ASSESSMENT OF POPULATION GENETIC STRUCTURE AND GENETIC DIVERSITY OF MULTIDRUG-RESISTANT *N. GONORRHOEAE* ISOLATES FROM KENYA USING GENOME-WIDE SNP ANALYSIS

HENRY KIEMA MUSEE

SB02/PU/40217/21

A Research Proposal Submitted in Partial Fulfilment of the Requirements for the Award of Bachelor of Science Degree in Biochemistry in the Department of Biochemistry and Biotechnology, Pwani University

2025

DECLARATION

This proposal is my original work and has not been presented for degree in any other university.

Signature …… Date…11th Feb, 2025……

Henry Kiema Musee

SB02/PU/40217/21

Supervisor's recommendation

We confirm that this thesis has been submitted with our approval.

Signature …… Date…13 February, 2025……

Prof. Wilton Mbinda

Department of Biochemistry and Biotechnology

Pwani University

DEDICATION

I dedicate this proposal to my entire family for the financial support, encouragement and emotional support during my entire undergraduate study.

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my supervisor, Prof Wilton Mbinda, for his invaluable guidance and support throughout this process. His expertise has been valuable in shaping my understanding towards bioinformatics.

I am also sincerely grateful to PUBReC for giving me the opportunity to study in the bioinformