is in accordance with recent data reported by the DGS [127, 129], the ECDC [125, 150] and the CDC [137]. Moreover, data show that isolates from MSM represent a high percentage of the collection, although most patients do not declare their sexual orientation. Conversely, the observed underrepresentation of women can rely either on men being usually more symptomatic than women [22] or on the potential higher rate of transmission within MSM communities.

Molecular typing data based on NG-MAST show that throughout the studies period, a vast number of ST have been detected, reflecting the genetic diversity of *N. gonorrhoeae*. This suggests a continuous introduction of novel strains in our country, which is a direct reflection of the sexually transmitted nature of the pathogen and its continuous exchange within sexual networks. Still, the observation that G1407 was the predominant circulating genogroup is consistent with data from 2013 retrieved from across Europe [232], which have reported this genogroup as harbouring a multidrug resistance profile. This is also the case for G1034 (also named G21) which was the predominant genogroup in two European countries (namely Latvia and Slovenia) and G225, which is widespread in Europe [232] and was previously associated with ciprofloxacin resistance in Portugal [315]. In contrast, data suggest that NG-MAST G7445 might be more frequent in Portugal as only one or two isolates were reported in the UK, Sweden and Malta in 2013 [232].

Regarding *N. gonorrhoeae* antimicrobial susceptibility, no cefixime and ceftriaxone resistance was observed, with only sporadic cases reaching the MIC breakpoint value. Although in 2016-2017 no ceftriaxone resistance had been reported by the EURO-GASP [135, 150, 154], novel reports suggest the potential re-emergence of resistant isolates in Europe in 2017-2018 [106, 107]. Nevertheless, this is not the case for cefixime, as a small percentage of isolates (~2%) exhibiting resistance are still reported in Europe [150, 154]. Concordantly, cephalosporin resistance seems to be observed in very few cases by year, as reported in the USA [137], Asia [319] and Africa [320]. Previous data from Europe showed that azithromycin resistance was detected in around 7.5% of isolates [135, 150], with data on the Portuguese isolates presenting the highest values (34.5%) in 2016 [150]. In fact, Portugal would be one of the European countries with higher reported azithromycin resistance, raising concerns of using this antimicrobial as a monotherapy. However, while this was observed within the dataset of 110 isolates reported to the ECDC in 2016, results from the whole Portuguese collection evidence that *N. gonorrhoeae* resistance to azithromycin in Portugal is likely below 10% (as seen in the last eight years). Also, the re-evaluation and reinterpretation of the EUCAST breakpoint for azithromycin, in the beginning of 2019 [94], eliminated the intermediate resistance category, which represented almost one third of the entire Portuguese isolate collection. Noteworthy, two cases of high level of

resistance for azithromycin (i.e., MIC $\geq$ 256) are comprised in our 2015 and 2017 datasets, although these were still fully susceptible to ceftriaxone. Several outbreak-associated high azithromycin resistant isolates were reported in England in 2015, with cases still being detected up to 2018 [156], which raises the concern of their spread in Europe. Still, Portugal needs to establish specific gonorrhoea treatment guidelines as is being done by other countries and the results of the present study should constitute a valid contribution to its elaboration.

Our data show that decreased susceptibility to ciprofloxacin has been constant throughout the study period, with ~40% of isolates displaying resistance. In fact, high level of ciprofloxacin resistance has been reported in Europe since 2006 [312], with the 2016-2017 reports showing that overall resistance is at 46.5% [135, 150]. Resistance to penicillin and tetracycline remain high in Portugal, while no resistance to spectinomycin was observed in our dataset. Nevertheless, as in other countries, spectinomycin is no longer available in Portugal. The gentamicin MIC distribution seems to follow the same pattern as in other European countries [150, 313] and the USA [321]. Although resistance does not show up, MIC values have increased during the study period with the relative number of isolates displaying potential intermediate resistance reaching ~70% within 2017 and 2018. Nonetheless, gentamicin remains an alternative for gonorrhoea treatment [113] as MIC values are still within those capable of effectively clearing genital gonorrhoea [322].

This global overview of the Portuguese *N. gonorrhoeae* AMR spanning fifteen years is unprecedented, as only a small subset of these isolates is reported in the ECDC EURO-GASP since 2009 [103, 150, 154]. These results stand as a snapshot of the actual trends, although potentially biased, as participation in the PTGonoNet is voluntary. However, the centralization of AMR provides the unique opportunity to have normalized countrywide data, obtained in accordance to EURO-GASP requirements, as INSA participates in its EQA program specific for *N. gonorrhoeae*. Moreover, like in many countries, a growing number of gonorrhoea diagnosis relies on molecular methods, hampering the availability of specimens for AMR evaluation. For instance, INSA's NRL laboratory diagnosed 125 and 290 cases (in 2017 and 2018, respectively) of *N. gonorrhoeae* by molecular methods, from which no strain was isolated and consequently not tested for antimicrobial susceptibility. The ongoing voluntary participation in PTGonoNET is allowing the genetic characterization of circulating strains and the continuous monitoring of antibiotic resistance emergence, which potentiates the early detection and facilitates the management of potential outbreaks. The maintenance and the enlargement of PTGonoNet, together with the ongoing centralized AMR testing, would surely strengthen the knowledge on *N. gonorrhoeae* AMR trends in Portugal, facilitating communication and concerted actions involving both National and International health authorities.

Acknowledgements: The author work like to acknowledge: 1) Cristina Furtado for her support in launching PTGonoNet (http://www.insa.min-saude.pt/category/areas-de-atuacao/doencas-infeciosas/rede-nacional-de-vigilancia-laboratorial-de-estirpes-de-n-gonorrhoeae/); 2) the laboratory technicians Lúcia Reis, Maria Arminda Ferreira, Albertina Paulino, and Ana Isabel Salas; 3) the support staff of INSA, in particular Sandra Pimenta, Rute Nunes, Augusta Correia, Susana Melo, Gustavo Duque, Lúcia Vieira, Conceição Moreira, Paula Grilo, Fernanda Martins and Elsa Graça; and, 4) all technicians and pathologists of PTGonoNet participants, in particular to José Manuel Amorim. Thanks to Jacinta Azevedo and Cândida Fernandes for fruitful discussions concerning gonorrhoea clinical presentation, epidemiology and therapeutics.

Funding information: M. Pinto was supported by the Portuguese Science and Technology Foundation (FCT) through grant SFRH/BD/109264/2015.

Availability of data and material: All generated or analyzed data are included in this published article and the supplementary information files.

Conflict of interest: The authors declare that they have no conflict of interest.

CHAPTER III

Neisseria gonorrhoeae clustering to reveal major European WGS-based genogroups in association with antimicrobial resistance

This Chapter was published in

Pinto M, Borges V, Isidro J, Rodrigues JC, Vieira L, Borrego MJ, Gomes JP. (2020) ***Neisseria gonorrhoeae*** **clustering to reveal major European WGS-based genogroups in association with antimicrobial resistance**. *Microbial Genomics* (*in press*). **DOI**: 10.1099/mgen.0.000481.

Personal contributions

MP contributed to the study design, performed wet-lab experiments and research, analyzed the data and wrote the manuscript.

3. *Neisseria gonorrhoeae* clustering to reveal major European WGS-based genogroups in association with antimicrobial resistance

3.1. Abstract

Neisseria gonorrhoeae, the bacterium responsible for the sexually transmitted disease gonorrhoea, has shown an extraordinary ability to develop antimicrobial resistance (AMR) to multiple classes of antimicrobials. With no available vaccine, managing *N. gonorrhoeae* infections demands effective preventive measures, antibiotic treatment and epidemiological surveillance. The latter two are progressively being supported by the generation of whole-genome sequencing (WGS) data on behalf of national and international surveillance programmes. In this context, this study aims to perform *N. gonorrhoeae* clustering into genogroups based on WGS data, for enhanced prospective laboratory surveillance. Particularly, it aims to identify the major circulating WGS-genogroups in Europe and to establish a relationship between these and AMR. Ultimately, it enriches public databases by contributing with WGS data from Portuguese isolates spanning 15 years of surveillance. A total of 3791 carefully inspected *N. gonorrhoeae* genomes from isolates collected across Europe were analysed using a gene-by-gene approach (i.e., using cgMLST). Analysis of cluster composition and stability allowed classifying isolates into a two-step hierarchical genogroup level determined by two allelic distance thresholds revealing cluster stability. Genogroup clustering in general agreed with available *N. gonorrhoeae* typing methods (i.e., MLST, NG-MAST and PubMLST core genome groups), highlighting the predominant genogroups circulating in Europe, and revealed that the vast majority of the genogroups present a dominant AMR profile. Additionally, a non-static gene-by-gene approach combined with a more discriminatory threshold for potential epidemiological linkage enabled us to match data with previous reports on outbreaks or transmission chains. In conclusion, this genogroup assignment allows a comprehensive analysis of *N. gonorrhoeae* genetic diversity and the identification of the WGS-based genogroups circulating in Europe, while facilitating the assessment (and continuously monitoring) of their frequency, geographic dispersion, and potential association with specific AMR signatures. This strategy may benefit Public Health actions through the prioritization of genogroups to be controlled, the identification of emerging resistance carriage, and the potential facilitation of data sharing and communication.

Keywords: *Neisseria gonorrhoeae*; Antimicrobial resistance; Whole-genome sequencing; Molecular epidemiology

3.2. Introduction

Gonorrhoea disease burden has increased in the last decade and remains a major public health concern worldwide. From 2012 up to 2016, the World Health Organization (WHO) observed an increase from 78.3 million to 87.0 million new cases of gonorrhoea worldwide [121, 323], with the European Centre for Disease Prevention and Control (ECDC) reporting around 90 000 notified cases in 2016 across 27 EU/EEA countries [135]. This sexually transmitted disease, caused by the bacterium *Neisseria gonorrhoeae*, usually affects the urogenital tract, causing urethritis in men and cervicitis in women, resulting in significant morbidity, but can also affect the anal canal and oropharyngeal tract causing proctitis and pharyngitis, respectively. Linked to disease burden, treatment strategies are being continuously challenged due to *N. gonorrhoeae*'s ability to acquire antimicrobial resistance (AMR) to multiple classes of antimicrobials, including β-lactams, tetracyclines, macrolides and quinolones, after the acquisition of genes or mutations conferring resistance [100, 120, 192, 232]. During the last decade, there has been a steady rise in minimum inhibitory concentrations (MICs) for cephalosporins, with resistant strains and clinical failures being described [102–104, 106, 107, 149, 324]. Worldwide increase in azithromycin resistance is also being reported [150–154], with resistant strains being associated with recent outbreaks [155, 156, 158].

With these last lines of treatment seemingly failing, and the fear that gonorrhoea might become untreatable, reaching a "superbug" status, the WHO, the ECDC and the Centers for Disease Control and Prevention (CDC) issued action plans since 2012, raising the awareness of healthcare professionals to define and monitor treatment failures of gonorrhoea worldwide [100]. In the absence of a vaccine, the management of *N. gonorrhoeae* mainly relies on preventive measures, an effective antibiotic treatment and epidemiological surveillance. As such, whole-genome sequencing (WGS) data are now being generated in order to monitor global *N. gonorrhoeae* epidemiology and AMR trends, with several countries already performing WGS on behalf of national or international surveillance programmes [130]. Nevertheless, during the last years, *N. gonorrhoeae* molecular epidemiology studies, using WGS data, have been mostly focused on the reporting of circulating genogroups defined by typing strategies relying on just a few genomic loci, namely the traditional two genes *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) scheme, the seven-loci-based Multi-Locus Sequence Typing (MLST), and the seven-loci *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) [303]. This has been performed, for instance, to group isolates based on the most prevalent sequence types (STs) [232] or to cluster isolates based on allelic differences of both NG-MAST genes [208, 231]. Therefore, these approaches do not take advantage of the full potential of using whole-genome data for typing as well demonstrated by the recent publication of insightful WGS-based studies relevant for the understanding of *N. gonorrhoeae* evolution [21], population

structure [311] and transmission of AMR determinants [300]. Nevertheless, *N. gonorrhoeae* exhibits a non-clonal population structure [325], due to frequent intra-species horizontal gene transfer, leading to the need of comprehensive identification and tracking of circulating lineages or genetic groups for this pathogen [21, 311]. Achieving this goal has been a recurring challenge in molecular epidemiology due to the lack of approaches to quantitatively evaluate cluster stability and therefore identify genetic thresholds that define stable clusters representing main circulating lineages. In this context, the present study, focused on the European panorama (> 3000 isolates from 21 countries), aims to perform WGS-based *N. gonorrhoeae* genogroup clustering for prospective WGS-based surveillance. Particularly, we aimed to identify the major circulating WGS-genogroups in Europe, assess their geographic spread and search for potential relationships between particular genogroups and specific AMR signatures. To achieve this, we used a large dataset of European gonococcal sequences available in public datasets, now enriched with 600 additional isolates spanning fifteen years of *N. gonorrhoeae* surveillance in Portugal.

3.3. Methods

3.3.1. Selection and characterization of Portuguese isolates

N. gonorrhoeae isolates analysed in the present study are part of the ongoing activity of the Portuguese National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNet) hosted at the National Reference Laboratory for Sexually Transmitted Infections of the Portuguese National Institute of Health (NRL) [326]. Six hundred Portuguese isolates (herein designated as "PT isolates") collected from 2003 up to 2017 from distinct specimens, spread across the country, and presenting different antimicrobial susceptibility profiles were selected for analysis. Regarding antimicrobial susceptibility testing, MICs for azithromycin, benzylpenicillin, cefixime, ceftriaxone, ciprofloxacin, gentamicin, spectinomycin and tetracycline were determined by E-test (BioMérieux, France), as previously described [326]. For 117 isolates from 2006 up to 2010, no MIC values were available for some antibiotics, as antibiotic susceptibility was determined by agar dilution breakpoint technique at the time by the Public Health England, which provided solely qualitative results [312] (**Supplementary Table S3.1.**). All but 65 isolates were tested for β-lactamase production using the chromogenic reagent Nitrocefin (Oxoid, UK) according to the manufacturer's instructions. Isolate antibiotic resistance was classified according to the current European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoint definitions [94]. From the 600 studied PT isolates, 491 isolates were subjected to WGS at the Portuguese NIH, and 109 isolates collected in 2013 were externally sequenced on behalf of the NRL participation in EURO-GASP [232]. Briefly, DNA was extracted from each isolate using the NucliSens easyMAG platform (BioMérieux, France) for total nucleic acid extraction according to the manufacturer's instructions. DNA was then subjected to Nextera XT library preparation (Illumina, USA) prior to

paired-end sequencing (2x250 bp or 2x150 bp) on either a MiSeq or a NextSeq 550 instrument (Illumina, USA), according to the manufacturer's instructions. Coded isolate designations, available anonymized metadata, antibiotic susceptibility data and WGS details for all isolates are reported in **Supplementary Table S3.1**.

3.3.2. Additional dataset for the European context and genome assembly

In order to obtain the current genetic diversity of *N. gonorrhoeae* circulating in European countries and to integrate the novel PT genomes in this global scenario, ultimately strengthening the analyses conducted in the present study, we took advantage of the public availability of 3263 genomes obtained from strains isolated in Europe. As such, WGS reads datasets were download from the European Nucleotide Archive (ENA) from the following BioProjects: PRJEB14933 [155]; PRJEB2124 [270]; PRJEB23008 [271]; PRJEB26560 [106]; PRJEB9227 [232]; PRJNA275092 [259]; PRJNA348107 [281]; PRJNA473385 [260]; PRJNA315363 [272]. Available data for the isolates used, including sample accession numbers, are detailed **Supplementary Table S3.2**.

All genome sequences were assembled using the INNUca v4.0.1 pipeline (https://github.com/B-UMMI/INNUca), an integrative bioinformatics pipeline for read quality analysis and *de novo* genome assembly [310]. Briefly, read quality analysis and improvement is performed using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic v0.36 [327] (with sample-specific read trimming criteria determined automatically based on FastQC report), respectively. Genomes were assembled with SPAdes v3.11 [328] and subsequently polished using Pilon v1.18 [329], with QA/QC statistics being monitored and reported throughout the analysis. For all isolates, assembly statistics including final genome assembly sizes, number of contigs and mean depth of coverage values are reported in **Supplementary Table S3.1.** and **Supplementary Table S3.1**. Only 28 out of the 3861 genomes (two PT and 26 publicly available genomes) were excluded after assembly, due to genome size larger than expected, mean depth of coverage below 15x or the detection of contamination.

3.3.3. *In silico* typing and antimicrobial resistance prediction

For all validated assemblies, *in silico* seven loci MLST prediction was performed using *mlst* v2.4 software (https://github.com/tseemann/mlst), which is integrated within the INNUca pipeline. NG-MAST was performed using the *ngmaster* v0.4 software [286], with novel alleles and sequence types being assigned after submission to the international database of the NG-MAST website (http://www.ng-mast.net/). Additionally, *rplF* fine-typing [330] was performed in order to confirm *Neisseria* species, upon query to the PubMLST *Neisseria* database (http://pubmlst.org). Identification of antimicrobial resistance determinants *in silico* was performed using ARIBA

v2.12.2 with the NG database [302], with all known variants analysed reported in **Supplementary Table S3.1.** and **Supplementary Table S3.2**. Additional variants not classified in the ARIBA database as "known variants" were also inspected [96, 192], namely: A39T and R44H in *mtrR*, 91K in *parC* and -35Adel in *mtrR* promoter. We also performed isolate typing with NG-STAR [303], upon query to the PubMLST *Neisseria* database, for interpretation purposes. Still, the results of this study were oriented towards the results provided by ARIBA, as it provides the ability to inspect more loci and, more importantly, variants that are carried heterogeneously.

3.3.4. Gene-by-gene analysis

Gene-by-gene analysis was performed by taking advantage of the publicly available panel of 1649 loci from the PubMLST *Neisseria* database (*N. gonorrhoeae* cgMLST v1.0 [311], available at http://pubmlst.org), which was prepared in the present study for the chewBBACA core suite v2.0.11 [331] using the *PrepExternalSchema* module. This was a quality control step to ensure that all loci are CDS as required by chewBBACA. Allele calling was performed on all assemblies using chewBBACA, with minimum Blast Score Ratio to 0.6 and size threshold adjusted to 0.3 (i.e., alleles with size variation of 30% to be tagged), and a training file generated by Prodigal v2.6.3 using the reference genome NCCP11945 (RefSeq accession number NC_011035). After inspection, the scheme was additionally curated by removing 16 loci that were flagged as "Repeated Loci" after allele calling and another 39 loci that were only successfully called in less than 1% of the genomes (**Supplementary Table S3.3.**). Exact and inferred matches were used to construct an allelic profile matrix, where other allelic classifications (see https://github.com/B-UMMI/chewBBACA/wiki) were assumed as "missing" loci. Assemblies with less than 1545 loci called (~97%) in the scheme were removed which occurred only for 44 assemblies (19 PT and 25 publicly available) under analysis (detailed in **Supplementary Table S3.1.** and **Supplementary Table S3.2.**). Of note, the application of a more stringent cut-off was meant to increase confidence in the cluster stability analysis, although a lower cut-off (namely a "traditional" 95% loci called) can be applied with this adapted scheme.

3.3.5. WGS-based genogroup classification

For WGS-based surveillance purposes, *N. gonorrhoeae* isolates were classified into two-step hierarchical WGS-based genogroups. For this, we used the maximum-shared loci of the PubMLST cgMLST scheme at 100% by all 3791 validated genomes (upon assembly- and cgMLST-based exclusion), which resulted in 822 loci [herein designated as Maximum-shared cgMLST scheme (MScgMLST); Table S4]. From the allelic profile matrix, the goeBURST algorithm [316, 332] implemented in the PHYLOViZ software [333] was used to generate clusters

at all possible allelic distance thresholds [here expressed as the number of allelic differences (AD) over the total number of loci under analysis]. Cluster concordance was then assessed by the Neighbourhood Adjusted Wallace Coefficient (nAWC), where all clusters partitioning at adjacent allelic distance thresholds (i.e., cut-off for allelic differences *n* and *n*+1) were compared using the *comparing_partitions.py* script (https://github.com/jacarrico/ComparingPartitions) [334, 335], as previously described [310, 336]. Following this, we searched for consecutive goeBURST thresholds over which cluster congruence is high (i.e., nAWC plateaus reflecting cluster stability) as a mean to define thresholds (or threshold ranges) useful for longitudinal surveillance, i.e., stable genogroups that represent major circulating lineages [310, 336]. Two stability phases, defined as the first and second intervals at which five or more consecutive thresholds yielded nAWC above 0.99, were used to cluster isolates into low- and high-level WGS-based genogroups, respectively. Each isolate was then classified into a specific genogroup designated as "GX.Y", where X and Y are non-redundant arbitrary three digit numbers associated, respectively, to the high- and low-level WGS-based genogroup clustering, i.e., constituting two-step hierarchical genogroup definition directly associated with the two determined allelic distance thresholds. We further identified the goeBURST threshold ranges with the highest typing concordance with MLST and NG-MAST classifications using the Adjusted Rand Index (ARI) and Adjusted Wallace Coefficient (AWC) [334, 335]. Using the same rationale, cluster composition at both threshold levels defining WGS-based genogroups were compared with cluster composition defined by NG-MAST and MLST, where the Shannon Index (SI) was used to measure cluster entropy with each partition. This was also performed to compare cluster composition between the MScgMLST scheme and the maximum shared loci obtained when using the cgMLST typing directly on the PubMLST website (i.e., 537 loci), with comparison only performed for the 1977 isolates where allelic profiles were available.

3.3.6. *Neisseria gonorrhoeae* genogroups and genetic clustering analysis

Gene-by-gene allelic profile matrix was used to construct a minimum spanning tree (MST) using the goeBURST algorithm [316] implemented in the PHYLOViZ online web-based tool [333] based on 100% shared loci between all isolates [332]. To take advantage of the maximum number of loci from the scheme (which may be key for discriminating potential epidemiological clusters), we used PHYLOViZ online 2.0 Beta version (http://online2.phyloviz.net/), which allows maximizing the shared genome in a dynamic manner [310, 337]. As such, for each subset of isolates under comparison, the maximum number of shared loci between them is automatically used for MST construction. AD thresholds for cluster inspection are expressed as percentages of AD over the total number of shared loci under comparison for every analysis of each subset of isolates at any given level. As there is no established cut-off range for *N. gonorrhoeae* outbreak detection and contact tracing, we applied a conservative approach to investigate (and characterize at AMR level)

sub-clusters with potential epidemiological links. In summary, after generating a sub-MST for each low-level genogroup, we applied a cut-off of 1.5% ADs, which represents half of the mean percentage of AD observed within each low-level genogroup of the full dataset.

3.3.7. Data availability

All reads generated for the present study were deposited in the ENA under the study accession number PRJEB36482 (individual run accession numbers are detailed in **Supplementary Table S3.1.**). The PubMLST *N. gonorrhoeae* cgMLST v1 scheme adapted for chewBBACA [331], the MScgMLST scheme, as well as allelic profile matrix for both schemes are available at https://doi.org/10.5281/zenodo.3946223. Additionally, a *.json* file is also made available for direct upload into the GrapeTree visualization software [338] in order to interactively explore data/metadata.

3.4. Results

3.4.1. *Neisseria gonorrhoeae* WGS-based genogroups

In the present study, 3791 *N. gonorrhoeae* genomes from isolates collected across Europe (21 EU/EEA countries) were analysed using the PubMLST *N. gonorrhoeae* cgMLST scheme. This scheme, enrolling 1594 loci, yielded 3623 unique allelic profiles (including loci not called). In order to determine WGS-based genogroups for long-term and large-scale surveillance, a shorter scheme was used, based on the loci shared (i.e., called) by all isolates. This scheme comprising 822 loci (**Supplementary Table S3.4.**) yielded 2422 unique allelic profiles (see Supplementary Material). After clustering using goeBURST, cluster composition was compared at all possible allelic distance thresholds using the nAWC (**Figure 3.1.**). Results showed that the two earliest cluster stability points are at 40 AD (4.87%; low-level) and at 79 AD (9.61%; high-level), respectively. Subsequently, we defined WGS-based genogroups at two-step hierarchical clustering levels taking into account these two thresholds. This assignment allows a comprehensive analysis of *N. gonorrhoeae* genetic diversity and the identification of the WGS-based genogroups circulating in Europe, while facilitating the assessment (and continuously monitoring) of their frequency, geographic dispersion, and potential association with specific AMR signatures. As such, we sought to assess the weight of each genogroup (at both high- and low-levels) in the whole dataset. The low-level stability point (i.e., more discriminatory) yielded 321 clusters, of which 208 are represented by one or two isolates and 38 by more than ten isolates, while the high-level cut-off yielded 180 clusters, of which 106 are composed by one or two isolates and 35 by more than ten isolates. Thirty-five high-level (**Table 3.1.**) and 38 low-level genogroups (**Table 3.2.**) composed

of at least ten isolates comprise 91.6 and 84.4% of the whole dataset analysed, respectively. Thus, we opted to perform all downstream analysis focusing particularly on these more represented genogroups. The mean allelic diversity within low-level genogroups composed of at least 10 isolates was of ~3% (AD=25), ranging from 0.2 (AD=2) up to 8.5% (AD=70) (**Figure 3.2.**), while for high-level genogroups it was ~5% (AD=41), ranging from 0.2 up to 10.7% (AD=88). When the number of shared loci was maximized within each genogroup (mean of 1.8-fold increase for all genogroups), the mean allelic diversity for low-level was observed at 3.5%, ranging from 0.2 up to 9.9% (AD=81), and 5.3% for high-level, ranging from 0.2 up to 12.4% (AD=102). These results suggest that diversity within each WGS-based genogroups remains similar when the number of loci under comparison is increased (i.e., when the number of shared loci approaches the entire cgMLST scheme), and that the typing resolution (i.e., degree of genetic relatedness between isolates) is not compromised by using the MScgMLST scheme. For instance, for genogroup G001.002, composed by 491 isolates, when the number of shared loci under analysis increased from 822 to 1408, the mean AD only increased from 4.12 to 4.76%.

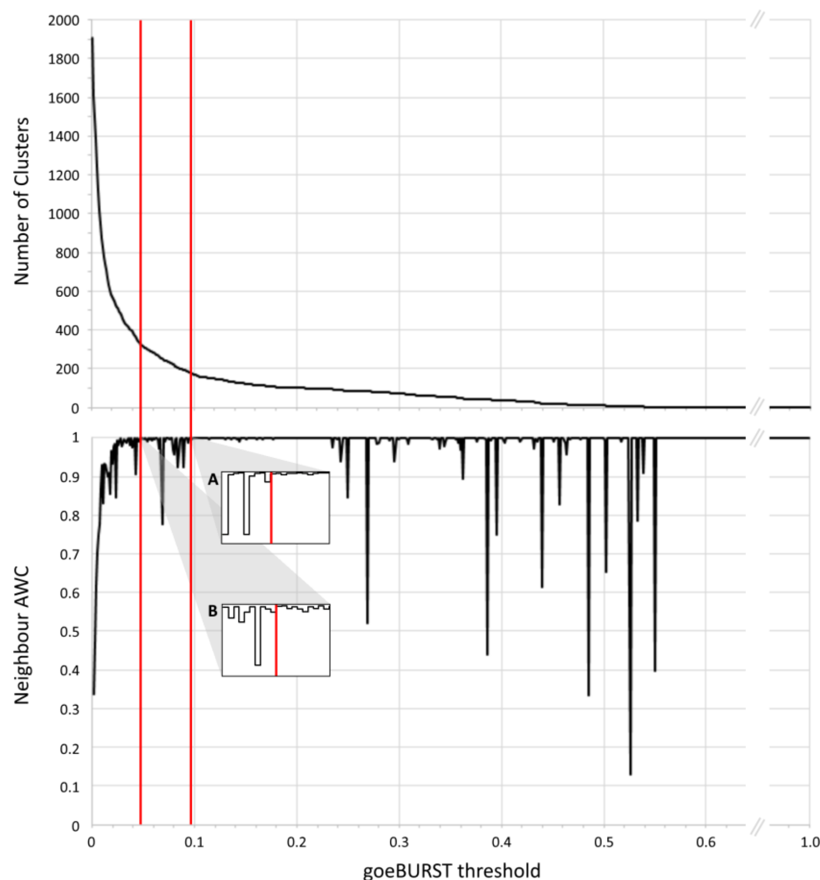

Figure 3.1. | Relationship between cluster stability (assessed by Neighbourhood Adjusted Wallace Coefficient) and goeBURST threshold based on the MScgMLST scheme for all 3791 *Neisseria gonorrhoeae* isolates analysed. **A** – Second point of cluster stability, corresponding to a goeBURST threshold of 0.096 (i.e., 79 allelic differences). **B** – First point of cluster stability, corresponding to a goeBURST threshold of 0.048 (i.e., 40 allelic differences).

Table 3.1. | Summary description of the high-level WGS-based genogroups enrolling at least ten isolates.

High-level WGS-based genogroup	Most common MLST ST (%)	2nd most common MLST ST (%)	Most common NG-MAST ST (%)	2nd most common NG-MAST ST (%)	Most common NG-STAR ST (%)	2nd most common NG-STAR ST (%)	# of enrolled isolates	# of enrolled countries	Isolate collection date interval	# of enrolled countries with at least 5 isolates
G001	9363 (38.2)	1580 (18.6)	2992 (40.1)	9768 (7.0)	63 (42.6)	301 (8.0)	817	20	2003-2017	15
G002	1901 (82.8)	1579 (6.3)	1407 (48.8)	2212 (3.0)	90 (72.5)	89 (2.3)	574	20	2004-2017	19
G005	1901 (98.5)	11992 (0.4)	225 (67.2)	5967 (3.6)	26 (70.4)	310 (6.9)	274	15	2004-2017	5
G004	7363 (97.6)	1587 (1.4)	2400 (55.9)	6360 (14.2)	158 (74.9)	419 (9.0)	211	17	2010-2017	8
G003	11990 (60.4)	1594 (35.6)	51 (35.1)	25 (34.2)	20 (59.4)	722 (9.4)	202	10	2005-2017	4
G006	1579 (100.0)	---	21 (32.0)	1034 (19.3)	139 (94.7)	856 (2.0)	150	13	2004-2017	7
G007	1918 (99.3)	10305 (0.7)	12 (57.3)	44 (32.9)	239 (89.5)	1510 (10.5)	143	1	1995-2014	1
G010	1584 (86.6)	11172 (9.0)	26 (62.7)	4528 (11.2)	178 (89.6)	410 (3.0)	134	7	2002-2016	3
G012	8122 (100.0)	---	292 (75.2)	210 (21.7)	299 (76.7)	85 (19.4)	129	6	2003-2015	3
G009	7822 (95.0)	11985 (1.7)	4995 (67.2)	10421 (19.3)	416 (89.1)	72 (2.5)	119	11	2012-2017	3
G008	1588 (76.7)	11247 (13.3)	3785 (11.1)	1582 (8.9)	247 (18.9)	870 (13.3)	90	16	2005-2017	6
G011	11516 (95.0)	11980 (5.0)	1780 (58.8)	5793 (16.3)	56 (67.5)	55 (27.5)	80	6	2008-2016	2
G014	8143 (96.7)	11968 (1.6)	5624 (45.9)	9918 (34.4)	426 (42.6)	436 (36.1)	61	13	2012-2017	4
G025	1893 (100.0)	---	8517 (92.9)	5993 (3.6)	142 (100.0)	---	56	3	2011-2016	1
G013	1582 (72.2)	11727 (13.0)	147 (24.1)	2997 (9.3)	186 (72.2)	896 (7.4)	54	9	2003-2014	3
G018	1596 (100.0)	---	384 (40.5)	190 (19.0)	955 (66.7)	307 (26.2)	42	4	2003-2014	1
G019	1588 (100.0)	---	1479 (40.5)	19083 (40.5)	271 (48.6)	433 (48.6)	37	5	2005-2017	2
G015	8135 (97.1)	11997 (2.9)	987 (55.9)	5120 (17.6)	729 (94.1)	New (2.9)	34	10	2013-2017	3
G016	1599 (68.8)	~1599 (31.3)	645 (62.5)	11461 (25.0)	520 (100.0)	---	32	4	2013-2017	3
G017	8163 (58.6)	11975 (31.0)	2 (65.5)	226 (10.3)	84 (86.2)	424 (6.9)	29	7	2005-2015	2
G020	7826 (57.1)	7359 (42.9)	2487 (57.1)	4186 (19.0)	992 (57.1)	231 (38.1)	21	6	2008-2016	1
G021	8156 (100.0)	---	5441 (68.4)	13489 (10.5)	442 (84.2)	982 (5.3)	19	8	2013-2017	1
G022	1892 (100.0)	---	6129 (36.8)	387 (26.3)	563 (94.7)	867 (5.3)	19	5	2008-2013	1
G027	7827 (86.7)	13489 (6.7)	2318 (20.0)	8845 (13.3)	38 (46.7)	1225 (13.3)	15	4	2013-2017	1
G039	8776 (100.0)	---	1285 (100.0)	---	950 (100.0)	---	15	4	2012-2014	1
G023	11986 (100.0)	---	8465 (85.7)	26 (7.1)	162 (85.7)	432 (7.1)	14	2	2012-2015	1
G029	1588 (100.0)	---	10801 (42.9)	11575 (42.9)	969 (50.0)	970 (21.4)	14	2	2013	2
G024	10932 (76.9)	11967 (15.4)	5004 (15.4)	4234 (15.4)	169 (46.2)	388 (30.8)	13	5	2008-2014	1
G031	11956 (91.7)	8130 (8.3)	New (83.3)	7414 (16.7)	2005 (83.3)	949 (16.7)	12	3	2013-2015	1
G026	7367 (91.7)	1579 (8.3)	40 (16.7)	18167 (16.7)	new (33.3)	733 (25.0)	12	2	2003-2009	1
G030	1590 (100.0)	---	684 (45.5)	5519 (18.2)	1559 (45.5)	190 (27.3)	11	2	2004-2013	1
G041	11177 (100.0)	---	1993 (100.0)	---	568 (100.0)	---	11	1	2013	1
G032	1585 (100.0)	---	471 (60.0)	752 (20.0)	153 (80.0)	913 (20.0)	10	3	2004-2013	1
G028	8112 (100.0)	---	8149 (40.0)	5560 (20.0)	352 (60.0)	468 (40.0)	10	2	2008-2017	1
G036	8114 (100.0)	---	4 (40.0)	~69 (30.0)	46 (100.0)	---	10	1	2004-2011	1

MLST – Multi-Locus Sequence Type; ST – Sequence type; NG-MAST – *Neisseria gonorrhoeae* Multi-Antigen Sequence Type; NG-STAR – *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance.

Table 3.2. | Summary description of the low-level WGS-based genogroups enrolling at least ten isolates.

Low-level WGS-based Genogroup	Most common MLST ST (%)	2nd most common MLST ST (%)	Most common NG-MAST ST (%)	2nd common NG-MAST ST (%)	Most common NG-STAR ST (%)	2nd common NG-STAR ST (%)	# of enrolled isolates	# of enrolled countries	Isolate collection date interval	# of enrolled countries with at least 5 isolates
G002.001	1901 (83.9)	1579 (6.6)	1407 (55.8)	2212 (3.4)	90 (82.7)	951 (2.6)	502	20	2007-2017	18
G001.002	9363 (52.7)	11428 (19.0)	2992 (58.5)	3935 (8.8)	63 (69.7)	67 (7.7)	491	19	2009-2017	14
G005.004	1901 (99.0)	11992 (0.4)	225 (67.4)	5967 (3.7)	26 (70.7)	310 (7.0)	273	15	2004-2014	5
G004.003	7363 (97.6)	1587 (1.4)	2400 (56.2)	6360 (14.3)	158 (79.5)	419 (8.6)	210	17	2010-2017	8
G006.007	1579 (100.0)	---	21 (32.0)	1034 (19.3)	139 (94.7)	856 (2.0)	150	13	2004-2017	7
G001.005	1580 (77.0)	8126 (19.6)	9768 (38.5)	359 (18.2)	1996 (34.5)	192 (27.7)	148	7	2003-2017	4
G007.008	1918 (99.2)	10305 (0.8)	12 (62.1)	44 (35.6)	239 (88.6)	1510 (11.4)	132	1	1995-2000	1
G003.006	11990 (93.1)	1594 (4.6)	25 (46.6)	51 (43.5)	20 (65.6)	722 (14.5)	131	5	2008-2017	1
G012.013	8122 (100.0)	---	292 (75.2)	210 (21.7)	299 (76.7)	85 (19.4)	129	6	2003-2015	3
G009.009	7822 (95.0)	11985 (1.7)	4995 (67.2)	10421 (19.3)	416 (89.1)	72 (2.5)	119	11	2012-2017	3
G010.010	1584 (88.1)	11172 (10.2)	26 (70.3)	4528 (12.7)	178 (95.8)	423 (1.7)	118	5	2002-2015	2
G011.011	11516 (95.0)	11980 (5.0)	1780 (58.8)	5793 (16.3)	56 (67.5)	902 (3.8)	80	6	2008-2016	2
G008.012	1588 (80.0)	11247 (20.0)	3785 (16.7)	1582 (13.3)	247 (23.3)	567 (13.3)	60	16	2005-2017	3
G001.015	1580 (59.0)	11999 (39.3)	995 (64.3)	1313 (10.7)	301 (42.9)	245 (42.9)	56	6	2005-2014	2
G025.030	1893 (100.0)	---	8517 (100.0)	---	142 (100.0)	---	52	3	2013-2016	1
G001.024	11864 (98.0)	12521 (2.0)	2992 (78.4)	~5230 (9.8)	439 (86.3)	453 (9.8)	51	3	2013-2015	1
G001.014	9363 (100.0)	---	7445 (76.6)	11500 (6.4)	301 (70.2)	965 (10.6)	47	5	2013-2017	1
G019.022	1588 (100.0)	---	1479 (41.7)	19083 (41.7)	433 (50.0)	271 (47.2)	36	5	2005-2014	2
G015.016	8135 (97.1)	11997 (2.9)	987 (55.9)	5120 (17.6)	729 (94.1)	~729 (2.9)	34	10	2013-2017	3
G002.017	1901 (91.2)	11107 (8.8)	3150 (14.7)	2018 (14.7)	127 (23.5)	150 (17.6)	34	7	2004-2017	2
G016.018	1599 (68.8)	~1599 (31.3)	645 (62.5)	11461 (25.0)	520 (100.0)	---	32	4	2013-2017	3
G014.020	8143 (96.7)	11971 (3.3)	5624 (93.3)	1691 (3.3)	426 (86.7)	890 (6.7)	30	9	2012-2017	3
G013.021	1582 (100.0)	---	147 (36.7)	2997 (16.7)	186 (66.7)	752 (10)	30	5	2003-2013	3
G018.025	1596 (100.0)	---	384 (56.7)	190 (26.7)	955 (93.3)	307 (6.7)	30	3	2003-2014	1
G017.019	8163 (55.6)	11975 (33.3)	2 (70.4)	226 (7.4)	84 (85.2)	424 (7.4)	27	7	2005-2015	2
G003.023	1594 (100.0)	---	51 (63.6)	8148 (22.7)	851 (77.3)	2014 (13.6)	22	2	2005-2017	1
G014.038	8143 (100.0)	---	9918 (95.5)	15909 (4.5)	436 (100.0)	---	22	2	2014-2017	1
G003.036	1594 (100.0)	---	10800 (68.4)	11042 (31.6)	20 (100.0)	---	19	3	2013	2
G021.029	8156 (100.0)	---	5441 (72.2)	13489 (11.1)	442 (88.9)	New (5.6)	18	7	2013-2017	1
G008.026	1588 (100.0)	---	3307 (43.8)	9171 (12.5)	249 (37.5)	53 (25)	16	6	2009-2013	1
G039.037	8776 (100.0)	---	1285 (100.0)	---	950 (100.0)	---	15	4	2012-2014	1
G023.027	11986 (100.0)	---	8465 (85.7)	26 (7.1)	162 (85.7)	432 (7.1)	14	2	2012-2015	1
G029.032	1588 (100.0)	---	10801 (42.9)	11575 (42.9)	969 (50.0)	970 (21.4)	14	2	2013	2
G013.028	11727 (53.8)	7363 (46.2)	1466 (30.8)	4333 (23.1)	186 (84.6)	1347 (7.7)	13	1	2005-2009	1
G020.031	7826 (100.0)	---	2487 (100.0)	---	992 (100.0)	---	12	3	2008-2013	1

MLST – Multi-Locus Sequence Type; ST – Sequence type; NG-MAST – *Neisseria gonorrhoeae* Multi-Antigen Sequence Type; NG-STAR – *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance.

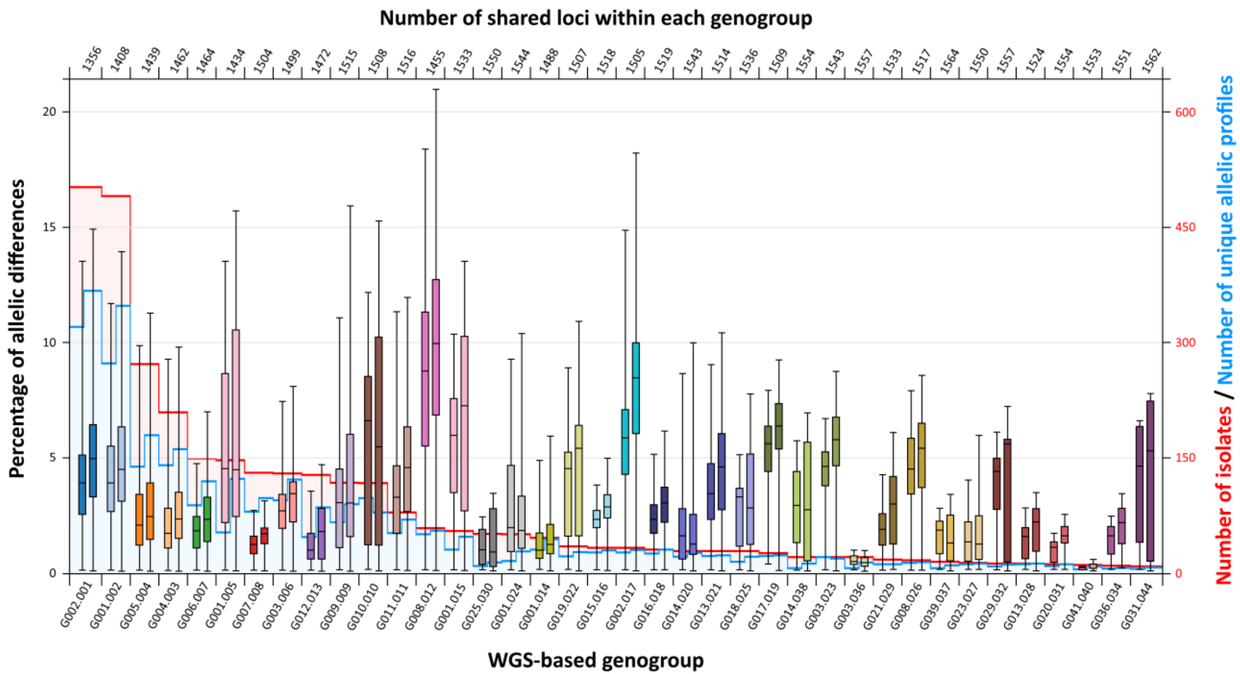

Figure 3.2. | Distribution of the percentage of pairwise allelic differences (AD) between isolates of all low-level WGS-based genogroups represented by at least 10 isolates. Data for each WGS-based genogroup is presented by pair of box plots with the same colour. Left and right box plots refers to AD distribution using the MScgMLST loci scheme or the maximum shared loci of the PubMLST cgMLST scheme within the genogroup, respectively. The respective total number of loci shared is presented on top of the graph. Red line refers to the number of isolates that constitute each low-level WGS-based genogroup, while the blue line refers to the number of unique allelic profiles detected within each genogroup using the two approaches.

3.4.2. From traditional to WGS-based typing

We then compared the relationship between the novel defined WGS-genogroups and the isolates' classification based on traditional typing methods for *N. gonorrhoeae* (**Figure 3.3.**), namely MLST (which yielded ~180 profiles; SI = 5.136) and NG-MAST (which yielded ~800 profiles; SI 7.431). Cluster congruence analysis between MLST types and WGS-genogroups showed that overall cluster congruence is slightly higher with the low-level WGS-genogroups than with the high-level (ARI = 0.584 and 0.524, respectively). Nevertheless, the greater cluster agreement between MLST and WGS-genogroups was observed at the high level, as we observed a 64.2% probability of isolates belonging to the same MLST type to also belong to the same high-level genogroup, and a probability of 54.6% when evaluating the low-level (based on AWC analysis). Still, the goeBURST threshold range that displayed simultaneously higher concordance and agreement with MLST was between 15.9 and 24.9% AD (i.e., 131 and 205 AD), where the highest AWC were observed (between 0.660 and 0.663) with ARI values above 0.520 (up to 0.529) (**Figure 3.4.**). Regarding NG-MAST typing, we observed that the goeBURST threshold range that displayed higher congruence and agreement was between 4.4 and 6.8% AD (i.e., 36 and 56 AD). When compared with MLST analysis, this lower threshold range could be explained by a higher

number of profiles observed in NG-MAST clustering. Nevertheless, although the results showed that the level of entropy in NG-MAST clustering is higher than with the other methods, the probability that a set of strains evidencing the same NG-MAST also belong to the same high- and low-level genogroups was very high (AWC of 0.979 and 0.887, respectively).

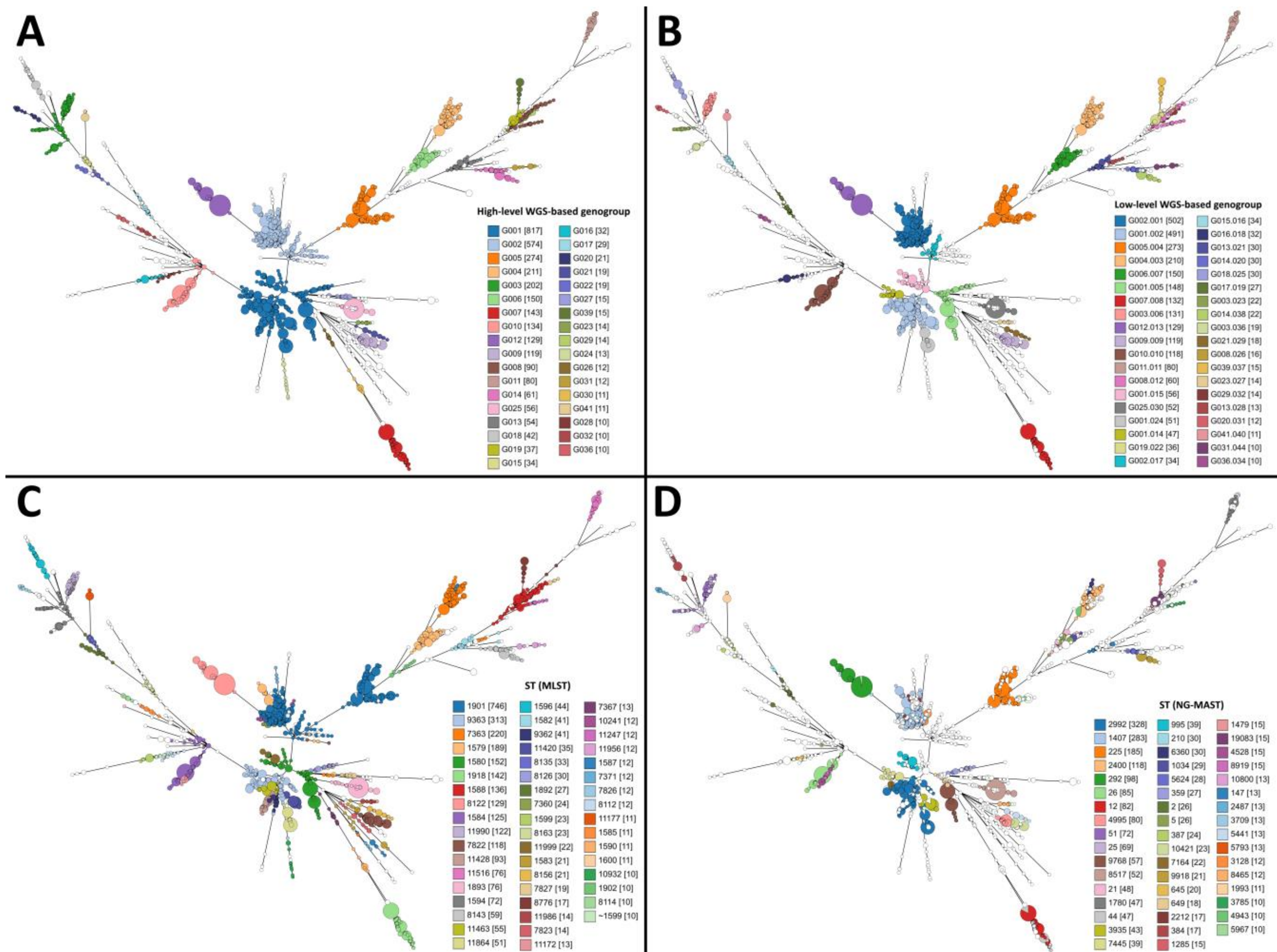

Figure 3.3. | Comparison of WGS-based genogroups, defined at two levels, and both traditional typing methods for *Neisseria gonorrhoeae*. Minimum spanning trees (MST) enrol all 3791 *N. gonorrhoeae* isolates based on the MScgMLST scheme (822 loci). Nodes, corresponding to a unique allelic profile, are coloured according to their corresponding (**A**) high-level WGS-based genogroup, (**B**) low-level WGS-based genogroup, (**C**) ST of the traditional seven loci MLST scheme and (**D**) ST of the two loci NG-MAST scheme. Numbers in parenthesis refer to the number of isolates comprising each genogroup or sequence type. MST were generated using the GrapeTree v1.5.0 software [338].

This overall trend is especially relevant for abundant genogroups (**Figure 3.3.**, **Table 3.1.** and **Table 3.2.**) as there is a strong correspondence between particular MLST and NG-MAST types and high- and low-level genogroups. For example, all 52 isolates from genogroup G025.030 are from MLST ST-1893 and NG-MAST ST-8517. In contrast, isolates from genogroup G006.007 are from MLST ST-1579 but present 25 distinct ST in NG-MAST typing (**Table 3.2.**). Additionally, comparison with PubMLST website results for 1977 isolates (discriminated in **Supplementary Table S3.2.**) showed that cluster congruence was high with the novel defined WGS-genogroups, as ARI values were equal to 0.815 and 0.931 when cluster composition was compared at the high

and low-level thresholds, respectively. Furthermore, we also compared the recently proposed core genome groups [311], defined at 400-locus differences threshold (Ng_cgc_400) using the PubMLST cgMLST scheme, for 1830 isolates for which data were available (discriminated in **Supplementary Table S3.2.**) and were present in our dataset (**Figure 3.5.**). We observed the Ng_cgc_400 groups have a less discriminatory resolution than the two-stable-level WGS-based genogroup defined in the present study (**Figure 3.5.B**). More importantly, results showed that when the same cut-off of ~24% (i.e., 400 AD in 1649 loci PubMLST scheme and 199 AD in 822 loci MScgMLST scheme) is applied to our data there is a remarkable agreement between isolate grouping (**Figure 3.5.A**), reinforcing that our threshold might be scaled to the entire scheme while keeping the robustness of the clustering. In summary, the novel defined genogroups correlated well with MLST and NG-MAST classification, which largely facilitates backwards compatibility in the transition to WGS-based typing, as associations between ST and AMR profiles have been previously reported [232].

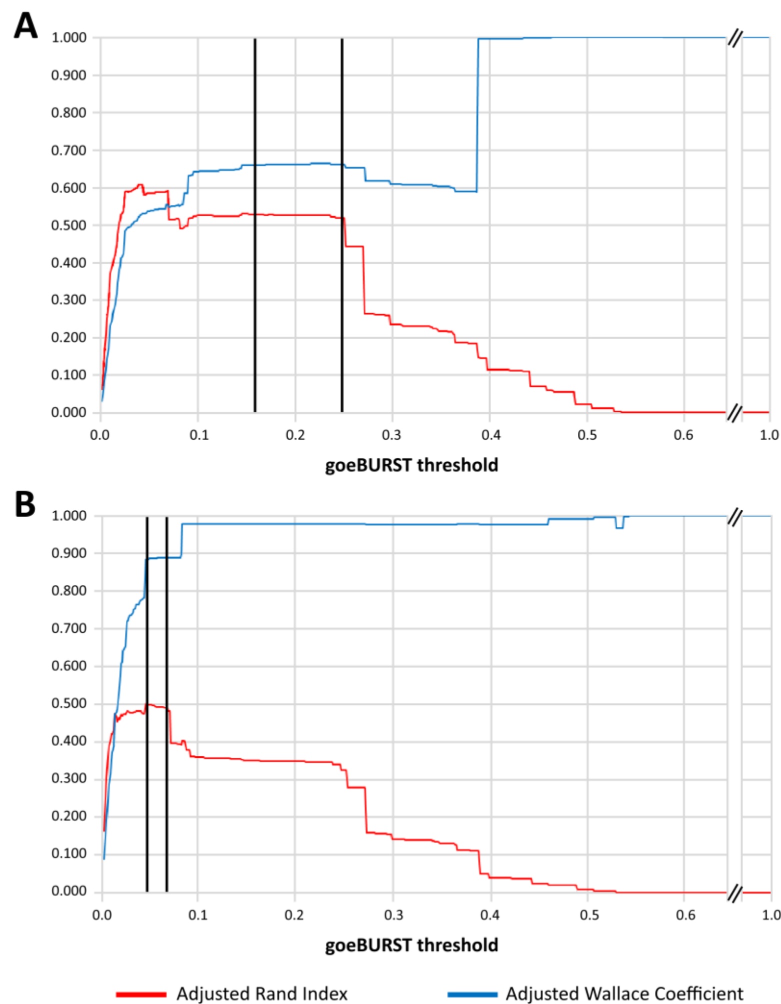

Figure 3.4. | Cluster congruence (Adjusted Rand Index) and agreement (Adjusted Wallace Coefficient) between traditional (**A**) MLST and (**B**) NG-MAST typing and goeBURST clustering at all thresholds (based on the MScgMLST scheme). Adjusted Wallace Coefficient values are directed from traditional typing to geoBURST clustering. Black lines outline the threshold ranges where higher congruence and agreement was simultaneously observed.

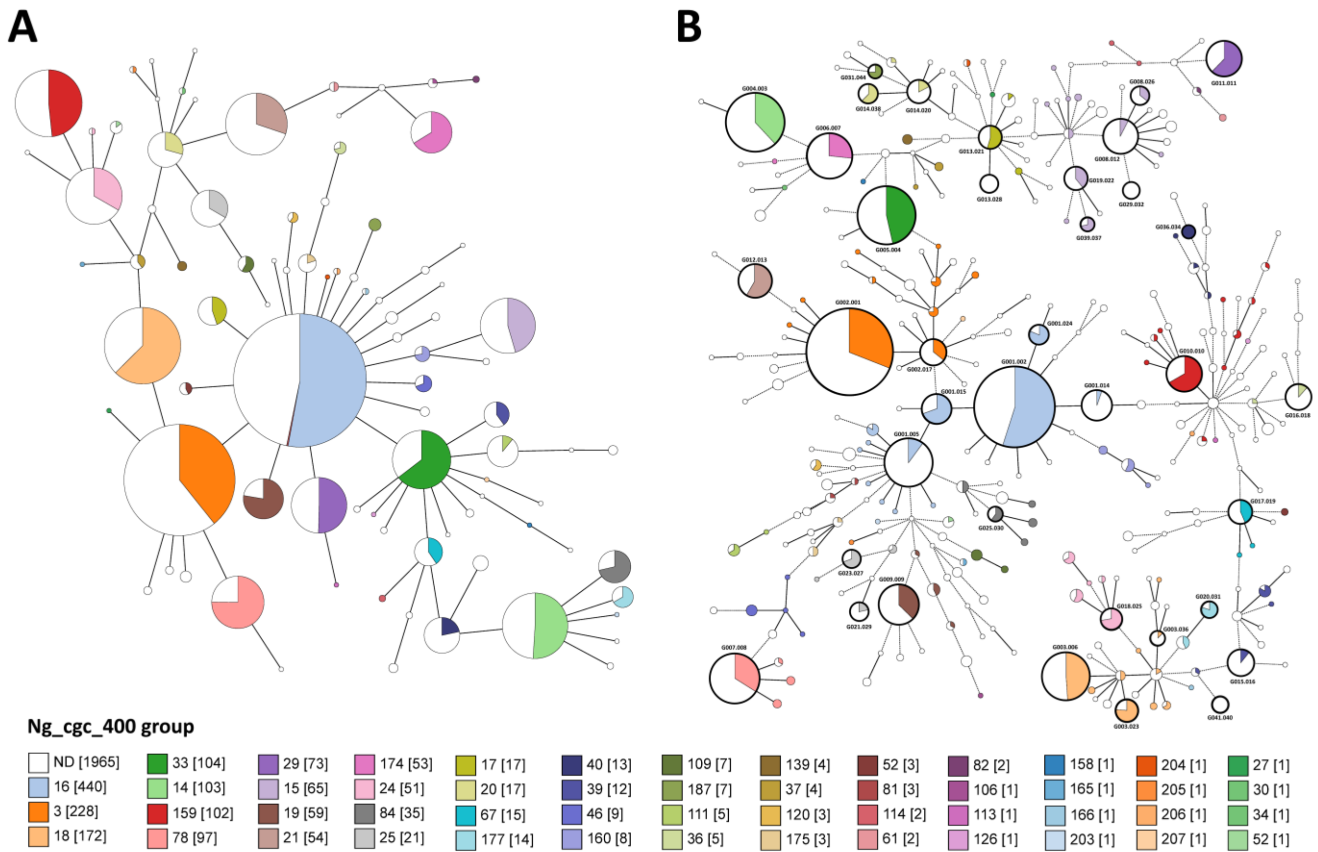

Figure 3.5. | Overlap between *Neisseria gonorrhoeae* WGS-based genogroups defined in this study and PubMLST core-genome groups. Minimum spanning trees (MST) were constructed based on allelic diversity found among the 822 genes shared by 100% of the isolates and nodes are coloured according to Ng_cgc_400 groups [311], for the sub-set of isolates used in both studies (N = 1830). **A** – MST nodes were collapsed at the pubMLST threshold (~24% AD), i.e., 199 AD in 822 loci MScgMLST scheme which correspond to 400 AD in 1649 loci pubMLST scheme, illustrating the huge clustering overlap between both strategies. **B** – MST nodes were collapsed at the low-level WGS-genogroup threshold (4.87%; 40 AD), with straight and dotted lines reflecting nodes linked with the AD below and above the threshold applied for the high-level (9.61%; 79 AD), illustrating that the stable threshold defined in the present study provide a higher discriminatory resolution than Ng_cgc_400. Numbers in parenthesis refer to the number of isolates comprising each Ng_cgc_400. MSTs were generated using the GrapeTree v1.5.0 software [338].

3.4.3. Distribution of *Neisseria gonorrhoeae* genogroups by European countries

All novel PT isolates were integrated within a dataset of publicly available *N. gonorrhoeae* genomes, in order to assess their genomic diversity and phylogenetic relationships within the European circulating strains. As it stands, the dataset is highly represented by isolates from the United Kingdom (n = 2263) and Portugal (n = 579) in comparison with other European countries, some of which have less than ten isolates sequenced at the date of the analysis (e.g., Iceland and Cyprus). Global phylogenetic analysis, based on the MScgMLST scheme, revealed that the PT isolates presented high genetic diversity, being dispersed within the European genomic panorama (**Figure 3.6.**). PT isolates belonged to 115 distinct WGS-based genogroups, of which 32 were mainly (>50%) constituted by PT isolates (e.g., G001.014, G016.018 and G013.028) and 60

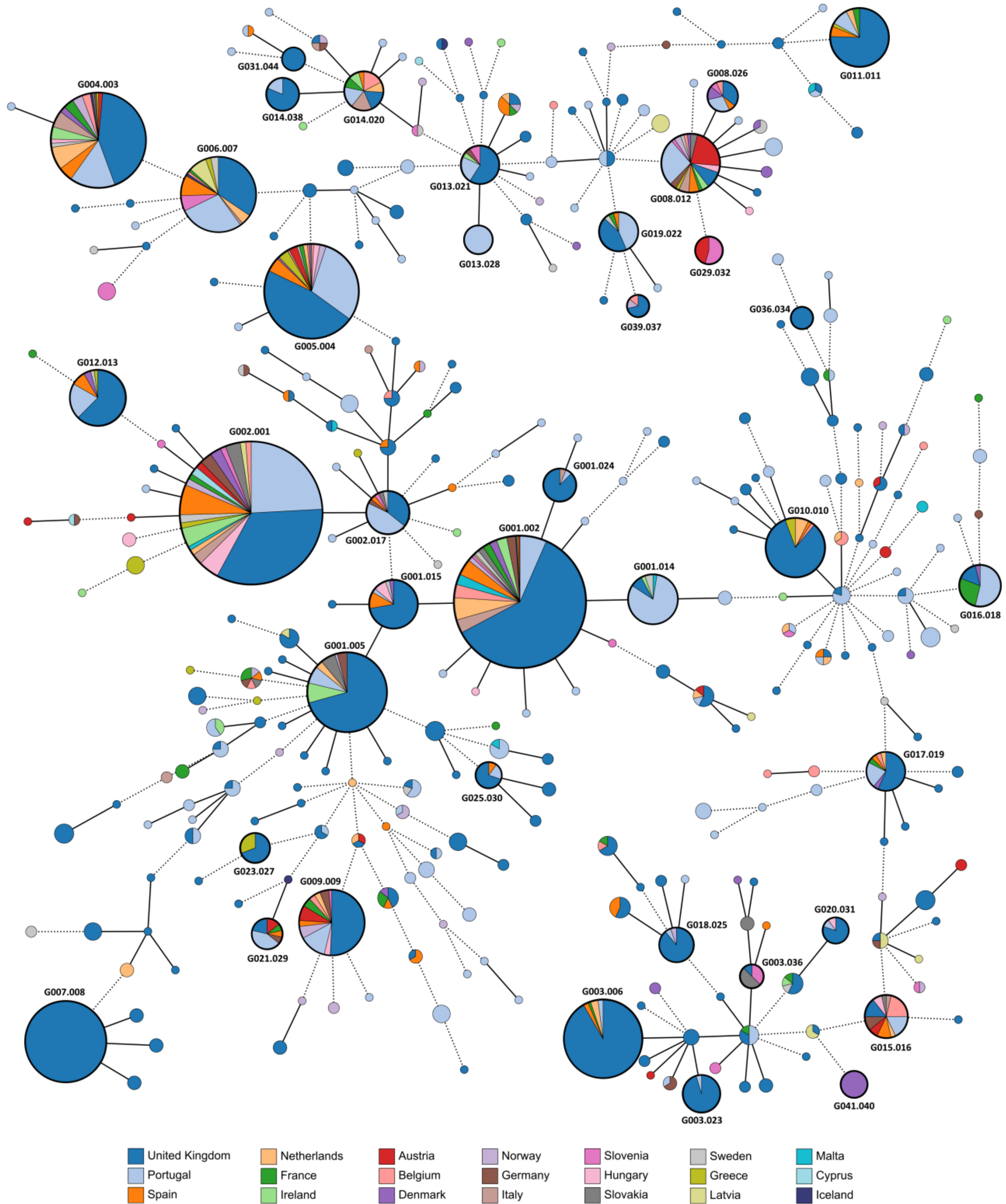

Figure 3.6. | Phylogeny of 3791 *Neisseria gonorrhoeae* isolates from Europe, based on a gene-by-gene approach using the MScgMLST scheme. Minimum spanning tree (MST) was constructed based on allelic diversity found among the 822 genes shared by 100% of the isolates. All nodes (which represent a unique allelic profile) presenting an allelic distance (AD) below 40, corresponding to the low-level genogroup threshold, have been collapsed for visualization purposes. Nodes are coloured according to different countries of origin. Straight and dotted lines reflect nodes linked with the allelic distances below and above the threshold applied for high-level WGS-based genogroup definition (79 AD), respectively. Low-level WGS-based genogroups comprised of more than ten isolates are highlighted by thicker black circles. MST was generated using the GrapeTree v1.5.0 software [338].

contained a single isolate. This is a reflection of *N. gonorrhoeae*'s global genetic diversity as a total of 161 low-level genogroups were represented by a single isolate, and 81 genogroups were composed by isolates from a single country (**Figure 3.7.**), for example, G007.008 (UK), G013.028 (Portugal), G041.40 (Denmark), G031.044 (UK), G036.034 (UK), G046.056 (Greece), G053.086 (Slovenia), G170.243 (Latvia), G002.050 (Portugal). Results showed that the more geographically widespread and abundant high-level genogroups were G001 and G002, enrolling strains from 20 countries, followed by G004, G008 and G005 enrolling isolates from 17, 16 and 15 countries, respectively (**Table 3.1.** and **Figure 3.6.**).

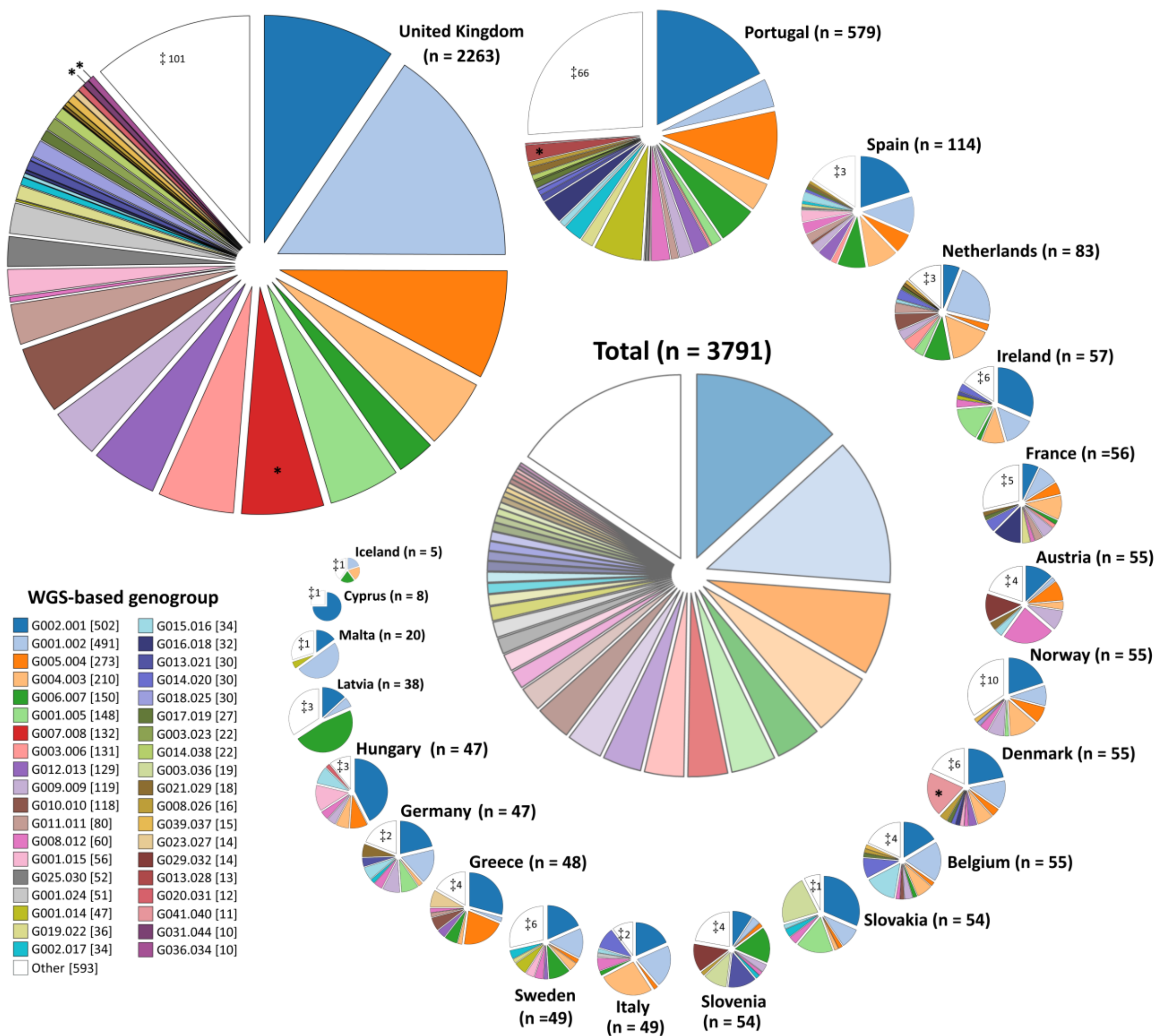

Figure 3.7. | Distribution of low-level WGS-based genogroups by country. Only genogroups with at least 10 isolates are coloured. Central pie chart refers to the total distribution of genogroups within the analysed dataset. *WGS-based genogroups only found in one country, represented by at least 10 isolates. ‡WGS-based genogroups only found in one country, represented by less than 10 isolates (adjacent number refers to the number of distinct genogroups).

Even when fragmented into low-level genogroups, these were still within the more prevalent genogroups. In fact, eleven low-level genogroups were represented by isolates of at least six countries, of which three were composed by at least five isolates, namely G002.001 (NG-MAST ST1407), G001.002 (ST 2992), G005.004 (ST 225), G004.003 (ST 2400), G006.007 (ST 21), G001.005 (ST 9768), G012.013 (ST 292), G009.009 (ST 4995), G008.012 (ST 3785), G015.016 (ST 977), G014.020 (ST 5624), with isolates being detected with time intervals of five up to 14 years apart (**Table 3.2.**).

3.4.4. WGS-based genogroups carrying antibiotic resistance determinants

We analysed 32 genetic determinants known to be involved in decreased susceptibility and resistance to antimicrobials in *N. gonorrhoeae* (namely for penicillin, tetracycline, ciprofloxacin, azithromycin, cephalosporins, spectinomycin, sulphonamides and rifampicin), in order to assess their relationship with WGS-based genogroups at high- and low-levels (**Figure 3.8.**). Overall, results showed two distinct sets of isolates displaying a contrasting pattern of AMR, i.e., one with isolates carrying several genetic determinants related with decreased susceptibility and resistance (set A) and another with more susceptible isolates (set B). In fact, 15 out of 19 high-level genogroups from set A contained isolates independently resistant to at least four distinct antimicrobial drugs, contrasting with only one out of 16 high-level genogroups from set B (**Figure 3.8.A**). The same trend was observed at the low-level, with 19 out of 24 genogroups from set A and only one out of 14 genogroups from set B (**Figure 3.8.B**). In order to understand if these resistant markers were carried together by the same isolates (i.e., multidrug resistance) we analysed the most dominant AMR profiles observed within each low-level genogroup (**Figure 3.9.**), as well as the dominant NG-STAR ST (**Table 3.2.**). We observed that for nine genogroups all isolates displayed the same specific AMR profile. Furthermore, results showed that there were 21 genogroups where the most dominant AMR profile was not the one exhibiting the largest repertoire of AMR determinants. In contrast, for eight genogroups, the dominant AMR profile was the one associated with more AMR determinants. Several scenarios can justify both observations, but sampling bias hampers a more in depth analysis of the evolutionary and epidemiological trajectory of the involved genogroups.

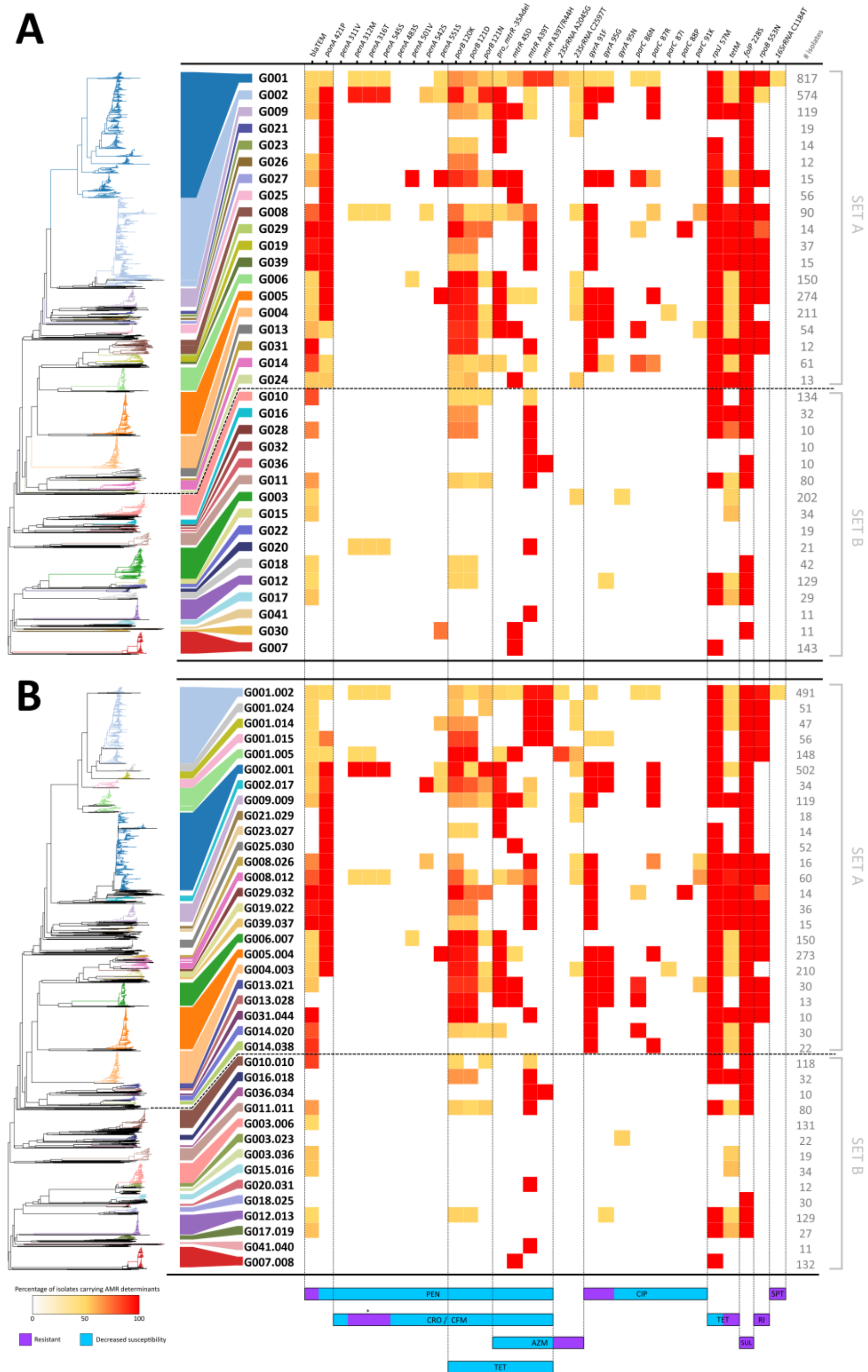

Figure 3.8. | Heatmap distribution and occurrence of the genetic determinants involved in antimicrobial resistance (AMR) by high- (**A**) and low-level (**B**) WGS-based genogroups. Genetic determinants are ordered by the affected antimicrobial drug class/antibiotic, with resistance or decreased susceptibility effect described at the bottom. Heatmap colour range correlates with the percentage of isolates carrying each genetic determinant within a given WGS-based genogroup. Number of isolates within each genogroup are presented on the right of each panel. The contextual neighbour-joining phylogenetic tree at the left side of each panel was generated based on the MScgMLST allelic profiles using the GrapeTree v1.5.0 software [338]. *Combination of these three mutations have been proposed as potentially inducing resistance to cephalosporins [232]. PEN – penicillins; TET – tetracycline; CIP – ciprofloxacin; AZM – azithromycin; CRO – ceftriaxone; CFM – cefixime; SPT – spectinomycin, SUL – sulphonamides; RI – rifampicin; AMR – Antimicrobial resistance.

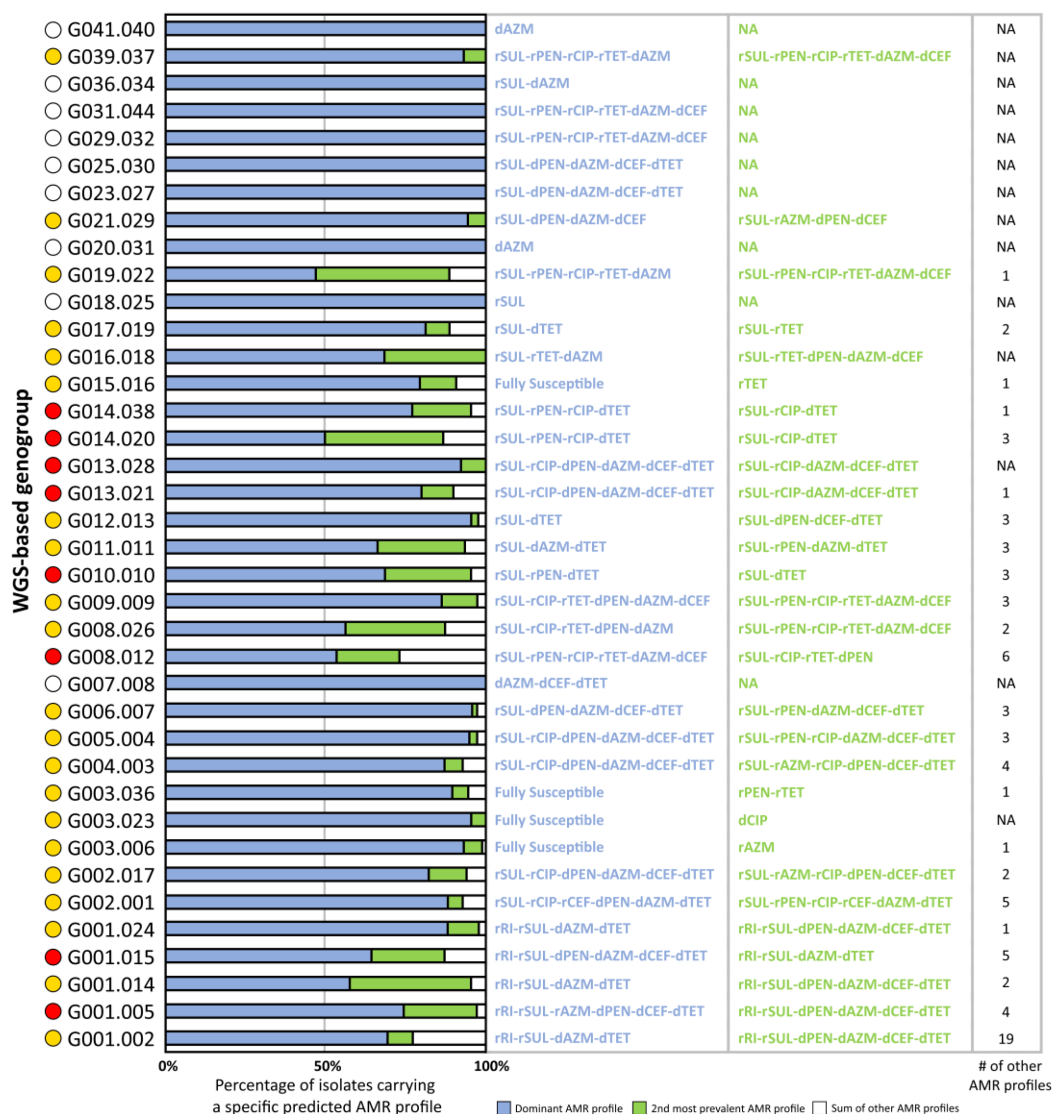

Figure 3.9. | Major predicted antimicrobial resistance (AMR) profiles observed within each low-level genogroup. Yellow circles indicate genogroups where the dominant AMR profile exhibits less genetic determinants associated with resistance than the second observed profile, for multiple classes of antimicrobials. Red circles indicate genogroups where the dominant AMR profile exhibit more genetic determinants associated with resistance for multiple classes of antimicrobials. White circle indicate genogroups with a unique AMR profile. PEN – penicillins; TET – tetracycline; CIP – ciprofloxacin; AZM – azithromycin; CEF – cephalosporins; SUL – sulphonamides; RI – rifampicin. rXXX – Resistant to; dXXX – Decreased susceptibility to.

Regarding specific genetic alterations, for both genogroup levels, resistance to ciprofloxacin was only observed in set A (with exception of one set B G012 resistant isolate from 2013), with all isolates from five different high-level genogroups simultaneously harbouring the resistance-mediating 91F and 95G amino acids in DNA-gyrase subunit A (encoded by *gyrA*; NG-STAR allele 1), and eight other displaying at least one of them. The same trend was observed for five mutations in the topoisomerase IV subunit C encoding-gene *parC* associated with decreased susceptibility to quinolones, which were exclusive of set A isolates (albeit dissimilarly spread across distinct high-level genogroups) (**Figure 3.8.A**). At the low-level (**Figure 3.8.B**), 15

genogroups were almost fully composed by ciprofloxacin-resistant isolates, with two others having only a few resistant isolates (G001.002 with 15 isolates from 2013 to 2015 and G001.0015 with one isolate from 2013). Amino acid alteration 553N in the RNA polymerase subunit B (*rpoB*), conferring resistance to rifampicin [96] was exclusively observed in set A. Likewise, the azithromycin resistance-associated mutation C2597T in *23SrRNA* was mostly observed in nine genogroups of the set A. Furthermore, only isolates from high-level genogroup G001 (**Figure 3.8.A**) carried the mutation A2045G (NG-STAR allele 1), conferring high-level resistance to azithromycin. More importantly, this seemed to be exclusive to low-level genogroups G001.005 and G001.002 (**Figure 3.8.B**), with isolates having a fixed mutation (i.e., present in all *23SrRNA* copies) or carrying it heterogeneously. Spectinomycin resistance-associated mutation in *16SrRNA* was only observed in one isolate from Sweden (G001.002) carrying it heterogeneously. Furthermore, resistance-associated markers for sulphonamides (dihydropteroate synthase encoding gene *folP*), tetracycline (presence of the *tetM*-carrying conjugative plasmid) and penicillin (presence of the *blaTEM* plasmid, associated with β-lactamase production) were widespread across most of the genogroups. Still, all markers seemed to be more frequent in genogroups from set A. For instance, all of them carried the mutation 228S in *folP*, while *blaTEM* and *tetM* were not found in only three and four high-level genogroups, respectively (**Figure 3.8.A**). Regarding the putative resistance to cephalosporins, results showed that most isolates from G002.001 (494 out of 502, ranging from 2007 to 2017) simultaneously possessed the three mutations (i.e., 312M, 316T and 545S) in the penicillin-binding protein 2 encoding gene (*penA*) that are known to potentially mediate resistance [232, 339], 97.0% of which precisely carry the mosaic *penA-XXXIV* allele (NG-STAR allele 266, mostly associated with ST 90) [105]. Besides G002.001 isolates, this SNP profile was only detected in 24 other isolates in the whole dataset (from 2013 to 2016), representing eight different genogroups). The whole *penA-XXXIV* allele was observed in 12 out these 24 isolates, which is suggestive of acquisition by recombination. Of note, other mutations in *penA* (inducing decreased susceptibility to penicillins and cephalosporins) were only observed in set A isolates. Concerning genetic markers associated with decreased susceptibility to penicillin, the mutation 421P in the penicillin-binding protein 1 encoding gene (*ponA*) was exclusively present in set A genogroups (**Figue 3.8.**); this mutation was carried by > 90% of isolates from 14 distinct low-level genogroups. Three SNPs associated with decreased susceptibility to penicillins, cephalosporins and tetracycline, targeting the major porin encoding-gene *porB* (type porB1b), were observed in 1590 isolates from set A contrasting with only 26 isolates from set B. Likewise, genetic alterations in the MtrR transcriptional regulator or its promoter, disrupting the MtrCDE efflux of substrate antimicrobials, were also mainly observed in set A. Of note, mutation A39T in *mtrR* was observed in most genogroups but the combination of mutations A39T and R44H, associated with decreased susceptibility [340, 341], was only observed in G001 and G036. Finally, it is noteworthy that our data suggests that genogroups G007, G022, G032 and G041 were susceptible to all antimicrobials.

3.4.5. Association between *Neisseria gonorrhoeae* genetic clustering and epidemiological data

In order to analyse genetic clusters with the potential to be epidemiologically linked, we applied a conservative threshold of 1.5% AD within each low-level genogroup, after sub-MST generation maximizing the shared loci between isolates at this level. As such, we observed a total of 315 genetic clusters spread out across 141 low-level genogroups. Results are presented for all clusters composed of more than ten isolates (**Figure 3.10.A**), with the phylogenetic distribution of all identified clusters with ≥ 5 isolates presented in **Figure 3.10.B** (full detailed data described in **Supplementary Table S3.1.** and **Supplementary Table S3.2.**). Expectedly, the more prevalent low-level genogroups presented the highest number of genetic clusters, namely G001.002, G002.001, G005.004 and G001.005 (**Figure 3.10.B**). Nevertheless, we also observed cases where a single large close-related genetic cluster was detected within a given genogroup (e.g., G023.027, G021.029, G041.040 and G001.014; see **Figure 3.10.B**). In another perspective, this analysis also highlighted genogroups disseminated at multi-country level with potential regional transmission chains. For instance, G015.016, which is composed by isolates from ten different countries, was fragmented into eight smaller genetic clusters with only two clusters presenting isolates from distinct countries. We also examined each cluster in a time scale by each isolate's collection date (**Figure 3.10.A**), which allowed us to discriminate between clusters that potentially circulated within a specific time interval (and have not since been detected) from clusters that potentially emerged and others that are still circulating, even though this is reliant on the dataset and time period analysed. For example, WGSCL0025 (G007.008, mean AD 1.6%) that enrols isolates spanning five years (1995 up to 2000) seems to have not been detected since, which relates to a described outbreak in Sheffield UK [270]. This was also the case for WGSCL0044 (G019.022, mean AD 1.5%) with isolates from 2005 up to 2009 and WGSCL0030 (G012.013, mean AD 1.6%) with isolates from 2003 up to 2009 (**Figure 3.10.A**), although these did not have a described epidemiological link as they contained PT isolates and isolates from the UK [270]. In contrast, WGSCL0024 (G001.005, mean AD 1.5%) and WGSCL0022 (G006.007, mean AD 1.9%) were composed by isolates that have been consistently detected from 2004 up to 2017 (the latest year within the dataset). Regarding clusters enriched by PT isolates, this could be observed for instance in WGSCL0020 (G005.004, mean AD 1.6%) with isolates from 2004 up to 2014, in WGSCL0042 (G001.014, mean AD 1.0%), a more recent cluster dating from 2013 up to 2017, and WGSCL0001, from 2011 up to 2017. Still, we believe that more genomic data is needed in order to reinforce these

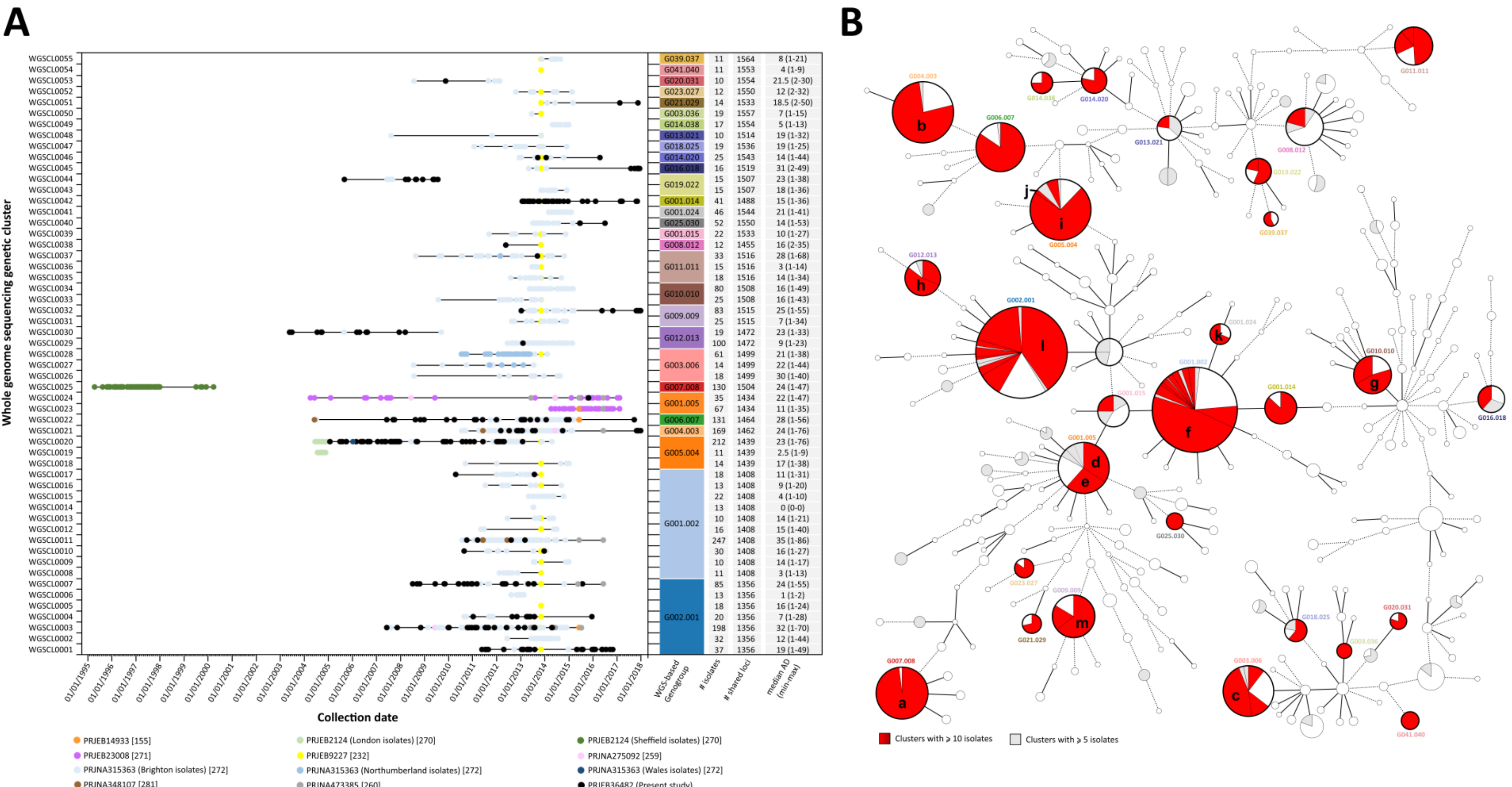

Figure 3.10. | Analysis of *Neisseria gonorrhoeae* WGS-based genetic clusters at low resolution level potentially concordant with epidemiological link. (**A**) Genetic cluster isolates' distribution by collection date, with detailed data of each cluster presented on the right. Each isolate colour refers to a specific study and black lines link the earliest and latest isolate detected. (**B**) Minimum spanning tree (also described in **Figure 3.6.**) of all isolates analysed in the present study highlighting WGS genetic clusters identified at conservative threshold of 1.5% AD. Nodes (which represent a unique allelic profile) presenting an allelic distance (AD) below 40, corresponding to the low-level genogroup threshold, have been collapsed for visualization purposes. **a** – WGSCL0025 Sheffield outbreak described in [270]. **b** – WGSCL0021 includes cluster ST2400 described in [272] and the ST2400 MSM-associated isolates described in [232]. **c** – WGSCL0028 Northeast England outbreak described in [272]. **d** – WGSCL0024 clade 2 described in [271]. **e** – WGSCL0023 Leeds outbreak [155] plus Clade 1 and 3 described by [271]. **f** – WGSCL0011 large cluster ST2992 described in [272]. **g** – WGSCL0034 cluster ST26 described in [272]. **h** – WGSCL0029 cluster ST292 described in [272]. **i/j** – WGSCL0019 and WGSCL0020 London outbreaks described in [270]. **k** – WGSCL0041 small cluster ST2992 described in [272]. **l** – WGSCL0003 cluster ST1407 which includes linked isolates from Brighton and London described in [272] and isolates described as cephalosporin resistant in [232]. **m** – WGSCL0032 the ST4995 MSM-associated isolates described in [232].

observations as they rely on the analysed database; for instance, data from the 2013 EURO-GASP survey [232] are a sub-sample restricted to the latter months of the year. Nevertheless, a relationship between the observed genetic clusters at this threshold and published data was clearly observed (**Figure 3.10.**). Eleven distinct subsets that were phylogenetically related (i.e. epidemiological verified outbreaks, or verified transmission chains) were consistent with the obtained genetic clusters at a threshold of 1.5%. This was the case for: WGSCL0028 (G003.006, mean AD 1.4%), an outbreak in Northeast England [272]; WGSCL0023 (G001.005, mean AD 0.8%), an outbreak of high-resistant azithromycin isolates in Leeds [155] linked to isolates from other published data [271]; and WGSCL0020, an outbreak in London [270], which could be linked to 49 PT isolates, 19 isolates from the EURO-GASP 2013 survey [232], 51 isolates from Brighton [272] and two isolates from the Netherlands [281]. Noteworthy, one PT isolate (NGPT15194) could be linked to the described high-resistant azithromycin *N. gonorrhoeae* isolates in Ireland and the UK (WGSCL0024), with sustained transmission [271].

3.5. Discussion

In the present study, we aimed to perform a comprehensive *N. gonorrhoeae* genogrouping, based on WGS data, for prospective WGS-based laboratory surveillance. Focused on the European panorama, we quantitatively evaluated cluster stability to identify the major circulating WGS-genogroups, assessed their geographic spread and searched for potential relationships between particular genogroups and specific AMR signatures. To achieve this we relied on a recently described method (nAWC) [310, 336] to unprecedently analyse cluster-partitioning concordance at all possible allelic distance thresholds for gonococci. Two cluster-stability points could be identified, which enabled the classification of isolates into hierarchical WGS-based genogroups. This assignment facilitates the assessment (and continuous monitoring) of the frequency, geographic dispersion, and potential association with specific AMR signatures of *N. gonorrhoeae* circulating genogroups, while providing a comprehensive analysis of the genetic diversity of this bacterium. Our approach resulted in the identification of 180 high-level and 321 low-level genogroups, several of which were composed by a single isolate or are represented by isolates from a single country. Both reflected the global genetic diversity and spread of *N. gonorrhoeae*, and showed the more prevalent and widespread strains circulating in Europe (**Figure 3.6**). At least five high-level genogroups (**Table 3.1.**) were found in more than 15 European countries and eleven low-level genogroups (**Table 3.2.**) were found in at least six countries (of which three are composed by at least five isolates). In an opposite scenario, we observed 81 genogroups that were composed by isolates from a single country (**Figure 3.7.**). Although *N. gonorrhoeae* infections lack a geographic structure, as revealed by its intercontinental spread [272, 277], our results also sustained the existence of *N. gonorrhoeae* transmission chains, likely confined at country/regional level.

Nevertheless, as WGS is becoming more accessible and the several gonococcal antimicrobial surveillance programmes are starting to generate substantial genomic data, we cannot discard that the inclusion of genomes from other continents may impact the genetic diversity landscape observed in this Europe-oriented study. This impact could be reflected in either the addition of entirely new genogroups or, for instance, the potential merger of genogroups with poorly represented genetic diversity (i.e., represented here by one or two isolates). In fact, several studies have pointed-out a contrasting genetic diversity in countries from other continents [21], such as Kenya [299], Canada [208, 231], Japan [296], New Zealand [284], Australia [283] or the USA [238]. Still, comparison of typing data showed that the novel defined genogroups were mostly populated by a dominant MLST or NG-MAST ST (being especially relevant for typing comparability) which largely facilitates backwards compatibility and data communication in the transition to WGS-based typing (**Figure 3.3.** and **Table 3.2.**). This observation is essential, as previous surveillance studies report particular NG-MAST genogroups or ST that not only need to be monitored but also display relationships with AMR profiles that cannot be dismissed [232]. Additionally, our study also showed that the quantitatively identified thresholds (4.97% and 9.61% AD; 40 and 79 AD in 822 loci, respectively), reflecting the earliest cluster stability (i.e., stable points with maximum resolution), provide a higher discriminatory resolution than the recently described PubMLST core-genome grouping threshold (400-AD in 1649 loci) [311] (**Figure 3.5.**). Still, the observed huge cluster agreement at the same threshold reinforces the likelihood that the robustness of the clustering will be kept if our high- and low-level thresholds are scaled to the entire PubMLST loci scheme.

Here we observed that several WGS-based genogroups seem to be associated with distinct profiles of AMR (**Figure 3.8.** and **Figure 3.9.**). In fact, one such genogroup was NG-MAST G1407 (identified as genogroup G002.001 in the present study), which has been detected worldwide [21, 236–241] and is one of the major types found across Europe [232, 233]. Most isolates from this genogroup have been shown to be associated with multidrug resistance, presenting resistance or an increased MIC (or both) to cefixime, ceftriaxone, azithromycin and ciprofloxacin [230, 232, 233]. Concordantly, genetic alterations with potential association with cephalosporin resistance [230, 232], i.e., simultaneous carriage of mutations 312M, 316T and 545S in *penA* (mainly represented by the *penA-XXXIV* allele [105]), were mostly observed in isolates from this genogroup. Still, we detected potential horizontal acquisition of this specific *penA* allele by isolates with different genome backgrounds (eight other genogroups). However, it is worth noting that all 102 PT isolates belonging to G002.001 (including 87 harbouring the *penA-XXXIV* allele) had low MIC values for both cefixime and ceftriaxone, whether or not belonging to NG-MAST 1407, thus challenging the association between these mutations in NG-MAST G1407 and resistance to cephalosporins. This also highlights the fact that some mechanisms of resistance in *N. gonorrhoeae* are still not fully disclosed, and further research is needed to understand the linkage between resistance and the

simultaneous carriage of genetic AMR determinants, namely the role of epistatic interactions on the level of antimicrobial susceptibility (i.e., differential MIC values). For instance, the first cephalosporin resistant isolate detected in Portugal (end of 2019) revealed a single alteration (542S) in *penA*, while presenting MICs of 0.19 and 0.38 mg/L to ceftriaxone and cefixime respectively [324]. Another genogroup that warrants particular attention is G001.005, where the carriage of the azithromycin resistance-associated mutation A2045G was mostly observed. This genogroup enrolled the isolates associated with high-level azithromycin resistance that have been linked to an outbreak in the UK [155, 271]. Noteworthy, we found a high-level azithromycin resistant PT isolate that was closely related to other resistant isolates from Ireland [259, 260] and the UK [271] (WGCL0024), suggesting its potential spread.

Overall, two distinct sets of genogroups displaying a contrasting pattern of AMR could be observed, i.e., one set with isolates carrying several genetic determinants related with decreased susceptibility and resistance, and another set mostly involving susceptible isolates (**Figure 3.8.**). For the vast majority of genogroups we observed a dominant AMR profile (of note, for nine genogroups, a single AMR profile was observed) and a lower fraction of one or more additional AMR profiles (**Figure 3.9.**). Since a predominant circulating genogroup may not be the one carrying an alarming AMR profile, the combination of these results may help prioritize which genogroups need to be subjected to more close surveillance, management and control at both country and continent levels. On the other hand, our results showed that, for some genogroups, the less frequently observed AMR profile was the one carrying a larger arsenal of genetic determinants for resistance, suggesting the potential emergence of novel resistant isolates within a WGS-subtype. Of note we detected genogroups where the dominant profile is the one associated with a larger repertoire of AMR determinants. This suggests a progressive accumulation of AMR determinants within these genogroups due to intensive antibiotic-driven selective pressures, although other scenarios cannot be discarded and warrant investigation (such as the presence of naturally resistant lineages). This information may be crucial for surveillance towards the identification of evolving strains whose emergence need to be alerted to public health authorities in order to control a local or cross-border spread in early stages.

The strategy described in the present study benefits from the comprehensive definition of WGS-subtypes for long-term surveillance (with backwards compatibility) and a conservative threshold range that can be used for contact tracing and outbreak investigation. While for long-term surveillance, the goal is usually to continuously record the bacterial types that are circulating in a particular geographic area, for outbreak investigation a genomic analysis with higher resolution is required, potentiating the identification of shared genomic signatures to infer transmission [277, 310]. Defining clusters composed of likely related isolates by using specific thresholds ranges is an important step in the application of any subtyping scheme, but optimizing these thresholds/parameters has been a recurring challenge in the field of molecular epidemiology [307,

310]. In our approach, we cannot discard that adjusting these thresholds and maximizing the dataset may affect cluster composition and stability, even though our results are highly concordant with previous observations (**Figure 3.6.** and **Figure 3.10.**). Nonetheless, this study provides a comprehensive genome-scale snapshot of the genetic diversity of circulating *N. gonorrhoeae* strains in Europe. Results suggest that the applied conservative threshold for potential epidemiological linkage may be suitable to highlight clusters for fine-tune analysis that could promote an epidemiological investigation, as cluster composition at this level seems to match with previous reports on outbreaks or transmission chains (**Figure 3.10.**). Additionally, the possibility to expand the number of loci under analysis in the scheme not only allows performing a dynamic analysis of epidemiological clusters, but also increases the confidence in results, as the link between strains is strengthened when AD decrease or remain unaltered after generating sub-MST with a higher number of shared loci (**Figure 3.10.**). This concept of dynamically reconstructing phylogenies for sub-sets of strains (by maximizing the number of shared loci) and the bioinformatics tools needed for its operationalization were set-up recently for foodborne pathogens [310] and were applied to respiratory pathogens [342] showing that they facilitate the use of WGS in routine surveillance and, consequently, epidemiological investigations.

Ultimately, in the frame of the demanding short-term transition from traditional genotyping methods, the current comparative study is an important step towards the implementation of a WGS-based laboratory workflow for *N. gonorrhoeae* surveillance in the Portuguese NRL. The laboratory surveillance and isolates collection built up on behalf of PTGonoNet, hosted at the Portuguese NRL [326], has already strengthened our contribution to the EURO-GASP [232]. This has been achieved through the enrichment of the current geo-temporal and genomic diversity of *N. gonorrhoeae*, by adding WGS data of isolates spanning 15 years of *N. gonorrhoeae* surveillance in Portugal. Together with the centralization of the AMR phenotypic characterization and molecular typing at the Portuguese NIH, and a strong articulation with the Public Health Authorities, the described framework should constitute the driving force towards a faster and robust prospective surveillance of cases, from antibiotic resistance prediction to transmission chain detection. Future studies based on this approach will be crucial to consolidate the benefits of this technological transition for Public Health through the prioritization of genogroups to be monitored, the identification of emerging resistance carriage, and the potential facilitation of data sharing and communication.

Data summary: All *Neisseria gonorrhoeae* reads generated in this study were deposited in the European Nucleotide Archive (ENA) (BioProject accession no. PRJEB36482; www.ebi.ac.uk/ena/data/view/PRJEB36482). ENA sample accession numbers are included in Supplementary Table S3.1. (available with the online version of the article). Read datasets used were downloaded from the ENA (BioProject accession no. PRJEB14933; PRJEB2124; PRJEB23008; PRJEB26560; PRJEB9227; PRJNA275092; PRJNA348107; PRJNA473385; PRJNA315363) with all sample accession numbers included in Supplementary Table S3.2. Supplementary material for this study was deposited in ZENODO and are available at https://doi.org/10.5281/zenodo.3946223.

Funding information: M. Pinto was supported by the Portuguese Science and Technology Foundation (FCT) through the grant SFRH/BD/109264/2015. This work was partially funded by the GenomePT project (POCI-01-0145-FEDER-022184), supported by COMPETE 2020 – Operational Programme for Competitiveness and Internationalisation (POCI), Lisboa Portugal Regional Operational Programme (Lisboa2020), Algarve Portugal Regional Operational Programme (CRESC Algarve2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF), and by the Portuguese Science and Technology Foundation (FCT).

Conflicts of interest: The authors declare that there are no conflicts of interest.

Ethical statement: All patient data associated with each *Neisseria gonorrhoeae* isolate were fully anonymized.

Acknowledgements: The authors would like to acknowledge all participants of the Portuguese National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNet - http://www.insa.min-saude.pt/category/areas-deatuacao/doencas-infeciosas/rede-nacional-de-vigilancia-laboratorial-de-estirpes-de-n-gonorrhoeae/).

CHAPTER IV

Insights into the *Neisseria gonorrhoeae* pan-genome: a preliminary assessment of genogroup-specific genetic signatures

---

Personal contributions

MP contributed to the study design, performed the research, analyzed all data and wrote the manuscript.

4. Insights into the *Neisseria gonorrhoeae* pan-genome: a preliminary assessment of genogroup-specific genetic signatures

4.1. Introduction

Like many bacteria, *N. gonorrhoeae* is known to have a repertoire of genetic elements that enable increased genomic plasticity, likely contributing to host colonization, invasion and survival. Within its ~2.2. Mbp circular chromosome with ~2000 coding sequences, *N. gonorrhoeae* possesses the necessary genetic machinery to display a complex inter- and intra-strain diversity, acquired through DNA transformation [166] and variation of displayed antigens [167]. In fact, one of the more outstanding features of *Neisseria* is the amount and variety of repetitive DNA sequences present in their genomes [168, 169]. Most of these repeat sequences are involved in intra-species DNA transformation, phage integration [170] and genome rearrangements [171]. *N. gonorrhoeae* also possesses a gonococcal genetic island (GGI) encoding a type IV secretion system (T4SS) that secretes single-stranded DNA into the extracellular environment providing substrate for DNA transformation [172, 173]. Besides these chromosomal-mediated mechanisms, *N. gonorrhoeae* can also carry plasmids, either the cryptic plasmid (it does not contain AMR genes or virulence factors) or plasmids associated with antimicrobial resistance (AMR) [163] (such as the *blaTEM* and *tetM* containing plasmids) that can be transferred between pathogenic and commensal species [164, 165]. In addition to this genomic plasticity, *N. gonorrhoeae* can infect and colonize several anatomic sites, sharing these environmental niches with other commensal *Neisseria* species, with which it can exchange genetic material. In the previous Chapter of this Thesis (Chapter III) we classified *N. gonorrhoeae* isolates into genogroups based on its core-genome, reflecting its major circulating lineages. In this context, the main objective of the present Chapter was to perform a draft pan-genomic analysis of *N. gonorrhoeae* in order to identify specific genetic signatures that could be associated with each prevalent genogroup. Particularly, we aimed to explore the presence or absence of genes within the *N. gonorrhoeae* accessory genome.

4.2. Methods

4.2.1. Construction of a draft *Neisseria gonorrhoeae* pan-genome

All 3791 *N. gonorrhoeae* genomes generated and analysed in Chapter III were annotated with Prokka v1.12 [343] using default parameters. From the General Feature Format files (*.gff* files), containing all annotated features, the *N. gonorrhoeae* pan-genome was generated with Roary v3.13.0 software [344], with a BLASTP+ (from the BLAST+ suite v2.9.0 [345]) minimum

percentage of identity set to 95% and without splitting paralogs (*-s* option). Genes were defined as core-genes when present in ≥ 99% of all genomes (i.e., ≥ 3754 genomes), as accessory genes when present in > 2 genomes and as singleton genes when present in a single genome. For functional diversity analysis, Clusters of Orthologous Groups of proteins (COGs) categories [346] were assigned to the pan-genome predicted proteins using the script *cdd2cog* v0.1 [347] after RPSBLAST+ (Reverse Position-Specific BLAST) (e-value cut-off of $1e^{-2}$), where only the best hit (i.e., lowest e-value) was considered.

4.2.2. Scoring the presence and absence of accessory genes

After generating the gene presence/absence table with Roary, we used Scoary v1.6.16 software [348] to identify accessory genes potentially discriminating WGS-based genogroups, both at the high- and low-level (as determined in Chapter III). Statistical significance of the presence of a specific genetic trait and a WGS-subtype (Fishers' exact test, performed by Scoary) was assumed when p-value was ≤ 0.005. All markers presenting p-value above the described were discarded. Only accessory genes associated with a specific WGS-subtype presenting simultaneously ≥ 80% Sensitivity (i.e., present in ≥ 80% of genomes of queried WGS-subtype) and ≥ 80% Specificity (i.e., absent in ≥ 80% of genomes of non-queried WGS-subtypes) were reported in the final analysis.

4.3. Results

4.3.1. Overview of the *Neisseria gonorrhoeae* pan-genome

Reconstruction of the *N. gonorrhoeae* pan-genome, based on a 3971 European genome dataset, revealed that it was composed of 4866 genes (all genes detailed in **Supplementary Table S4.1.**), of which 1696 comprised the core-genome, 1871 the accessory-genome and 1299 were singletons. Core-genome was achieved with ~100 genomes and data suggested that *N. gonorrhoeae* has an open pan-genome, as the addition of novel genomes continuously increased the number of genes not shared in the core-genome (**Figure 4.1.**). *N. gonorrhoeae*'s core- and accessory-genomes differed greatly in the distribution and proportion of functional categories (**Figure 4.2.**). The core-genome was enriched with genes associated with housekeeping functions, such as energy production, translation, amino acid transport and metabolism (**Figure 4.2A**). In contrast, the accessory-genome presented a high proportion of genes with an unknown function (61.7%, contrasting with 23.5% in core genome), and an expectedly lower number of housekeeping genes (**Figure 4.2B**).

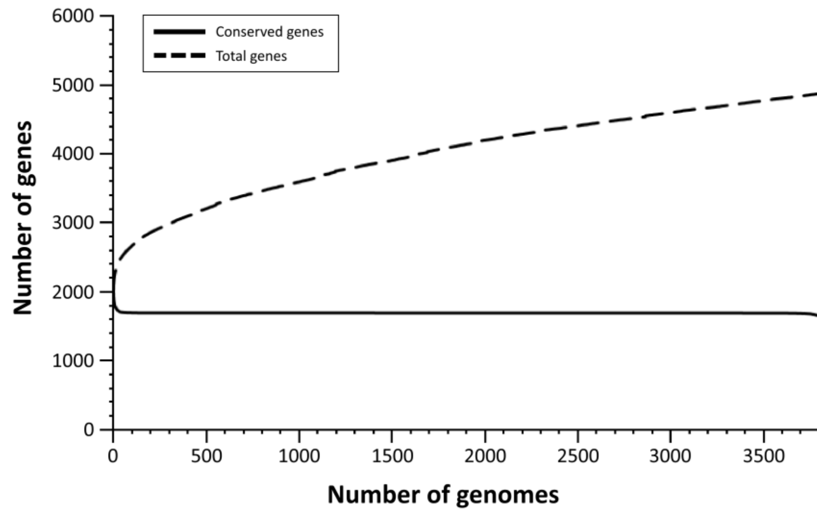

Figure 4.1. | *Neisseria gonorrhoeae* pan- and core-genomes. Progression in the number of genes in the pan-genome (dotted line) and core-genome (straight line) as more *N. gonorrhoeae* genomes are randomly added to the comparison until a total 3791 genomes.

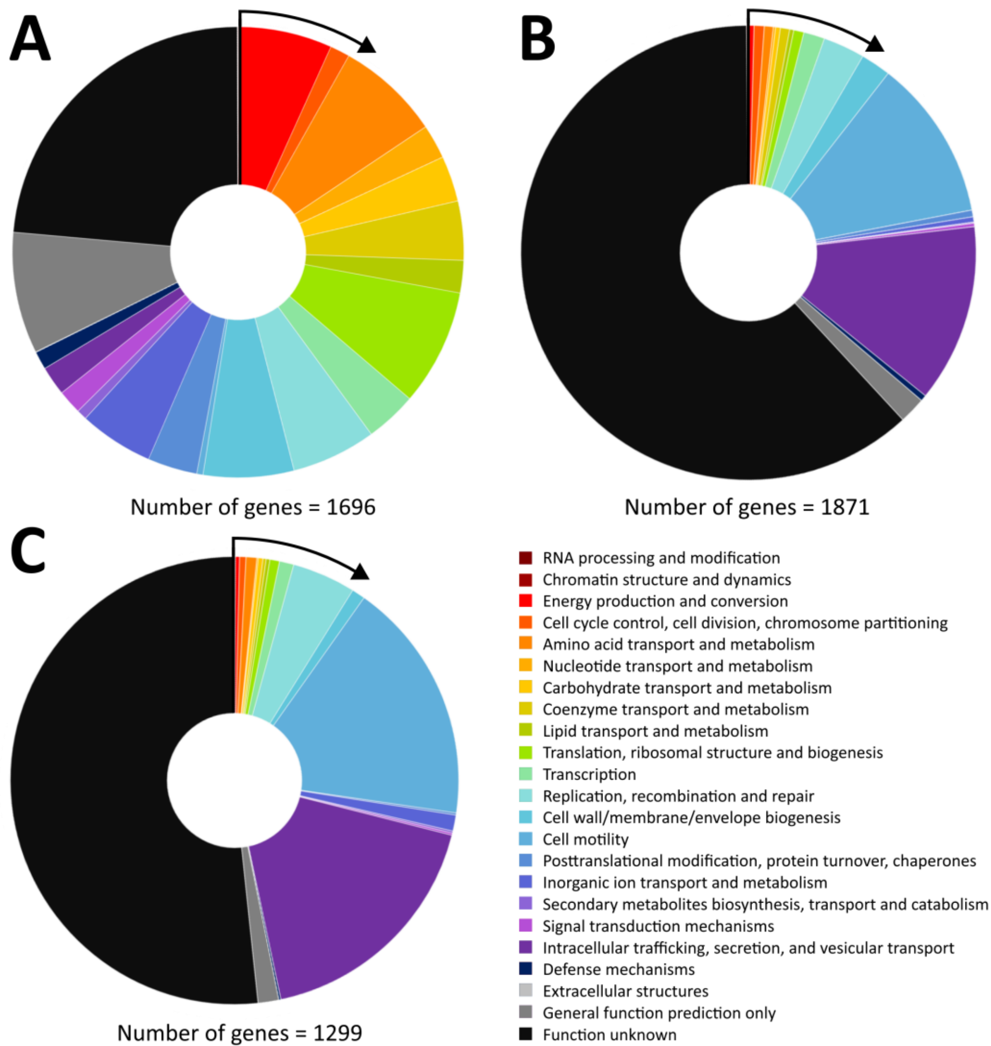

Figure 4.2. | Functional diversity of core- and accessory-genomes of *Neisseria gonorrhoeae*. Clusters of orthologous groups (COGs) of the predicted proteins of the (**A**) core-genome genes, (**B**) accessory genes and (**C**) singletons. Data are expressed as percentages over the total number of genes in each category.

This trend was also observed when analysing singleton genes' functional categories, where 51.7% of genes had an unknown function. Noteworthy, we observed that genes related to cell motility and intracellular trafficking and transport were overrepresented in the accessory-genome, as well as within the singletons. In fact, cell motility-associated genes represented 11.5% and 17.5% of the accessory-genome and singletons, respectively, contrasting with 0.4% in the core-genome. Likewise, transport-associated genes constituted 12.7% and 17.7% of the accessory-genome and singletons, respectively, while the in the core-genome these represented only 2.1%.

4.3.2. Association between accessory genes and WGS-based genogroups

Distribution of accessory genes, with ≥ 80% sensitivity and specificity, by high- or low-level WGS-based genogroup (represented by at least 10 isolates) is presented in **Figure 4.3.**, with genes' full description detailed in **Supplementary Table S4.2**. At the high-level, we observed a mean of 20 genes (ranging from three to 46 genes) that were associated with each of the 35 analysed genogroups. When taking into account redundant accessory genes, i.e., highlighted genes associated simultaneously with distinct genogroups, a set of 277 different genes were obtained (**Figure 4.3.A**). This redundancy was due to the presence of genogroups with few isolates, together with the parameters used. For instance, we detected three genes (Gene_1890, Gene_2433 and Gene_2419; see **Supplementary Table S4.2.A** and **Supplementary Table S4.3.A)** that displayed > 80% sensitivity and specificity for 14 different genogroups, twelve of which were comprised by less than 35 isolates. More importantly, of the 277 genes, 119 encoded for hypothetical proteins, 64 were pilin or pilin-associated genes, 14 were phage-associated genes and eight genes encoded for membrane/transport-associated proteins (**Supplementary Table S4.2.A**). Regarding accessory genes associated with low-level genogroups, results showed that 282 genes could be differentially linked to each of the 38 analysed genogroups (**Figure 4.3.B** and **Supplementary Table S4.2.B**). When compared with genogroups at the high-level, the same pattern in gene function was observed, with 117 genes encoding hypothetical proteins, 73 were pilin or pilin-associated genes, 14 were phage-associated genes and eight genes encoded for membrane/transport-associated proteins. In fact, there was an overlap of 254 genes between the analyses at both genogroup levels, with genes associated with cell motility or intracellular trafficking, secretion, and vesicular transport being overrepresented (excluding genes with unknown function). Regarding the 73 detected pilin or pilin-associated genes, these enrolled 37 out of the 38 low-level genogroups (**Supplementary Table 4.3.B**). We observed a mean of four genes in association with each of the genogroups, which suggests that different genogroups present distinct genetic profiles regarding the presence of pili genes. Within these, several genes associated with PilW and PilE assembly were observed, regardless of the genogroup size, suggesting distinct patterns of pilin formation involved in *N. gonorrhoeae* immune evasion and

pathogenesis, which may warrant future investigation. Of note, genes presenting 100% sensitivity and specificity (i.e., exclusive genes associated with a genogroup) were rarely observed, with only 13 and 5 genes detected at the high- and low-level, respectively (**Supplementary Table S4.2.**).

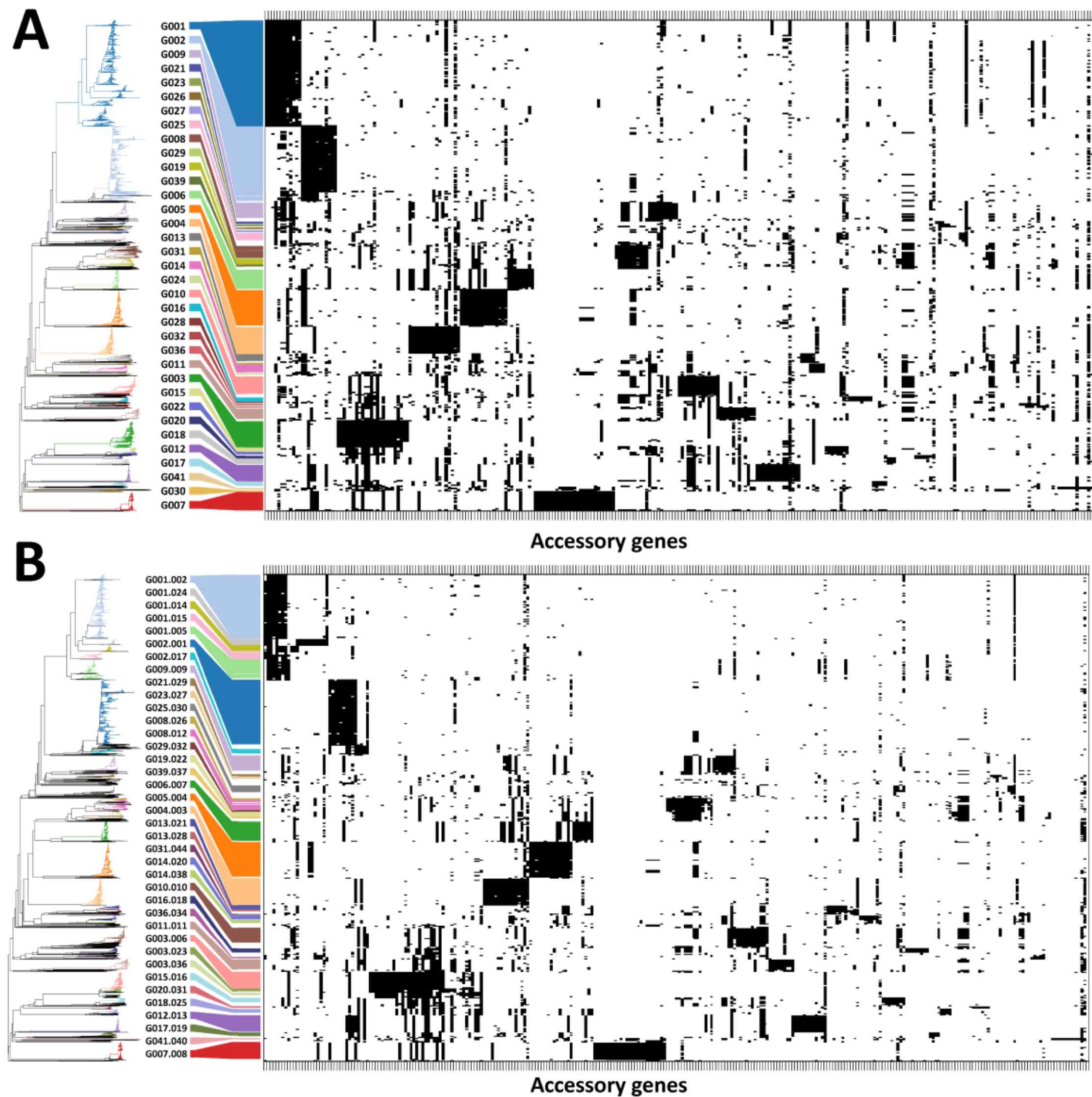

Figure 4.3. | Distribution of accessory-genome genes presenting more than 80% specificity and sensitivity across WGS-based genogroups. **A** – Distribution of 277 accessory genes by high-level genogroups. **B** – Distribution of 282 accessory genes by low-level genogroups. Gene order in the Figure is detailed in **Supplementary Table S4.3**. The contextual neighbour-joining phylogenetic tree at the left side of each panel was generated based on the MScgMLST allelic profiles using the GrapeTree v1.5.0 software [338].

Within these accessory genes, regardless of the genogroup level, we detected the plasmid-carried *blaTEM* and *tetM* genes, associated with antimicrobial resistance to β-lactams and tetracycline, respectively. Moreover, *blaTEM* and *tetM* were highlighted as specific genetic markers for three and eight distinct low-level genogroups, respectively (**Supplementary Table 4.3.B**). Nevertheless, for *blaTEM*, we observed that this gene was singled out only because genogroups were composed of ≤ 15 isolates (namely G029.032, G031.044, and G039.037), which

allow it to pass the 80% cut-off that was applied for sensitivity/specificity. In fact, we previously observed that this gene was carried by 509 isolates, spanning several low-level genogroups (**Figure 3.8.** of Chapter III), using a read-based approach. Still, here, using an assembly-based approach, data showed that 386 isolates carried *blaTEM* (**Supplementary Table S4.2.**). This exemplifies the importance of having large representative groups when determining associations between specific traits (either phenotypic or genetic) and the carriage of genetic markers, and caution is advisable when performing such inferences. On the other hand, for *tetM*, although the identified genogroups were composed by up to 60 isolates, these corresponded to the ones previously detected, as most of the enrolled isolates carried this gene, namely G008.012, G008.026, G009.009, G016.018, G019.022, G029.032, G031.044, and G039.037 (**Figure 3.8.** of Chapter III). Curiously, genogroup G002.001 (enrolling 502 isolates) was associated with specific penicillin-binding protein 2-encoding gene (i.e., *penA*), being present in 493 genomes and in only nine others from non-G002.001 genogroups. This genogroup had been previously highlighted for carrying the *penA* type XXXIV mosaic allele associated with decreased susceptibility to cephalosporin [105] (See Chapter III). Of note, among the other seven detected membrane proteins-encoding genes, four encoded porins, being associated with nine and twelve distinct high- and low-level genogroups, respectively. Still, it is important to state that all results should be analysed taking into account that the pan-genome was inherently generated based on gene clustering at a minimum percentage of identity set at 95%. As such, it would be expected that highly polymorphic genes would be split into different gene groups, which does not necessarily reflect the present/absence of a gene but more likely a divergent allelic profile.

4.4. Discussion

In the present study we explored a draft pan-genome of *N. gonorrhoeae*, constructed based on the genetic diversity of 3971 genomes collected from across Europe. We aimed to identify specific genetic signatures within the accessory-genome of *N. gonorrhoeae* that could be associated with the WGS-based genogroups determined in the previous Chapter (Chapter III). Based on the used dataset, we observed that *N. gonorrhoeae* may possess an open pan-genome, as the subsequent addition of genomes in the analysis continuously increased the number of genes not shared in the core-genome (**Figure 4.1.**), even though a high degree of singletons was still observed. This was expected, as this bacterium is known to possess several genetic mechanisms involved in DNA transformation, with HGT being one of the major mechanisms to generate genetic diversity towards immune evasion, survival and pathogenesis [166, 168]. Although our results showed that the accessory-genome was mostly composed by genes of unknown function, we observed that it was also enriched with genes associated with cell motility and transport (**Figure 4.2.**). Within these, several membrane protein-encoding genes and pilin or pilin-

associated genes were differentially present throughout the distinct low-level genogroups, regardless of the genogroup size. Considering the surface exposed character of the proteins encoded by these genes, it would be expected that they present higher genetic diversity, as a means to adapt to the host environment. In fact, the pilin are known to be involved in adherence, colonization and invasion in *N. gonorrhoeae* [65], being also targets of antigenic variation towards immune evasion [168]. As these genes are singled-out in association with distinct WGS-based genogroups, we hypothesize that they may possess distinct adherence/pathogenic profiles, which in turn may underlie more advantageous transmission, and the higher prevalence observed for some of them. Moreover, if the genetic diversification of the pilin genes only occurred during the infection process, the potential association of pilin-associated genes with specific genogroups would not have been highlighted, pointing to an established mechanism of adherence in the early stages of infection. Our analysis also highlights genes associated with antimicrobial resistance, namely the plasmid-carried *blaTEM* and *tetM*. Nevertheless, we observed that in the case of *blaTEM*, it was only detected in genogroups composed by few isolates. Curiously, this was not the case for *tetM*, which reinforces the association between carriage of this AMR determinant with specific genogroups observed in the previous Chapter.

Overall, our data suggest the existence of several genes that seem to be associated with our defined WGS-based genogroups within *N. gonorrhoeae*'s accessory genome. This issue should be subjected to further investigation, as most hits correspond to genes with unknown function. Still, these results must be regarded just as overall trends, suggesting potential shared genetic mechanisms within circulating isolates of the same genogroups. In fact, we found that there was a lack of genes presenting 100% sensitivity and specificity for a given genogroup that could be used as markers for genogroup classification beyond the core-genome, as previously performed for other pathogens [349]. The present study constitutes an exploratory work into the pan-genome of *N. gonorrhoeae* and presents some limitations. For instance, results are highly dependent on the dataset used, namely the size of each enrolled WGS-genogroup and the percentage of identity chosen to cluster genes. In fact, regarding the latter, changing the percentage of identity to 70% would reduce the pan-genome size from 4866 to 3043 genes (with little impact in the core-genome; i.e., from 1696 to 1723), which suggests that the genes detected in the accessory genome with association with genogroups may not reflect a true presence/absence but, more likely, distinct allelic profiles of the same genes. As for the dataset, a higher representation of isolates from each genogroup could aid in consolidating the observed association, in order to disclose if specific genetic signatures could underlie the adaptation of circulating strains towards increased transmission or pathogenesis.

CHAPTER V

First case of a cephalosporin resistant *Neisseria gonorrhoeae* isolate detected in Portugal

---

This Chapter was published in

Pinto M, Matias R, Rodrigues JC, Duarte S, Vieira L, Gonçalves I, Gonçalves MJ, Ramos MH, Gomes JP, Borrego MJ. (2020) **Cephalosporin-Resistant *Neisseria gonorrhoeae* Isolated in Portugal, 2019**. *Sexually Transmitted Diseases*; 47(11):e54–e56.

DOI: 10.1097/OLQ.0000000000001218.

Personal contributions

MP contributed to the study design, performed the research, analyzed the data and wrote the manuscript.

5. First case of a cephalosporin resistant isolate detected in Portugal

5.1. Abstract

We report a multidrug resistant *Neisseria gonorrhoeae* exhibiting resistance to ceftriaxone and cefixime, isolated in Portugal in 2019. Whole-genome sequencing was performed for typing and identification of genetic determinants of antimicrobial resistance. Due to its antimicrobial susceptibility profile, awareness should be raised for the circulation of this strain.

Keywords: *Neisseria gonorrhoeae*; surveillance; multidrug resistance; Whole-genome sequencing; sexually transmitted infections.

5.2. Introduction

Neisseria gonorrhoeae disease burden and antimicrobial resistance remain a global concern. The current empirical treatment of uncomplicated gonorrhoea relies on ceftriaxone and azithromycin. During the last decade, a steady rise in minimum inhibitory concentrations (MICs) of cephalosporins has been observed worldwide [104, 106, 135, 152, 153], with some reporting clinical failures upon treatment. Increase in azithromycin resistance is also being reported [135, 152, 153, 259, 271, 281], and resistant strains have been associated with recent outbreaks [155, 158]. In Europe, although some countries have consistently reported resistance to azithromycin [155, 259, 271, 281], including Portugal, resistance to cephalosporins is still rare [135, 232, 260]. In the context of the European Centre for Disease Prevention and Control's call to action for surveillance and management of resistant strains [135, 232], the current report describes the first case detected in Portugal (in late 2019) of a multidrug resistant *N. gonorrhoeae* isolate exhibiting simultaneous resistance to ceftriaxone and cefixime.

5.3. Case description

At the end of September 2019, a 27 years old male patient with unremarkable past medical history developed a clinical picture of dysuria and whitish urethral discharge. On the third day of symptoms, he attended an emergency room and was prescribed with ciprofloxacin for a urinary tract infection. As no improvement was observed, on eleventh day of symptoms he attended the emergency room of the tertiary Hospital of Oporto, Portugal. He reported unprotected MSM (men who have sex with men) contacts with five occasional Portuguese partners during the last six months, in Oporto and Lisbon, the last of which occurring two weeks before the onset of

symptoms. He denied commercial sex and attendance to chemsex parties or saunas, and reported no travels to other countries in the last 12 months. Swabs collected from the urethra, rectum and oropharyngeal tract tested positive for *N. gonorrhoeae*, by Real-Time PCR (SACACE, Biotechnology, Como, Italy). Other sexually transmitted infections were screened (HIV, hepatitis B and C, syphilis) and tested negative. He was empirically treated with 500 mg ceftriaxone intramuscularly and 1 g azithromycin orally. *N. gonorrhoeae* was isolated from the urethra (no culture was performed for the pharynx and rectum swabs) and sent to the National Reference Laboratory for Sexually Transmitted Infections (NRL) at the Portuguese National Institute of Health (INSA), for antimicrobial susceptibility testing, as part of the ongoing Portuguese National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNet - http://www.insa.min-saude.pt/category/areas-de-atuacao/doencas-infeciosas/rede-nacional-de-vigilancia-laboratorial-de-estirpes-de-n-gonorrhoeae/). As the patient missed the follow up consultation, clinical re-observation could only occur two months after the initial treatment, upon a new request from the patient. Although he denied new sexual contacts after the emergency room episode, the patient reported relapse of symptoms and new swabs were collected from the urethra, rectum and pharynx. Only the urethral swab tested positive by Real-Time PCR for *N. gonorrhoeae* but an isolate culture could not be obtained. Since the patient was asymptomatic, no new treatment was administered and it was assumed that the PCR positive results was due to remnant gonococcal DNA. One month later, swabs from the three anatomic locations were collected, all testing negative for *N. gonorrhoeae*.

5.4. Characterization of *Neisseria gonorrhoeae* isolate NGPT19535

All antibiotic susceptibility testing was performed in duplicate using Etest (bioMerieux, Marcy-l' Étoile, France), with MICs being determined for penicillin, ceftriaxone, cefixime, azithromycin, tetracycline, ciprofloxacin, spectinomycin and gentamicin, and results interpreted using the current European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines [94]. Isolate NGPT19535 exhibited resistance to ceftriaxone, cefixime, ciprofloxacin and tetracycline, and intermediate resistance to penicillin (**Table 5.1.**). Isolate was susceptible to both azithromycin and spectinomycin. β-lactamase production tested negative [chromogenic reagent Nitrocefin (Oxoid Limited, Hampshire, UK)]. After extraction, DNA was subjected to the Nextera XT library preparation protocol (Illumina, San Diego, California, USA) prior to paired-end sequencing (2x250 bp) on a MiSeq instrument (Illumina, San Diego, California, USA), according to the manufacturer's instructions. Genome was assembled using the INNuca v4.0.1 pipeline (https://github.com/B-UMMI/INNUca) [310]. Multi-locus sequence typing (MLST) and *N. gonorrhoeae* multi-antigen sequence typing (NG-MAST) were performed *in silico*, with NGPT19535 belonging to ST-7822 and ST17370, respectively. A gene-by-gene analysis was

performed using the PubMLST *N. gonorrhoeae* cgMLST scheme v1.0 adapted to the chewBBACA suite v2.0.11 [331] on a dataset of 3791 genomes from European countries [106, 155, 232, 260, 270–272, 281]. Global phylogenetic analysis revealed no close allelic profile match, with the closest isolate (Portugal, 2017, accession #ERR3851705) having 154 allelic differences in 1537 loci and the closest published genome having 436 allelic differences (Spain, 2013, #ERR1560929). Raw read data were deposited in the European Nucleotide Archive under the accession #ERR3965143**.**

Table 5.1. | Antimicrobial susceptibility and genetic characterization of *Neisseria gonorrhoeae* strain NGPT19535, isolated in Portugal in 2019.

Isolate (ID, country, year of isolation	**NGPT19535, Portugal, 2019**	**ECDC_ES016, Spain, 2013** [232]
Minimal inhibitory concentrations (mg/L)		
Benzylpenicillin	1	NA
Ceftriaxone	0.19	0.12
Cefixime	0.38	0.12
Azithromycin	0.75	0.5
Tetracycline	1.5	NA
Spectinomycin	8	NA
Ciprofloxacin	4	16
Gentamicin	4	NA
Sequence type		
MLST	7822	1901
NG-MAST	17370	2493
Genetic determinants for antimicrobial resistance (gene, antimicrobial affected and genetic alteration)		
penA (CFM; CRO; PEN)	G542S, Allele 5.002 (non-mosaic)	G542S, Allele 5.002 (non-mosaic)
porB1b (CFM; CRO; PEN; TET)	G120K, A121D	G120K, A121D
mtrR (CFM; CRO; PEN; TET; AZM)	A39T	A39T
ponA (CFM; CRO; PEN)	L421P	L421P
gyrA (CIP)	S91F (D95A)	S91F (D95G)
parC (CIP)	S87R	S87R
rpsJ (TET)	V57M	V57M
folP (SUL)	R229S	R229S
tetM (TET)	Absent	Present
blaTEM (PEN)	Absent	Present

Mutation in parenthesis in *gyrA* indicate a different substitution in both isolates but in the same position. MLST: Multi-Locus Sequence Typing; NG-MAST: *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing. NA: not available; CFM: cefixime; CRO: ceftriaxone; PEN: penicillin; AZM: azithromycin; TET: tetracycline; CIP: ciprofloxacin; SUL: sulphonamide.

Genetic determinants for antimicrobial resistance were assessed by the ARIBA v2.12.2 software (**Table 5.1.**) [302]. NGPT19535 exhibited a non-mosaic *penA* allele with amino acid alteration G542S, potentially reducing β-lactam acetylation of penicillin-binding protein 2 (PBP2). Alterations G120K and A121D in porin encoding gene *porB1b* and L421P in *penA*

(PBP1) were observed, which contribute to overall resistance to cephalosporins and penicillins [120, 192, 232, 302]. No mutations in 23S rRNA gene (i.e., A2059G and C2611T) were present but we observed SNP A39T in *mtrR,* which could underlie the MIC value obtained for azithromycin by increasing the expression of the MtrCDE efflux pump system [120, 192, 302]. Two amino acid substitutions in GyrA (DNA gyrase subunit A), known to be responsible for ciprofloxacin resistance by reducing quinolone binding to DNA gyrase were detected, along with alteration S97R in *parC*, reducing binding to topoisomerase IV [120, 192, 232, 302]. Amino acid alteration V57M in ribosomal protein S10 encoding gene *rpsJ* contributed for tetracycline resistance [192, 232]. Although susceptibility was not tested, NGPT19535 also possessed the alteration R229S in *folP* associated with resistance to sulphonamides [192, 232]. Plasmid carrying *tetM* (for tetracycline resistance) or *blaTEM* (for β-lactamase production) and mutations in *16S rRNA* gene (for spectinomycin resistance) were not observed. Within our dataset [106, 155, 232, 259, 260, 270–272, 281], the only isolate displaying the same mutational profile in antimicrobial resistant determinants was collected in Spain in 2013 (accession #ERR1469570), while exhibiting MICs for both cephalosporins close to the EUCAST breakpoint for resistance of 0.125 mg/L (**Table 5.1.**).

5.5. Discussion

This report describes the first case detected in Portugal of a multidrug resistant *N. gonorrhoeae* strain with combined resistance to ceftriaxone and cefixime, isolated in late 2019. *N. gonorrhoeae* antimicrobial resistance monitoring and surveillance in Portugal has been performed since 2003 at the NRL of INSA, in collaboration with private and public laboratories as part of the PTGonoNet. To date, no isolate presenting MIC for cefixime and ceftriaxone above the EUCAST breakpoint definition for resistance had been detected in Portugal, although resistance to azithromycin had been continuously observed [135]. In Europe, although cefixime resistant isolates have been described since 2009 in several countries, only a single case of ceftriaxone resistance was reported up to 2015 by the European Gonococcal Antimicrobial Surveillance Programme (EURO-GASP) [135], which is the network responsible for carrying out sentinel surveillance of gonococcal antimicrobial resistance in the European Union/European Economic Area. As the dual-therapy of ceftriaxone and azithromycin is still a recommended empirical treatment for uncomplicated gonorrhoea, it is essential to report and monitor emerging resistant isolates to ceftriaxone, as gonorrhoea treatment could become compromised. Additionally, in many European countries, including Portugal, a vast number of gonorrhoea diagnosis rely on molecular methods, hampering the availability of specimens for antimicrobial resistance evaluation and spread control.

Isolate NGPT19535 revealed an uncommon sequence type combination (MLST/ST-7822 and NG-MAST/ST-17370) since no isolate with such NG-MAST sequence type was found in the PubMLST database (https://pubmlst.org/). WGS analysis revealed that it is genetically distinct from publically available genomes [106, 155, 232, 259, 260, 270–272, 281], even those presenting MLST/ST-7822, namely isolates from the UK [272] and from other European countries [232, 259]. Still, we found that the antimicrobial resistance-associated mutational profile observed was similar to a single isolate from Spain, although the latter was highly distinct genetically. We suggest that the resistance observed in NGPT19535 was most likely associated with the combination of alterations observed in the PBP1, PBP2, PorB and MtrR encoding genes. Of note, although other genetic alterations currently described as associated with resistance to cephalosporins in *penA* [120, 192] were not observed here, alteration G542S in *penA* has already been reported as being statistically associated with elevated MICs, even though its effect on resistance has not been proven [120]. The identification of this mutation in the present study reinforces the likelihood of its association with resistance to cephalosporins. Still, although the patient had denied interim sex, we cannot completely discard the possibility of re-infection two months after treatment. In summary, notification of multidrug resistant cases with increased MIC for cephalosporins is crucial at an international level, as the spread of these isolates, along with their potential to evolve, might threaten the current recommended dual antimicrobial therapy.

Conflict of Interest and Sources of Funding: M.P. was supported by the Portuguese Science and Technology Foundation through grant SFRH/BD/109264/2015. For the remaining authors, none were declared.

CHAPTER VI

Neisseria gonorrhoeae laboratory surveillance in Portugal in the whole-genome sequencing era

Personal contributions

MP contributed to the study design, performed wet-lab experiments and research, analyzed all data and wrote the manuscript.

6. *Neisseria gonorrhoeae* laboratory surveillance in Portugal in the whole-genome sequencing era

6.1. Introduction

Whole-genome sequencing (WGS), coupled with bioinformatics, has allowed to widen our knowledge on *Neisseria gonorrhoeae*'s origin, evolution, population structure, pathogenesis, antimicrobial resistance (AMR) mechanisms and epidemiology [21, 270, 272, 277, 300]. Most national or international gonococcal AMR surveillance programs are now generating these data in order to monitor and control the spread of resistant *N. gonorrhoeae* strains [130]. In fact, with the transition to a WGS-based routine laboratory surveillance being a reality, laboratories whose main mission is to perform epidemiological surveillance are now facing the need to apply novel genomic approaches to provide detailed epidemiological analysis of *N. gonorrhoeae* transmission, outbreak detection and contact tracing, including AMR data. There is also the need to complement other laboratory techniques in order to improve diagnostics and treatment in the clinic and inform public health policies to limit the spread of antibiotic resistant strains [277]. Several WGS studies have already allowed to unveil the current AMR genetic trends of circulating *N. gonorrhoeae* strains in several countries and how the genetic determinants for AMR are transmitted [21, 232, 282, 291, 300], as well as disclose local transmission chains and outbreaks [155, 270, 272, 287, 288]. The ever evolving bioinformatics tools are making it possible to rapidly perform *in silico* isolate typing from genome assemblies either by NG-MAST (*Neisseria gonorrhoeae* Multi-Antigen Sequence Typing) [286] or MLST (Multi-Locus Sequencing Typing) [301], as well as to identify the presence/absence of AMR genetic determinants directly from WGS reads [302] or genome assemblies [303].

In this context, the genetic diversity of the circulating strains in Portugal remains to be disclosed, and the *modus operandi* of the transition to a WGS-based laboratory surveillance of *N. gonorrhoeae* needs to be defined. As such, the National Reference Laboratory for Sexually Transmitted Infections (NRL-STI) of the Portuguese National Institute of Health (NIH) is currently in the process of performing such transition. In a previous Chapter of this dissertation (Chapter III), we described a gene-by-gene approach to classify isolates into WGS-based genogroups. Here, we applied the described methodology in order to disclose potential epidemiological links between isolates, i.e., potential transmission chains occurring in Portugal, based on all sequenced isolates to date. Particularly, we applied this approach to a novel *N. gonorrhoeae* subset for WGS-based surveillance consisting of isolates collected in 2019*, and

*__Author's note__: In this Chapter, we intended to include 110 isolates from 2018 that were sent to the Welcome Trust Institute on behalf of the European Gonococcal Antimicrobial Surveillance Programme (EURO-GASP), but due to the COVID-19 pandemic in 2019, these were not made available in time to be included in this work.

integrated the obtained genomes with the previously sequenced isolates (See Chapter III). For context, in this Chapter we also provide a description of the complete 2019 isolate collection of the Portuguese National Laboratory Network for *Neisseria gonorrhoeae* Collection (PTGonoNET).

6.2. Methods

6.2.1. Portuguese *Neisseria gonorrhoeae* genomic dataset

N. gonorrhoeae isolates analysed in this work are part of the ongoing activity of the PTGonoNet hosted at the NRL-STI of the Portuguese NIH [326]. A total of 734 Portuguese isolates (herein designated as "PT isolates") were enrolled in the present study, including: 589 previously sequenced PT isolates (see Chapter III Section 3.3.1.); 134 PT isolates collected in 2019 (selected to constitute the 2019 WGS-based surveillance dataset; see below); and 11 isolates from 2017 up to 2020 that either presented alarming AMR profiles, were associated with invasive disease or were collected from unusual anatomic sites. Antimicrobial susceptibility testing was performed as previously described (See Chapter III Section 3.3.1.). Briefly, minimum inhibitory concentrations (MICs) for azithromycin, benzylpenicillin, cefixime, ceftriaxone, ciprofloxacin, gentamicin, spectinomycin and tetracycline were determined by E-test (BioMérieux, France) and isolate antibiotic resistance was classified according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoint definitions [94]. Coded isolate designations, antimicrobial susceptibility data and all available anonymized metadata are reported in **Supplementary Table S6.1.** (see also **Supplementary Table S3.1.** from Chapter III for previously sequenced isolates).

To frame the 2019 WGS-based surveillance dataset, we selected isolates from the 2019 PTGonoNET collection based on specimen of origin, geographic location, and antimicrobial susceptibility profiles. The selected dataset includes all isolates hailing from the Center, Alentejo, Algarve and Azores regions, as well as 41 isolates from the Lisbon Metropolitan Area and 39 from the North. It also covers 31 out of the 61 azithromycin resistant isolates, all cephalosporin resistant and potential gentamicin resistant isolates, and isolates presenting either resistance or elevated MICs to all other tested antimicrobials. Noteworthy, this dataset also included the first resistant isolates to ceftriaxone, cefixime and gentamicin detected in Portugal. For specimen type, we included all isolates hailing from the rectum (n=22), articular fluid (n=6), conjunctiva (n=3), blood (n=2), pharynx (n=1) and neovagina (n=1), as well as 9 isolates from the cervix/vagina and 90 isolates from the urethra.

6.2.2. Whole-genome sequencing, genome assembly and typing

DNA was extracted from each isolate using the NucliSens easyMAG platform (BioMérieux, France) for total nucleic acid extraction according to the manufacturer's instructions. DNA was then subjected to Nextera XT library preparation (Illumina, USA) prior to paired-end sequencing (2x250 bp or 2x150 bp) on either a MiSeq or a NextSeq 550 instrument (Illumina, USA), according to the manufacturer's instructions. All genome sequences were assembled using the INNUca v4.0.1 pipeline (https://github.com/B-UMMI/INNUca) [310], as previously described (detailed in Chapter III Section 3.3.2). *In silico* seven loci MLST prediction was performed using *mlst* v2.4 software (https://github.com/tseemann/mlst), NG-MAST was performed using the *ngmaster* v0.4 software [286], and *rplF* fine-typing [330] was performed in order to confirm *Neisseria* species, upon query to the PubMLST *Neisseria* database (http://pubmlst.org). Identification of AMR determinants *in silico* was performed using ARIBA v2.12.2 [302] as previously described (see Chapter III Section 3.3.3.). NG-STAR typing [303] was also performed for interpretation purposes, upon query to the PubMLST Neisseria database.

All isolates were classified into the two-step hierarchical WGS-based genogroups, as implemented in Chapter III (See Section 3.3.5. for details). Briefly, after allele calling using the Maximum shared cgMLST (MScgMLST) scheme, with chewBBACA v2.0.16 [331], the novel genomes' allelic profiles were integrated with the previous 3791 genome dataset (from Chapter III). After minimum spanning tree (MST) reconstruction, using PHYLOViZ [316, 332, 333], the previously determined goeBURST thresholds were applied [i.e., 4.97% and 9.61% allelic differences (AD), corresponding to 40 and 79 AD in 822 loci, respectively] and novel genomes were assigned to genogroups based on their clustering. A novel genogroup code was attributed to isolates that did not cluster with previously determined genogroups. For all isolates, assembly statistics and typing data are reported in **Supplementary Table S6.1**.

6.2.3. cgMLST and cluster analysis

Gene-by-gene analysis was performed using the 1594 loci cgMLST scheme prepared for chewBBACA, with allele calling performed as previously described (See Chapter III Section 3.3.4. for details). For each subset of isolates under comparison, the maximum number of shared loci between them was maximized using PHYLOViZ online 2.0 Beta version (http://online2.phyloviz.net) before MST construction [310, 337]. Allelic distance thresholds for cluster inspection were expressed as percentages of AD over the total number of shared loci under

comparison for every analysis of each subset of isolates at any given level. Genomes with less than 95% of all called loci were excluded from the analysis. Of note, this cut-off resulted in the exclusion of one 2019 PT isolate (NGPT19039) but allowed the recovery of 10 previously excluded PT genomes (e.g. NGPT03012; NGPT04006; NGPT04008; NGPT06012; NGPT06024; NGPT10025; NGPT11094; NGPT12074; NGPT13061; NGPT14055). To investigate sub-clusters with potential epidemiological link, after generating a sub-MST for each low-level genogroup, we applied a cut-off of 1.5% AD (as described in Chapter III Section 3.3.6.), and re-analysed each resulting cluster by maximizing the number of shared loci between them.

6.2.4. Data availability

All reads generated for the present study were deposited in the European Nucleotide Archive under the study accession number PRJEB36482 (individual run accession numbers are detailed in **Supplementary Table S6.1.**). A Microreact project [317] was created to update data with the PTGonoNet 2019 isolates (https://microreact.org/project/q8vwip8qPLJAhRwz3RedFE/ec303d1f).

6.3. Results

6.3.1. Description of the PTGonoNet 2019 collection

For context, as we described the full PTGonoNet isolate collection up to 2018 in Chapter II [326], the following section aims to update the dataset with a short description of the *N. gonorrhoeae* isolates collected during 2019 (see Author's Notes Section). All detailed data are presented in **Supplementary Table S6.2**. Briefly, during 2019, PTGonoNet collected 682 *N. gonorrhoeae* isolates**. Most isolates were collected from the Lisbon Metropolitan Area (n=517) followed by the North (n=110), Center (n=35), Alentejo (n=9), Algarve (n=7), and the Azores Autonomous Region (n=4). The vast majority of isolates were collected from men (621/682; 91.1%), of which 6.6% (41/621) were men who have sex with men (MSM), although sexual orientation was largely undeclared (572/621; 92.1% of all men). Most isolates originated from urogenital tract swabs (669/682; 98.1%), including from the urethra, cervix (or vagina, including a neovagina) and rectum, and 1.9% of isolates were collected from blood, conjunctiva, pharynx and articular fluids. Regarding patient age distribution, most isolates hailed from young adults from 20 to 39 years of age (424/682; 62.2%), followed by adults aged above 40 (128/682; 18.8%) and adolescents and children below 19 years (107/682; 15.7%).

****Author's note**: This corresponds to the total number of isolates that were characterized phenotypically for AMR. The complete collection enrols 213 additional isolates that were received at the beginning of 2020 and due to the COVID-19 pandemic in 2019 were not characterized in time for this work.

Regarding AMR, global trends are presented in **Figure 6.1.**, for all tested antimicrobials. Noteworthy, this dataset included the first isolates detected in Portugal exhibiting resistance to ceftriaxone and cefixime, and potential resistance to gentamicin. Decreased susceptibility to penicillin, tetracycline, gentamicin and ciprofloxacin remained high, with more than 50% of isolates presenting resistance or intermediate resistance to these antimicrobials (**Figure 6.1.**). All isolates were susceptible to spectinomycin. As for azithromycin, results show that ~10% of isolates were resistant.

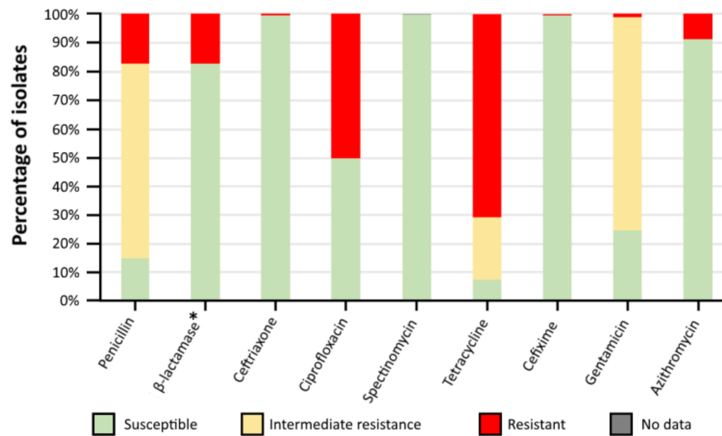

Figure 6.1. | Antimicrobial susceptibility trends of the 2019 PTGonoNet isolate collection. The dataset is composed of 682 isolates. *for β-lactamase data presented refers to positive (red) or negative (green) production.

6.3.2. Description of the Portuguese genomic dataset

The full PT isolate genomic dataset included at least 30 isolates *per* collection year, with the exception of the following periods: i) 2003 up to 2006, where few isolates were available and thus all of them were used; ii) 2013, 2017 and 2019, where more than 100 were sequenced in order to have a broader snapshot within the recent years. Most isolates were collected from the Lisbon Metropolitan Area (n=484) followed by the North (n=139), Center (n=64), Algarve (n=27), Alentejo (n=13), and the Azores Autonomous Region (n=5), with only two isolates having untraceable data. The vast majority of isolates were collected from men (626/734; 85.3%), of which 29.7% were MSM (186/626), although sexual orientation was frequently undeclared. Isolates mostly originated from urogenital tract swabs (683/734; 93.0%), including from the urethra, cervix (or vagina), rectum and scrotum, and 7.0% of isolates were collected from blood, conjunctiva, pharynx, pelvic biopsy, articular and ascitic fluids. Regarding patient age distribution, most isolates hailed from young adults from 20 to 39 years of age (496/734; 67.6%), followed by adults aged above 40 (144/733; 19.6%) and adolescents and children below 19 years (86/734; 11.7%). Global resistance trends of the genomic dataset are presented in **Figure 6.2.A**, which overall reflect those observed for the entire PTGonoNet Collection presented in detail in

Chapter II (**Figure 6.2.B**) [326]. For all antimicrobials tested, 76 isolates were fully susceptible, 650 isolates presented resistance to at least one antimicrobial, and 349 presented simultaneous resistance to three or more antimicrobials.

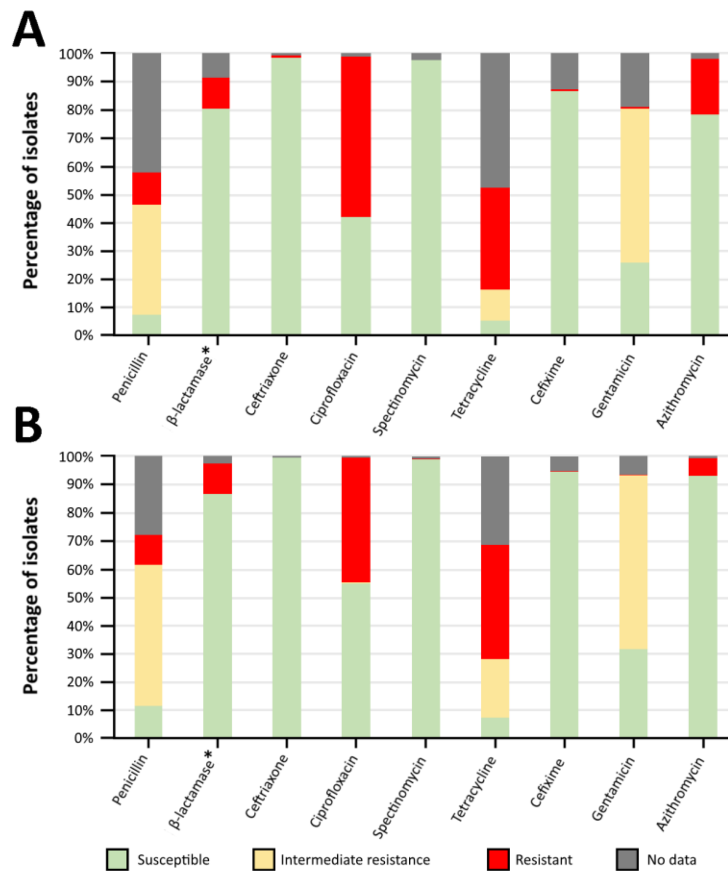

Figure 6.2. | Comparison of antimicrobial susceptibility data of (**A**) the Portuguese genomic dataset and (**B**) the PTGonoNet Collection from 2003 to 2019. The Portuguese genomic dataset and the PTGonoNet Collection enrol 734 and 3278 isolates, respectively. *for β-lactamase data presented refers to positive (red) or negative (green) production.

6.3.3. Overview of *Neisseria gonorrhoeae* genetic diversity circulating in Portugal from 2013 to 2019

A total of 733 validated *N. gonorrhoeae* genomes, of strains collected in Portugal from 2013 to 2019, were analysed in the present study. Of note, the inclusion of the novel sequenced genomes decreased the number of shared loci at 100% in the MScgMLST from 822 to 713 loci. As such, we opted to apply a double approach to classify novel genomes into WGS-based genogroup, i.e., by applying the previously defined goeBURST thresholds of 4.87% and 9.61% to both the MST based on the 100% shared loci (i.e., 713 loci) and the MST based on the 822 MScgMLST loci scheme. Results showed that both approaches yielded identical classifications, reflecting the robustness of cluster stability at these thresholds as described in Chapter III. Overall, PT isolates integrated 82 high-level and 143 low-level genogroups. Novel sequenced PT isolates

were found to be dispersed in the phylogenetic tree (**Figure 6.3.**), indicating that the genetic diversity of the chosen subset most likely represents the 2019 isolates' collection. Low-level genogroups with ≥ five isolates enrolled 561 isolates (76.5% of all PT dataset). The most commonly found genogroups were: G002.001 (105 isolates); G005.004 (57); G001.002 (43); G001.014 (41); G004.003 (32); G006.007 (31); G016.018 (26); G079.129 (20); G002.017 (18); G008.012 (15); G009.009 (13); G012.013 (13); G013.028 (13); G021.029 (12); G127.207 (10); G019.022 (10). Additionally, we observed that 94 low-level genogroups were composed exclusively of PT isolates (enrolling 203 isolates), of which 60 were represented by a single isolate (8.2% of the PT dataset). Regarding the PT 2019 sub-set, isolates were distributed throughout 52 different low-level genogroups (**Figure 6.3.**). Twenty-six novel genogroups were identified, reflecting a previously undisclosed diversity in our broad dataset. Furthermore, the introduction of these novel genomes allowed connecting some previously unlinked genogroups that enrolled one or two isolate, namely: high-level genogroup G068 (G068.119, composed by two isolates) was merged with G034 (i.e., as G034.119); G142, G164 and G166, enrolling each one isolate, were merged with G127, G034 and G112, respectively. The main circulating genogroups identified in this dataset, reflecting the circulating strains in 2019, were: G001.002 (n=19); G079.129 (n=15); G127.207 (n=8); G190.338 (n=7); G004.003 (n=7); G021.029 (n=6); G016.018 (n=5); G127.185 (n=5).

In order to analyse the potential circulating trends of genogroups through time, we examined the relative frequency by year of the ten most commonly detected genogroups observed in our PT genomic dataset (**Figure 6.4.**). Throughout the studied time period, results showed that some of these genogroups predominated in a certain year and then diminished in circulation being detected only punctually later, while for others circulation was maintained in time. For instance, G005.004, which was the predominantly observed genogroup in 2007, was decreasingly detected up to 2012, with only a few cases observed up to 2019. On the other hand, G002.001 and G001.002 were continuously detected in circulation since 2007 and 2010, respectively, with G002.001 being the most detected genogroup between 2010 and 2016 (with the exception of 2014). Noteworthy, results suggest the emergence of genogroup G079.129, as it was the second most commonly found genogroup in our 2019 genomic dataset (**Figure 6.4.**). More importantly, this genogroup was exclusively composed by PT isolates, being firstly detected in 2017 (with only five isolates), and now being represented by an additional 15 novel isolates. Data also indicated that this genogroup was widespread in the country, enrolling isolates from the Lisbon Metropolitan Area, North, Center, Algarve, Alentejo and the Azores Autonomous Region (**Supplementary Table S6.1.** and **Supplementary Table S3.1.** from Chapter III). Still, these overall observations are highly dependent on the isolates sub-sets that are selected for WGS to represent each collection year.

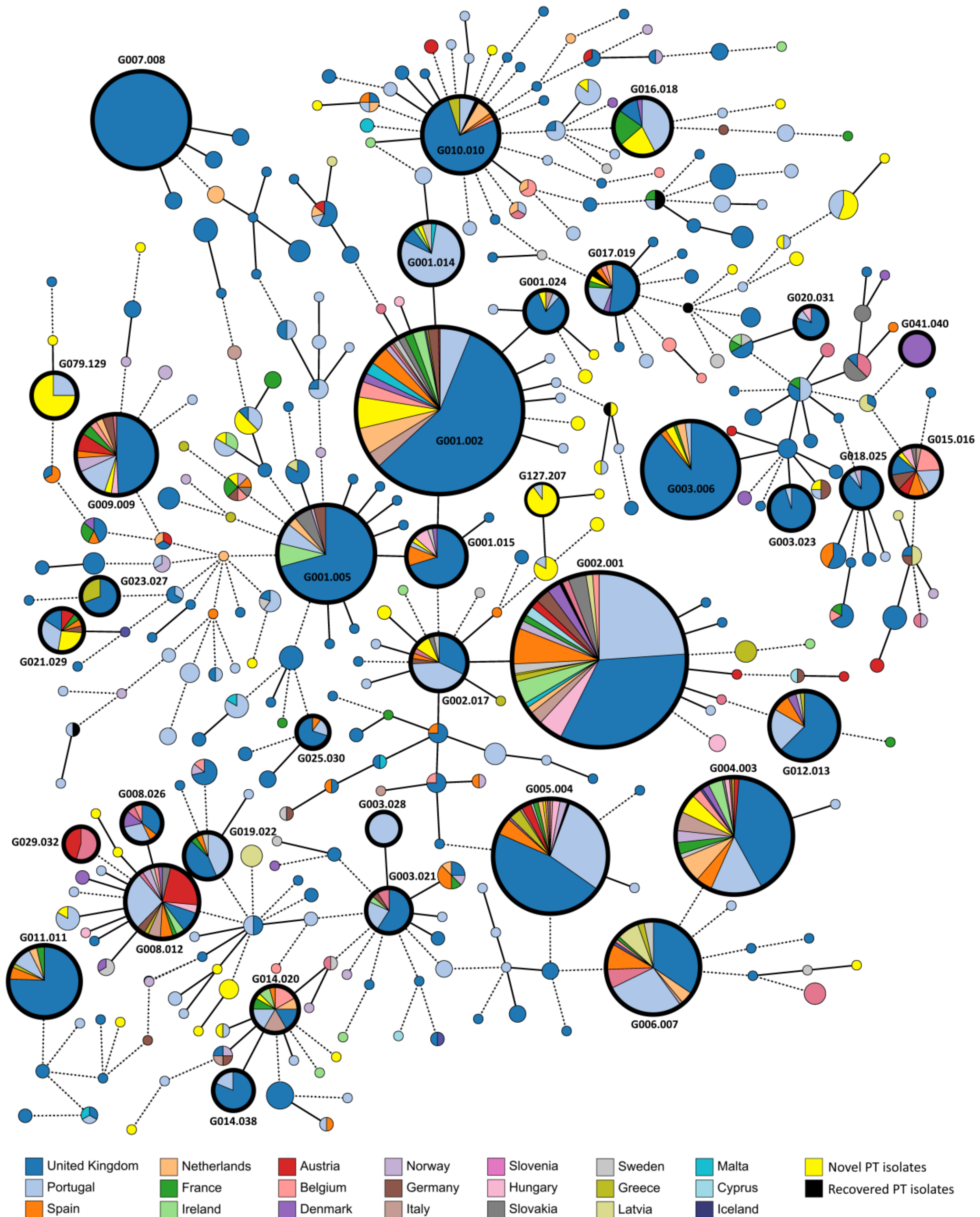

Figure 6.3. | Phylogenetic distribution of novel *Neisseria gonorrhoeae* genomes from Portugal. Minimum spanning tree (MST) was constructed based on allelic diversity found among the MScgMLST scheme (822 loci), enrolling a total of 3945 isolates collected across Europe. All nodes presenting an allelic distance (AD) below 40, corresponding to the low-level genogroup threshold, have been collapsed for visualization purposes. Nodes are coloured according to different countries of origin, novel sequenced PT genomes and recovered PT genomes. Straight and dotted lines reflect nodes linked with the ADs below and above the threshold applied for high-level WGS-based genogroup definition (79 AD), respectively. Low-level WGS-based genogroups comprised of more than ten isolates are highlighted by thicker black circles. MST was generated using the GrapeTree v1.5.0 software [338].

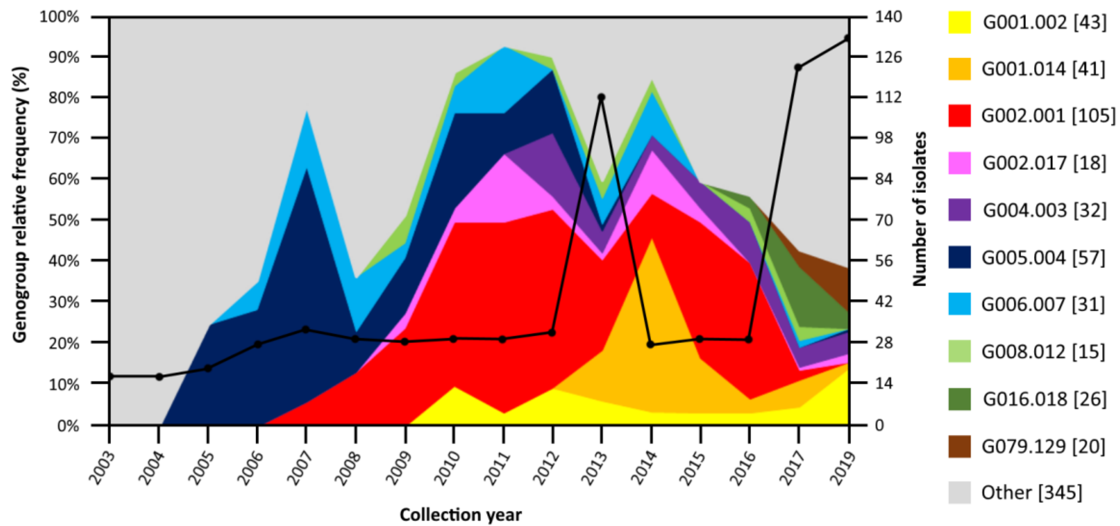

Figure 6.4. | Relative frequency by year of all top ten predominant genogroups observed in our PT genomic dataset. Black line refers to the total number of isolates representing each collection year. Values in parenthesis refer to the total number of PT isolates in each genogroup.

6.3.4. Antimicrobial resistance within *Neisseria gonorrhoeae* genogroups predominant in Portugal

Analysis of the association between AMR carriage and genogroups revealed that the top five most represented genogroups (G002.001, G005.004, G001.002, G001.014 and G004.003) enrol differentially resistant *N. gonorrhoeae* PT isolates (dominant AMR profiles for these genogroups can be explored in **Figure 3.9.** from Chapter III). In fact, G005.004 and G004.003 were found to be mostly associated with ciprofloxacin resistance (with > 50% isolates also presenting resistance to penicillin and tetracycline), which is in line with the overall AMR trends observed in Portugal [326]. As previously mentioned, G005.004 was largely detected in 2007 (**Figure 6.4.**) and 56 out of the 57 PT isolates within this genogroup presented phenotypic resistance to ciprofloxacin, carrying both the *gyrA* (i.e., with alterations 91F and 95G) and the *parC* (i.e., with alteration 87R) alleles associated with resistance [96, 120, 192]. Of note, following these results, the remaining isolate was retested and shown to be resistant (namely NGPT09060, now presenting MIC of 32 mg/L instead of 0.002 mg/L). Additionally, when looking at the entire dataset, of all 274 isolates that compose this genogroup (**Figure 6.3.**), 271 also simultaneously carry both ciprofloxacin resistance-associated alleles. The same trend was observed for G004.003, that emerged in 2012 (**Figure 6.4.**), where all PT isolates present ciprofloxacin resistance. G001.002 enrols 43 isolates hailing from Portugal, of which 29 were phenotypically resistant to azithromycin. We previously observed that azithromycin-resistant isolates could be emerging in this genogroup as, from the 511 G001.002 isolates, 14 of these possess the *23SrRNA* alteration C2597T and one possesses the A2045G associated with high-level azithromycin resistance [96, 120, 192]. Mostly composed of PT isolates (41 out of 49)

(**Figure 6.3.**), G001.014 includes 33 azithromycin-resistant isolates. Here, resistance could not be explained by alterations in *23SrRNA* but may be rather due to the mutational profile of the *mtrR* gene, although the involvement of another mechanism cannot be discarded as 589 isolates that were found to have the same *mtrR* allelic profile do not present azithromycin resistance. Finally, G002.001, which is mostly composed by isolates carrying the *penA* type XXXIV mosaic allele (481 isolates out of 505) associated with resistance to cephalosporins [105], was the most prevalent genogroup in our dataset, enrolling 105 PT isolates. Nevertheless, we observe that the 89 PT isolates that possess this *penA* allele did not present decreased susceptibility to cephalosporins, suggesting that resistance is not exclusively mediated by this mechanism. These isolates may however possess the genetic backbone necessary to acquire the undisclosed genetic alterations driving cephalosporin resistance. Of note, since G079.129 was observed to be a recently emerging genogroup in Portugal, analysis of the AMR profiles of its isolates revealed that these presented simultaneous resistance to ciprofloxacin, tetracycline and penicillin, but were susceptible to the first-line treatment antimicrobials (i.e., azithromycin and cephalosporins).

6.3.5. Identification of potential epidemiologically linked isolates

In order to analyse genetic clusters enrolling PT isolates with the potential to be epidemiologically linked, we applied a conservative threshold of 1.5% AD within each low-level genogroup, after sub-MST generation maximizing the shared loci between isolates at this level. At this high-resolution level, we observed a total of 120 genetic clusters (including pairs of linked isolates) including at least one PT isolate (**Figure 6.5**). These enrolled 536 PT isolates (73.1% of the PT dataset) and 1421 isolates collected across Europe, while 197 PT isolates were found to be unlinked to any other within our dataset. All cluster compositions are detailed in **Supplementary Table S6.3**. Results showed that, at the applied threshold, 302 PT isolates were linked with other isolates collected throughout Europe (i.e., from all 20 countries of our dataset) within 21 different low-level genogroups, forming 37 different clusters (composed of ≥ three isolates), 14 of which had only a single PT isolate (**Figure 6.5**. and **Supplementary Table S6.3.**). The overall size of clusters enrolling PT and European isolates ranged from 3 up to 248 isolates (**Figure 6.6.A**), with five clusters being composed by more than 100 isolates from more than 10 countries (i.e., WGSCL0003, WGSCL0011, WGSCL0020, WGSCL0021, WGSCL0022). Moreover, the mean time interval between the oldest and most recent isolate in each of these clusters was ~4 years (**Figure 6.6.B**), ranging from a few days up to more that thirteen years (**Supplementary Table S6.3.**). Within these European clusters, several PT isolates clustered together (**Figure 6.7.**). In fact, results showed that most clustered PT isolates hailed from the same geographic region, mainly from the Lisbon Metropolitan Area (**Figure 6.7.A and Figures C-I**) or from the North (**Figure 6.7.A and Figure 6.7. C**). More importantly, results also revealed that

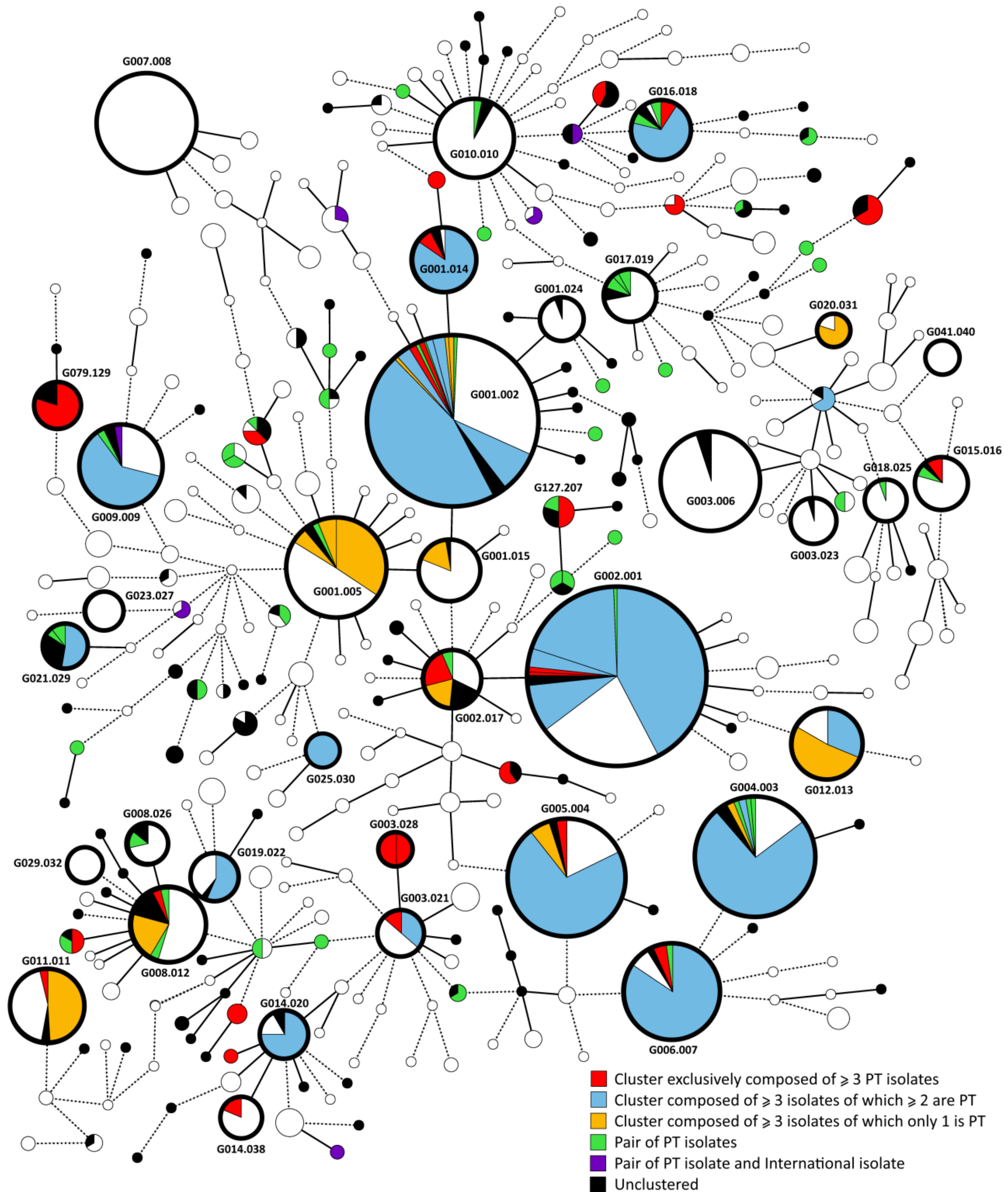

Figure 6.5. | Phylogenetic distribution of *Neisseria gonorrhoeae* genomic clusters that include isolates from Portugal. Minimum spanning tree (MST) was constructed based on allelic diversity found among the MScgMLST scheme (822 loci), enrolling a total of 3945 isolates collected across Europe. All nodes presenting an allelic distance (AD) below 40, corresponding to the low-level genogroup threshold, have been collapsed for visualization purposes. Nodes are coloured according to different types of clusters enrolling PT genomes. Straight and dotted lines reflect nodes linked with the ADs below and above the threshold applied for high-level WGS-based genogroup definition (79 AD), respectively. Low-level WGS-based genogroups comprised of more than ten isolates are highlighted by thicker black circles. MST was generated using the GrapeTree v1.5.0 software [338].

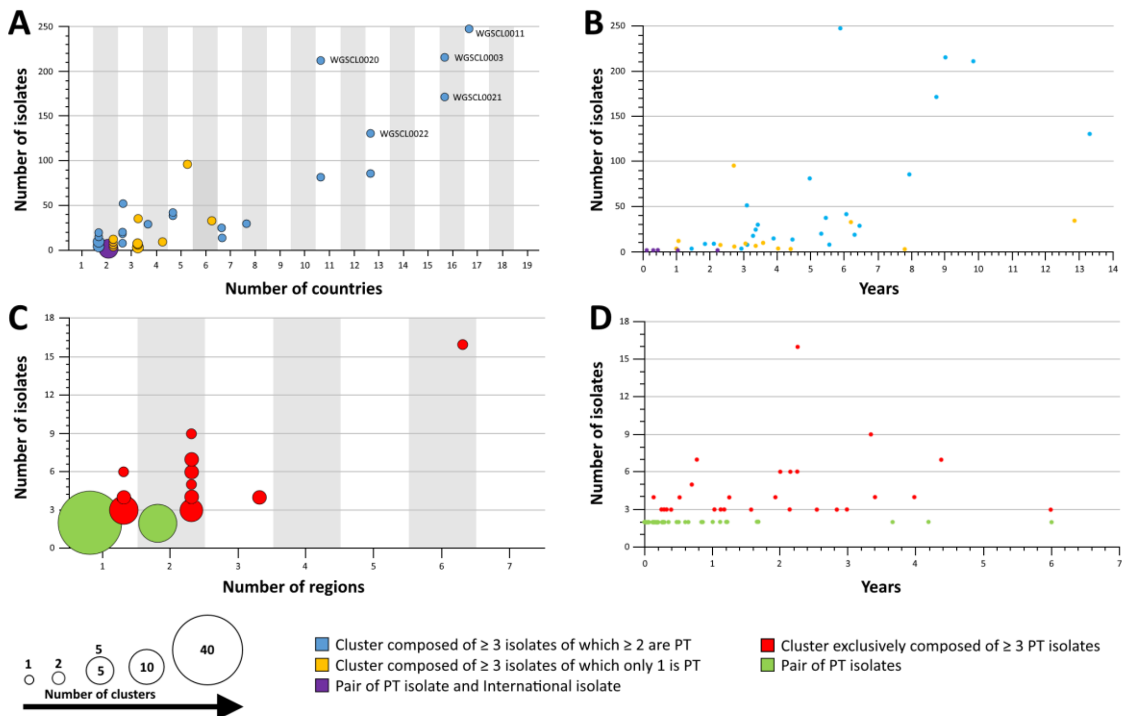

Figure 6.6. | Distribution of genetic clusters enrolling Portuguese (PT) isolates by geographic region and time interval. **A** – Distribution of clusters composed of PT and international isolates by number of enrolled countries. Genetic cluster enrolling more than 100 isolates are discriminated in this panel. **B** – Distribution of clusters composed of PT and international isolates by the time interval (in years) between the oldest and most recent isolate in the cluster. **C** – Distribution of clusters composed exclusively of PT isolates by number of enrolled regions of the country. **D** – Distribution of clusters composed exclusively of PT isolates by the time interval (in years) between the oldest and most recent isolate in the cluster. *y*-axis refers to the number of isolates within each cluster. For panel A and C, dots' sizes' are proportionate to the number of genetic clusters.

isolates belonging to the predominant genogroups observed in Portugal, which were associated with worrying AMR profiles (see Sections above), clustered with other European isolates at this more discriminatory threshold, namely within G002.001 (**Figure 6.7.A**), G005.004 (**Figure 6.7.C**), G001.002 (**Figure 6.7.J**), G001.014 (**Figure 6.7.E**) and G004.003 (**Figure 6.7.B**), enrolling a total of 203 PT isolates. These results suggest the existence of widespread transmission chains in Europe with potential links to cases detected in Portugal. An example of these potential links was detected within the cluster WGSCL0032 (**Figure 6.7.F** and **Supplementary Table S6.3.**) that enrols 82 isolates from 11 different countries (including seven PT isolates), where the PT isolate NGPT13003 was linked at 3 AD to two isolates hailing from the UK. Concordantly, the patient associated with this isolate declared a contact with an *N. gonorrhoeae*-positive sexual partner from the United Kingdom (UK). Additionally, we observed that 106 PT isolates clustered in pairs, six of which were paired with international isolates from Spain, the UK and Slovenia (**Supplementary Table S6.3.**). When analysing the time interval between collection dates, AD, and geographic region, results suggested that some of these paired isolates might represent direct

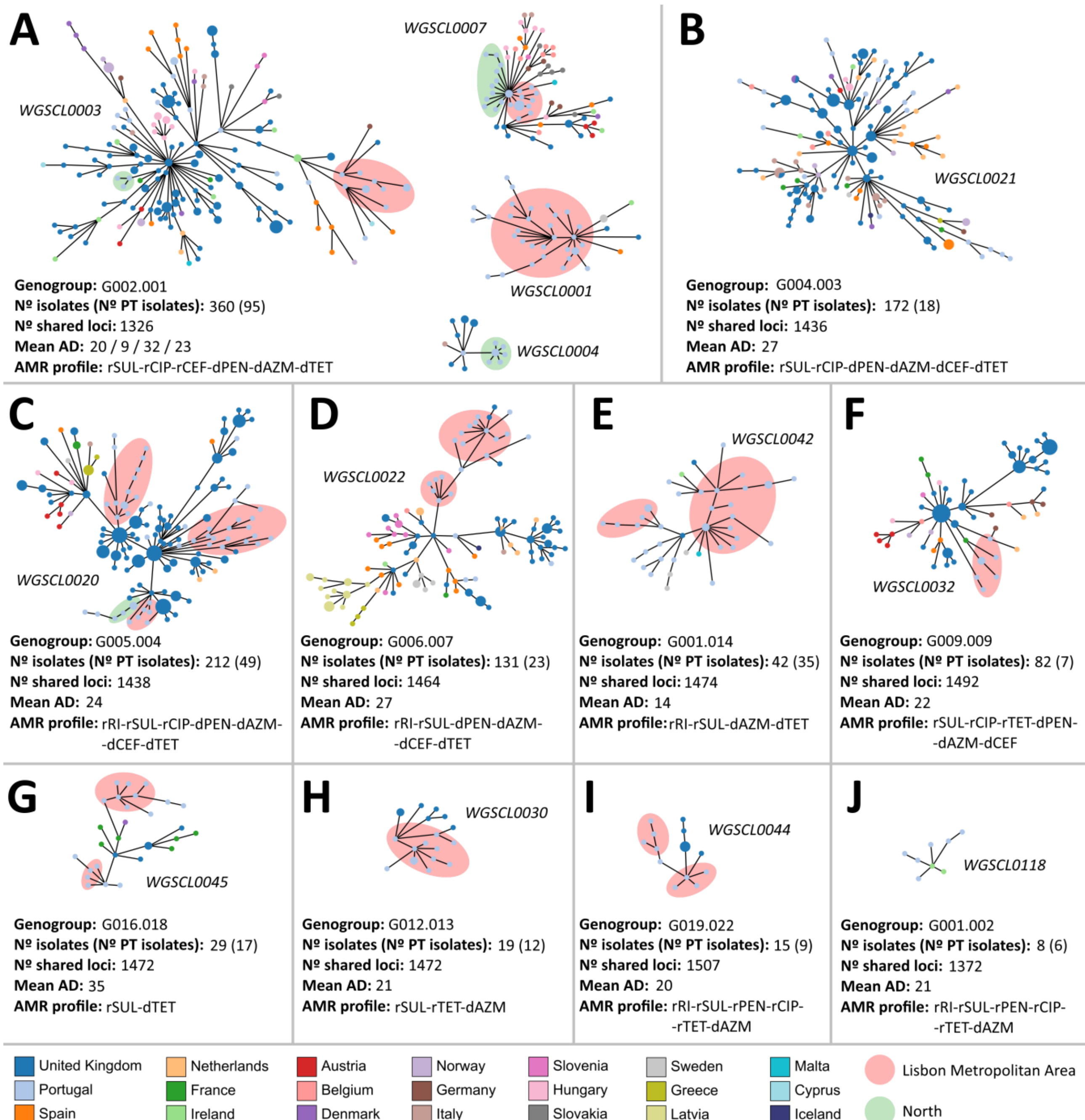

Figure 6.7. | *Neisseria gonorrhoeae* clusters composed by European isolates and more than five Portuguese isolates. Nodes are coloured by country of origin. Clustered PT isolates hailing from the same region according to the second level of the Nomenclature of Territorial Units for Statistics (NUTII) are highlighted. Antimicrobial resistance (AMR) profiles refer to the dominant profile observed within each cluster based on data obtained using ARIBA v2.12.2 [302]. AZM – Azithromycin; CIP – Ciprofloxacin; PEN – Penicillin; RI – Rifampicin; TET – Tetracycline; CEF – cephalosporins; GEN – Gentamicin; SUL – Sulphonamides; rXXX – Resistant to; dXXX – Decreased susceptibility to. MSTs were generated using the GrapeTree v1.5.0 software [338].

sexual contacts. For instance, isolates composing clusters WGSCL0149, WGSCL0230, WGSCL0288 and WGSCL0343 were collected from the same geographic region, with less than six days apart and present up to five ADs (**Supplementary Table S6.3.**). On the other hand, other paired isolates might be related to a more prolonged transmission chain within a particular

geographic region. For example, WGSCL0107 and WGSCL0319 contained two close related isolates each (11 and 5 AD, respectively), collected from the Lisbon Metropolitan Area with a one-year interval. Of note, whenever data was available regarding sexual contacts (which was not collected systematically as the PTGonoNET is an opportunistic collection), all isolates collected from confirmed sexual partners clustered together at this threshold (detailed in **Supplementary Table S6.3.**).

Additionally, data showed the existence of 27 clusters composed exclusively by PT isolates (enrolling a total of 125 isolates and 24 distinct genogroups). Clusters composed of more than five isolates are presented in **Figure 6.8.**, and smaller clusters are presented in **Figure 6.9**. Overall, results suggested the existence of several potential transmission chains in Portugal, with clusters confined to a short period of time (i.e., all isolates collected within the same year) (**Figure 6.6.B**, **Figure 6.8.F** and **Figure 6.8.H**) and others to a geographic region (**Figure 6.6.A**, **Figure 6.8.C**). Still, with the exception of one (WGSCL0086; **Figure 6.8.C**), all larger PT clusters enrolled isolates from at least two distinct regions of the country, suggesting that transmission of *N. gonorrhoeae* is not confined to specific local communities (**Figure 6.6.A**). This was particularly exemplified in the detected G079.129 cluster (WGSCL0237; **Figure 6.8.E**) that, as previously mentioned, involves isolates from six distinct regions of the country, with transmission potentially occurring since 2017, hypothetically originating from the Lisbon Metropolitan Area, as the earliest detected isolates hail from this region. On the other hand, results for two other clusters (WGSCL0258 and WGSCL0335; **Figure 6.8.G** and **Figure 6.8.H**, respectively) suggested the existence of transmission chains in the North, with a potential spread to other regions of the country. Of note, the mean time period within each presented cluster in **Figure 6.8.** was approximately two years, ranging from seven months up to more than four years (**Figure 6.6.B**). Regarding the smaller PT isolate clusters (i.e., enrolling less than five isolates), results showed that 10 out of the 19 clusters were composed of isolates from the same region of the country (**Figure 6.9.**), while two others enrolled isolates from three distinct regions (**Figure 6.9.D** and **Figure 6.9.S**). As observed in the larger cluster, the mean time period within each smaller cluster was approximately two years, ranging from one month up to six years. Still, 11 out of the 27 observed clusters exclusively composed of PT isolates with potential epidemiological link possessed isolates from the latest collection year of the dataset (i.e., 2019), which could suggest that these transmission chains are still active in Portugal.

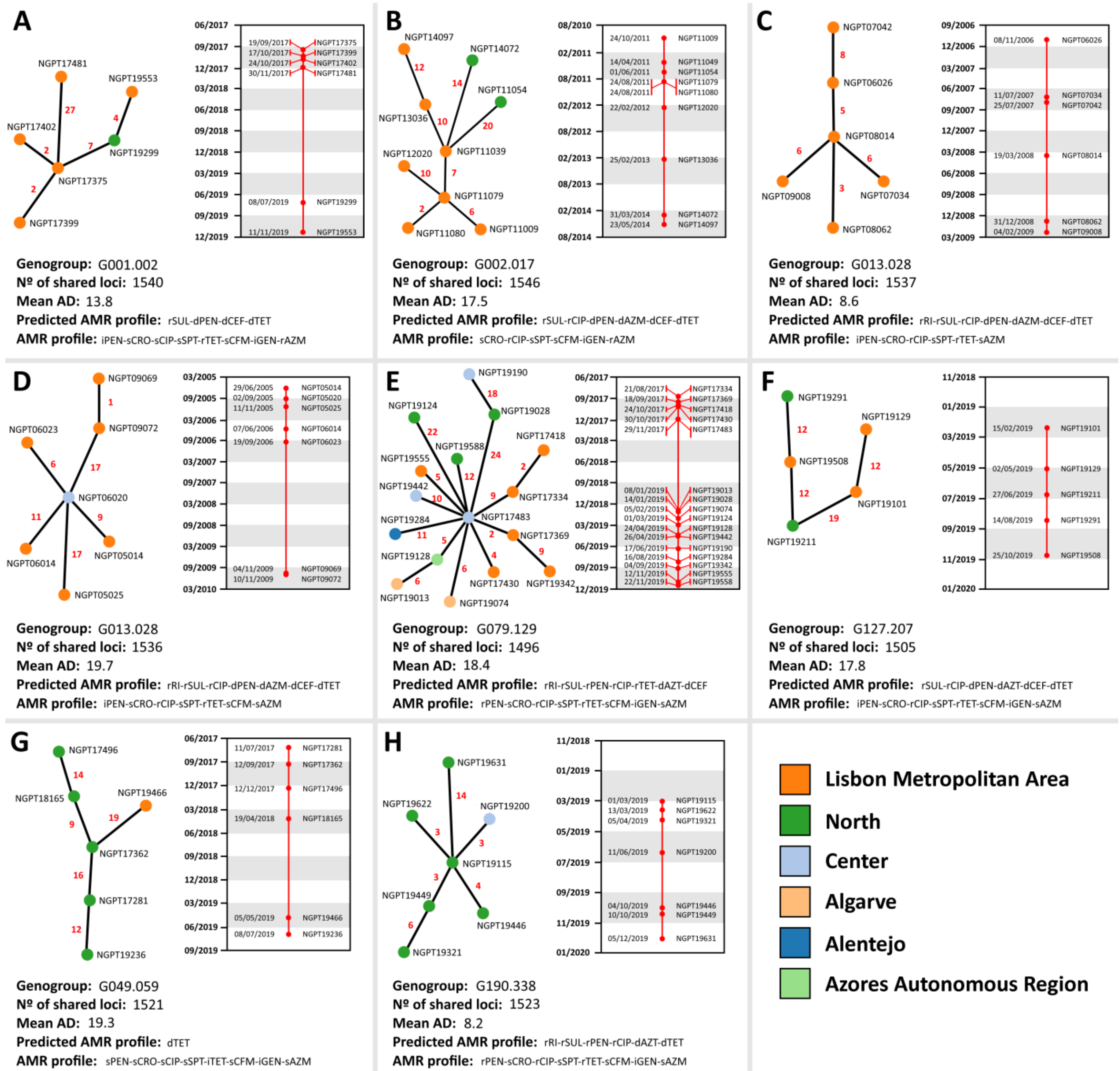

Figure 6.8. | *Neisseria gonorrhoeae* clusters exclusively composed by at least five Portuguese isolates. Nodes are coloured by region according the second level of the Nomenclature of Territorial Units for Statistics (NUTII). The numbers in red on the connecting lines represent the allele differences (AD) between strains. Antimicrobial resistance (AMR) profiles refer to the dominant profile observed within each cluster. Predicted AMR profiles are based on data obtained using ARIBA v2.12.2 [302]. Graphs on the right of the minimum spanning tree (MST) of each panel represent the collection timeline of each isolate involved in the cluster. AZM – Azithromycin; CIP – Ciprofloxacin; CFM – Cefixime; CRO – Ceftriaxone; PEN – Penicillin; RI – Rifampicin; SPT – Spectinomycin; TET – Tetracycline; CEF – cephalosporins; GEN – Gentamicin; SUL – Sulphonamides; rXXX – Resistant to; dXXX – Decreased susceptibility to; sXXX – Susceptible to; iXXX – Intermediate resistant to. MSTs were generated using the GrapeTree v1.5.0 software [338].

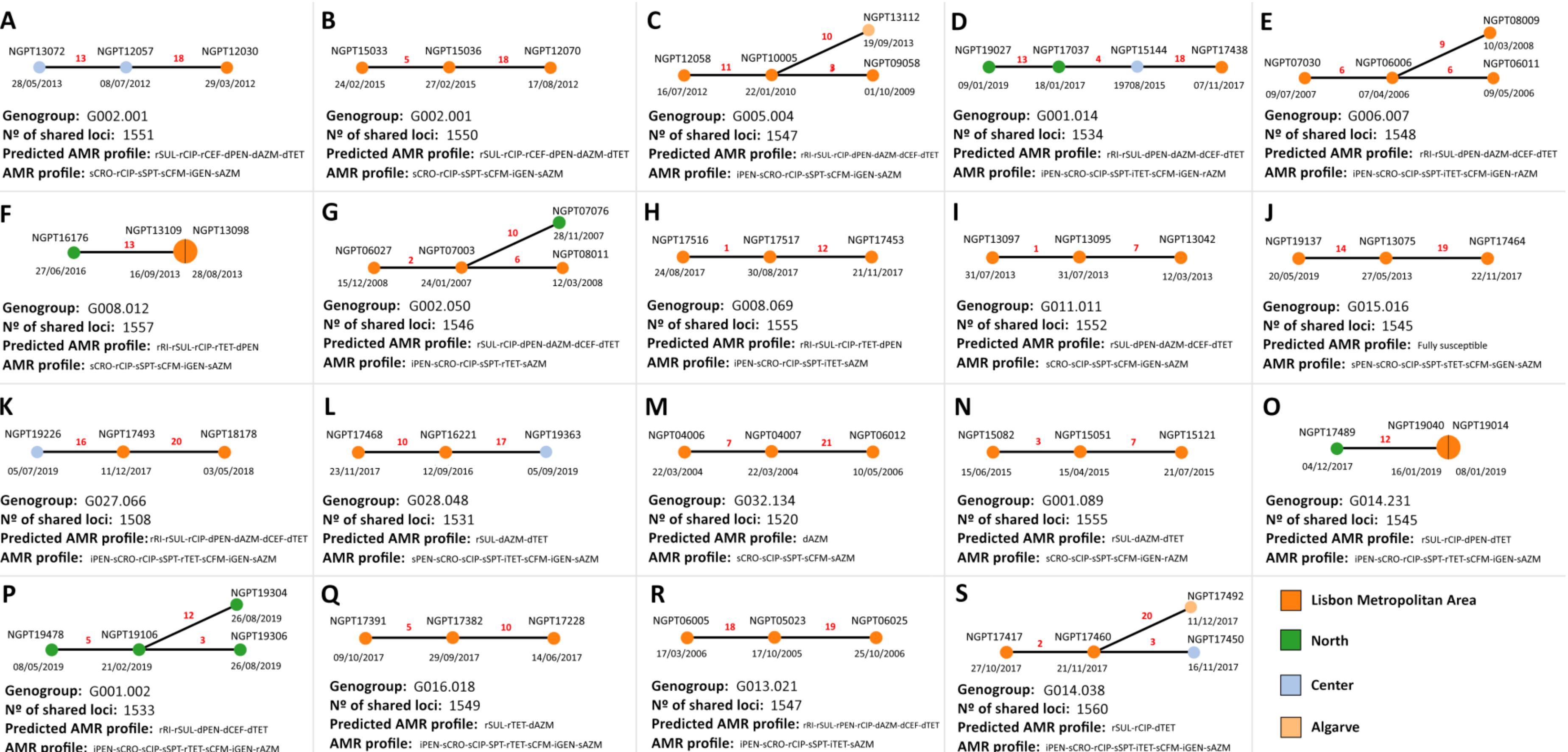

Figure 6.9. | *Neisseria gonorrhoeae* clusters exclusively composed by three or four Portuguese isolates. Nodes are coloured by region according the second level of the Nomenclature of Territorial Units for Statistics (NUTII). The numbers in red on the connecting lines represent the allele differences (AD) between strains. Antimicrobial resistance (AMR) profiles refer to the dominant profile observed within each cluster. Predicted AMR profiles are based on data obtained using ARIBA v2.12. 2 [302]. AZM – Azithromycin; CIP – Ciprofloxacin; CFM – Cefixime; CRO – Ceftriaxone; PEN – Penicillin; RI – Rifampicin; SPT – Spectinomycin; TET – Tetracycline; CEF – cephalosporins; GEN – Gentamicin; SUL – Sulphonamides; rXXX – Resistant to; dXXX – Decreased susceptibility to; sXXX – Susceptible to; iXXX – Intermediate resistant to. Minimum spanning trees were generated using the GrapeTree v1.5.0 software [338]

6.3.6. Highlights of *Neisseria gonorrhoeae* isolates needing close monitoring in Portugal

Within our PT genomic dataset, we observed isolates that warrant a highlight either because they presented an alarming AMR profile or were associated with uncommon clinical manifestations. One such case is the cephalosporin-resistant isolate previously described in Chapter IV. At the time it was analysed and reported, this isolate (NGPT19535) was unlinked to any other in our European-oriented dataset [324]. With the inclusion of the 2019 dataset, we observed that this isolate now integrated genogroup G127.207 (of note, other isolates from this genogroup also clustered together at the applied cut-off of 1.5% AD; **Figure 6.8.F**) and was closely related with isolate NGPT19536, collected on the same date and region, and differing only in the *porB* allele of the NG-STAR typing scheme (**Supplementary Table S6.1.** and **Supplementary Tablse S6.3.**). In fact, both isolates were separated by 16 AD in 1574 shared loci, although the latter presented neither the same overall phenotypic resistance profile nor the same cephalosporin resistance profile (MIC for ceftriaxone and cefixime was ~0.016 mg/L), reinforcing the previous observation that other determinants may be involved in the resistance mechanism of isolate NGPT9535 [324].

We also detected a novel cefixime-resistant isolate (NGPT19511), which was unlinked with any other isolate from our dataset, but belonged to the same genogroup (G035.033) as seven isolates from other European countries from 2013, namely ERR1469575/Spain, ERR1469696/Slovakia, ERR1469739/Norway, ERR1514756/Belgium, ERR1549706/Germany, ERR1602749/France and ERR1602761/France [232] (**Supplementary Table S3.2.** from Chapter III). Phylogenetically, isolates from France and NGPT19511 presented 49 AD and 56 AD from all others, respectively, i.e., these isolates were unlinked when applying our cut-off for potential epidemiological investigation. More importantly, all eight isolates possess the same *penA* type X mosaic allele (not observed in any other PT isolates) and all seven non-PT isolates also presented phenotypic resistance to cefixime (i.e., MIC ≥ 0.125 mg/L, ranging from 0.125 up to 0.5 mg/L) [232]. These results suggest a potential introduction of this isolate in the Lisbon Metropolitan Area in 2019, so awareness needs to be raised to monitor its circulation.

In addition, within our dataset we observed two genogroups that enrolled isolates associated with recent cases of invasive gonorrhoea (isolates collected from blood or articular fluid, from 2017 up to 2019), namely G045.059 (NGPT17281, NGPT18165, NGPT19268 and NGPT19466) and G127.207 (NGPT18321, NGPT19101, NGPT19129 and NGPT19291). More importantly, data showed that within each genogroup these isolates were linked at the 1.5% AD threshold, with isolates from G045.059 (namely NGPT17281, NGPT18165 and NGPT19466) being linked to other isolates detected in the North region (**Figure 6.8.G**) and isolates from G127.207 (namely NGPT19101, NGPT19129, NGPT19291) being associated with both the

Lisbon Metropolitan Area and the North regions (**Figure 6.8.F**). Considering these results and the absence of a confirmed epidemiological linkage between these cases, we hypothesize that these isolates may possess an increased ability to progress towards invasiveness, so further investigation is required in order to assess this.

6.4. Discussion

We report the findings of the first nationwide genomic epidemiology *N. gonorrhoeae* study in the era of WGS-based laboratory surveillance performed in Portugal, enrolling isolates collected on behalf of the PTGonoNet [326]. Here, we analysed 733 *N. gonorrhoeae* genomes from isolates collected from 2003 up to 2019, with countrywide coverage and presenting distinct AMR profiles, which were integrated in a European-oriented genomic dataset. We aimed to strengthen the epidemiological surveillance of *N. gonorrhoeae* in Portugal alongside a more detailed analysis of the circulating AMR determinants. Results showed that the overall genetic diversity of *N. gonorrhoeae* circulating in Portugal is high, as isolates from our genomic dataset were classified into 142 distinct low-level genogroups (**Figure 6.3.**). Still, within the studied time period, we found that different genogroups were predominantly detected in a certain year, with some diminishing in circulation throughout time and others whose circulation were maintained in the population (**Figure 6.4.**). These results reflect the existence of likely confined transmission chains at country or regional level, while also illustrating the continuous cross-border spread of *N. gonorrhoeae* that introduces novel strains into a country due, for instance, to large international sexual networks. In fact, the most predominant genogroups observed in our PT genomic dataset, namely G002.001, G005.004, G001.002, G004.003 and G006.007, were also the ones most detected in Europe (**Figure 6.3.** and **Table 3.2.** from Chapter III). Of note, G002.001 (associated with NG-MAST genogroup G1407) is a commonly reported *N. gonorrhoeae* lineage in association with AMR with worldwide coverage [21, 232, 233, 236–241, 324]. Overall, data also showed that these genogroups were associated with decreased susceptibility to antimicrobials, particularly to azithromycin, cephalosporins and ciprofloxacin, although sampling bias cannot be discarded (i.e., we aimed to have a diversified coverage, but selection also privileged isolates that presented AMR profiles of interest). These results allow the rapid identification of specific genogroups whose local spread needs to be controlled, which in turn enhances *N. gonorrhoeae* surveillance by promoting the prospective monitoring of genogroup frequency and geographic spread, towards more oriented Public Health actions to control the spread of *N. gonorrhoeae* AMR. In fact, the ultimate purpose of a robust molecular surveillance system (based on discriminatory genome sequencing data) is to provide the basis for a more oriented surveillance and action by Public Health Authorities, i.e., it to give the opportunity for public health efforts/resources to be prioritized according to the pathogen circulation landscape in each country.

For example, knowing that a particular multi-drug resistant genogroup is circulating in a given geographic region or specific communities may allow the implementation of prophylactic/treatment measures more tailored to control the spread of such genogroups.

Insights on the clustering at a higher discriminatory level (i.e., at a threshold below the ones representing cluster stability) further supported the suitability of gene-by-gene typing strategies to detect/track outbreaks and transmission chains as a means to promote and/or support epidemiological investigation [310]. Here, we observed 120 clusters enrolling at least one PT isolate (**Figure 6.5.**), 43 of which involving other isolates collected in Europe and, more importantly (**Figure 6.7.**), 77 that were composed exclusively of PT isolates (**Figure 6.8.**, **Figure 6.9.** and **Supplementary Table S6.3.**). These results provide a comprehensive breakdown of the *N. gonorrhoeae* transmission chains in Portugal, some of which seem to be still active given the data from the latest collected isolates. As such, the applied strategy allows the identification of specific clusters that should be subjected to fine-tuned genomic analysis and ultimately to launch epidemiological investigations. We hope that the release of the ongoing 2018 EURO-GASP genomic survey data will aid in understanding how the circulation landscape of *N. gonorrhoeae* in Europe has evolved in more recent years. In addition, the integration of these data will be essential to understand if some of the alarming cases detected in Portugal, highlighted throughout this Chapter, have also been observed in other European countries. Within our PT genomic dataset, we found several genetic profiles that need short-term close monitoring as they were associated either with uncommon clinical manifestations or presented alarming AMR profiles. For instance, we detected two recent cases of cephalosporin-resistant isolates, the first detected in our country since 2003 [324, 326], whose potential spread needs to be monitored and ultimately controlled. Our results also suggest, in several instances, that the mechanisms underlying *N. gonorrhoeae* AMR are still not fully disclosed. In addition, some isolates may possess the genetic backbone necessary to rapidly acquire the genetic alterations driving resistance, either through recombination or point mutation, given the high genetic relatedness to resistant isolates (e.g., isolates from the cephalosporin-associated genogroup G002.001). Further research is required to understand the linkage between resistance and the simultaneous carriage of genetic AMR determinants, particularly the role of epistatic interactions at the level of antimicrobial susceptibility (i.e., driving differential MIC values), as the example of those observed between *gyrA* and *parC* ciprofloxacin-associated resistant variants [350]. Additionally, the inclusion of the 2019 dataset allowed us to identify clusters enrolling genetically close related isolates associated with invasive gonorrhoea. Particularly, in one of these clusters (**Figure 6.8.G**), most isolates were collected in the North region of the country, raising concerns about a potential local spread, and the hypothesis that these isolates may possess an increased ability to progress towards invasiveness. In fact, a recently published study has also reported the existence of a *N. gonorrhoeae* strain associated with endocarditis in the North (in Penafiel) [351]. Reports from the

Centers for Disease Control and Prevention in 2019 also described an increased cluster of *N. gonorrhoeae* related with Disseminated Gonococcal Infections in the USA [352], which is in line with our hypothesis, although WGS data will be necessary to understand if these cases fall within the same genogroup as our PT isolates.

Finally, in the present study we focused on providing a comprehensive snapshot of the main *N. gonorrhoeae* genogroups circulating in Portugal and their association with AMR. This has been achieved through the enrichment of the current geo-temporal and genomic diversity of *N. gonorrhoeae*, by adding WGS data of isolates spanning 15 years of *N. gonorrhoeae* surveillance in Portugal. The integration of the novel PT genomes and classification into genogroups reflects the robustness of the cluster stability observed at the previously described thresholds (Chapter III), reinforcing that the genogroups defined at these thresholds mirror the major circulating *N. gonorrhoeae* lineages. Overall, this effort constituted a turning point to consolidate the genomic epidemiology of gonococci in Portugal that will certainly need to be maintained and carried forward at the Portuguese NRL-STI. At this time, the best approach to accomplish the latter is still under consideration. For instance, we aim to take advantage of the promising novel developments of chewBBACA, namely the release of the Chewie Nomenclature Server [353]. This approach may allow maintaining the allelic profiles and loci definition of our chewBBACA-adapted cgMLST scheme, the rapid identification of the defined genogroups, as well as facilitating data harmonization and communication at national and international level. As WGS increasingly becomes more cost-effective, it is important to be able to provide the necessary bioinformatics tools (preferably freely-available, which is the case of chewBBACA [331]), to perform automated analysis of WGS data and communicate harmonized gonococci surveillance data at national and international levels. Our results show that the applied framework can be used to identify (and dynamically visualize) clusters with associated epidemiological and AMR data, with benefits for clinical and public health purposes, such as to orient testing and screening, tailor treatment strategies, or promote public health actions for infection control in risk groups at national or international levels. We anticipate that the major future hurdle in the application of this approach will be the optimization of the underlying WGS/bioinformatics methods directly to biological samples, considering the fact that, very soon, *N. gonorrhoeae* culture will be replaced by nucleic acids detection in most labs.

CHAPTER VII

Novel approach to capture the genome of *Neisseria gonorrhoeae* directly from clinical samples

---

Personal contributions

MP contributed to the study design, performed wet-lab experiments and research, analyzed all data and wrote the manuscript.

7. Novel approach to capture the genome of *Neisseria gonorrhoeae* directly from clinical samples

7.1. Introduction

Neisseria gonorrhoeae is generally transmitted from infected individuals during sexual activity, by direct contact between the mucosal membranes of the urogenital tract, anal canal, and oropharynx. Occasionally it also causes infections in other anatomic sites, such as the eye, joints/tendons, the heart or even the brain, causing severe complications like arthritis, endocarditis or meningitis [22, 23]. The ability of the gonococcal genome to undergo continual mutation and internal recombination is well known, resulting in rapidly evolving gonococcal populations, and several mutations have been linked to antimicrobial resistance (AMR) or to decreased susceptibility to a given antimicrobial [100, 120, 192, 232]. *N. gonorrhoeae* diagnosis is currently performed either by culture or nucleic acids amplification technologies (NAAT). Nevertheless, culture is progressively being abandoned in favour of NAAT, as studies have shown that these: i) offer higher sensitivity; ii) can be performed successfully on diverse specimen types (and the bacteria does not need to be viable); iii) provide faster results, with less hands-on time; and iv) can be scalable and automated for high throughput testing, making it cheaper [24]. Noteworthy, these tests are particularly important for oropharyngeal samples as there is an increasing number of confirmed cases of infection on these specimens, and *N. gonorrhoeae* isolation from this anatomic location is challenging and often unsuccessful [24, 85]. Still, this transition has a huge impact on AMR surveillance as these tests cannot yet provide information on antimicrobial susceptibility, which still requires isolation of the bacteria by culture. Adding to this scenario, the notification of gonorrhoea cases (either confirmed by culture or NAAT) cannot include AMR data due to the lack of standardized cultural methods and to the unavailability of molecular AMR tests.

In the last decade, there has been a call for novel strategies and technologies that could aid in the surveillance and management of *N. gonorrhoeae* AMR [276, 277, 300]. As such, Next-Generation Sequencing (NGS) coupled with bioinformatics has the potential to be invaluable tools in this field. In fact, whole-genome sequencing (WGS) data are now being generated in order to monitor global *N. gonorrhoeae* epidemiology, transmission chains and AMR trends [21, 106, 155, 192, 232, 272, 300]. Additionally, WGS data, either using Illumina [192] or Oxford Nanopore technologies [306], and bioinformatics are now starting to be used to predict AMR *in silico*, based on the combined knowledge of *N. gonorrhoeae* AMR genetic drivers and previously obtained phenotypic data. Deep-sequencing technologies can aid in rapidly disclosing the current trends in AMR, as they are scalable for high-throughput testing, can be applied directly on clinical

samples or bacterial cultures and provide genomic data on fixed and potentially emerging mutations [277].

In recent years, we have successfully applied a culture-independent approach to capture and sequence the whole genome of bacteria directly from clinical samples, namely for the sexually transmitted pathogens *Treponema pallidum* [354] and *Chlamydia trachomatis* [355], and for the respiratory pathogen *Mycobacterium tuberculosis* [356]. These studies have allowed to bypass the culture bottleneck and to gain insight into the genetic diversity of intra- and inter-patient infecting bacterial populations [354, 355], as well as to analyse genetic AMR determinants [354, 356]. For additional context, every year, the National Reference Laboratory for Sexually Transmitted Infections (NRL-STI) at the Portuguese National Institute of Health (INSA) performs routine diagnosis of *N. gonorrhoeae* (along with other sexually transmitted pathogens) on thousands of clinical specimens using NAAT. As such, since 2017, the NRL-STI started an ongoing collection of clinical samples, enrolling specimens that tested positive for *N. gonorrhoeae*. Up to 2019, this collection included more than 1000 samples, namely urine samples and swabs from the urethra, cervix, pharynx and rectum, some of which relate to the same patients, in pairs or triplets (i.e., collected from distinct anatomic sites) (data summarized in **Table 7.1.**). As such, the main objective of the present Chapter is to apply this culture-independent approach to *N. gonorrhoeae,* to obtain WGS data directly from complex clinical specimens from distinct anatomic sites. Particularly, we aimed to bypass the culture step prior to WGS, and assess the ability to perform *in silico* typing, phylogenetic inferences and predict AMR profiles based on the obtained genomic data.

Table 7.1. | Distribution of *Neisseria gonorrhoeae*-positive clinical samples available at INSA.

	Year			
	2017	**2018**	**2019**	**Total**
# Clinical Samples	78	397	552	1027
Sample Pairs	9	61	81	151
Sample Triplet	1	9	14	24
Gender				
M	65	374	516	955
F	12	23	31	66
Unknown	1	.	5	6
Specimen type				
Vaginal swab	9	10	17	36
Urethral swab	13	40	56	109
Anal swab	25	117	154	296
Oropharyngeal swab	19	145	194	358
Urine	12	85	131	228

Sample Pairs and Sample Triplet refer to samples collected from distinct anatomic sites of the same patient.

7.2. Material and methods

7.2.1. *Neisseria gonorrhoeae*-positive clinical samples

The selected clinical samples for the present study are part of an ongoing collection of positive samples (more than 1000 from 2017 up to 2019) that is being created at INSA (**Table 7.1.**), enrolling specimens that tested positive for *N. gonorrhoeae* during the NRL-STI's diagnosis duty (using routine NAAT). For the present study, 35 clinical samples were selected, fifteen of which were pairs from the same patient, with sample description presented in **Table 7.2.** Total genomic DNA was extracted from all clinical specimens using the QIAamp DNA Mini Kit (Qiagen, Venlo, Netherlands) according to the manufacturer's instructions.

Table 7.2. | Description of all clinical samples enrolled in the present study

Sample	Patient gender	Patient age	Sampling date	Geographic Region	Specimen type
NGCL001_A	M	34	20/09/2017	Lisbon MA	Anal swab
NGCL002_A	M	38	21/05/2018	Lisbon MA	Anal swab
NGCL003_Ap1	M	32	31/10/2017	Lisbon MA	Anal swab
NGCL003_Op1	M	32	31/10/2017	Lisbon MA	Oropharyngeal swab
NGCL004_Ap2	M	39	29/11/2017	Lisbon MA	Anal swab
NGCL004_Op2	M	39	29/11/2017	Lisbon MA	Oropharyngeal swab
NGCL005_Ap3	M	37	12/12/2017	Lisbon MA	Anal swab
NGCL005_Op3	M	37	12/12/2017	Lisbon MA	Oropharyngeal swab
NGCL006_Ap4	M	No data	19/12/2017	Lisbon MA	Anal swab
NGCL006_Op4	M	No data	19/12/2017	Lisbon MA	Oropharyngeal swab
NGCL007_Ap5	M	37	04/01/2018	Lisbon MA	Anal swab
NGCL007_Op5	M	37	04/01/2018	Lisbon MA	Oropharyngeal swab
NGCL008_Ap6	M	43	21/02/2018	Lisbon MA	Anal swab
NGCL008_Op6	M	43	21/02/2018	Lisbon MA	Oropharyngeal swab
NGCL009_Ap7	M	18	14/03/2018	Lisbon MA	Anal swab
NGCL009_Op7	M	18	14/03/2018	Lisbon MA	Oropharyngeal swab
NGCL010_Ap8	M	22	14/03/2018	Lisbon MA	Anal swab
NGCL010_Op8	M	22	14/03/2018	Lisbon MA	Oropharyngeal swab
NGCL011_U	M	No data	11/01/2019	Lisbon MA	Urine
NGCL012_Ap9	M	22	31/10/2018	Lisbon MA	Anal swab
NGCL012_Op9	M	22	31/10/2018	Lisbon MA	Oropharyngeal swab
NGCL013_Ap10	M	40	26/11/2018	Lisbon MA	Anal swab
NGCL013_Op10	M	40	26/11/2018	Lisbon MA	Oropharyngeal swab
NGCL014_Ap11	M	20	26/11/2018	Lisbon MA	Anal swab
NGCL014_Op11	M	20	26/11/2018	Lisbon MA	Oropharyngeal swab
NGCL015_Ap12	M	36	05/12/2018	Lisbon MA	Anal swab
NGCL015_Op12	M	36	05/12/2018	Lisbon MA	Oropharyngeal swab
NGCL016_Ap13	M	No data	20/02/2019	Lisbon MA	Anal swab
NGCL016_Op13	M	No data	20/02/2019	Lisbon MA	Oropharyngeal swab
NGCL017_Ap14	M	30	12/12/2018	Lisbon MA	Anal swab
NGCL017_Op14	M	30	12/12/2018	Lisbon MA	Oropharyngeal swab
NGCL018_Ap15	M	45	17/12/2018	Lisbon MA	Anal swab
NGCL018_Op15	M	45	17/12/2018	Lisbon MA	Oropharyngeal swab
NGCL019_UR*	M	22	30/11/2017	Lisbon MA	Urethral swab
NGCL020_UR**	M	No data	19/12/2017	Lisbon MA	Urethral swab

Sample ID tags are composed by the sample name, specimen type (A – anal; O – Oropharyngeal; U – Urine; UR – Urethral) and the pair code (p#). MA – Metropolitan Area. *associated isolate NGPT17487 (accession #ERS4271282). **associated isolate NGPT17512 (#ERS4271296).

7.2.2. Generation of standard curves for real-time quantitative PCR (qPCR)

To quantify the number of *N. gonorrhoeae* genomes in each sample, a plasmid standard curve was generated as previously described for other pathogens [354, 357]. Primers for the conserved gene *NGO_0870* (currently updated to *NGO_RS04315*) were designed based on constant regions (primers NGO_870-A CCGACCTACGGCAGAATCC and NGO_870-B GCTGCGGAACTGGTCGAA were designed using Primer Express software; Applied Biosystems, MA, USA), according to the reference genome sequence of *N. gonorrhoeae* strain FA1090 (Genbank accession #AE004969.1). Briefly, an amplified fragment (66 bp) of *NGO_0870* was cloned into the pJET1.2 Cloning Vector (Thermo Scientific, MA, USA) according to the manufacturer's instructions. After transformation of DH5α *E. coli* with the cloned vector and subsequent overnight propagation, plasmid DNA was purified and transformation was confirmed by PCR and Sanger sequencing. The plasmid copy number was calculated according to the formula: Nº plasmid/mL = [Avogadro's Nº x Plasmid conc. (g/mL)] / MW of 1 mol of plasmids (g). The standard curve consisted of eight-serial plasmid dilutions ($1x10^1$ to $1x10^8$ plasmid copies/µL). The number of human cells per sample was quantified by a similarly generated plasmid standard curve using an amplified fragment (73 bp) of a single copy human gene (*β-actin*) cloned in a similar vector, as previously described [357] (primers β-actin-3 GGTGCATCTCTGCCTTACAGATC and β-actin-4 ACAGCCTGGATAGCAACGTACAT).

7.2.3. qPCR for quantification of *Neisseria gonorrhoeae* and human cells in clinical samples

The real-time quantification was performed using the Light-Cycler® 480 SYBR Green chemistry and optical plates (Roche Diagnostics, Basel, Switzerland). The qPCR reagents consisted of 2 x SYBR Green I Master Mix, 400 nM of each primer and 5 µL of DNA sample in a final volume of 25 µL. The thermocycling profile was: 10 min/95 °C followed by 40 cycles of 15 s/95 °C and 1 min/ 60 °C. Specificity was checked by generating the dissociation melting curves. Absolute quantification of bacterial and human genomes was calculated in relation to the respective plasmid standard curve. Final number of *N. gonorrhoeae* and human genome copies was calculated in relation to the volume of input DNA used downstream (see next section).

7.2.4. Design and synthesis of RNA "baits" for DNA capture

In order to capture *N. gonorrhoeae* DNA directly from clinical samples, complementary RNA oligonucleotide "baits", 120 nucleotide in size, were designed to span the ~2.1 Mb of the *N. gonorrhoeae* genome, as previously described [354]. As such, the complete reference genome sequences of *N. gonorrhoeae* strains FA1090 (RefSeq accession #NC_002946), FDAARGOS_205 (#NZ_CP020418) and NCCP11945 (#NC_011035) were fragmented *in silico* into 120 bp sequences, as well as fourteen plasmid sequences [pAustralian (#NC_025191), pCmGFP (#NC_011521); pEM1 (#NC_019211); pEP5289 (#NC_014105), pJD1 (#NC_001377), pJD4 (#NC_002098), pNG869_1 (#NZ_CM003346), pNG869_2 (#NZ_CM003347), pNG869_3 (#NZ_CM003348), pNGK (#NC_011034), pSJ5.2 (#NC_010881), punamed_MS11 (#NC_022243), punamded_1 (#NZ_CP020416), punamed_2 (#NZ_CP020417)]. All resulting sequences were BLASTn searched against the Human Genomic + Transcript database to excluded homologous sequences to the human genome. A total of 49814 RNA probes were generated, and this custom bait library was then uploaded to the SureDesign software (https://earray.chem.agilent.com/suredesign/) and synthesized by Agilent Technologies (Agilent Technologies, CA, USA). Before enrichment and WGS, DNA samples were quantified using Qubit HS kit (Invitrogen, MA, USA) to calibrate the input to 10-200 ng DNA in 7 µL.

7.2.5. SureSelect$^{XT\ HS}$ *Neisseria gonorrhoeae* enrichment and WGS directly from clinical samples

N. gonorrhoeae whole-genome capture from DNA samples was performed using Agilent Technologies' SureSelect$^{XT\ HS}$ target enrichment system for Illumina paired-end multiplexed sequencing library protocol (G9702-90000, version C0, September 2018; Agilent Technologies, CA, USA) upon enzymatic fragmentation with SureSelect low input enzymatic fragmentation kit (G9702-90050, Revision A0, September 2018; Agilent Technologies, CA, USA), according to the manufacturer's instructions. *N. gonorrhoeae* enriched libraries were subjected to cluster generation and paired-end sequencing (2×150 bp) in an Illumina NextSeq 550 equipment, according to the manufacturer's instructions.

7.2.6. *Neisseria gonorrhoeae* read filtering and genome assembly

All raw reads produced were taxonomically classified with Kraken v2 [358] with default parameters, using the pre-built MiniKraken 8Gb database (available at https://ccb.jhu.edu/software/kraken/). A first draft assembly was constructed using INNUca

v4.0.1 pipeline [310] (see Chapter III Section 3.3.2 for details) and assemblies filtered by contig length (contigs with ≥ 200 bp), GC content (contigs with GC content between 5.0% and 95%) and k-mer coverage (k-mer coverage of ≤ 2) were re-analysed using Kraken. All contigs classified up to the genus *Neisseria* or up to the species *Neisseria gonorrhoeae* were then filtered into a new draft assembly. Finally, trimmed reads (produced during the INNUca pipeline using Trimmomatic v0.36 [327] were mapped to the resulting draft assembly using Bowtie2 v2.2.9 [359], extracted and re-analysed using Kraken. All mapping reads were then used to construct a final assembly using the INNUca pipeline. First and final draft assembly statistics for all samples are presented in **Table 7.3**. Additionally, raw reads were mapped to a human reference genome (RefSeq accession #NC_000001-NC_000024 and #NC_012920), using Bowtie2 v2.2.9, in order to assess the success of bacterial DNA enrichment.

Table 7.3. | Assembly statistics for the first and final draft assemblies.

Sample	First draft assembly					Final draft assembly		
	Mean Depth of Coverage	# of contigs	Size (bp)	# of contigs kept after Kraken	% reads mapped to 1rst draft assembly [a]	Mean Depth of Coverage	# of contigs	Size (bp)
NGCL001_A	354.5	416	2357204	188	97.87	258.7	116	2168187
NGCL002_A	134.9	945	3374293 [b]	179	94.51	76.76	153	2099945
NGCL003_Ap1	562.8	2650	3398020 [b]	398	97.93	230.63	229	2173600
NGCL003_Op1	687.9	11082	6956993 [b]	2373	79.81	198.12	3864	1946311 [c]
NGCL004_Ap2	601.0	5156	4083730 [b]	164	57.76	87.28	171	2141240
NGCL004_Op2	568.1	2283	3249746 [b]	573	31.96	271.98	223	2154660
NGCL005_Ap3	537.8	386	2239727	163	28.42	371.56	135	2121348
NGCL005_Op3	575.7	5045	5149829 [b]	1287	68.85	197.58	1190	2850896 [c]
NGCL006_Ap4	845.2	12838	6484078 [b]	214	88.66	105.38	215	2148349
NGCL006_Op4	800.5	2143	2022901	497	65.57	1252.29	861	401355 [c]
NGCL007_Ap5	689.7	727	2395294	148	21.85	433.21	126	2130021
NGCL007_Op5	967.2	10013	4363432 [b]	777	27.7	578.06	950	391360 [c]
NGCL008_Ap6	600.0	673	2397589	148	83.82	363.48	138	2135550
NGCL008_Op6	615.5	4585	5305862 [b]	1867	12.02	148.18	2554	3385176 [c]
NGCL009_Ap7	484.9	578	2316448	224	82.39	314.23	129	2126143
NGCL009_Op7	520.1	11080	6748869 [b]	3718	61.27	71.9	3465	2234942 [c]
NGCL010_Ap8	565.9	1622	2799584	243	85.73	308.74	152	2140971
NGCL010_Op8	463.5	3493	3983042 [b]	1074	30.14	139.95	2019	1803425 [c]
NGCL011_U	430.2	213	2202711	150	72.02	313.47	117	2163400
NGCL012_Ap9	400.6	332	2236250	158	25.45	265.33	135	2127223
NGCL012_Op9	545.5	4493	4757693 [b]	932	93.03	150.47	1296	2810346 [c]
NGCL013_Ap10	459.8	274	2178473	160	86.59	343.91	138	2104785
NGCL013_Op10	643.4	3308	4372497 [b]	731	56.44	200.36	1324	2774503 [c]
NGCL014_Ap11	886.5	10298	5597744 [b]	165	91.58	363.35	147	2142073
NGCL014_Op11	473.2	10309	7959214 [b]	3700	63.88	116.67	2920	3576009 [c]
NGCL015_Ap12	389.8	453	2308592	150	54.39	252.86	146	2126112
NGCL015_Op12	521.2	7833	6771077 [b]	1903	63.07	148.34	3837	1710268 [c]
NGCL016_Ap13	561.7	13893	5775520 [b]	2369	80.06	106.54	2237	1239392 [c]
NGCL016_Op13	628.7	8827	8030652 [b]	2753	26.08	165.03	2869	3706557 [c]
NGCL017_Ap14	667.9	8844	5845805 [b]	184	15.42	326.39	158	2081167
NGCL017_Op14	511.8	3105	2818062 [b]	856	65.61	442.68	1635	808111 [c]
NGCL018_Ap15	731.2	10030	6064209 [b]	344	62.86	303.59	150	2126313
NGCL018_Op15	565.6	6671	5919365 [b]	1403	32.17	139.44	3780	1812991 [c]
NGCL019_UR	76.6	192	2140307	186	55.45	53.98	154	2123931
NGCL020_UR	100.0	226	2204493	181	24.07	64.15	158	2169061

[a] refers to the % of mapped reads after trimming using Trimmomatic v0.36 [327]. [b] first draft assemblies flagged due to genome size. [c] final draft assemblies flagged due to genome size and total number of contigs.

7.2.7. Isolate typing, antimicrobial resistance prediction and phylogenetic analysis

In silico seven loci Multi-locus Sequence Typing (MLST) prediction was performed using *mlst* v2.4 software (https://github.com/tseemann/mlst). *Neisseria gonorrhoeae* Multi-Antigen Sequence Typing (NG-MAST) was performed using the *ngmaster* v0.4 software [286] (http://www.ng-mast.net/). Fine-typing for *Neisseria* species confirmation was performed using the *rplF* gene, upon query to the PubMLST *Neisseria* database (http://pubmlst.org). Identification of antimicrobial resistance determinants (see Chapter III for full list of determinants) *in silico* was performed using ARIBA v2.12.2 with the NG database [302]. Typing with the seven-loci *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance (NG-STAR) scheme [303] was also performed for interpretation purposes, upon query to the PubMLST Neisseria database.

Gene-by-gene analysis was performed using chewBBACA software v2.0.11 [331] by taking advantage of the cgMLST *N. gonorrhoeae* cgMLST v1.0 scheme [311] (1594 loci) and the Maximum-shared cgMLST scheme (822 loci) as described in Chapter III. Genomes exhibiting less than 95% of loci called in the cgMLST scheme were excluded from the phylogenetic analysis. Isolates were sub-typed using the WGS-based genogroup strategy described in Chapter III. Briefly, the novel genomes' allelic profiles were integrated with the previous 3945 genome dataset (See Chapter VI). For each subset of isolates under comparison, the maximum number of shared loci between them was maximized using PHYLOViZ online 2.0 Beta version (http://online2.phyloviz.net/) before minimum spanning tree (MST) construction [310, 332, 337]. Allelic distance thresholds for cluster inspection are expressed as percentages of allelic differences (AD) over the total number of shared loci under comparison for every analysis of each subset of isolates at any given level. To investigate sub-clusters with potential epidemiological link to the novel genomes, after generating a sub-MST for each low-level genogroup, we applied a cut-off of 1.5% AD (as described in Chapter III Section 3.3.6.), and re-analysed each resulting cluster by maximizing the number of shared loci between them.

7.3. Results

7.3.1. WGS of *Neisseria gonorrhoeae* directly from clinical samples

In order to perform whole-genome sequencing of *N. gonorrhoeae* directly from clinical samples, we applied the SureSelect$^{XT\ HS}$ target enrichment system to total genomic DNA samples extracted from 35 *N. gonorrhoeae*-positive clinical samples. Results showed that the targeted capture of Bacterial DNA was successful for all samples, regardless of the level of eukaryotic DNA "contamination" in each samples (**Figure 7.1.A**). The mean percentage of mapped reads of

all samples to the reference human genome was ~2%, while the mean percentage of mapped reads to reference *N. gonorrhoeae* genomes was ~88%. Still, first draft assembly data revealed that 22 samples had genomes larger than the expected (i.e., above 2.1 Mbp), and ten assemblies produced more than 1000 contigs (**Table 7.3.**), which is indicative of an overall "contamination" (i.e., non-*N. gonorrhoeae* sequencing data).

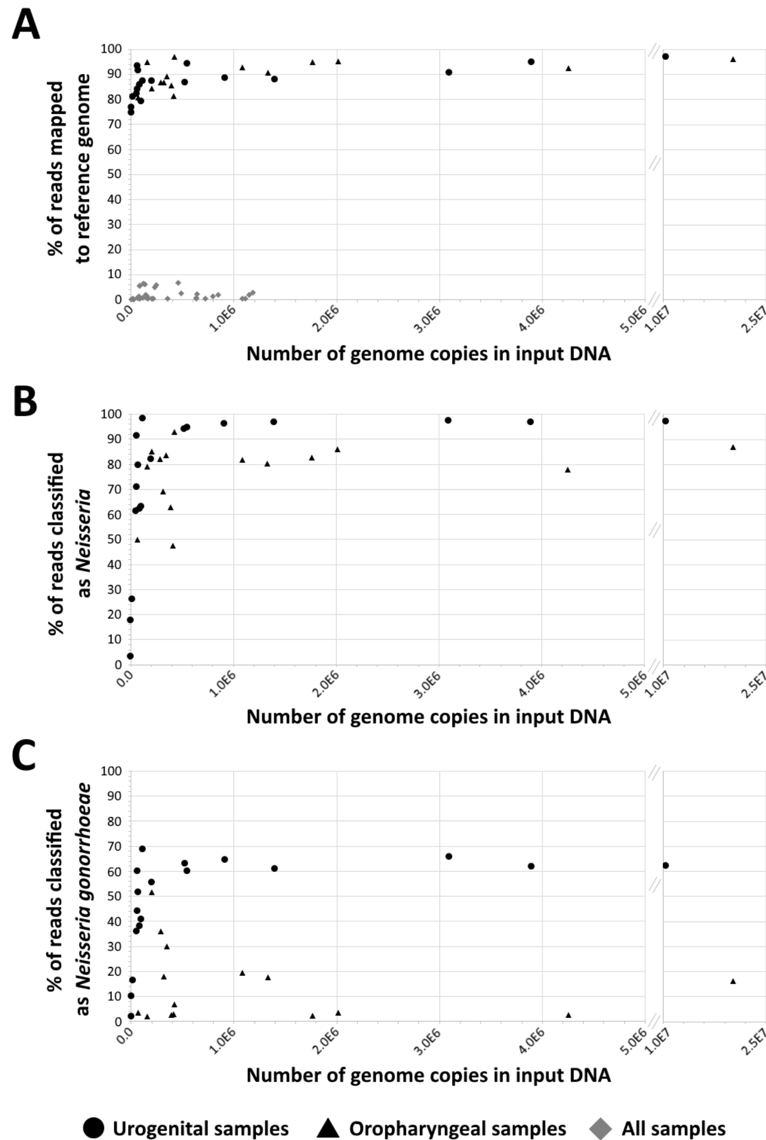

Figure 7.1. | Enrichment success *versus* the *Neisseria gonorrhoeae* and human genome copies in the input DNA. Genome copies for human and *N. gonorrhoeae* were assessed by qPCR. (**A**) Relationship between the number of genome copies and the percentage of mapped reads. Black circles and triangles refer to the percentage of mapped reads of urogenital and oropharyngeal samples, respectively, to *N. gonorrhoeae* reference genome NCCP11945 (Refseq accession #NC_011035) *versus* the number of gonococcal genomes copies. Grey diamonds refer to the percentage of mapped reads to the human reference genome (RefSeq accession #NC_000001-NC_000024 and #NC_012920) *versus* the number of human genome copies. Relationship between the number of *N. gonorrhoeae* genome copies and the percentage of classified reads up to the Genus *Neisseria* (**B**) or as *N. gonorrhoeae* (**C**) using Kraken v2 [358].

In fact, due to the close genetic relatedness of *Neisseria* species, targeted capture revealed not to be specific for *N. gonorrhoeae* (**Figure 7.1.B**, **Figure 7.1.C** and **Figure 7.2.A**). Taxonomic classification up to the *Neisseria* species of all generated raw reads showed that twelve samples exhibited more than 50% of classified reads from other *Neisseria* species (**Figure 7.2.A**), particularly *N. meningitidis* and *N. subflava*, both commensal species of the human respiratory tract. Additionally, ten genomes did not present *N. gonorrhoeae* reads as the predominant taxonomic hit (**Supplementary Table S7.1.**). This was particularly underlined in samples associated with oropharyngeal swabs (**Figure 7.1.B** and **Figure 7.1.C**), an environment where several commensal *Neisseria* species are present. Following this, we attempted to remove reads classified as *Neisseria* species other than *N. gonorrhoeae*, in order to recover highly contaminated samples, which resulted in a mean loss of ~60% of reads across samples (**Supplementary Table S7.2.**). Based on genome sizes, assemblies seemed to have improved for eight samples (**Table 7.3.**). Still, although the predominant taxonomic hit was now *N. gonorrhoeae* in 29 samples, this did not overall improve the most problematic genomes (**Figure 7.2.B** and **Supplementary Table S7.2.**) as, in order to obtain reads for subsequent genome assembly, we had to resort to reference-based mapping, which most likely carried over non-*N. gonorrhoeae* reads due to high inter-species homology.

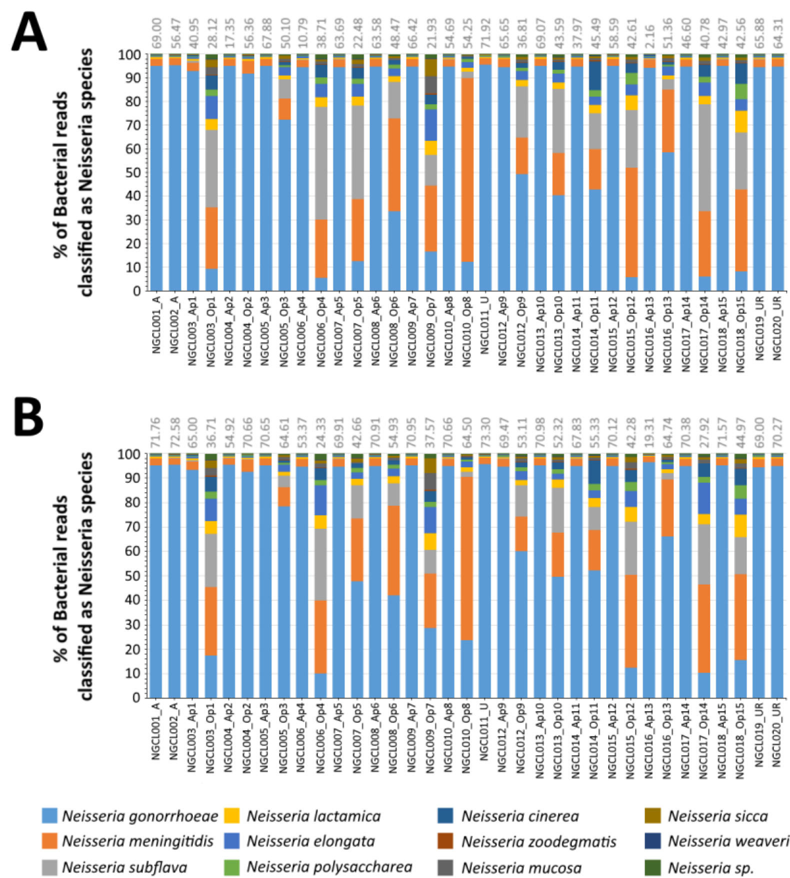

Figure 7.2. | Taxonomic classification within *Neisseria* species of raw (**A**) and filtered (**B**) reads generated for all clinical samples. Data refer to the percentage of reads classified as belonging to *Neisseria spp.* within those classified as Bacterial reads by Kraken v2 [358]. Values above the graphs refer to the total number of classified reads up to the *Neisseria* species over the total number of generated reads.

7.3.2. Global strain typing and antimicrobial resistance profile

After improving genome assemblies by filtering reads, we proceeded to perform strain typing for all clinical samples and *in silico* extract the genetic determinants for AMR. The exclusion criteria for the total number of loci called in the cgMLST scheme (i.e., < 95% loci called), together with the previously described results (see section above), resulted in the exclusion of 16 genomes from all downstream analysis. Of note, this exclusion enrolled all oropharyngeal-associated samples (**Table 7.4.**). We then performed strain typing on all 19 validated genomes, based on the traditional MLST and NG-MAST schemes, as well as on our defined WGS-based genogroups (described in Chapter III), with results presented in **Table 7.4**. Results showed that typing was overall successful for all final genome assemblies, with both traditional typing schemes. Novel genomes were integrated into ten distinct WGS-based genogroups, namely G009.009, G014.038, G142.207, G015.016, G079.129, G127.185, G040.043, G021.029, G004.003, and a new one linked to the high-level genogroups G001 (**Figure 7.3.**). Noteworthy, results showed that seven out of the 19 genomes were classified into two distinct genogroups, namely G142.207 (n=3) and G079.129 (n=4).

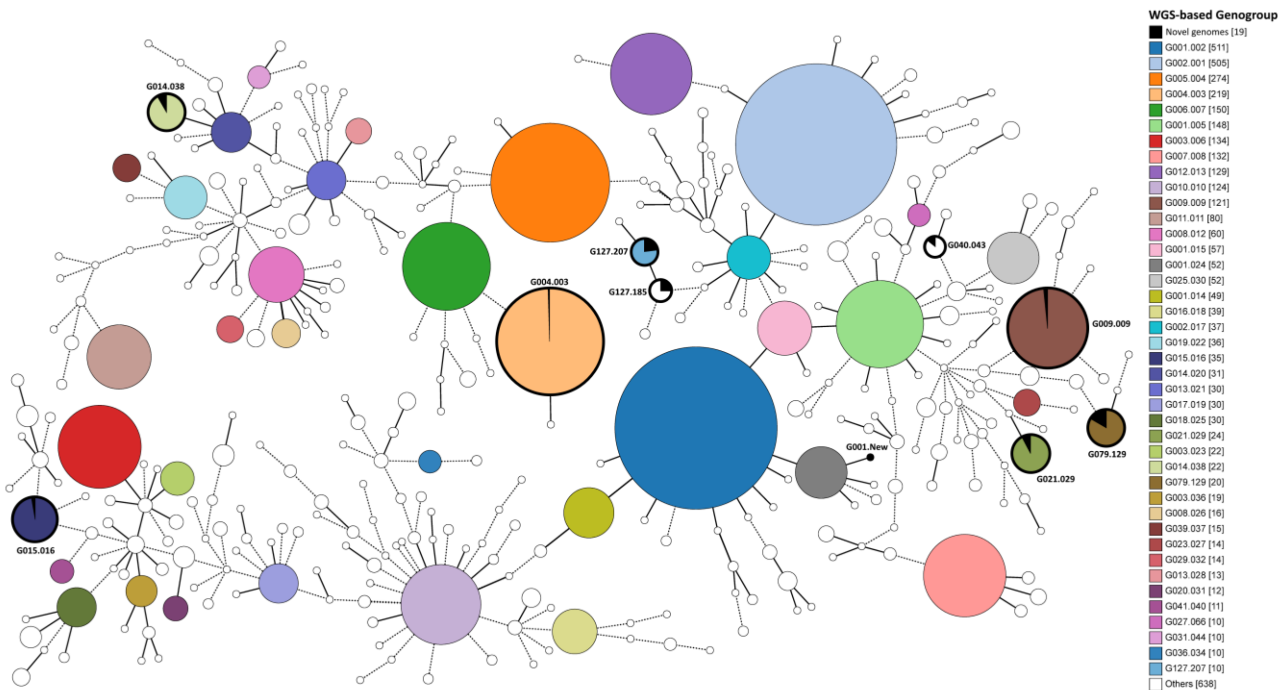

Figure 7.3. | Phylogenetic distribution of novel validated *Neisseria gonorrhoeae* genomes captured directly from clinical samples. Minimum spanning tree (MST) was constructed based on allelic diversity found among the 822 genes (MScgMLST scheme), enrolling a total of 3964 isolates collected across Europe. All nodes presenting an allelic distance (AD) below 40, corresponding to the low-level genogroup threshold, have been collapsed for visualization purposes. Nodes are coloured according to different low-level genogroups and novel genomes. Straight and dotted lines reflect nodes linked with the ADs below and above the threshold applied for high-level WGS-based genogroup definition (79 AD), respectively. WGS-based genogroups enrolling novel genomes are highlighted by thicker black circles. MST was generated using the GrapeTree v1.5.0 software [338].

Table 7.4. | Strain typing statistics for all *Neisseria gonorrhoeae* genomes captured from clinical samples.

Sample	WGS-based genogroup	MLST (ST)	NG-MAST (ST)	# of loci called with the cgMLST schema (1594 loci)	# of loci called with the MScgMLST schema (822 loci)
NGCL001_A	G009.009	7822	4995	1563	821
NGCL002_A	G014.038	8143	9918	1559	820
NGCL003_Ap1	G142.207	7822	14994	1538	812
NGCL003_Op1*	ND	ND	ND	101	61
NGCL004_Ap2	G015.016	8135	387	1551	821
NGCL004_Op2*	ND	8135	387	1500	800
NGCL005_Ap3	G079.129	1583	15589	1561	822
NGCL005_Op3*	ND	ND	15589	1444	760
NGCL006_Ap4	G127.185	10314	New	1554	821
NGCL006_Op4*	ND	ND	ND	23	12
NGCL007_Ap5	G040.043	13781	5441	1563	821
NGCL007_Op5*	ND	ND	ND	11	6
NGCL008_Ap6	G021.029	8156	5441	1564	822
NGCL008_Op6*	ND	ND	ND	624	311
NGCL009_Ap7	G079.129	1583	15589	1558	822
NGCL009_Op7*	ND	ND	ND	324	184
NGCL010_Ap8	G142.207	7822	14994	1562	822
NGCL010_Op8*	ND	ND	ND	261	93
NGCL011_U	G004.003	7363	New	1574	822
NGCL012_Ap9	G021.029	8156	18974	1564	821
NGCL012_Op9*	ND	ND	ND	1451	761
NGCL013_Ap10	G014.038	8143	8426	1563	822
NGCL013_Op10*	ND	8143	ND	1405	748
NGCL014_Ap11	G127.185	7822	12547	1561	821
NGCL014_Op11*	ND	ND	ND	1001	524
NGCL015_Ap12	G079.129	1583	15589	1558	820
NGCL015_Op12*	ND	ND	ND	49	34
NGCL016_Ap13*	ND	ND	ND	182	106
NGCL016_Op13*	ND	ND	ND	753	370
NGCL017_Ap14	G001.New	13292	9208	1566	822
NGCL017_Op14*	ND	ND	ND	47	20
NGCL018_Ap15	G079.129	1583	15589	1564	822
NGCL018_Op15*	ND	ND	ND	76	44
NGCL019_UR	G142.207	7822	New	1547	821
NGCL020_UR	G009.009	7822	4995	1555	822

MLST – Multi-locus Sequence Type; NG-MAST – *Neisseria gonorrhoeae* Multi-antigen Sequence Type; cgMLST – core-genome Multi-locus Sequence Typing; MScgMLST – Maximum-shared core-genome Multi-locus Sequence Typing; ND – Not determined. *Samples failing the minimum number of loci called criteria (i.e., ≥ 95% loci called).

In silico read-based detection of AMR determinants for *N. gonorrhoeae* and NG-STAR typing results are presented in **Table 7.5**. Most genomes presented alterations 228S in *folP* (encoding for dihydropteroate synthase) and 57M in *rpsJ* (ribosomal protein S10), associated with resistance to sulphonamides and tetracycline, respectively, while ten carried *ponA* (penicillin-binding protein 1) alteration 421P associated with decreased susceptibility to penicillins. Fifteen genomes presented alterations in *gyrA* associated with ciprofloxacin resistance, with 14 carrying simultaneously alterations 91F and 95A, and one possessing the combination 91F/95G. Additionally, nine genomes also carried alteration 87R in *parC* (topoisomerase IV subunit C) associated with ciprofloxacin resistance. Only one genome presented azithromycin resistance-associated alteration C2597T in *23SrRNA*, although it was only observed at ~20%, suggesting

Table 7.5. | Summary of all *in silico* detected antimicrobial resistance determinants.

Sample	NG-STAR (ST)	*folP*	*rpsJ*	*ponA*	*porB*				*rpoB*	*pro_mtrR*	*mtrR*			*gyrA*			*parC*	*penA*						*23SrRNA*	*16SrRNA*	*tetM*	*blaTEM*
		228S	57M	421P	type	120K	121D	121N	553N	35Adel	45D	A39T	R44H	91F	95G	95A	87R	Allele type	312M	316T	501V	542S	545S	C2597T (%)	C1184T (%)	P/A	P/A
NGCL001_A	416	√	√	√	*porB1b*	.	.	.	.	√	√	.	.	√	.	√	√	Type II non-mosaic	.	.	.	.	.	.	.	.	√
NGCL002_A	436	√	√	.	*porB1b*	.	.	.	.	.	.	.	.	√	.	√	√	Type II non-mosaic	.	.	.	.	.	.	.	.	.
NGCL003_Ap1	~1387	√	√	√	*porB1a*	NA	NA	NA	.	.	.	√	.	√	.	√	√	Type V non-mosaic	.	.	.	√	.	.	.	.	.
NGCL004_Ap2	729	.	.	.	*porB1b*	.	.	.	.	.	.	.	.	.	.	.	.	Type XV non-mosaic	.	.	.	.	.	√ (19.4)	.	.	.
NGCL005_Ap3	1340	√	√	.	*porB1b*	.	.	.	√	.	.	√	.	√	.	√	.	Type XLIII non-mosaic	.	.	√	.	.	.	.	√	√
NGCL006_Ap4	~1387	√	√	√	*porB1b*	.	.	.	.	.	.	√	.	√	.	√	√	Type V non-mosaic	.	.	.	√	.	.	.	.	.
NGCL007_Ap5	New	√	√	.	*porB1b*	.	.	.	.	.	√	.	.	.	.	.	.	Type IX non-mosaic	.	.	.	.	.	.	.	.	.
NGCL008_Ap6	442	√	.	√	*porB1b*	.	.	.	.	√	.	.	.	.	.	.	.	Type II non-mosaic	.	.	.	.	.	.	.	.	.
NGCL009_Ap7	1340	√	√	.	*porB1b*	.	.	.	√	.	.	√	.	√	.	√	.	Type XLIII non-mosaic	.	.	√	.	.	.	.	√	√
NGCL010_Ap8	1387	√	√	√	*porB1a*	NA	NA	NA	.	.	.	√	.	√	.	√	√	Type V non-mosaic	.	.	.	√	.	.	√ (95.8)	.	.
NGCL011_U	480	√	√	√	*porB1b*	√	√	.	.	√	.	.	.	√	√	.	.	Type XLIV non-mosaic	.	.	.	.	.	.	.	.	.
NGCL012_Ap9	~442	√	.	√	*porB1b*	.	.	.	.	√	.	.	.	.	.	.	.	Type II non-mosaic	.	.	.	.	.	.	.	.	.
NGCL013_Ap10	~1169	√	√	.	*porB1b*	√	.	√	.	.	.	.	.	√	.	√	√	Type 67 mosaic	√	√	.	.	√	.	.	.	.
NGCL014_Ap11	~1387	√	√	√	*porB1a*	NA	NA	NA	.	.	.	√	.	√	.	√	√	Type V non-mosaic	.	.	.	√	.	.	.	.	.
NGCL015_Ap12	1340	√	√	.	*porB1b*	.	.	.	√	.	.	√	.	√	.	√	.	Type XLIII non-mosaic	.	.	√	.	.	.	.	.	.
NGCL017_Ap14	439	√	√	.	*porB1b*	.	.	.	√	.	.	√	√	.	.	.	.	Type II non-mosaic	.	.	.	.	.	.	.	.	.
NGCL018_Ap15	~1340	√	√	.	*porB1b*	.	.	.	√	.	.	√	.	√	.	√	.	Type XLIII non-mosaic	.	.	√	.	.	.	.	√	.
NGCL019_UR	~1387	√	√	√	*porB1b*	.	.	.	.	.	.	√	.	√	.	√	√	Type V non-mosaic	.	.	.	√	.	.	.	.	.
NGCL020_UR	416	√	√	√	*porB1b*	.	.	.	.	√	√	.	.	√	.	√	√	Type II non-mosaic	.	.	.	.	.	.	.	.	√

NG-STAR – *Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance; √ – Presence of antimicrobial resistant determinant. NA – Not applicable. Surveyed antimicrobial resistance determinants were chosen based on previous reports [120, 192, 232]

that only one copy of this gene possesses this SNP, which may not be enough to trigger resistance [120, 192]. Noteworthy, one genome (NGCL010_Ap8) carried the alteration C1184T in the *16SrRNA* gene, associated with Spectinomycin resistance, which was never observed in any sequenced genome from Portugal to date (See Chapter III and VI). No genome carried a *penA* (penicillin-binding protein 2) allele type associated with decreased susceptibility to cephalosporins.

7.3.3. Phylogenetic analysis

As previously mentioned, cgMLST calling data showed that 19 genomes could be included in our WGS dataset without compromising the maximum number of shared loci within sub-sets, since more than 95% of loci were called (**Table 7.4.**). This highlighted the possibility to perform phylogenetic inferences from genomic data captured directly from clinical samples, i.e., that genome assemblies constructed from complex WGS data provided the same resolution power as those obtained from culture. As such, after the integration of the novel genomes within our WGS dataset, we assessed if these integrated genetic clusters with potential epidemiological linkage, by applying the conservative threshold of 1.5% AD within each low-level genogroup, after sub-MST generation maximizing the shared loci between isolates at this level. At this threshold, results showed that 12 out of 19 genomes could be linked with other isolates from our dataset (**Figure 7.4.**). Those unlinked were NGCL017_Ap14 that was shown to be classified into a novel WGS-genogroup (**Figure 7.3.**), and five other genomes that, although integrating known WGS-genogroups, displayed AD above the applied cut-off, namely NGCL015_Ap12 and NGCL018_Ap15 (**Figure 7.4.B**), NGCL014_Ap11 (**Figure 7.4.E**), NGCL013_Ap10 (**Figure 7.4.F**), and NGCL007_Ap5 (**Figure 7.4.H**). More importantly, with exception of NGCL001_A, all remaining 11 genomes were directly linked to other isolates collected in Portugal (**Figure 7.4.**). Noteworthy, nine novel genomes integrated three genogroups that were previously highlighted as emerging in recent years (See Chapter VI), namely G079.129 (n=4) (**Figure 7.4.B**), G127.207 (n=3) (**Figure 7.4.C**) and G127.185 (n=2) (**Figure 7.4.E**), five of which clustering with other isolates at the applied threshold.

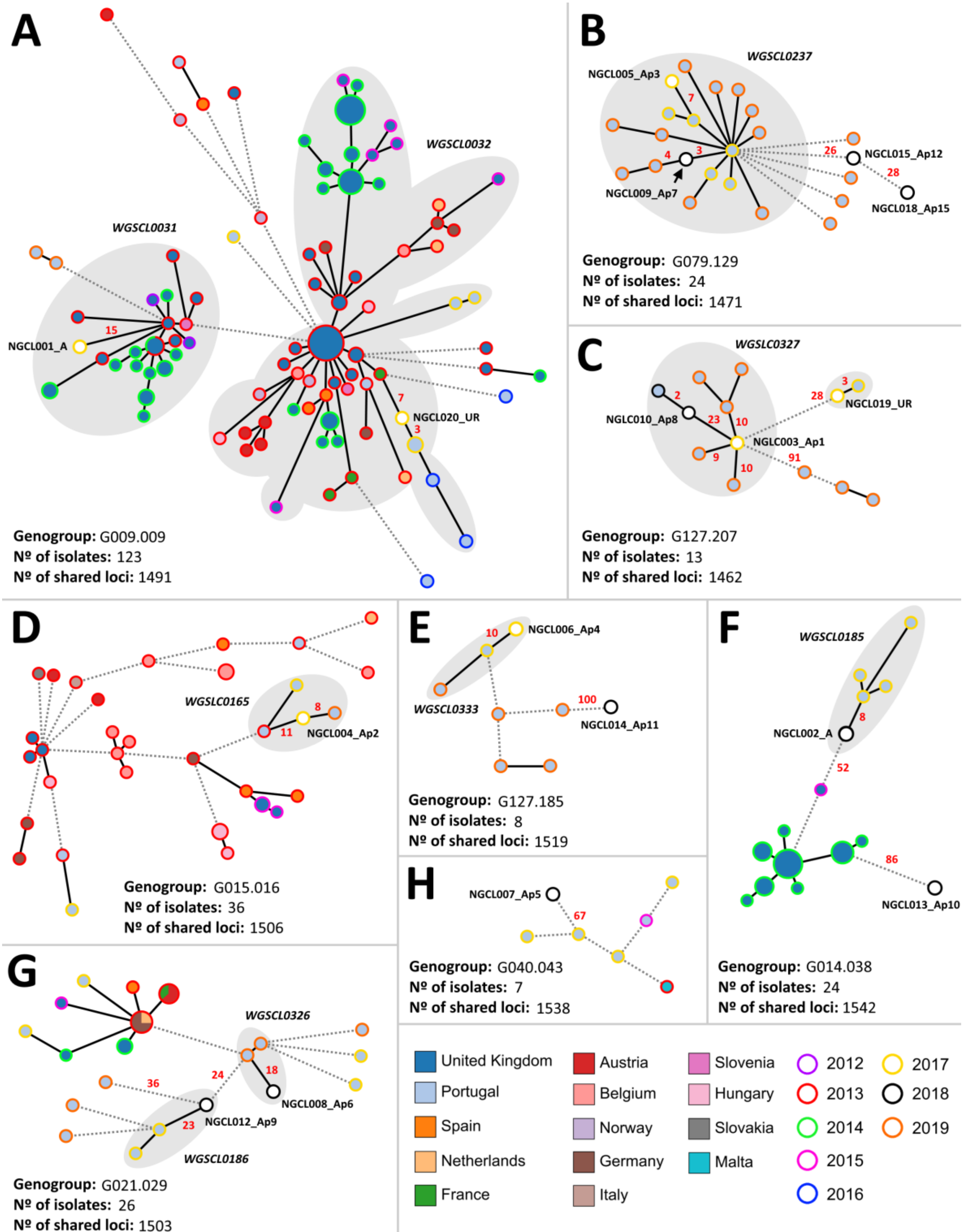

Figure 7.4. | *Neisseria gonorrhoeae* clusters enrolling genomes captured directly from clinical samples. Minimum spanning trees (MSTs) were constructed after maximizing the number of shared loci between isolates from the same genogroup. Nodes are coloured by country of origin and outlines by collection year. The numbers in red on the connecting lines represent the allele differences (AD) between novel genomes and their closest links. Clusters defined at a threshold of 1.5% AD within each genogroup and enrolling novel genomes are highlighted in grey, and WGSCL number refers to the ones determined in Chapter III and Chapter VI. MSTs were generated using the GrapeTree v1.5.0 software [338].

For clinical samples NGCL019_UR and NGCL020_UR, *N. gonorrhoeae* had already been isolated from culture and sequenced for comparison purposes (NGPT17487 and NGPT17512, respectively). In both cases, results showed that both genomes (i.e., obtained from culture and culture-independent methods) were linked together, displaying only 3 AD (**Figure 7.4.A** and **Figure 7.4.C**). Additionally, when looking at the entire 1594 loci scheme, both pairs only exhibited a total of 6 AD within the maximum number of shared/called loci (i.e., 1547 loci for NGCL019_UR/NGPT17487 and 1555 loci for NGCL020_UR/NGPT17512). This could be due to the inherent higher genetic diversity present in the culture-independent captured genomes, since all the gonococcal population present in the samples was sequenced (i.e., closely mirroring the one infecting a patient) unimpacted by the culture bottleneck. Regarding missing loci that were differentially observed in each genome pair (i.e., loci called in one but missing from the other), data showed that the genomes obtained by the culture-independent strategy were not impacted by these. In fact, genome pair NGCL019_UR/NGPT17487 displayed 10 and 16 differentially missing loci, respectively, while NGCL020_UR/NGPT17512 displayed 9 and 19, respectively.

7.4. Discussion

In the present study, we applied a culture-independent approach to capture and sequence *N. gonorrhoeae* genomes directly from clinical specimens collected from distinct anatomic sites. In particular, we aimed to bypass the culture step for WGS and assess the ability to carry out *in silico* typing, to perform phylogenetic inferences and to predict AMR profiles based on the obtained genomic data. As culture is progressively being abandoned in favour of NAAT, it is important to implement alternative approaches to report *N. gonorrhoeae* AMR and molecular epidemiological data. Our results showed that the applied strategy was successful in the capture of bacterial DNA, independently of the degree of human DNA present in each complex clinical sample (**Figure 7.1.**). Nevertheless, due to both the presence of commensal *Neisseria* species and the high genetic relatedness of these, capturing *N. gonorrhoeae* from oropharyngeal samples was unsuccessful. In fact, even when attempting to filter reads by taxonomic classification coupled with reference-based read mapping, it was impossible to extract useful *N. gonorrhoeae* WGS data from these samples. The capture of genomic data from oropharyngeal samples would be of particular importance as *N. gonorrhoeae* isolation from these is recognized as challenging and most often unsuccessful [24], and pharyngeal infections are considered to be more difficult to treat [360]. Our data (**Table 7.1.**) together with previous reports [24, 85, 361–364] have shown that diagnosis based on these specimens increases the number of confirmed cases of gonococcal infection, as oropharyngeal infections are mostly asymptomatic and, consequently, undiagnosed.

Additionally, it has been suggested that this site may be important for the development of AMR because the presence in the pharyngeal mucosa of non-gonococcal *Neisseria* species may promote the transfer of resistance to *N. gonorrhoeae* [365, 366].

Still, in this work, we were able to capture *N. gonorrhoeae* WGS data from 19 distinct clinical samples hailing from the urogenital tract. This allowed us to predict AMR profiles (**Table 7.5.**) and perform phylogenetic inferences (**Figure 7.4.**) with high discriminatory power. In fact, this culture-independent approach had already allowed to perform AMR prediction, for instance, for *M. tuberculosis* [356]. Even though we could not successfully apply this approach to all types of clinical samples, we showed that it is possible to obtain and integrate *N. gonorrhoeae* WGS data with data generated from isolated strains. Here, we were able to integrate most of the novel obtained genomes into genetic clusters linked at a threshold with potential epidemiological linkage. This opens the possibility of using this approach in response to particular settings where strains could not be isolated, namely in judicial settings, i.e., in cases of sexual assault or sexual abuse (if gonorrhoea infections are present and transmitted), or in outbreak settings, as illustrated recently for *C. trachomatis* [355]. Additionally, it may allow to gain insights into the genetic diversity of intra- and inter-patient infecting *N. gonorrhoeae* populations, with particular importance in cases of invasive disease. The ability to analyse the genetic diversity of *N. gonorrhoeae* directly from clinical samples (i.e., to analyse the evolving bacterial population during infection) may aid in understanding how this pathogen can progress to invasiveness during the host-pathogen arms race. This was recently achieved for the sexually transmitted pathogen *T. pallidum*, the causative agent of syphilis [354], which can also infect multiple anatomic sites and has complex disease progression stages. As this chapter corresponds to an ongoing work, it is our aim to explore these data for genomic regions displaying inter- and intra-strain genetic heterogeneity, namely through the inspection of both single nucleotide sites with allelic variation and in-length variable DNA homopolymeric tracts (poly-G/C and poly-A/T sites), as previously performed for *T. pallidum* [354] and *C. trachomatis* [367].

Ultimately, considering the new era of laboratory diagnosis where molecular technologies are readily replacing culture, it is crucial to build the necessary tools for a sustainable surveillance of *N. gonorrhoeae* AMR. This becomes particularly important as, in most countries, the diagnosis of gonococcal infections is being based on modern day molecular assays applied to multiple non-invasive specimens from the same patients (e.g. urine and extragenital swabs), which cannot, at this stage, be used to determine AMR in *N. gonorrhoeae*. As such, to overcome the inability to analyse gonorrhoea-positive oropharyngeal samples observed here, we anticipate the need to implement a Next Generation Sequencing (NGS) strategy based on amplicons (rather than on the whole genome) to perform the necessary prediction of AMR profiles and molecular epidemiological inferences. This strategy will require targeting the major AMR genetic drivers,

as well as genetic markers reflecting the major gonococcal lineages. Moreover, as it is based on deep-sequencing, it may allow to simultaneously observe the emergence of AMR mutations (i.e., the analysis of minor variants) during infection, as it also bypasses the culture bottleneck (see next Chapter for more details).

Acknowledgements: The author work like to acknowledge Dr. Jacinta Azevedo, Dr. João Roxo and Dr. Mafalda Salvador of the Sexually Transmitted Diseases Clinic of the Lapa Health Centre, and the "Grupo de Ativistas em Tratamentos", namely the CheckpointLX for the collection of all clinical samples enrolled in the present study.

CHAPTER VIII

Final overview, concluding remarks and future directions

---

8. Final overview, concluding remarks and future directions

8.1. Final discussion

The global goal of this PhD dissertation was to strengthen *N. gonorrhoeae* epidemiological and antimicrobial resistance (AMR) surveillance using whole-genome sequencing (WGS) approaches. This work was developed within the scope of the National Reference Laboratory for Sexually Transmitted Infections (NRL-STI), with access to a large ongoing *N. gonorrhoeae* culture collection enrolling a countrywide laboratory network (Portuguese National Laboratory Network for *Neisseria gonorrhoeae* Collection – PTGonoNet), officially formalized in 2010, and hosted and supported by the Portuguese National Institute of Health (INSA). At the time of writing, this opportunistic collection already constitutes 17 years of *N. gonorrhoeae* AMR surveillance performed by the NRL-STI (from 2003 up to 2020), totalizing 3877 isolates (last update in October 2020). In addition, we took advantage of the existence of a Next-Generation Sequencing (NGS) core-facility based at the Department of Human Genetics of INSA, and a Bioinformatics Unit based at the Department of Infectious Diseases, where the author of this PhD dissertation currently develops his scientific activity.

In this context, we started by analysing the data associated with the vast culture collection of *N. gonorrhoeae* isolates at the NRL-STI (Chapter II). The main driver for this work was the lack of information regarding the overall AMR trends of *N. gonorrhoeae* in Portugal. This collection enrolled 2596 isolates, collected since 2003, from distinct clinical specimens, with countrywide coverage, and that had been phenotypically characterized for AMR for eight distinct antimicrobials. We observed that the vast majority of isolates (~88%) presented decreased susceptibility or resistance to at least one of the tested antimicrobials. Data revealed that, up to 2018, resistance to penicillin, tetracycline and ciprofloxacin were widespread, and that there was a potential increase in gentamicin intermediate resistance, based on interpretive results. Still, as the guidelines for resistance classification were updated in the beginning of 2019, resistance to azithromycin was found to be overall low in our country, contrary to what was previously reported. A later extended analysis to 2019, introducing 682 isolates to the dataset (described in Chapter VI), showed that these AMR trends were preserved. More importantly, this additional dataset allowed us to identify the first cephalosporin-resistant *N. gonorrhoeae* isolate detected in Portugal (corresponding to a published paper described in Chapter V) that was notified to Public Health Authorities and reported to the scientific community, along with an additional cefixime-resistant isolate, whose circulation also warrants notification. This work constituted the first large compilation of AMR data for *N. gonorrhoeae* in our country, which revealed that most of this pathogen's "older" treatment strategies have been compromised, in accordance with what has been observed worldwide. Particularly, with azithromycin resistance being maintained

throughout the years (even if at a low frequency) and the introduction of cephalosporin-resistance isolates in our country, the current treatment guidelines may face challenges in the near future, prompting the need to control the spread of these strains.

National and international surveillance programmes are increasingly promoting the use of genomic data to track *N. gonorrhoeae* main circulating lineages and the emergence and spread of AMR. Nevertheless, most studies still rely on traditional typing methodologies to report circulating strains, either through sequence types (ST) or agglomerated ST-based genogroups, which do not take advantage of genome-scale data. As such, in a second stage, we reported a comprehensive WGS-based genogroup assignment for *N. gonorrhoeae*, based on the identification of the maximum discriminatory genetic thresholds reflecting cluster stability (Chapter III), which represent main circulating lineages. This deeply rooted bioinformatics work took advantage of a European-oriented genomic dataset of 3791 genomes (including novel sequenced Portuguese isolates) to define a two-step hierarchical genogroup classification, based on the two earliest cluster stability threshold observed. Our results showed that these WGS-based genogroups correlated well with other typing techniques (i.e., MLST, NG-MAST, and the recently proposed core-genome groups) and further linked to specific AMR signatures. In fact, several genogroups carrying alarming AMR genetic profiles were found to be predominant and widespread in Europe. Additionally, a preliminary exploration of a draft *N. gonorrhoeae* pan-genome pointed that the defined genogroups may also carry specific genetic mechanisms within their accessory-genomes towards adaptation, pathogenesis and survival (Chapter IV). Insights on the clustering at a higher discriminatory level (at a threshold compatible with epidemiological linkage) further supported the suitability of gene-by-gene typing strategies to detect/track outbreaks and transmission chains as a means to promote and/or support epidemiological investigation. In this regard, this study enhances *N. gonorrhoeae* surveillance by promoting the prospective monitoring of genogroup frequency and geographic spread, towards more oriented Public Health actions to control the spread of *N. gonorrhoeae* AMR, while maintaining backwards compatibility with previous used typing methods.

Following this output, we applied this approach to disclose the genetic diversity of *N. gonorrhoeae* in Portugal, using a dataset of 733 genomes from isolates of the PTGonoNET, collected throughout our country, from 2003 to 2020 and displaying different AMR profiles (described in Chapter VI). This constituted the first nationwide genomic epidemiology *N. gonorrhoeae* study in the era of WGS-based laboratory surveillance performed in Portugal. Within the studied time period, we observed that the overall genetic diversity of this pathogen in our country is high, illustrating the continuous cross-border spread of gonorrhoea that introduces novel strains into a country as part of large sexual transmission networks. It also reflects the existence of likely confined transmission chains at country or regional level. Regarding the later, at a higher discriminatory level, our results provided a comprehensive breakdown of the *N.*

gonorrhoeae transmission chains in Portugal and suggested the emergence of novel WGS-based genogroups. Some potential transmission chains span several years and seem to be still active given the data from the latest year of surveillance. Additionally, we observed the existence of European-spread genogroups (previously described in Chapter III) carrying alarming AMR profiles in our country, as well as the emergence of *N. gonorrhoeae* strains with potentially increased invasiveness capabilities. As it stands, this work constituted a turning point to consolidate the genomic epidemiology of gonococci in Portugal, as the implemented gene-by-gene approach will be carried forward at the Portuguese NRL-STI of INSA.

Finally, considering that diagnosis is being progressively based on nucleic acid amplification techniques, which do not allow to disclose AMR data and epidemiological links, we assessed the possibility of capturing WGS data directly from clinical samples via a culture-independent approach (Chapter VII). We took advantage of our previous successful experience with other pathogens and enrolled in a study relying on both wet-lab and dry-lab stages. This relied on the bioinformatics design of 120-nucleotide RNA "baits" to hybridize and capture *N. gonorrhoeae* genomes from complex clinical samples such as oropharyngeal, urethral and anal swabs. We observed that the applied methodology was unsuccessful for oropharyngeal samples, due to the presence of commensal *Neisseria* species, presenting high genetic homology to *N. gonorrhoeae*. On the other hand, the major goal was still achieved in urogenital swabs, as it was possible to determine AMR profiles, perform traditional and WGS-based subtyping and phylogenetically link, with high resolution, culture-independent sequenced genomes with WGS data generated from cultured isolates.

In summary, we believe the work described in this PhD dissertation stands as an important contribution towards a more effective WGS-based surveillance of gonococcal infections, essentially through the development of a *N. gonorrhoeae* WGS-based classification system at a global level. This comprehensive assignment facilitates the assessment and continuous monitoring of genogroup frequency, geographic dispersion, and, more importantly their association with specific AMR signatures, which constitutes one of the major Public Health concerns around this sexually transmitted disease-causing bacterium. This described strategy may benefit Public Health actions through the prioritization of genogroups to be controlled, the identification of emerging resistance carriage, and the potential facilitation of data sharing and communication. In this context, this work strengthened the capacity building of the National Reference Laboratory for Sexually Transmitted Infections of the Portuguese National Institute of Health, in order to promote a more effective WGS-based surveillance of infections caused by *N. gonorrhoeae*.

8.2. Future perspectives

As is the inherit nature of the pursuit of scientific knowledge, answering a question only leads to several others being raised. The present PhD dissertation is no exception, and throughout the different Chapters described in this work several scientific questions were raised, and future investigation is required to address them, for instance:

i) How does the European-oriented scenario of genogroup distribution translate at the worldwide scale?

ii) How will the proposed gene-by-gene approach be maintained in order to provide consistent harmonized and communicable data for Public Health purposes?

iii) How can we obtain reliable epidemiological and AMR data directly from all *N. gonorrhoeae*-positive complex clinical specimens towards enhanced molecular surveillance?

With WGS now widespread, enriching the present work by taking advantage of the more than 15 000 *N. gonorrhoeae* genomes currently published, collected from across the World, will aid in clarifying the first point. We also anticipate that very recent developments of the chewBBACA suite, namely the release of the Chewie Nomenclature Server, will allow us to share our WGS-based gene-by-gene typing strategy while maintaining a common nomenclature, and thus, simplifying the comparison and communication of results. Finally, the development of a targeted deep-sequencing amplicon-based methodology, in opposition to sequencing the whole genome, could be a viable alternative to obtain AMR data directly from clinical samples. This could be achieved by creating a NGS-based amplicon scheme targeting the major genes known to be involved in *N. gonorrhoeae* AMR, which could be coupled with a minimum loci panel that permits a robust inference of WGS-based genogroups. Ultimately, this NGS-based approach should be designed for rapid response (aiding therapeutic decisions) and high-throughput testing (providing simultaneous results on AMR for multiple antimicrobials and samples), providing a faster and cheaper alternative to culture and traditional AMR testing, thus constituting a foundational step towards the future application of point-of-care molecular tests to *N. gonorrhoeae*.

References

[1] Tonjum T. (2005) ***Neisseriales***. In: Brenner D, Krieg N, Staley J, Garrity G (editors). *Bergey's Manual of Systematic Bacteriology Volume Two: The Proteobacteria (Part C)*. New York: Springer-Verlag. pp. 774–797.

[2] Parte AC. (2018) **LPSN – List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on**. *Int J Syst Evol Microbiol* 68(6):1825–1829.

[3] Elias J, Frosch M, Vogel U. (2015) ***Neisseria***. In: Jorgensen J, Pfaller M, Carroll K, Funke G, Landry M, et al. (editors). *Manual of Clinical Microbiology 11th Edition*. Washington DC: American Society of Microbiology Press. pp. 635–651.

[4] Diallo K, MacLennan J, Harrison OB, Msefula C, Sow SO, Daugla DM, Johnson E, Trotter C, MacLennan CA, Parkhill J, Borrow R, Greenwood BM, Maiden MCJ. (2019) **Genomic characterization of novel *Neisseria* species**. *Sci Rep* 9(1):13742.

[5] Liu G, Tang CM, Exley RM. (2015) **Non-pathogenic *Neisseria*: members of an abundant, multi-habitat, diverse genus**. *Microbiology* 161(7):1297–1312.

[6] Heydecke A, Andersson B, Holmdahl T, Melhus A. (2013) **Human wound infections caused by *Neisseria animaloris* and *Neisseria zoodegmatis*, former CDC Group EF-4a and EF-4b**. *Infect Ecol Epidemiol* 3:20312.

[7] Harrison OB, Claus H, Jiang Y, Bennett JS, Bratcher HB, Jolley KA, Corton C, Care R, Poolman JT, Zollinger WD, Frasch CE, Stephens DS, Feavers I, Frosch M, Parkhill J, Vogel U, Quail MA, Bentley SD, Maiden MC. (2013) **Description and nomenclature of *Neisseria meningitidis* Capsule Locus**. *Emerg Infect Dis* 19(4):566–573.

[8] Stein D. (2016) ***Neisseria***. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K, Stackebrandt E (editors). *The Procaryotes*. Singapore: Springer. pp. 602–647.

[9] Luys G. (1912) **Traite de la blennorragie et de ses complications**. Paris: Librairie Octave Doin.

[10] Christodoulides M. (2019) **Preface**. In: Christodoulides M (editor). *Neisseria gonorrhoeae - Methods and Protocols*. Southampton: Humana Press. pp. v–xii.

[11] Morton RS. (1995) **Sexual attitudes, preferences and infections in Ancient Egypt**. *Sex Transm Infect* 71(3):180–186.

[12] Willcox RR. (1949) **Venereal Disease in the Bible**. *Sex Transm Infect* 25(1):28–33.

[13] Gruber F, Lipozenčić J, Kehler T. (2015) **History of venereal diseases from antiquity to the renaissance**. *Acta Dermatovenerol Croat* 23(1):1–11.

[14] Neisser A. (1879) **Uber eine der gonorrhoe eigentümliche micrococcenform**. Cent Med Wissensch 17:497–500.

[15] Ligon BL. (2005) **Albert Ludwig Sigesmund Neisser: Discoverer of the cause of gonorrhea**. *Semin Pediatr Infect Dis* 16(4):336–341.

[16] Oriel JD. (1989) **Eminent venereologists**. 1. Albert Neisser. *Sex Transm Infect* 65(4):229–234.

[17] Kampmeier RH. (1983) **Introduction of sulfonamide therapy for gonorrhea**. *Sex Transm Dis* 10(2):81–84.

[18] Dunlop E. (1949) **Gonorrhoea and the sulphonamides**. *Br J Vener Dis* 25(2):81–83.

[19] Unemo M, Shafer WM. (2011) **Antibiotic resistance in *Neisseria gonorrhoeae*: origin, evolution, and lessons learned for the future**. *Ann N Y Acad Sci*1230:E19–E28.

[20] Unemo M, del Rio C, Shafer WM. (2015) **Antimicrobial resistance expressed by *Neisseria gonorrhoeae*: a major global Public Health problem in the 21st century**. *Microbiol Spectr* 4(3).

[21] Sánchez-Busó L, Golparian D, Corander J, Grad YH, Ohnishi M, Flemming R, Parkhill J, Bentley SD, Unemo M, Harris SR. (2019) **The impact of antimicrobials on gonococcal evolution**. *Nat Microbiol* 4(11):1941–1950.

[22] Shaughnessy J, Ram S, Rice PA (2019) **Biology of the gonococcus: disease and pathogenesis**. In: *Methods in Molecular Biology*. Humana Press Inc. pp. 1–27.

[23] Humbert MV, Christodoulides M. (2019) **Atypical, yet not infrequent, infections with *Neisseria* species**. *Pathogens* 9(1):10.

[24] Meyer T, Buder S. (2020) **The laboratory diagnosis of *Neisseria gonorrhoeae*: current testing and future demands**. *Pathogens* 9(2):91.

[25] Edwards JL, Apicella MA. (2004) **The molecular mechanisms used by *Neisseria gonorrhoeae* to initiate infection differ between Men and Women**. *Clin Microbiol Rev* 17(4):965–981.

[26] Sweet R, Walker. (2011) **Gonorrhea infection in women: prevalence, effects, screening, and management**. *Int J Womens Health* 3:197–206.

[27] Lee JS, Choi HY, Lee JE, Lee SH, Oum BS. (2002) **Gonococcal keratoconjunctivitis in adults**. *Eye (Lond.)* 16(5):646–649.

[28] Lessing JN, Slingsby TJ, Betz M. (2019) **Hyperacute gonococcal keratoconjunctivitis**. *J Gen Intern Med* 34(3):477–478.

[29] Odegaard K, Gundersen T. (1973) **Gonococcal pharyngeal infection**. *Sex Transm Infect* 49(4):350–352.

[30] Stolz E, Schuller J. (1974) **Gonococcal oro- and nasopharyngeal infection**. *Sex Transm Infect* 50(2):104–108.

[31] Yavelow SL, Wiznia A, Brennessel DJ, Glaser JH. (1984) **Disseminated gonorrhoea from a pharyngeal infection in a prepubertal child**. *Int J Pediatr Otorhinolaryngol* 7(3):297–300.

[32] Ratnatunga CS. (1972) **Gonococcal pharyngitis**. *Sex Transm Infect* 48(3):184–186.

[33] Wiesner PJ, Tronca E, Bonin P, Pedersen AHB, Holmes KK. (1973) **Clinical spectrum of pharyngeal gonococcal infection**. *N Engl J Med* 288(4):181–185.

[34] Fiumara NJ, Wise HM, Many M. (1967) **Gonorrheal pharyngitis**. *N Engl J Med* 276(22):1248–1250.

[35] Schaefer RA, Enzenauer RJ, Pruitt A, Corpe RS. (1992) **Acute gonococcal flexor tenosynovitis in an adolescent male with pharyngitis. A case report and literature review**. *Clin Orthop Relat Res* (281):212–215.

[36] Woods CR. (2005) **Gonococcal infections in neonates and young children**. *Semin Pediatr Infect Dis* 16(4):258–270.

[37] Austin TW, Yang W, Pattison FM. (1977) **Oropharyngeal gonorrhea: disseminated gonococcal disease**. *Can Med Assoc J* 117(5):438.

[38] Bro-Jorgensen A, Jensen T. (1971) **Gonococcal tonsillar tnfections**. *BMJ* 4(5788):660–661.

[39] Bodsworth NJ, Price R, Nelson MJ. (1993) **A case of gonococcal mastitis in a male**. *Sex Transm Infect* 69(3):222–223.

[40] Ceniceros A, Galen B, Madaline T. (2019) **Gonococcal breast abscess**. *IDCases* 18:e00620.

[41] Bateman AC. (2017) **Unusual cause of a wound infection**. *J Appl Lab Med An AACC Publ* 2(3):444–448.

[42] Pendle S, Barnes T. (2016) ***Neisseria gonorrhoeae* isolated from an unexpected site**. *Sex Health* 13(6):593–594.

[43] Korting HC. (1987) [**Uncomplicated gonorrhea and disseminated gonococcal infections - clinical aspects, diagnosis and therapy**]. *Urologe A* 26(5):237–245.

[44] Thayer WS, Blumer G. (1896) Ulcera**tive endocarditis due to the gonococcus: gonorrheal septicemia**. *Bull Johns Hopkins Hosp* 7:57–63.

[45] Thayer WS, Lazear JW. (1899) **A Second case of gonorrhoeal speticaemia and ulcerative endocarditis with observations upon the cardiac complications of gonorrhoea**. *J Exp Med* 4(1):81–116.

[46] Schoolnik GK, Buchanan TM, Holmes KK. (1976) **Gonococci causing disseminated gonococcal infection are resistant to the bactericidal action of normal human sera**. *J Clin Invest* 58(5):1163–1173.

[47] Brooks GF, Israel KS, Petersen BH. (1976) **Bactericidal and opsonic activity against *Neisseria gonorrhoeae* in sera from patients with disseminated gonococcal infection**. *J Infect Dis* 134(5):450–462.

[48] Petersen BH, Graham JA, Brooks GF. (1976) **Human deficiency of the eighth component of complement. The requirement of C8 for serum *Neisseria gonorrhoeae* bactericidal activity**. *J Clin Invest* 57(2):283–290.

[49] Alexander ER. (1988) **Gonorrhea in the Newborn**. *Ann N Y Acad Sci* 549:180–186.

[50] Masi AT, Eisenstein BI. (1981) **Disseminated gonococcal infection (DGI) and gonococcal arthritis (GCA): II. Clinical manifestations, diagnosis, complications, treatment, and prevention**. *Semin Arthritis Rheum* 10(3):173–197.

[51] Rice PA. (2005) **Gonococcal arthritis (disseminated gonococcal infection)**. *Infect Dis Clin North Am* 19(4):853–861.

[52] Henderson G, Ritchie WT. (1909) **Gonococcal meningitis**. *Rev Neurol Psychiatry* 7:57–87.

[53] Smith D. (1922) **Gonococcal Meningitis**. *Lancet* 199(5155):1217.

[54] Bradford WL, Kelley HW. (1933) **Gonococcic meningitis in a new born infant: with review of the literature**. *Am J Dis Child* 46(3):543–549.

[55] Wall TC, Peyton RB, Corey GR. (1989) **Gonococcal endocarditis: a new look at an old disease**. *Medicine (Baltimore)* 68(6):375–380.

[56] Shetty A. (2004) **Gonococcal endocarditis: a rare complication of a common disease**. *J Clin Pathol* 57(7):780–781.

[57] Nie S, Wu Y, Huang L, Pincus D, Tang YW, Lu X. (2014) **Gonococcal endocarditis: a case report and literature review**. *Eur J Clin Microbiol Infect Dis* 33(1):23–27.

[58] Beatrous S V, Grisoli SB, Riahi RR, Matherne RJ, Matherne RJ. (2017) **Cutaneous manifestations of disseminated gonococcemia**. *Dermatol Online J* 23(1): 13030/qt33b24006.

[59] Cowan L. (1969) **Gonococcal ulceration of the tongue in the gonococcal dermatitis syndrome**. *Sex Transm Infect* 45(3):228–231.

[60] Ghosn SH, Kibbi AG. (2004) **Cutaneous gonococcal infections**. *Clin Dermatol* 22(6):476–480.

[61] Brunham RC, Gottlieb SL, Paavonen J. (2015) **Pelvic inflammatory disease**. *N Engl J Med* 372(21):2039–2048.

[62] Reekie J, Donovan B, Guy R, Hocking JS, Kaldor JM, Mak DB, Pearson S, Preen D, Stewart L, Ward J, Liu B; Chlamydia and Reproductive Health Outcome Investigators; Chlamydia and Reproductive Health Outcome Investigators. (2017) **Risk of pelvic inflammatory disease in relation to chlamydia and gonorrhea testing, repeat testing, and positivity: a population-based cohort study**. *Clin Infect Dis* 66(3):437–443.

[63] Heumann CL, Quilter LAS, Eastment MC, Heffron R, Hawes SE. (2017) **Adverse birth outcomes and maternal *Neisseria gonorrhoeae i*nfection: a population-based cohort study in Washington State**. *Sex Transm Dis* 44(5):266–271.

[64] Bardin T. (2003) **Gonococcal arthritis**. *Best Pract Res Clin Rheumatol* 17(2):201–208.

[65] Quillin SJ, Seifert HS. (2018) ***Neisseria gonorrhoeae* host adaptation and pathogenesis**. *Nat Rev Microbiol* 16(4):226–240.

[66] Craig L, Pique ME, Tainer JA. (2004) **Type IV pilus structure and bacterial pathogenicity**. *Nat Rev Microbiol* 2(5):363–378.

[67] Obergfell KP, Seifert HS. (2016) **The pilin N-terminal domain maintains *Neisseria gonorrhoeae* transformation competence during pilus phase variation**. *PLOS Genet* 12(5):e1006069.

[68] Lambden PR, Heckels JE, James LT, Watt PJ. (1979) **Variations in surface protein composition associated with virulence properties in opacity types of *Neisseria gonorrhoeae***. *J Gen Microbiol* 114(2):305–312.

[69] Stern A, Brown M, Nickel P, Meyer TF. (1986) **Opacity genes in *Neisseria gonorrhoeae*: Control of phase and antigenic variation**. *Cell* 47(1):61–71.

[70] Song W, Ma L, Chen R, Stein DC. (2000) **Role of lipooligosaccharide in opa-independent invasion of *Neisseria gonorrhoeae* into human epithelial cells**. *J Exp Med* 191(6):949–960.

[71] van Vliet SJ, Steeghs L, Bruijns SC, Vaezirad MM, Snijders Blok C, Arenas Busto JA, Deken M, van Putten JP, van Kooyk Y. (2009) **Variation of *Neisseria gonorrhoeae* lipooligosaccharide directs dendritic cell–induced T helper responses**. *PLoS Pathog* 5(10):e1000625.

[72] Wetzler LM, Barry K, Blake MS, Gotschlich EC. (1992) **Gonococcal lipooligosaccharide sialylation prevents complement-dependent killing by immune sera**. *Infect Immun* 60(1):39–43.

[73] Mavrogiorgos N, Mekasha S, Yang Y, Kelliher MA, Ingalls RR. (2014) **Activation of NOD receptors by *Neisseria gonorrhoeae* modulates the innate immune response**. *Innate Immun* 20(4):377–389.

[74] Kaparakis M, Turnbull L, Carneiro L, Firth S, Coleman HA, Parkington HC, Le Bourhis L, Karrar A, Viala J, Mak J, Hutton ML, Davies JK, Crack PJ, Hertzog PJ, Philpott DJ, Girardin SE, Whitchurch CB, Ferrero RL. (2010) **Bacterial membrane vesicles deliver peptidoglycan to NOD1 in epithelial cells**. *Cell Microbiol* 12(3):372–385.

[75] Zhou X, Gao X, Broglie PM, Kebaier C, Anderson JE, Thom N, Apicella MA, Sempowski GD, Duncan JA. (2014) **Hexa-acylated lipid A is required for host inflammatory response to *Neisseria gonorrhoeae* in experimental gonorrhea**. *Infect Immun* 82(1):184–192.

[76] Singleton TE, Massari P, Wetzler LM. (2005) **Neisserial Porin-Induced Dendritic Cell Activation Is MyD88 and TLR2 Dependent**. *J Immunol* 174(6):3545–3550.

[77] Criss AK, Seifert HS. (2012) **A bacterial siren song: intimate interactions between *Neisseria* and neutrophils**. *Nat Rev Microbiol* 10(3):178–190.

[78] Papp JR, Schachter J, Gaydos CA, Van Der Pol B, Centers for Disease Control and Prevention. (2014) **Recommendations for the laboratory-based detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* - 2014**. *MMWR Recomm Reports* 63(RR-02):1–19.

[79] Martin DH, Cammarata C, Van Der Pol B, Jones RB, Quinn TC, Gaydos CA, Crotchfelt K, Schachter J, Moncada J, Jungkind D, Turner B, Peyton C. (2000) **Multicenter Evaluation of AMPLICOR and Automated COBAS AMPLICOR CT/NG Tests for *Neisseria gonorrhoeae***. *J Clin Microbiol* 38(10):3544–3549.

[80] Cook RL, Hutchison SL, Østergaard L, Braithwaite RS, Ness RB. (2005) **Systematic Review: Noninvasive Testing for *Chlamydia trachomatis* and *Neisseria gonorrhoeae***. *Ann Intern Med* 142(11):914–925.

[81] Van Dyck E, Ieven M, Pattyn S, Van Damme L, Laga M. (2001) **Detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* by enzyme immunoassay, culture, and three nucleic acid amplification Tests**. *J Clin Microbiol* 39(5):1751–1756.

[82] Serra-Pladevall J, Caballero E, Roig G, Juvé R, Barbera MJ, Andreu A. (2015) **Comparison between conventional culture and NAATs for the microbiological diagnosis in gonococcal infection**. *Diagn Microbiol Infect Dis* 83(4):341–343.

[83] Bromhead C, Miller A, Jones M, Whiley D. (2013) **Comparison of the cobas 4800 CT/NG test with culture for detecting *Neisseria gonorrhoeae* in genital and nongenital specimens in a low-prevalence population in New Zealand**. *J Clin Microbiol* 51(5):1505–1509.

[84] Van Der Pol B, Hook EW, Williams JA, Smith B, Taylor SN. (2015) **Performance of the BD CTQx and GCQx amplified assays on the BD Viper LT compared with the BD Viper XTR system**. *Sex Transm Dis* 42(9):521–523.

[85] Cornelisse VJ, Chow EP, Huffam S, Fairley CK, Bissessor M, De Petra V, Howden BP, Denham I, Bradshaw CS, Williamson D, Chen MY. (2017) **Increased detection of pharyngeal and rectal gonorrhea in men who have sex with men after transition from culture to nucleic acid amplification testing**. *Sex Transm Dis* 44(2):114–117.

[86] Cheng A, Kirby JE. (2014) **Evaluation of the Hologic Gen-Probe PANTHER, APTIMA Combo 2 assay in a tertiary care teaching hospital**. *Am J Clin Pathol* 141(3):397–403.

[87] Marlowe EM, Hardy D, Krevolin M, Gohl P, Bertram A, Arcenas R, Seiverth B, Schneider T, Liesenfeld O. (2017) **High-throughput testing of urogenital and extragenital specimens for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* with cobas ® CT/NG**. *Eur J Microbiol Immunol (Bp.)* 7(3):176–186.

[88] Van Der Pol B, Williams JA, Fuller D, Taylor SN, Hook EW 3rd. (2017) **Combined testing for chlamydia, gonorrhea, and trichomonas by use of the BD Max CT/GC/TV assay with genitourinary specimen types**. *J Clin Microbiol* 55(1):155–164.

[89] Spence JM, Wright L, Clark VL. (2008) **Laboratory maintenance of *Neisseria gonorrhoeae***. *Curr Protoc Microbiol* Chaper 4:Unit 4A.1.

[90] Alexander S, Ison C. (2005) **Evaluation of commercial kits for the identification of *Neisseria gonorrhoeae***. *J Med Microbiol* 54(Pt 9):827–831.

[91] Dillon JR, Carballo M, Pauze M. (1988) **Evaluation of eight methods for identification of pathogenic *Neisseria* species: Neisseria-Kwik, RIM-N, Gonobio-Test, Minitek, Gonochek II, GonoGen, Phadebact Monoclonal GC OMNI Test, and Syva MicroTrak test**. *J Clin Microbiol* 26(3):493–497.

[92] Carannante A, De Carolis E, Vacca P, Vella A, Vocale C, De Francesco MA, Cusini M, Del Re S, Dal Conte I, Cristaudo A, Ober P, Sanguinetti M, Stefanelli P. (2015) **Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) for identification and clustering of *Neisseria gonorrhoeae***. *BMC Microbiol* 15:142.

[93] Buchanan R, Ball D, Dolphin H, Dave J. (2016) **Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry for the identification of *Neisseria gonorrhoeae***. *Clin Microbiol Infect* 22(9):815.e5-815.e7.

[94] The European Committee on Antimicrobial Susceptibility Testing. (2019) **Breakpoint tables for interpretation of MICs and zone diameters. Version 9.0, 2019**. [Accessed: 28 October 2020] Available at: http://www.eucast.org.

[95] Unemo M, Fasth O, Fredlund H, Limnios A, Tapsall J. (2009) **Phenotypic and genetic characterization of the 2008 WHO *Neisseria gonorrhoeae* reference strain panel intended for global quality assurance and quality control of gonococcal antimicrobial resistance surveillance for public health purposes**. *J Antimicrob Chemother* 63(6):1142–1151.

[96] Unemo M, Golparian D, Sánchez-Busó L, Grad Y, Jacobsson S, Ohnishi M, Lahra MM, Limnios A, Sikora AE, Wi T, Harris SR. (2016) **The novel 2016 WHO *Neisseria gonorrhoeae* reference strains for global quality assurance of laboratory investigations: phenotypic, genetic and reference genome characterization**. *J Antimicrob Chemother* 71(11):3096–3108.

[97] Liu H, Taylor TH, Pettus K, Trees D. (2014) **Assessment of Etest as an alternative to agar dilution for antimicrobial susceptibility testing of *Neisseria gonorrhoeae***. *J Clin Microbiol* 52(5):1435–1440.

[98] Papp JR, Rowlinson MC, O'Connor NP, Wholehan J, Razeq JH, Glennen A, Ware D, Iwen PC, Lee LV, Hagan C. (2018) **Accuracy and reproducibility of the Etest to detect drug-resistant *Neisseria gonorrhoeae* to contemporary treatment**. *J Med Microbiol* 67(1):68–73.

[99] Bolan GA, Sparling PF, Wasserheit JN. (2012) **The emerging threat of untreatable gonococcal infection**. *N Engl J Med* 366(6):485–487.

[100] Goire N, Lahra MM, Chen M, Donovan B, Fairley CK, Guy R, Kaldor J, Regan D, Ward J, Nissen MD, Sloots TP, Whiley DM. (2014) **Molecular approaches to enhance surveillance of gonococcal antimicrobial resistance**. *Nat Rev Microbiol* 12(3):223–229.

[101] Hook EW 3rd, Kirkcaldy RD. (2018) **A brief history of evolving diagnostics and therapy for gonorrhea: lessons learned**. *Clin Infect Dis* 67(8):1294–1299.

[102] Ohnishi M, Golparian D, Shimuta K, Saika T, Hoshina S, Iwasaku K, Nakayama S, Kitawaki J, Unemo M. (2011) **Is *Neisseria gonorrhoeae* initiating a future era of untreatable gonorrhea?: detailed characterization of the first strain with high-level resistance to ceftriaxone**. *Antimicrob Agents Chemother* 55(7):3538–3545.

[103] Cole MJ, Spiteri G, Chisholm SA, Hoffmann S, Ison CA, Unemo M, Van de Laar M. (2014) **Emerging cephalosporin and multidrug resistant gonorrhoea in Europe**. *Eurosurveillance* 19(45):20955.

[104] Yu RX, Yin Y, Wang GQ, Chen SC, Zheng BJ, Dai XQ, Han Y, Li Q, Zhang GY, Chen X. (2014) **Worldwide susceptibility rates of *Neisseria gonorrhoeae* isolates to cefixime and cefpodoxime: a systematic review and meta-analysis**. *PLoS One* 9(1):e87849.

[105] Unemo M, Golparian D, Nicholas R, Ohnishi M, Gallay A, Sednaoui P. (2012) **High-level cefixime- and ceftriaxone-resistant *Neisseria gonorrhoeae* in France: novel *penA* mosaic allele in a successful international clone causes treatment failure**. *Antimicrob Agents Chemother* 56(3):1273–1280.

[106] Eyre DW, Sanderson ND, Lord E, Regisford-Reimmer N, Chau K, Barker L, Morgan M, Newnham R, Golparian D, Unemo M, Crook DW, Peto TE, Hughes G, Cole MJ, Fifer H, Edwards A, Andersson MI. (2018) **Gonorrhoea treatment failure caused by a *Neisseria gonorrhoeae* strain with combined ceftriaxone and high-level azithromycin resistance, England, February 2018**. *Eurosurveillance* 23(27):1800323

[107] Poncin T, Fouere S, Braille A, Camelena F, Agsous M, Bebear C, Kumanski S, Lot F, Mercier-Delarue S, Ngangro NN, Salmona M, Schnepf N, Timsit J, Unemo M, Bercot B. (2018) **Multidrug-resistant *Neisseria gonorrhoeae* failing treatment with ceftriaxone and doxycycline in France, November 2017**. *Eurosurveillance* 23(21):1800264.

[108] Golparian D, Ohlsson A, Janson H, Lidbrink P, Richtner T, Ekelund O, Fredlund H, Unemo M. (2014) **Four treatment failures of pharyngeal gonorrhoea with ceftriaxone (500 mg) or cefotaxime (500 mg), Sweden, 2013 and 2014**. *Eurosurveillance* 19(30):20862.

[109] Suay-García B, Pérez-Gracia M. (2018) **Future prospects for *Neisseria gonorrhoeae* treatment**. *Antibiotics (Basel)* 7(2):49.

[110] World Health Organization. (2012) **Global action plan to control the spread and impact of antimicrobial resistance in *Neisseria gonorrhoeae***. Geneva. [Accessed: 28 October 2020] Available at: https://apps.who.int/iris/bitstream/handle/10665/44863/9789241503501_eng.pdf

[111] European Centre for Disease Prevention and Control. (2012) **Response plan to control and manage the threat of multidrug- resistant gonorrhoea in Europe**. Stockholm. [Accessed: 28 October 2020] Available at: https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/1206-ECDC-MDR-gonorrhoea-response-plan.pdf

[112] Centers for Disease Control and Prevention. (2012) **Cephalosporin-resistant *Neisseria gonorrhoeae* public health response plan**. [Accessed: 28 October 2020] Available at: https://www.cdc.gov/std/treatment/ceph-r-responseplanjuly30-2012.pdf

[113] World Health Organization. Reproductive Health and Research. (2016) **WHO guidelines for the treatment of *Neisseria gonorrhoeae***. [Accessed: 28 October 2020] Available at: http://apps.who.int/iris/bitstream/10665/246114/1/9789241549691-eng.pdf

[114] Australasia Sexual Health Alliance. (2019) **Gonorrhoea - Australian STI management guidelines**. [Accessed: 28 October 2020] Available at: http://www.sti.guidelines.org.au/sexually-transmissible-infections/gonorrhoea

[115] Public Health Agency of Canada. (2020) **Canadian guidelines on sexually transmitted infections**. [Accessed: 28 October 2020] Available at: https://www.canada.ca/en/public-health/services/infectious-diseases/sexual-health-sexually-transmitted-infections/canadian-guidelines/sexually-transmitted-infections.html

[116] Workowski KA, Bolan GA; Centers for Disease Control and Prevention. (2015) **Sexually transmitted diseases treatment guidelines, 2015**. *MMWR Recomm Rep*. 64(RR.03):1–137.

[117] Bignell C, Fitzgerald M; Guideline Development Group; British Association for Sexual Health and HIV UK. (2011) **UK national guideline for the management of gonorrhoea in adults, 2011**. *Int J STD AIDS* 22(10):541–547.

[118] Bignell C, Unemo M; European STI Guidelines Editorial Board. (2013) **2012 European guideline on the diagnosis and treatment of gonorrhoea in adults**. *Int J STD AIDS* 24(2):85–92.

[119] Nicol M, Roke C, Obstet D. (2014) **New Zealand guideline for the management of gonorrhoea, 2014, and response to the threat of antimicrobial resistance members of the gonorrhoea guideline writing group**. [Accessed: 28 October 2020] Available at: https://www.nzshs.org/docman/guidelines/best-practice-guidelines/142-new-zealand-guideline-for-the-management-of-gonorrhoea-2014-and-response-to-the-threat-of-antimicrobial-resistance/file

[120] Unemo M, Golparian D, Eyre DW. (2019) **Antimicrobial resistance in *Neisseria gonorrhoeae* and treatment of gonorrhea**. In: *Methods in Molecular Biology*. Humana Press Inc. pp. 37–58.

[121] World Health Organization. (2018) **Report on global sexually transmitted infection surveillance, 2018**. Geneva. [Accessed: 28 October 2020] Available at: https://apps.who.int/iris/bitstream/handle/10665/277258/9789241565691-eng.pdf

[122] Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, Chico RM, Smolak A, Newman L, Gottlieb S, Thwin SS, Broutet N, Taylor MM. (2019) **Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016**. *Bull World Health Organ* 97(8):548-562P.

[123] Centers for Disease Control and Prevention. (2019) **Sexually transmitted diseases surveillance 2018**. Atlanta, GA. [Accessed: 28 October 2020] Available at: https://stacks.cdc.gov/view/cdc/79370

[124] Stenger MR, Pathela P, Anschuetz G, Bauer H, Simon J, Kohn R, Schumacher C, Torrone E. (2017) **Increases in the rate of *Neisseria gonorrhoeae* among gay, bisexual and other men who have sex with men - findings from the sexually transmitted disease surveillance network 2010–2015**. *Sex Transm Dis* 44(7):393–397.

[125] European Centre for Disease Prevention and Control. (2019) **Gonorrhoea - Annual Epidemiological Report for 2017**. Stockholm. [Accessed: 28 October 2020] Available at: https://www.ecdc.europa.eu/sites/default/files/documents/gonorrhoea-annual-epidemiological-report-2017.pdf

[126] European Centre for Disease Prevention and Control. (2018) **Gonorrhoea - Annual epidemiological report for 2016**. Stockholm. [Accessed: 28 October 2020] Available at: https://www.ecdc.europa.eu/sites/default/files/documents/AER_for_2016-gonorrhoea.pdf

[127] Direção-Geral da Saúde. (2017) **Doenças de Declaração Obrigatória 2013-2016, Volume I - Portugal**. (2017) Lisboa: Direção-Geral da Saúde. [Accessed: 28 October 2020] Available at: http://hdl.handle.net/10400.26/22529

[128] Direção-Geral da Sáude. (2017) **Doenças de Declaração Obrigatória 2013-2016, Volume II - Regiões**. Lisboa: Direção-Geral da Saúde. [Accessed: 28 October 2020] Available at: http://hdl.handle.net/10400.26/22530

[129] Jardim Santos C, Gomes B, Ribeiro AI. (2020) **Mapping geographical patterns and high rate areas for sexually transmitted infections in Portugal**. *Sex Transm Dis* 47(4):261–268.

[130] Weston EJ, Wi T, Papp J. (2017) **Strengthening global surveillance for antimicrobial drug–resistant *Neisseria gonorrhoeae* through the enhanced gonococcal antimicrobial surveillance program**. *Emerg Infect Dis* 23(13):S47–S52.

[131] Public Health England. (2017) **Surveillance of antimicrobial resistance in *Neisseria gonorrhoeae* in England and Wales Key findings from the Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP)**. London. [Accessed: 28 October 2020] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/651636/GRASP_Report_2017.pdf

[132] Public Health England. (2018) **Surveillance of antimicrobial resistance in *Neisseria gonorrhoeae* in England and Wales Key findings from the Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP)**. London. [Accessed: 28 October 2020] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/746261/GRASP_2017_report.pdf

[133] Public Health England. (2019) **Antimicrobial resistance in *Neisseria gonorrhoeae* in England and Wales**. London. [Accessed: 28 October 2020] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/834924/GRASP__2018_report.pdf

[134] European Centre for Disease (2018) **Prevention and Control. Gonococcal antimicrobial susceptibility surveillance in Europe 2016**. Stockholm. [Accessed: 28 October 2020] Available at: https://www.ecdc.europa.eu/sites/default/files/documents/EURO-GASP-report-2016.pdf

[135] European Centre for Disease Prevention and Control. **Gonococcal antimicrobial susceptibility surveillance in Europe – Results summary 2017**. Stockholm. [Accessed: 28 October 2020] Available at: https://www.ecdc.europa.eu/sites/default/files/documents/Euro-GASP-2017.pdf

[136] Kirkcaldy RD, Kidd S, Weinstock HS, Papp JR, Bolan GA. (2013) **Trends in antimicrobial resistance in *Neisseria gonorrhoeae* in the USA: the Gonococcal Isolate Surveillance Project (GISP), January 2006–June 2012**. *Sex Transm Infect* 89 Suppl:iv5–iv10.

[137] Kirkcaldy RD, Harvey A, Papp JR, Del Rio C, Soge OO, Holmes KK, Hook EW 3rd, Kubin G, Riedel S, Zenilman J, Pettus K, Sanders T, Sharpe S, Torrone E. (2016) **Neisseria gonorrhoeae antimicrobial susceptibility surveillance — The Gonococcal Isolate Surveillance Project, 27 Sites, United States, 2014**. *MMWR Surveill Summ* 65(7):1–19.

[138] Public Health Agency of Canada. (2016) **National surveillance of antimicrobial susceptibilities of *Neisseria gonorrhoeae* in Canada - Annual summary 2016**. [Accessed: 28 October 2020] Available at: https://www.canada.ca/en/public-health/services/publications/drugs-health-

products/national-surveillance-antimicrobial-susceptibilities-neisseria-gonorrhoeae-annual-summary-2016.html

[139] Public Health Agency of Canada. (2017) **National surveillance of antimicrobial susceptibilities of *Neisseria gonorrhoeae* in Canada: Annual summary 2017**. [Accessed: 28 October 2020] Available at: https://www.canada.ca/en/public-health/services/publications/drugs-health-products/national-surveillance-antimicrobial-susceptibilities-neisseria-gonorrhoeae-annual-summary-2017.html

[140] Lahra MM, Enriquez R, George CRR. (2019) **Australian Gonococcal Surveillance Programme Annual Report, 2017**. *Commun Dis Intell* 43.

[141] Lahra MM, Enriquez RP, George CRR. (2020) **Australian Gonococcal Surveillance Programme Annual Report, 2018**. *Commun Dis Intell* 44.

[142] Carannante A, Prignano G, Cusini M, Matteelli A, Dal Conte I, Ghisetti V, D'Antuono A, Cavrini F, Antonetti R, Stefanelli P. (2012) **Cefixime and ceftriaxone susceptibility of *Neisseria gonorrhoeae* in Italy from 2006 to 2010**. *Clin Microbiol Infect* 18(6):558–564.

[143] Centers for Disease Control and Prevention. (2011) **Cephalosporin susceptibility among *Neisseria gonorrhoeae* Isolates-United States, 2000-2010**. *MMWR Morb Mortal Wkly Rep* 60(26):873–877.

[144] Chisholm SA, Alexander S, Desouza-Thomas L, Maclure-Webster E, Anderson J, Nichols T, Lowndes CM, Ison CA; GRASP Collaborative Group. (2011) **Emergence of a *Neisseria gonorrhoeae* clone showing decreased susceptibility to cefixime in England and Wales**. *J Antimicrob Chemother* 66(11):2509–2512.

[145] Whiley DM, Goire N, Lambert SB, Nissen MD, Sloots TP, Tapsall JW. (2011) **Reduced susceptibility to ceftriaxone in *Neisseria gonorrhoeae* is spread internationally by genetically distinct gonococcal populations**. *J Antimicrob Chemother* 66(5):1186–1187.

[146] Akasaka S, Muratani T, Yamada Y, Inatomi H, Takahashi K, Matsumoto T. (2001) **Emergence of cephem- and aztreonam-high-resistant *Neisseria gonorrhoeae* that does not produce β-lactamase**. *J Infect Chemother* 7(1):49–50.

[147] Tanaka M, Nakayama H, Huruya K, Konomi I, Irie S, Kanayama A, Saika T, Kobayashi I. (2006) **Analysis of mutations within multiple genes associated with resistance in a clinical isolate of *Neisseria gonorrhoeae* with reduced ceftriaxone susceptibility that shows a multidrug-resistant phenotype**. *Int J Antimicrob Agents* 27(1):20–26.

[148] van Dam AP, van Ogtrop ML, Golparian D, Mehrtens J, de Vries HJ, Unemo M. (2014) **Verified clinical failure with cefotaxime 1g for treatment of gonorrhoea in the Netherlands: a case report**. *Sex Transm Infect* 90(7):513–514.

[149] Cámara J, Serra J, Ayats J, Bastida T, Carnicer-Pont D, Andreu A, Ardanuy C. (2012) **Molecular characterization of two high-level ceftriaxone-resistant *Neisseria gonorrhoeae* isolates detected in Catalonia, Spain**. *J Antimicrob Chemother* 67(8):1858–1860.

[150] Day MJ, Spiteri G, Jacobsson S, Woodford N, Amato-Gauci AJ, Cole MJ, Unemo M; Euro-GASP network. (2018) **Stably high azithromycin resistance and decreasing ceftriaxone susceptibility in *Neisseria gonorrhoeae* in 25 European countries, 2016**. *BMC Infect Dis* 18(1):609.

[151] Thakur SD, Levett PN, Horsman GB, Dillon J-AR. (2018) **High levels of susceptibility to new and older antibiotics in *Neisseria gonorrhoeae* isolates from Saskatchewan (2003–15): time to consider point-of-care or molecular testing for precision treatment?** *J Antimicrob Chemother* 73(1):118–125.

[152] Liang JY, Cao WL, Li XD, Bi C, Yang RD, Liang YH, Li P, Ye XD, Chen XX, Zhang XB. (2016) **Azithromycin-resistant *Neisseria gonorrhoeae* isolates in Guangzhou, China (2009–2013): coevolution with decreased susceptibilities to ceftriaxone and genetic characteristics**. *BMC Infect Dis* 16:152.

[153] Katz AR, Komeya AY, Kirkcaldy RD, Whelen AC, Soge OO, Papp JR, Kersh EN, Wasserman GM, O'Connor NP, O'Brien PS, Sato DT, Maningas EV, Kunimoto GY, Tomas JE. (2017) **Cluster of *Neisseria gonorrhoeae* isolates with high-level azithromycin resistance and decreased ceftriaxone susceptibility, Hawaii, 2016**. *Clin Infect Dis* 65(6):918–923.

[154] Cole MJ, Spiteri G, Jacobsson S, Woodford N, Tripodo F, Amato-Gauci AJ, Unemo M; Euro-GASP network. (2017) **Overall low extended-spectrum cephalosporin resistance but high azithromycin resistance in *Neisseria gonorrhoeae* in 24 European countries, 2015**. *BMC Infect Dis* 17(1):617.

[155] Chisholm SA, Wilson J, Alexander S, Tripodo F, Al-Shahib A, Schaefer U, Lythgow K, Fifer H. (2016) **An outbreak of high-level azithromycin resistant *Neisseria gonorrhoeae* in England**. *Sex Transm Infect* 92(5):365–367.

[156] Smolarchuk C, Wensley A, Padfield S, Fifer H, Lee A, Hughes G. (2018) **Persistence of an outbreak of gonorrhoea with high-level resistance to azithromycin in England, November 2014–May 2018**. *Eurosurveillance* 23(23):1800287.

[157] Lahra MM, Ryder N, Whiley DM. (2014) **A new multidrug-resistant strain of *Neisseria gonorrhoeae* in Australia**. *N Engl J Med* 371(19):1850–1851.

[158] Lahra MM, Ward A, Trembizki E, Hermanson J, Clements E, Lawrence A, Whiley D. (2017) **Treatment guidelines after an outbreak of azithromycin-resistant *Neisseria gonorrhoeae* in South Australia**. *Lancet Infect Dis* 17(2):133–134.

[159] Terkelsen D, Tolstrup J, Johnsen CH, Lund O, Larsen HK, Worning P, Unemo M, Westh H. (2017) **Multidrug-resistant *Neisseria gonorrhoeae* infection with ceftriaxone resistance and intermediate resistance to azithromycin, Denmark, 2017**. *Eurosurveillance* 22(42):17–00659.

[160] Golparian D, Rose L, Lynam A, Mohamed A, Bercot B, Ohnishi M, Crowley B, Unemo M. (2018) **Multidrug-resistant *Neisseria gonorrhoeae* isolate, belonging to the internationally spreading Japanese FC428 clone, with ceftriaxone resistance and intermediate resistance to azithromycin, Ireland, August 2018**. *Eurosurveillance* 23(47):1800617.

[161] Chen SC, Hu LH, Zhu XY, Yin YP. (2009) **Gonococcal urethritis caused by a multidrug resistant *Neisseria gonorrhoeae* strain with high-level resistance to spectinomycin in China**. *Emerg Microbes Infect* 9(1):517–519.

[162] Yuan Q, Li Y, Xiu L, Zhang C, Fu Y, Jiang C, Tang L, Peng J. (2019) **Identification of multidrug-resistant *Neisseria gonorrhoeae* isolates with combined resistance to both ceftriaxone and azithromycin, China, 2017–2018**. *Emerg Microbes Infect* 8(1):1546–1549.

[163] Pachulec E, van der Does C. (2010) **Conjugative plasmids of *Neisseria gonorrhoeae***. *PLoS One* 5(4):e9962.

[164] van Passel MWJ, Bart A, Luyf ACM, van Kampen AHC, van der Ende A. (2006) **Identification of acquired DNA in *Neisseria lactamica***. *FEMS Microbiol Lett* 262(1):77–84.

[165] van Passel MWJ, van der Ende A, Bart A. (2006) **Plasmid diversity in *Neisseriae***. *Infect Immun* 74(8):4892–4899.

[166] Hamilton HL, Dillard JP. (2006) **Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination**. *Mol Microbiol* 59(2):376–385.

[167] Hill SA, Masters TL, Wachter J. (2016) **Gonorrhea – an evolving disease of the new millennium**. *Microb Cell* 3(9):371–389.

[168] Rotman E, Seifert HS. (2014) **The genetics of *Neisseria* species**. *Annu Rev Genet* 48:405–431.

[169] Marri PR, Paniscus M, Weyand NJ, Rendón MA, Calton CM, Hernández DR, Higashi DL, Sodergren E, Weinstock GM, Rounsley SD, So M. (2010) **Genome sequencing reveals widespread virulence gene exchange among human *Neisseria* species**. *PLoS One* 5(7):e11835.

[170] Kawai M, Uchiyama I, Kobayashi I. (2005) **Genome comparison in silico in *Neisseria* suggests integration of filamentous bacteriophages by their own transposase**. *DNA Res* 12(6):389–401.

[171] Spencer-Smith R, Varkey EM, Fielder MD, Snyder LAS. (2012) **Sequence features contributing to chromosomal rearrangements in *Neisseria gonorrhoeae*.** *PLoS One* 7(9):e46023.

[172] Dillard JP, Seifert HS. (2001) **A variable genetic island specific for *Neisseria gonorrhoeae* is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates**. *Mol Microbiol* 41(1):263–277.

[173] Hamilton HL, Schwartz KJ, Dillard JP. (2001) **Insertion-duplication mutagenesis of Neisseria: use in characterization of DNA transfer genes in the gonococcal genetic island**. *J Bacteriol* 183(16):4718–4726.

[174] Moxon R, Bayliss C, Hood D. (2006) **Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation**. *Annu Rev Genet* 40:307–333.

[175] Snyder LAS, Butcher SA, Saunders NJ. (2001) **Comparative whole-genome analyses reveal over 100 putative phase-variable genes in the pathogenic *Neisseria* spp**. *Microbiology* 147(Pt 8):2321–2332.

[176] Jordan PW, Snyder LAS, Saunders NJ. (2005) **Strain-specific differences in *Neisseria gonorrhoeae* associated with the phase variable gene repertoire**. *BMC Microbiol* 5:21.

[177] Martin P, van de Ven T, Mouchel N, Jeffries AC, Hood DW, Moxon ER. (2003) **Experimentally revised repertoire of putative contingency loci in *Neisseria meningitidis* strain MC58: evidence for a novel mechanism of phase variation**. *Mol Microbiol* 50(1):245–257.

[178] Zelewska MA, Pulijala M, Spencer-Smith R, Mahmood HA, Norman B, Churchward CP, Calder A, Snyder LAS. (2016) **Phase variable DNA repeats in *Neisseria gonorrhoeae* influence transcription, translation, and protein sequence variation**. *Microb Genomics* 2(8):e000078.

[179] Seib KL, Jen FE-C, Scott AL, Tan A, Jennings MP. (2017) **Phase variation of DNA methyltransferases and the regulation of virulence and immune evasion in the pathogenic *Neisseria***. *Pathog Dis* 75(6):80.

[180] Wanford JJ, Green LR, Aidley J, Bayliss CD. (2018) **Phasome analysis of pathogenic and commensal *Neisseria* species expands the known repertoire of phase variable genes, and highlights common adaptive strategies**. *PLoS One* 13(5):e0196675.

[181] Braun DC, Stein DC. (2004) **The *lgtABCDE* gene cluster, involved in lipooligosaccharide biosynthesis in *Neisseria gonorrhoeae*, contains multiple promoter sequences**. *J Bacteriol* 186(4):1038–1049.

[182] Shafer WM, Datta A, Kolli VS, Rahman MM, Balthazar JT, Martin LE, Veal WL, Stephens DS, Carlson R. (2002) **Phase variable changes in genes *lgtA* and *lgtC* within the *lgtABCDE* operon of *Neisseria gonorrhoeae* can modulate gonococcal susceptibility to normal human serum**. *J Endotoxin Res* 8(1):47–58.

[183] Banerjee A, Wang R, Uljon SN, Rice PA, Gotschlich EC, Stein DC. (1998) **Identification of the gene (*lgtG*) encoding the lipooligosaccharide chain synthesizing glucosyl transferase from *Neisseria gonorrhoeae***. *Proc Natl Acad Sci* 95(18):10872–10877.

[184] Danaher RJ, Levin JC, Arking D, Burch CL, Sandlin R, Stein DC. (1995) **Genetic basis of *Neisseria gonorrhoeae* lipooligosaccharide antigenic variation**. *J Bacteriol* 177(24):7275–7279.

[185] Bhat KS, Gibbs CP, Barrera O, Morrison SG, Jähnig F, Stern A, Kupsch EM, Meyer TF, Swanson J. (1991) **The opacity proteins of *Neisseria gonorrhoeae* strain MS11 are encoded by a family of 11 complete genes**. *Mol Microbiol* 5(8):1889–1901.

[186] Dempsey JA, Litaker W, Madhure A, Snodgrass TL, Cannon JG. (1991) **Physical map of the chromosome of *Neisseria gonorrhoeae* FA1090 with locations of genetic markers, including *opa* and *pil* genes**. *J Bacteriol* 173(17):5476–5486.

[187] Sadarangani M, Pollard AJ, Gray-Owen SD. (2011) **Opa proteins and CEACAMs: pathways of immune engagement for pathogenic *Neisseria***. *FEMS Microbiol Rev* 35(3):498–514.

[188] Cahoon LA, Seifert HS. (2011) **Focusing homologous recombination: pilin antigenic variation in the pathogenic *Neisseria***. *Mol Microbiol* 81(5):1136–1143.

[189] Davies JK, Harrison PF, Lin YH, Bartley S, Khoo CA, Seemann T, Ryan CS, Kahler CM, Hill SA. (2014) **The use of high-throughput DNA sequencing in the investigation of antigenic variation: application to *Neisseria* speci**es. *PLoS One* 9(1):e86704.

[190] Hill SA, Davies JK. (2009) **Pilin gene variation in *Neisseria gonorrhoeae*: reassessing the old paradigms**. *FEMS Microbiol Rev* 33(3):521–530.

[191] Forest KT, Bernstein SL, Getzoff ED, So M, Tribbick G, Geysen HM, Deal CD, Tainer JA. (1996) **Assembly and antigenicity of the *Neisseria gonorrhoeae* pilus mapped with antibodies**. *Infect Imm*un 64(2):644–652.

[192] Eyre DW, De Silva D, Cole K, Peters J, Cole MJ, Grad YH, Demczuk W, Martin I, Mulvey MR, Crook DW, Walker AS, Peto TEA, Paul J. (2017) **WGS to predict antibiotic MICs for *Neisseria gonorrhoeae***. *J Antimicrob Chemother* 72(7):1937–1947.

[193] Fiebelkorn KR, Crawford SA, Jorgensen JH. (2005) **Mutations in *folP* associated with elevated sulfonamide MICs for *Neisseria meningitidis* clinical isolates from five continents**. *Antimicrob Agents Chemother* 49(2):536–540.

[194] Swedberg G, Fermér C, Sköld O. (1993) **Point mutations in the dihydropteroate synthase gene causing sulfonamide resistance**. *Adv Exp Med Biol.* 338:555–558.

[195] Morse SA, Johnson SR, Biddle JW, Roberts MC. (1986) **High-level tetracycline resistance in *Neisseria gonorrhoeae* is result of acquisition of streptococcal *tetM* determinant**. *Antimicrob Agents Chemother* 30(5):664–670.

[196] Gascoyne DM, Heritage J, Hawkey PM, Turner A, van Klingeren B. (1991) **Molecular evolution of tetracycline-resistance plasmids carrying TetM found in *Neisseria gonorrhoeae* from different countries**. *J Antimicrob Chemother* 28(2):173–183.

[197] Hu M, Nandi S, Davies C, Nicholas RA. (2005) **High-level chromosomally mediated tetracycline resistance in *Neisseria gonorrhoeae* results from a point mutation in the *rpsJ* gene encoding ribosomal protein S10 in combination with the *mtrR* and *penB* resistance determinants**. *Antimicrob Agents Chemother* 49(10):4327–4334.

[198] Nolte O, Müller M, Reitz S, Ledig S, Ehrhard I, Sonntag HG. (2003) **Description of new mutations in the *rpoB* gene in rifampicin-resistant *Neisseria meningitidis* selected in vitro in a stepwise manner**. *J Med Microbiol* 52(Pt 12):1077–1081.

[199] Unemo M, Nicholas RA, Jerse AE, Davies C, Shafer WM. (2014) **Pathogenic *Neisseria*: genomics, molecular biology and disease intervention**. In: Davies JK, Kahler CM (editors). *Pathogenic Neisseria: Genomics, Molecular Biology and Disease Intervention*. London: Caister Academic Press. pp. 161–192.

[200] Zhao S, Duncan M, Tomberg J, Davies C, Unemo M, Nicholas RA. (2009) **Genetics of chromosomally mediated intermediate resistance to ceftriaxone and cefixime in *Neisseria gonorrhoeae***. *Antimicrob Agents Chemother* 53(9):3744–3751.

[201] Bharat A, Demczuk W, Martin I, Mulvey MR. (2015) **Effect of variants of penicillin-binding protein 2 on cephalosporin and carbapenem susceptibilities in *Neisseria gonorrhoeae***. *Antimicrob Agents Chemother* 59(8):5003–5006.

[202] Spratt BG. (1988) **Hybrid penicillin-binding proteins in penicillin-resistant strains of *Neisseria gonorrhoeae***. *Nature* 332(6160):173–176.

[203] Unemo M, Shafer WM. (2014) **Antimicrobial resistance in *Neisseria gonorrhoeae* in the 21st century: past, evolution, and future**. *Clin Microbiol Rev* 27(3):587–613.

[204] Zarantonelli L, Borthagaray G, Lee EH, Veal W, Shafer WM. (2001) **Decreased susceptibility to azithromycin and erythromycin mediated by a novel *mtr(R)* promoter mutation in *Neisseria gonorrhoeae***. *J Antimicrob Chemother* 47(5):651–654.

[205] Cousin SL. (2003) **Acquired macrolide resistance genes and the 1 bp deletion in the *mtrR* promoter in *Neisseria gonorrhoeae***. J Antimicrob Chemother 51(1):131–133.

[206] Ohneck EA, Zalucki YM, Johnson PJ, Dhulipala V, Golparian D, Unemo M, Jerse AE, Shafer WM. (2011) **A novel mechanism of high-level, broad-spectrum antibiotic resistance caused by a single base pair change in *Neisseria gonorrhoeae***. *MBio* 2(5):e00187-11

[207] Jacobsson S, Golparian D, Cole M, Spiteri G, Martin I, Bergheim T, Borrego MJ, Crowley B, Crucitti T, Van Dam AP, Hoffmann S, Jeverica S, Kohl P, Mlynarczyk-Bonikowska B, Pakarna G, Stary A, Stefanelli P, Pavlik P, Tzelepi E, Abad R, Harris SR, Unemo M. (2016) **WGS analysis and molecular resistance mechanisms of azithromycin-resistant (MIC >2 mg/L) *Neisseria gonorrhoeae* isolates in Europe from 2009 to 2014**. *J Antimicrob Chemother* 71(11):3109–3116.

[208] Demczuk W, Martin I, Peterson S, Bharat A, Van Domselaar G, Graham M, Lefebvre B, Allen V, Hoang L, Tyrrell G, Horsman G, Wylie J, Haldane D, Archibald C, Wong T, Unemo M, Mulvey MR. (2016) **Genomic epidemiology and molecular resistance mechanisms of azithromycin-resistant *Neisseria gonorrhoeae* in Canada from 1997 to 2014**. *J Clin Microbiol* 54(5):1304–1313.

[209] Warner DM, Shafer WM, Jerse AE. (2008) **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(2):462–478.

[210] Olesky M, Hobbs M, Nicholas RA. (2002) **Identification and analysis of amino acid mutations in porin IB that mediate intermediate-level resistance to penicillin and tetracycline in *Neisseria gonorrhoeae***. *Antimicrob Agents Chemother* 46(9):2811–2820.

[211] Ohnishi M, Ono E, Shimuta K, Watanabe H, Okamura N. (2010) **Identification of TEM-135 β-lactamase in penicillinase-producing *Neisseria gonorrhoeae* strains in Japan**. *Antimicrob Agents Chemother* 54(7):3021–3023.

[212] Palmer HM, Leeming JP, Turner A. (2000) **A multiplex polymerase chain reaction to differentiate β-lactamase plasmids of *Neisseria gonorrhoeae***. *J Antimicrob Chemother* 45(6):777–782.

[213] Chisholm SA, Dave J, Ison CA. (2010) **High-level azithromycin resistance occurs in *Neisseria gonorrhoeae* as a result of a single point mutation in the *23S rRNA* genes**. *Antimicrob Agents Chemother* 54(9):3812–3816.

[214] Ng LK, Martin I, Liu G, Bryden L. (2002) **Mutation in *23S rRNA* associated with macrolide resistance in *Neisseria gonorrhoeae***. *Antimicrob Agents Chemother* 46(9):3020–3025.

[215] Belland RJ, Morrison SG, Ison C, Huang WM. (1994) ***Neisseria gonorrhoeae* acquires mutations in analogous regions of *gyrA* and *parC* in fluoroquinolone-resistant isolates**. *Mol Microbiol* 14(2):371–380.

[216] Zhao S, Tobiason DM, Hu M, Seifert HS, Nicholas RA. (2005) **The *penC* mutation conferring antibiotic resistance in *Neisseria gonorrhoeae* arises from a mutation in the PilQ secretin that interferes with multimer stability**. *Mol Microbiol* 57(5):1238–1251.

[217] Roberts MC, Chung WO, Roe D, Xia M, Marquez C, Borthagaray G, Whittington WL, Holmes KK. (1999) **Erythromycin-resistant *Neisseria gonorrhoeae* and oral commensal *Neisseria* spp. carry known rRNA methylase genes**. *Antimicrob Agents Chemother* 43(6):1367–1372.

[218] Rouquette-Loughlin CE, Balthazar JT, Shafer WM. (2005) **Characterization of the MacA–MacB efflux system in *Neisseria gonorrhoeae***. *J Antimicrob Chemother* 56(5):856–860.

[219] Luna VA, Cousin S, Whittington WLH, Roberts MC. (2000) **Identification of the conjugative *mef* gene in clinical *Acinetobacter junii* and *Neisseria gonorrhoeae* Isolates**. *Antimicrob Agents Chemother* 44(9):2503–2506.

[220] Rouquette-Loughlin C, Dunham SA, Kuhn M, Balthazar JT, Shafer WM. (2003) **The NorM efflux pump of *Neisseria gonorrhoeae* and *Neisseria meningitidis* recognizes antimicrobial cationic compounds**. *J Bacteriol* 185(3):1101–1106.

[221] Lindback E, Rahman M, Jalal S, Wretlind B. (2002) **Mutations in *gyrA*, *gyrB*, *parC*, and *parE* in quinolone-resistant strains of *Neisseria gonorrhoeae***. *APMIS* 110(9):651–657.

[222] Galimand M, Gerbaud G, Courvalin P. (2000) **Spectinomycin resistance in *Neisseria* spp. due to mutations in *16S rRNA***. *Antimicrob Agents Chemother* 44(5):1365–1366.

[223] Unemo M, Golparian D, Skogen V, Olsen AO, Moi H, Syversen G, Hjelmevoll SO. (2013) ***Neisseria gonorrhoeae* strain with high-level resistance to spectinomycin due to a novel resistance mechanism (mutated ribosomal protein S5) verified in Norway**. *Antimicrob Agents Chemother* 57(2):1057–1061.

[224] Ilina EN, Malakhova MV, Bodoev IN, Oparina NY, Filimonova AV, Govorun VM. (2013) **Mutation in ribosomal protein S5 leads to spectinomycin resistance in *Neisseria gonorrhoeae***. *Front Microbiol* 4:186.

[225] Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG. (1998) **Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms**. *Proc Natl Acad Sci* 95(6):3140–3145.

[226] Maiden MCJ. (2006) **Multilocus sequence typing of Bacteria**. *Annu Rev Microbiol* 60:561–588.

[227] Martin IMC, Ison CA, Aanensen DM, Fenton KA, Spratt BG. (2004) **Rapid sequence-based identification of gonococcal transmission clusters in a large metropolitan area**. *J Infect Dis* 189(8):1497–1505.

[228] Unemo M, Dillon J-AR. (2011) **Review and international recommendation of methods for typing *Neisseria gonorrhoeae* isolates and their implications for improved knowledge of gonococcal epidemiology, treatment, and biology**. *Clin Microbiol Rev* 24(3):447–458.

[229] Młynarczyk-Bonikowska B, Malejczyk M, Majewski S, Unemo M. (2018) **Antibiotic resistance and NG-MAST sequence types of *Neisseria gonorrhoeae* isolates in Poland compared to the world**. *Adv Dermatology Allergol* 35(6):546–551.

[230] Młynarczyk-Bonikowska B, Majewska A, Malejczyk M, Młynarczyk G, Majewski S. (2020) **Multiresistant *Neisseria gonorrhoeae*: a new threat in second decade of the XXI century**. *Med Microbiol Immunol* 209(2):95–108.

[231] Demczuk W, Lynch T, Martin I, Van Domselaar G, Graham M, Bharat A, Allen V, Hoang L, Lefebvre B, Tyrrell G, Horsman G, Haldane D, Garceau R, Wylie J, Wong T, Mulvey MR. (2015) **Whole-genome phylogenomic heterogeneity of *Neisseria gonorrhoeae* isolates with decreased cephalosporin susceptibility collected in Canada between 1989 and 2013**. *J Clin Microbiol* 53(1):191–200.

[232] Harris SR, Cole MJ, Spiteri G, Sánchez-Busó L, Golparian D, Jacobsson S, Goater R, Abudahab K, Yeats CA, Bercot B, Borrego MJ, Crowley B, Stefanelli P, Tripodo F, Abad R, Aanensen DM, Unemo M; Euro-GASP study group. (2018) **Public health surveillance of multidrug-resistant clones of *Neisseria gonorrhoeae* in Europe: a genomic survey**. *Lancet Infect Dis* 18(7):758–768.

[233] Chisholm SA, Unemo M, Quaye N, Johansson E, Cole MJ, Ison CA, Van de Laar MJ. (2013) **Molecular epidemiological typing within the European Gonococcal Antimicrobial Resistance Surveillance Programme reveals predominance of a multidrug-resistant clone**. *Eurosurveillance* 18(3):20358.

[234] European Centre for Disease Prevention and Control. (2018) **Molecular typing of *Neisseria gonorrhoeaea* study of 2013 isolates**. Stockholm. [Accessed: 28 October 2020] Available at: https://www.ecdc.europa.eu/sites/default/files/documents/Molecular-typing-N-gonorrhoeae-web.pdf

[235] Public Health Agency of Canada. (2015) **National surveillance of antimicrobial susceptibilities of *Neisseria gonorrhoeae* annual summary 2014**. [Accessed: 28 October 2020] Available at: https://www.canada.ca/content/dam/canada/health-canada/migration/healthy-canadians/publications/drugs-products-medicaments-produits/2014-neisseria/alt/surveillance-gonorrhoeae-2014-eng.pdf

[236] Martin I, Sawatzky P, Liu G, Allen V, Lefebvre B, Hoang L, Lovgren M, Haldane D, Caeseele PV, Horsman G, Garceau R, Ratnam S, Wong T, Gilmour M. (2013) **Antimicrobial susceptibilities and distribution of sequence types of *Neisseria gonorrhoeae* isolates in Canada: 2010**. *Can J Micro*biol 59(10):671–678.

[237] Gose S, Nguyen D, Lowenberg D, Samuel M, Bauer H, Pandori M. (2013) ***Neisseria gonorrhoeae* and extended-spectrum cephalosporins in California: surveillance and molecular detection of mosaic *penA***. *BMC Infect Dis* 13:570.

[238] Grad YH, Kirkcaldy RD, Trees D, Dordel J, Harris SR, Goldstein E, Weinstock H, Parkhill J, Hanage WP, Bentley S, Lipsitch M. (2014) **Genomic epidemiology of *Neisseria gonorrhoeae* with reduced susceptibility to cefixime in the USA: a retrospective observational study**. *Lancet Infect Dis* 14(3):220–226.

[239] Shimuta K, Unemo M, Nakayama S, Morita-Ishihara T, Dorin M, Kawahata T, Ohnishi M; Antibiotic-Resistant Gonorrhea Study Group. (2013) **Antimicrobial resistance and molecular typing of *Neisseria gonorrhoeae* isolates in Kyoto and Osaka, Japan, 2010 to 2012: intensified surveillance after identification of the first strain (H041) with high-level ceftriaxone resistance**. *Antimicrob Agents Chemother* 57(11):5225–5232.

[240] Gianecini RA, Golparian D, Zittermann S, Litvik A, Gonzalez S, Oviedo C, Melano RG, Unemo M, Galarza P; Gonococcal Antimicrobial Susceptibility Surveillance Programme-Argentina (GASSP-AR) Working Group. (2019) **Genome-based epidemiology and antimicrobial resistance determinants of *Neisseria gonorrhoeae* isolates with decreased susceptibility and resistance to extended-spectrum cephalosporins in Argentina in 2011–16**. *J Antimicrob Chemother* 74(6):1551–1559.

[241] Chen SC, Yin YP, Dai XQ, Unemo M, Chen XS. (2016) **First nationwide study regarding ceftriaxone resistance and molecular epidemiology of *Neisseria gonorrhoeae* in China**. *J Antimicrob Chemother* 71(1):92–99.

[242] Zheng Z, Liu L, Shen X, Yu J, Chen L, Zhan L, Chen H, Lin C, Jiang Y, Xia H, Wang L, Yu F. (2019) **Antimicrobial resistance and molecular characteristics among *Neisseria gonorrhoeae c*linical isolates in a Chinese tertiary hospital**. *Infect Drug Resist* 12:3301–3309.

[243] Qin X, Zhao Y, Chen W, Wu X, Tang S, Li G, Yuqi Y, Cao W, Liu X, Huang J, Yang J, Chen W, Tang W, Zheng H. (2019) **Changing antimicrobial susceptibility and molecular characterisation of *Neisseria gonorrhoeae* isolates in Guangdong, China: in a background of rapidly rising epidemic**. *Int J Antimicrob Agents* 54(6):757–765.

[244] Yan J, Xue J, Chen Y, Chen S, Wang Q, Zhang C, Wu S, Lv H, Yu Y, van der Veen S. (2019) **Increasing prevalence of *Neisseria gonorrhoeae* with decreased susceptibility to ceftriaxone and resistance to azithromycin in Hangzhou, China (2015–17)**. *J Antimicrob Chemother* 74(1):29-37

[245] Attram N, Agbodzi B, Dela H, Behene E, Nyarko EO, Kyei NNA, Larbi JA, Lawson BWL, Addo KK, Newman MJ, Duplessis CA, Adams N, Unemo M, Letizia AG. (2019) **Antimicrobial resistance (AMR) and molecular characterization of *Neisseria gonorrhoeae* in Ghana, 2012-2015**. *PLoS One* 14(10):e0223598.

[246] Chen SC, Yin YP, Dai XQ, Unemo M, Chen XS. (2014) **Antimicrobial resistance, genetic resistance determinants for ceftriaxone and molecular epidemiology of *Neisseria gonorrhoeae* isolates in Nanjing, China**. *J Antimicrob Chemother* 69(11):2959–2965.

[247] Jiang FX, Lan Q, Le WJ, Su XH. (2017) **Antimicrobial susceptibility of *Neisseria gonorrhoeae* isolates from Hefei (2014–2015): genetic characteristics of antimicrobial resistance**. *BMC Infect Dis* 17(1):366.

[248] Sood S, Mahajan N, Singh R, Agrawal SK, Shende T, Kapil A, Kar HK, Sharma VK. (2019) **Typing of *Neisseria gonorrhoeae* isolates by phenotypic and genotypic techniques in New Delhi, India**. *J Lab Physicians* 11(1):045–050.

[249] Mahajan N, Sood S, Singh R, Kapil A, Das BK, Sreenivas V, Kar HK, Sharma VK. (2016) **Antimicrobial resistance and *Neisseria gonorrhoeae* multiantigen sequence typing profile of *Neisseria gonorrhoeae* in New Delhi, India**. *Sex Transm Dis* 43(8):506–516.

[250] Hananta IPY, van Dam AP, Schim van der Loeff MF, Dierdorp M, Wind CM, Soebono H, de Vries HJC, Bruisten SM. (2018) **Molecular epidemiology of *Neisseria gonorrhoeae* strains circulating in Indonesia using multi-locus variable number tandem repeat analysis (MLVA) and *Neisseria gonorrhoeae* multi-antigen sequence typing (NG-MAST) techniques**. *BMC Infect Dis* 18(1):7.

[251] Tanaka M, Furuya R, Kobayashi I, Kanesaka I, Ohno A, Katsuse AK. (2019) **Antimicrobial resistance and molecular characterisation of *Neisseria gonorrhoeae* isolates in Fukuoka, Japan, 1996–2016**. *J Glob Antimicrob Resist* 17:3–7.

[252] Jabeen K, Bhawan Mal P, Khan E, Chandio S, Jacobsson S, Unemo M. (2016) **Antimicrobial resistance and *Neisseria gonorrhoeae* multiantigen sequence typing (NG-MAST) genotypes in *N. gonorrhoeae* during 2012–2014 in Karachi, Pakistan**. *BMC Infect Dis* 16:353.

[253] Lee H, Suh YH, Lee S, Kim YK, Han MS, Bae HG, Unemo M, Lee K. (2019) **Emergence and spread of cephalosporin-resistant *Neisseria gonorrhoeae* with mosaic *penA* alleles, South Korea, 2012–2017**. *Emerg Infect Dis* 25(3):416–424.

[254] Olsen B, Pham TL, Golparian D, Johansson E, Tran HK, Unemo M. (2013) **Antimicrobial susceptibility and genetic characteristics of *Neisseria gonorrhoeae* isolates from Vietnam, 2011**. *BMC Infect Dis* 13:40.

[255] O'Reilly LC, Goire N, Fisk RE, Speers DJ. (2015) **Molecular epidemiology of *Neisseria gonorrhoeae* using multi-antigen sequence typing and pulse-field gel electrophoresis in highly endemic Western Australian populations**. *BMC Infect Dis* 15:272.

[256] Lebedzeu F, Golparian D, Titov L, Pankratava N, Glazkova S, Shimanskaya I, Charniakova N, Lukyanau A, Domeika M, Unemo M. (2015) **Antimicrobial susceptibility/resistance and NG-MAST characterisation of *Neisseria gonorrhoeae* in Belarus, Eastern Europe, 2010–2013**. *BMC Infect Dis* 15:29.

[257] Micaëlo M, Goubard A, La Ruche G, Denamur E, Tenaillon O, Cambau E, Jacquier H, Bercot B. (2017) **Molecular epidemiology of penicillinase-producing *Neisseria gonorrhoeae* isolates in France**. *Clin Microbiol Infect* 23(12):968–973.

[258] Brunner A, Nemes-Nikodem E, Jeney C, Szabo D, Marschalko M, Karpati S, Ostorhazi E. (2006) **Emerging azithromycin-resistance among the *Neisseria gonorrhoeae* strains isolated in Hungary**. *Ann Clin Microbiol Antimicrob* 15(1):53.

[259] Mac Aogáin M, Fennelly N, Walsh A, Lynagh Y, Bekaert M, Lawlor B, Walsh P, Kelly B, Rogers TR, Crowley B. (2017) **Fourteen draft genome sequences for the first reported cases of azithromycin-resistant *Neisseria gonorrhoeae* in Irela**nd. *Genome Announc* 5(23):e00403-17.

[260] Ryan L, Golparian D, Fennelly N, Rose L, Walsh P, Lawlor B, Mac Aogáin M, Unemo M, Crowley B. (2018) **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* 37(9):1661–1672.

[261] de Laat MM, Wind CM, Bruisten SM, Dierdorp M, de Vries HJC, Schim van der Loeff MF, van Dam AP. (2019) **Ceftriaxone reduced susceptible *Neisseria gonorrhoeae* in the Netherlands, 2009 to 2017**. *Sex Transm Dis* 46(9):594–601.

[262] Mlynarczyk-Bonikowska B, Serwin AB, Golparian D, Walter de Walthoffen S, Majewski S, Koper M, Malejczyk M, Domeika M, Unemo M. (2014) **Antimicrobial susceptibility/resistance and genetic characteristics of *Neisseria gonorrhoeae* isolates from Poland, 2010-2012**. *BMC Infect Dis* 14:65.

[263] Kubanova A, Kubanov A, Frigo N, Solomka V, Semina V, Vorobyev D, Khairullin R, Unemo M. (2014) **Russian gonococcal antimicrobial susceptibility programme (RU-GASP) – resistance in *Neisseria gonorrhoeae* during 2009–2012 and NG-MAST genotypes in 2011 and 2012**. *BMC Infect Dis* 14:342.

[264] Shaskolskiy B, Dementieva E, Kandinov I, Chestkov A, Kubanov A, Deryabin D, Gryadunov D. (2020) **Genetic diversity of *Neisseria gonorrhoeae* multi-antigen sequence types in Russia and Europe**. *Int J Infect Dis* 93:1–8.

[265] Kubanov A, Vorobyev D, Chestkov A, Leinsoo A, Shaskolskiy B, Dementieva E, Solomka V, Plakhova X, Gryadunov D, Deryabin D. (2016) **Molecular epidemiology of drug-resistant *Neisseria gonorrhoeae* in Russia (Current Status, 2015)**. *BMC Infect Dis* 16:389.

[266] Jeverica S, Golparian D, Matičič M, Potočnik M, Mlakar B, Unemo M. (2014) **Phenotypic and molecular characterization of *Neisseria gonorrhoeae* isolates from Slovenia, 2006-12: rise and fall of the multidrug-resistant NG-MAST genogroup 1407 clone?** *J Antimicrob Chemother* 69(6):1517–1525.

[267] Guerrero-Torres MD, Menéndez MB, Guerras CS, Tello E, Ballesteros J, Clavo P, Puerta T, Vera M, Ayerdi O, Carrio JC, Mozo I, Del Romero J, Vázquez JA, Abad R. (2019) **Epidemiology, molecular characterisation and antimicrobial susceptibility of *Neisseria gonorrhoeae* isolates in Madrid, Spain, in 2016**. *Epidemiol Infect* 147:e274.

[268] Cobo F, Cabezas-Fernández MT, Avivar C. (2019) **Typing and antimicrobial susceptibility of 134 *Neisseria gonorrhoeae* strains from Southern Spain**. *Rev Esp Quimioter* 32(2):114–120.

[269] Endimiani A, Guilarte YN, Tinguely R, Hirzberger L, Selvini S, Lupo A, Hauser C, Furrer H. (2014) **Characterization of *Neisseria gonorrhoeae* isolates detected in Switzerland (1998–2012): emergence of multidrug-resistant clones less susceptible to cephalosporins**. *BMC Infect Dis* 14:106.

[270] Didelot X, Dordel J, Whittles LK, Collins C, Bilek N, Bishop CJ, White PJ, Aanensen DM, Parkhill J, Bentley SD, Spratt BG, Harris SR. (2016) **Genomic analysis and comparison of two gonorrhea outbreaks**. *mBio* 7(3):e00525-16.

[271] Fifer H, Cole M, Hughes G, Padfield S, Smolarchuk C, Woodford N, Wensley A, Mustafa N, Schaefer U, Myers R, Templeton K, Shepherd J, Underwood A. (2018) **Sustained transmission of high-level azithromycin-resistant *Neisseria gonorrhoeae* in England: an observational study**. *Lancet Infect Dis* 18(5):573–581.

[272] De Silva D, Peters J, Cole K, Cole MJ, Cresswell F, Dean G, Dave J, Thomas DR, Foster K, Waldram A, Wilson DJ, Didelot X, Grad YH, Crook DW, Peto TE, Walker AS, Paul J, Eyre DW. (2016) **Whole-genome sequencing to determine transmission of *Neisseria gonorrhoeae*: an observational study**. *Lancet Infect Dis* 16(11):1295–1303.

[273] Martin I, Sawatzky P, Allen V, Hoang L, Lefebvre B, Mina N, Wong T, Gilmour M. (2012) **Emergence and characterization of *Neisseria gonorrhoeae* isolates with decreased susceptibilities to ceftriaxone and cefixime in Canada**. *Sex Transm Dis* 39(4):316–323.

[274] Pedersen ML, Poulsen P, Berthelsen L, Nørgaard C, Hoffmann S, Jensen JS. (2016) **Rapid spread of *Neisseria gonorrhoeae* ciprofloxacin resistance due to a newly introduced resistant strain in Nuuk, Greenland, 2012–2015: a community-based prospective cohort study**. *BMJ Open* 6(8):e011998.

[275] Gianecini R, Romero MLM, Oviedo C, Vacchino M, Galarza P; Gonococcal Antimicrobial Susceptibility Surveillance Programme-Argentina (GASSP-AR) Working Group. (2017) **Emergence and spread of *Neisseria gonorrhoeae* isolates with decreased susceptibility to extended-spectrum cephalosporins in Argentina, 2009 to 2013**. *Sex Transm Dis* 44(6):351–355.

[276] Blank S, Daskalakis DC. (2018) ***Neisseria gonorrhoeae* – rising infection rates, dwindling treatment options**. *N Engl J Med* 379(19):1795–1797.

[277] Mortimer TD, Grad YH. (2019) **Applications of genomics to slow the spread of multidrug resistant *Neisseria gonorrhoeae***. *Ann N Y Acad Sci* 1435(1):93–109.

[278] Alfsnes K, Eldholm V, Olsen AO, Brynildsrud OB, Bohlin J, Steinbakk M, Caugant DA. (2020) **Genomic epidemiology and population structure of *Neisseria gonorrhoeae* in Norway, 2016–2017**. *Microb Genomics* 6(4):e000359.

[279] Osnes MN, Didelot X, de Korne-Elenbaas J, Alfsnes K, Brynildsrud OB, Syversen G, Nilsen Ø, de Blasio BF, Caugant DA, Eldholm V (2020) **The sudden emergence of a *Neisseria gonorrhoeae* strain with reduced susceptibility to extended-spectrum cephalosporins, Norway**. *bioRxiv* 2020.02.07.935825.

[280] Town K, Field N, Harris SR, Sánchez-Busó L, Cole MJ, Pitt R, Fifer H, Mohammed H, Hughes G. (2020) **Phylogenomic analysis of *Neisseria gonorrhoeae* transmission to assess sexual mixing and HIV transmission risk in England: a cross-sectional, observational, whole-genome sequencing study**. *Lancet Infect Dis* 20(4):478–486.

[281] Wind CM, de Vries E, Schim van der Loeff MF, van Rooijen MS, van Dam AP, Demczuk WHB, Martin I, de Vries HJC. (2017) **Decreased azithromycin susceptibility of *Neisseria gonorrhoeae* isolates in patients recently treated with azithromycin**. *Clin Infect Dis* 65(1):37–45.

[282] Golparian D, Harris SR, Sánchez-Busó L, Hoffmann S, Shafer WM, Bentley SD, Jensen JS, Unemo M. (2020) **Genomic evolution of *Neisseria gonorrhoeae* since the preantibiotic era (1928–2013): antimicrobial use/misuse selects for resistance and drives evolution**. *BMC Genomics* 21(1):116.

[283] Williamson DA, Chow EPF, Gorrie CL, Seemann T, Ingle DJ, Higgins N, Easton M, Taiaroa G, Grad YH, Kwong JC, Fairley CK, Chen MY, Howden BP. (2019) **Bridging of *Neisseria gonorrhoeae* lineages across sexual networks in the HIV pre-exposure prophylaxis era**. *Nat Commun* 10(1):3988.

[284] Lee RS, Seemann T, Heffernan H, Kwong JC, Gonçalves da Silva A, Carter GP, Woodhouse R, Dyet KH, Bulach DM, Stinear TP, Howden BP, Williamson DA. (2018) **Genomic epidemiology and antimicrobial resistance of *Neisseria gonorrhoeae* in New Zealand**. *J Antimicrob Chemother* 73(2):353–364.

[285] Lahra MM, Martin I, Demczuk W, Jennison AV, Lee KI, Nakayama SI, Lefebvre B, Longtin J, Ward A, Mulvey MR, Wi T, Ohnishi M, Whiley D. (2018) **Cooperative recognition of internationally disseminated ceftriaxone-resistant *Neisseria gonorrhoeae* strain**. *Emerg Infect Dis* 24(4):735–740.

[286] Kwong JC, Gonçalves da Silva A, Dyet K, Williamson DA, Stinear TP, Howden BP, Seemann T. (2016) **NGMASTER: in silico multi-antigen sequence typing for *Neisseria gonorrhoeae***. *Microb Genomics* 2(8):e000076.

[287] Kwong JC, Chow EPF, Stevens K, Stinear TP, Seemann T, Fairley CK, Chen MY, Howden BP. (2018) **Whole-genome sequencing reveals transmission of gonococcal antibiotic resistance among men who have sex with men: an observational study**. *Sex Transm Infect* 94(2):151–157.

[288] Buckley C, Forde BM, Trembizki E, Lahra MM, Beatson SA, Whiley DM. (2018) **Use of whole genome sequencing to investigate an increase in *Neisseria gonorrhoeae* infection among women in urban areas of Australia**. *Sci Rep* 8(1):1503.

[289] Whiley DM, Jennison A, Pearson J, Lahra MM. (2018) **Genetic characterisation of *Neisseria gonorrhoeae* resistant to both ceftriaxone and azithromycin**. *Lancet Infect Dis* 18(7):717–718.

[290] Grad YH, Harris SR, Kirkcaldy RD, Green AG, Marks DS, Bentley SD, Trees D, Lipsitch M. (2016) **Genomic epidemiology of gonococcal resistance to extended-spectrum cephalosporins, macrolides, and fluoroquinolones in the United States, 2000–2013**. *J Infect Dis* 214(10):1579–1587.

[291] Schmerer MW, Abrams AJ, Seby S, Thomas JC 4th, Cartee J, Lucking S, Vidyaprakash E, Pham CD, Sharpe S, Pettus K, St Cyr SB, Torrone EA, Kersh EN; Antimicrobial-Resistant *Neisseria gonorrhoeae* Working Group, Gernert KM; Antimicrobial-Resistant *Neisseria gonorrhoeae* Working Group. (2020) **Genomic characterization of *Neisseria gonorrhoeae* strains from 2016 U.S. sentinel surveillance displaying reduced susceptibility to azithromycin**. *Antimicrob Agents Chemother* 64(5):e02420-19.

[292] Ezewudo MN, Joseph SJ, Castillo-Ramirez S, Dean D, Del Rio C, Didelot X, Dillon JA, Selden RF, Shafer WM, Turingan RS, Unemo M, Read TD. (2015) **Population structure of *Neisseria gonorrhoeae* based on whole genome data and its relationship with antibiotic resistance**. *PeerJ* 3:e806.

[293] Thomas JC, Seby S, Abrams AJ, Cartee J, Lucking S, Vidyaprakash E, Schmerer M, Pham CD, Hong J, Torrone E, Cyr SS, Shafer WM, Bernstein K, Kersh EN, Gernert KM; Antimicrobial-Resistant *Neisseria gonorrhoeae* Working Group. (2019) **Evidence of recent genomic evolution in gonococcal strains with decreased susceptibility to cephalosporins or azithromycin in the United States, 2014–2016**. *J Infect Dis* 220(2):294–305.

[294] Mortimer TD, Pathela P, Crawley A, Rakeman JL, Lin Y, Harris SR, Blank S, Schillinger JA, Grad YH. (2020) **The distribution and spread of susceptible and resistant *Neisseria gonorrhoeae* across demographic groups in a major metropolitan center**. *Clin Infect Dis* 23:ciaa1229.

[295] Lan PT, Golparian D, Ringlander J, Van Hung L, Van Thuong N, Unemo M. (2020) **Genomic analysis and antimicrobial resistance of *Neisseria gonorrhoeae* isolates from Vietnam in 2011 and 2015–16**. *J Antimicrob Chemother* 75(6):1432–1438.

[296] Yahara K, Nakayama SI, Shimuta K, Lee KI, Morita M, Kawahata T, Kuroki T, Watanabe Y, Ohya H, Yasuda M, Deguchi T, Didelot X, Ohnishi M. (2018) **Genomic surveillance of *Neisseria gonorrhoeae* to investigate the distribution and evolution of antimicrobial-resistance determinants and lineages**. *Microb Genomics* 4(8):e000205.

[297] Wang H, Wang Y, Yong G, Li X, Yu L, Ma S, Luo T. (2020) **Emergence and genomic characterization of the ceftriaxone-resistant *Neisseria gonorrhoeae* FC428 clone in Chengdu, China**. *J Antimicrob Chemother* 75(9):2495-2498.

[298] Peng JP, Yin YP, Chen SC, Yang J, Dai XQ, Zheng HP, Gu WM, Zhu BY, Yong G, Zhong N, Hu LH, Cao WL, Zheng ZJ, Wang F, Zhi Q, Zhang C, Xiu LS, Liu B, Dong J, Sun LL, Zhu YF, Chen XS, Jin Q. (2019) **A whole-genome sequencing analysis of *Neisseria gonorrhoeae* isolates in China: an observational study**. *EClinicalMedicine.* 14;7:47-54.

[299] Cehovin A, Harrison OB, Lewis SB, Ward PN, Ngetsa C, Graham SM, Sanders EJ, Maiden MCJ, Tang CM. (2018) **Identification of novel *Neisseria gonorrhoeae* lineages harboring resistance plasmids in coastal Kenya**. *J Infect Dis* 218(5):801–808.

[300] Sánchez-Busó L, Harris SR. (2019) **Using genomics to understand antimicrobial resistance and transmission in *Neisseria gonorrhoeae***. *Microb Genomics* 5(2):e000239.

[301] Jolley KA, Maiden MCJ. (2010) **BIGSdb: Scalable analysis of bacterial genome variation at the population level**. *BMC Bioinformatics* 11:595.

[302] Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, Harris SR. (2017) **ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads**. *Microb Genomics* 3(10):e000131.

[303] Demczuk W, Sidhu S, Unemo M, Whiley DM, Allen VG, Dillon JR, Cole M, Seah C, Trembizki E, Trees DL, Kersh EN, Abrams AJ, de Vries HJC, van Dam AP, Medina I, Bharat A, Mulvey MR, Van Domselaar G, Martin I. (2017) ***Neisseria gonorrhoeae* Sequence Typing for Antimicrobial Resistance, a novel antimicrobial resistance multilocus typing scheme for tracking global dissemination of *N. gonorrhoeae* Strains**. *J Clin Microbiol* 55(5):1454–1468.

[304] Eyre DW, Golparian D, Unemo M. (2019) **Prediction of minimum inhibitory concentrations of antimicrobials for *Neisseria gonorrhoeae* using whole-genome sequencing**. In: *Methods in Molecular Biology*. Humana Press Inc. pp. 59–76.

[305] Demczuk W, Martin I, Sawatzky P, Allen V, Lefebvre B, Hoang L, Naidu P, Minion J, VanCaeseele P, Haldane D, Eyre DW, Mulvey MR. (2020) **Equations to predict antimicrobial MICs in *Neisseria gonorrhoeae* using molecular antimicrobial resistance determinants**. *Antimicrob Agents Chemother* 64(3):e02005-19.

[306] Břinda K, Callendrello A, Ma KC, MacFadden DR, Charalampous T, Lee RS, Cowley L, Wadsworth CB, Grad YH, Kucherov G, O'Grady J, Baym M, Hanage WP. (2020) **Rapid inference of antibiotic resistance and susceptibility by genomic neighbour typing**. *Nat Microbiol* 5(3):455–464.

[307] Schürch AC, Arredondo-Alonso S, Willems RJL, Goering RV. (2018) **Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches**. *Clin Microbiol Infect* 24(4):350–354.

[308] Maiden MC, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND. (2013) **MLST revisited: the gene-by-gene approach to bacterial genomics**. *Nat Rev Microbiol* 11(10):728–736.

[309] Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, Gilpin B, Smith AM, Man Kam K, Perez E, Trees E, Kubota K, Takkinen J, Nielsen EM, Carleton H; FWD-NEXT Expert Panel. (2017) **PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance**. *Eurosurveillance* 22(23):30544.

[310] Llarena A-K, Ribeiro-Gonçalves BF, Nuno Silva D, Halkilahti J, Machado MP, Da Silva MS, Jaakkonen A, Isidro J, Hämäläinen C, Joenperä J, Borges V, Viera L, Gomes JP, Correia C, Lunden J, Laukkanen-Ninios R, Fredriksson-Ahomaa M, Bikandi J, San Millan R, Martinez-Ballesteros

I,Laorden L, Mäesaar M, Grantiņa-Ieviņa L, Hilbert F, Garaizar J, Oleastro M, Nevas M, Salmenlinna S, Hakkinen M, Carriço JA, Rossi M. (2018) **INNUENDO: A cross-sectoral platform for the integration of genomics in the surveillance of food-borne pathogens**. *EFSA supporting publication* 15(11):EN-1498. pp. 142

[311] Harrison OB, Cehovin A, Skett J, Jolley KA, Massari P, Genco CA, Tang CM, Maiden MCJ. (2020) ***Neisseria gonorrhoeae* population genomics: use of the gonococcal core genome to improve surveillance of antimicrobial resistance**. *J Infect Dis*. jiaa002.

[312] European Centre for Disease Prevention and Control. (2017) **Gonococcal antimicrobial susceptibility surveillance in Europe - 2015**. Stockholm. [Accessed: 28 October 2020] Available at: https://www.ecdc.europa.eu/sites/default/files/documents/gonococcal-antimicrobial-susceptibility-surveillance-Europe-2015.pdf

[313] Boiko I, Golparian D, Krynytska I, Bezkorovaina H, Frankenberg A, Onuchyna M, Jacobsson S, Unemo M. (2019) **Antimicrobial susceptibility of *Neisseria gonorrhoeae* isolates and treatment of gonorrhoea patients in Ternopil and Dnipropetrovsk regions of Ukraine, 2013–2018**. *APMIS* 127(7):503–509.

[314] Brown LB, Krysiak R, Kamanga G, Mapanje C, Kanyamula H, Banda B, Mhango C, Hoffman M, Kamwendo D, Hobbs M, Hosseinipour MC, Martinson F, Cohen MS, Hoffman IF. (2010) ***Neisseria gonorrhoeae* Antimicrobial Susceptibility in Lilongwe, Malawi, 2007**. *Sex Transm Dis* 37(3):169–172.

[315] Florindo C, Pereira R, Boura M, Nunes B, Paulino A, Gomes JP, Borrego MJ. (2010) **Genotypes and antimicrobial-resistant phenotypes of *Neisseria gonorrhoeae* in Portugal (2004-2009)**. *Sex Transm Infect* 86(6):449–453.

[316] Francisco AP, Vaz C, Monteiro PT, Melo-Cristino J, Ramirez M, Carriço JA. (2012) **PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods**. *BMC Bioinformatics* 13:87.

[317] Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, Feil EJ, Holden MTG, Yeats CA, Grundmann H, Spratt BG, Aanensen DM. (2016) **Microreact: visualizing and sharing data for genomic epidemiology and phylogeography**. *Microb genomics* 2(11):e000093.

[318] The European Committee on Antimicrobial Susceptibility Testing. (2009) **Breakpoint tables for interpretation of MICs and zone diameters. Version 1.0**. [Accessed: 28 October 2020] Available at: http://www.eucast.org

[319] Yin YP, Han Y, Dai XQ, Zheng HP, Chen SC, Zhu BY, Yong G, Zhong N, Hu LH, Cao WL, Zheng ZJ, Wang F, Zhi Q, Zhu XY, Chen XS. (2018) **Susceptibility of *Neisseria gonorrhoeae* to azithromycin and ceftriaxone in China: A retrospective study of national surveillance data from 2013 to 2016**. *PLoS Med* 15(2):e1002499

[320] Kularatne R, Maseko V, Gumede L, Kufa T. (2018) **Trends in *Neisseria gonorrhoeae* antimicrobial resistance over a ten-year surveillance period, Johannesburg, South Africa, 2008–2017**. *Antibiotics (Basel) 7(3):58*.

[321] Mann LM, Kirkcaldy RD, Papp JR, Torrone EA. (2018) **Susceptibility of *Neisseria gonorrhoeae* to gentamicin-Gonococcal Isolate Surveillance Project, 2015-2016**. *Sex Transm Dis* 45(2):96–98.

[322] Ross JDC, Brittain C, Cole M, Dewsnap C, Harding J, Hepburn T, Jackson L, Keogh M, Lawrence T, Montgomery AA, Roberts TE, Sprange K, Tan W, Thandi S, White J, Wilson J, Duley L; G-ToG

trial team. (2019) **Gentamicin compared with ceftriaxone for the treatment of gonorrhoea (G-ToG): a randomised non-inferiority trial**. *Lancet* 393(10205):2511–2520.

[323] Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, Stevens G, Gottlieb S, Kiarie J, Temmerman M. (2015) **Global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting**. *PLoS One* 10(12):e0143304.

[324] Pinto M, Matias R, Rodrigues JC, Duarte S, Vieira L, Gonçalves I, Gonçalves MJ, Ramos MH, Gomes JP, Borrego MJ. (2020) **Cephalosporin-Resistant *Neisseria gonorrhoeae* Isolated in Portugal, 2019**. *Sex Transm Dis* 47(11):e54–e56.

[325] O'Rourke M, Stevens E. (1993) **Genetic structure of *Neisseria gonorrhoeae* populations: A non-clonal pathogen**. *J Gen Microbiol* 139(11):2603–2611.

[326] Pinto M, Rodrigues JC, Matias R, Água-Doce I, Cordeiro D, Correia C, Gomes JP, Borrego MJ; PTGonoNet. (2020) **Fifteen years of a nationwide culture collection of *Neisseria gonorrhoeae* antimicrobial resistance in Portugal**. *Eur J Clin Microbiol Infect Dis* 39(9):1761–1770.

[327] Bolger AM, Lohse M, Usadel B. (2014) **Trimmomatic: a flexible trimmer for Illumina sequence data**. *Bioinformatics* 30(15):2114–2120.

[328] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. (2012) **SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing**. *J Comput Biol* 19(5):455–477.

[329] Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. (2014) **Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement**. *PLoS One* 9(11):e112963.

[330] Bennett JS, Watkins ER, Jolley KA, Harrison OB, Maiden MCJ. (2014) **Identifying *Neisseria* species by use of the 50S Ribosomal protein L6 (*rplF*) gene**. *J Clin Microbiol* 52(5):1375–1381.

[331] Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J, Santos S, Ramirez M, Carriço JA.2018) **chewBBACA: A complete suite for gene-by-gene schema creation and strain identification**. Microb Genomics 4(3):e000166.

[332] Francisco AP, Bugalho M, Ramirez M, Carriço JA. (2009) **Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach**. *BMC Bioinformatics* 10:152.

[333] Ribeiro-Gonçalves B, Francisco AP, Vaz C, Ramirez M, Carriço JA. (2016) **PHYLOViZ Online: web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees**. *Nucleic Acids Res* 44(W1):W246–W251.

[334] Carriço JA, Silva-Costa C, Melo-Cristino J, Pinto FR, de Lencastre H, Almeida JS, Ramirez M. (2006) **Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*.** *J Clin Microbiol* 44(7):2524–2532.

[335] Severiano A, Pinto FR, Ramirez M, Carriço JA. (2011) **Adjusted Wallace coefficient as a measure of congruence between typing methods**. *J Clin Microbiol* 49(11):3997–4000.

[336] Barker DO, Carriço J, Kruczkiewicz P, Palma F, Rossi M, Taboada EN (2018) **Rapid identification of stable clusters in bacterial populations using the adjusted Wallace coefficient**. *bioRxiv* 299347.

[337] Nascimento M, Sousa A, Ramirez M, Francisco AP, Carriço JA, Vaz C. (2017) **PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods**. *Bioinformatics* 33(1):128–129.

[338] Zhou Z, Alikhan NF, Sergeant MJ, Luhmann N, Vaz C, Francisco AP, Carriço JA, Achtman M. (2018) **GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens**. *Genome Res* 28(9):1395–1404.

[339] Deng X, Klausner JD. (2020) **Six *penA* codons accurately and reliably predict cefixime-decreased susceptibility in *Neisseria gonorrhoeae***. *J Infect Dis* 221(5):851–852.

[340] Belkacem A, Jacquier H, Goubard A, Mougari F, La Ruche G, Patey O, Micaëlo M, Semaille C, Cambau E, Bercot B. (2016) **Molecular epidemiology and mechanisms of resistance of azithromycin-resistant *Neisseria gonorrhoeae* isolated in France during 2013–14**. *J Antimicrob Chemother* 71(9):2471–2478.

[341] Calado J, Castro R, Lopes Â, Campos MJ, Rocha M, Pereira F. (2019) **Antimicrobial resistance and molecular characteristics of *Neisseria gonorrhoeae* isolates from men who have sex with men**. *Int J Infect Dis* 79:116–122.

[342] Macedo R, Pinto M, Borges V, Nunes A, Oliveira O, Portugal I, Duarte R, Gomes JP. (2009) **Evaluation of a gene-by-gene approach for prospective whole-genome sequencing-based surveillance of multidrug resistant *Mycobacterium tuberculosis***. Tuberculosis 115:81–88.

[343] Seemann T. (2014) **Prokka: rapid prokaryotic genome annotation**. *Bioinformatics* 30(14):2068–2069.

[344] Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. (2015) **Roary: rapid large-scale prokaryote pan genome analysis**. *Bioinformatics* 31(22):3691–3693.

[345] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. (2009) **BLAST+: architecture and applications**. *BMC Bioinformatics* 10:421.

[346] Tatusov RL, Galperin MY, Natale DA, Koonin EV. (2000) **The COG database: a tool for genome-scale analysis of protein functions and evolution**. *Nucleic Acids Res* 28(1):33–36.

[347] Leimbach A. (2016) **bac-genomics-scripts: Bovine *E. coli* mastitis comparative genomics edition (Version bovine_ecoli_mastitis)**. *Zenodo*. http://doi.org/10.5281/zenodo.215824.

[348] Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. (2016) **Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary**. *Genome Biol* 17(1):238.

[349] Pinto M, González-Díaz A, Machado MP, Duarte S, Vieira L, Carriço JA, Marti S, Bajanca-Lavado MP, Gomes JP. (2019) **Insights into the population structure and pan-genome of *Haemophilus influenzae***. *Infect Genet Evol* 67:126–135.

[350] Schubert B, Maddamsetti R, Nyman J, Farhat MR, Marks DS. (2019) **Genome-wide discovery of epistatic loci affecting antibiotic resistance in *Neisseria gonorrhoeae* using evolutionary couplings**. *Nat Microbiol* 4(2):328–338.

[351] Neto A, Sevilha J, Seabra D, Oliveira I, Dos Santos RP, Andrade A, Pinho P, da Costa PM, Viana M, Pinto P. (2020) **Acute aortic regurgitation due to endocarditis caused by disseminated gonococcal infection – a case report**. *Sex Transm Dis* Epub ahead of print. **DOI**: 10.1097/OLQ.0000000000001251.

[352] Nettleton WD, Kent JB, Macomber K, Brandt MG, Jones K, Ridpath AD, Raphael BH, Wells EV. (2020) **Notes from the Field: Ongoing Cluster of Highly Related Disseminated Gonococcal Infections — Southwest Michigan, 2019**. *MMWR Morb Mortal Wkly Rep* 69(12):353–354.

[353] Mamede R, Vila-Cerqueira P, Silva M, Carriço JA, Ramirez M. (2020) **Chewie Nomenclature Server (chewie-NS): a deployable nomenclature server for easy sharing of core and whole genome MLST schemas**. *Nucleic Acids Res* gkaa889.

[354] Pinto M, Borges V, Antelo M, Pinheiro M, Nunes A, Azevedo J, Borrego MJ, Mendonça J, Carpinteiro D, Vieira L, Gomes JP. (2016) **Genome-scale analysis of the non-cultivable *Treponema pallidum* reveals extensive within-patient genetic variation**. *Nat Microbiol* 2:16190.

[355] Borges V, Cordeiro D, Salas AI, Lodhia Z, Correia C, Isidro J, Fernandes C, Rodrigues AM, Azevedo J, Alves J, Roxo J, Rocha M, Côrte-Real R, Vieira L, Borrego MJ, Gomes JP. (2019) ***Chlamydia trachomatis*: when the virulence-associated genome backbone imports a prevalence-associated major antigen signature**. *Microb Genomics* 5(11):e000313.

[356] Macedo R. (2019) **Tuberculosis: new era for diagnosis and surveillance using whole-genome sequencing-based approaches**. (PhD Thesis) *Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa*.

[357] Gomes JP, Borrego MJ, Atik B, Santo I, Azevedo J, Brito de Sá A, Nogueira P, Dean D. (2006) **Correlating Chlamydia trachomatis infectious load with urogenital ecological success and disease pathogenesis**. *Microbes Infect* 8(1):16–26.

[358] Wood DE, Salzberg SL. (2014) **Kraken: Ultrafast metagenomic sequence classification using exact alignments**. *Genome Biol* 15(3):R46.

[359] Langmead B, Salzberg SL. (2012) **Fast gapped-read alignment with Bowtie 2**. *Nat Methods* 9(4):357–359.

[360] Fifer H, Natarajan U, Jones L, Alexander S, Hughes G, Golparian D, Unemo M. (2016) **Failure of dual antimicrobial therapy in treatment of gonorrhea**. *N Engl J Med* 374(25):2504–2506.

[361] Dudareva-Vizule S, Haar K, Sailer A, Wisplinghoff H, Wisplinghoff F, Marcus U; PARIS study group. (2014) **Prevalence of pharyngeal and rectal *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections among men who have sex with men in Germany**. *Sex Transm Infect* 90(1):46–51.

[362] Kent CK, Chaw JK, Wong W, Liska S, Gibson S, Hubbard G, Klausner JD. (2005) **Prevalence of rectal, urethral, and pharyngeal chlamydia and gonorrhea detected in 2 clinical settings among men who have sex with men: San Francisco, California, 2003**. *Clin Infect Dis* 41(1):67–74.

[363] Weston EJ, Kirkcaldy RD, Stenger M, Llata E, Hoots B, Torrone EA. (2018) **Narrative review: Assessment of *Neisseria gonorrhoeae* infections among men who have sex with men in national and sentinel surveillance systems in the United States**. *Sex Transm Dis* 45(4):243–249.

[364] Allen VG, Mitterni L, Seah C, Rebbapragada A, Martin IE, Lee C, Siebert H, Towns L, Melano RG, Low DE. (2013) ***Neisseria gonorrhoeae* treatment failure and susceptibility to cefixime in Toronto, Canada**. *JAMA* 309(2):163–170.

[365] Spratt B, Bowler L, Zhang Q-Y, Zhou J, Smith J. (1992) **Role of interspecies transfer of chromosomal genes in the evolution of penicillin resistance in pathogenic and commensal *Neisseria* species**. *J Mol Evol* 34(2):115–125.

[366] Wadsworth CB, Arnold BJ, Sater MRA, Grad YH. (2018) **Azithromycin resistance through interspecific acquisition of an epistasis-dependent efflux pump component and transcriptional regulator in *Neisseria gonorrhoeae***. *mBio* (4):e01419-18.

[367] Lodhia Z. (2017) **Contribution for the knowledge on curable sexually transmitted infections with special emphasis on *Chlamydia trachomatis***. (Masters Thesis) *Faculdade de Ciências, Universidade de Lisboa*.

Supplementary Material

Supplementary Table S2.1. | Associated patient data, molecular typing and antibiotic susceptibility data of all viable *Neisseria gonorrhoeae* isolates collected from 2003 up to 2018 in Portugal on behalf of the PTGonoNET.

MS Excel file available at: https://ndownloader.figshare.com/files/25370801

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S2_1_/13182038

Supplementary Table S3.1. | Metadata, genome assembly statistics and antibiotic resistance phenotype and genotype data of the Portuguese *Neisseria gonorrhoeae* isolates enrolled in the present study.

MS Excel file available at: https://ndownloader.figshare.com/files/25370864

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S3_1_/13182050

Supplementary Table S3.2. | Metadata, genome assembly statistics and antibiotic resistance genotype data of the publicly available European *Neisseria gonorrhoeae* isolates enrolled in the present study.

MS Excel file available at: https://ndownloader.figshare.com/files/25371116

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S3_2_/13182068

Supplementary Table S3.3. | List of loci removed from the pubMLST *Neisseria gonorrhoeae* cgMLST v1 scheme, after preparation for chewBBACA.

MS Excel file available at: https://ndownloader.figshare.com/files/25371287

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S3_3_/13182071

Supplementary Table S3.4. | List of loci enrolled in the Maximum Shared cgMLST scheme used for clustering analysis.

MS Excel file available at: https://ndownloader.figshare.com/files/25371347

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S3_4_/13182074

Supplementary Table S4.1. | *Neisseria gonorrhoeae* pan-genome composition based on 3971 genomes.

MS Excel file available at: https://ndownloader.figshare.com/files/25371374

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S4_1_/13182080

Supplementary Table S4.2. | Accessory genes associated with high- and low-level WGS-based genogroups, presenting simultaneously more than 80% sensitivity and specificity.

MS Excel file available at: https://ndownloader.figshare.com/files/25371485

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S4_2_/13182110

Supplementary Table S4.3. | Distribution of accessory genes presenting simultaneously more than 80% sensitivity and specificity by high- and low-level WGS-genogroups.

MS Excel file available at: https://ndownloader.figshare.com/files/25371491

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S4_3_/13182116

Supplementary Table S6.1. | Metadata, genome assembly statistics and antibiotic resistance phenotype and genotype data of the novel sequenced Portuguese *Neisseria gonorrhoeae* isolates enrolled in the present study.

MS Excel file available at: https://ndownloader.figshare.com/files/25371494

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S6_1_/13182119

Supplementary Table S6.2. | Associated patient data and antibiotic susceptibility data of all viable *Neisseria gonorrhoeae* isolates collected in 2019 in Portugal on behalf of the PTGonoNET.

MS Excel file available at: https://ndownloader.figshare.com/files/25371518

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S6_2_/13182125

Supplementary Table S6.3. | Composition of genetic clusters enrolling Portugusese isolates.

MS Excel file available at: https://ndownloader.figshare.com/files/25373450

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S6_3_/13182353

Supplementary Table S7.1. | Summary of the taxonomic classification of all raw reads for all clinical samples.

MS Excel file available at: https://ndownloader.figshare.com/files/25373453

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S7_1_/13182356

Supplementary Table S7.2. | Summary of the taxonomic classification of all recovered reads (after mapping) for all clinical samples.

MS Excel file available at: https://ndownloader.figshare.com/files/25373456

Online version available at: https://figshare.com/articles/dataset/Supplementary_Table_S7_2_/13182359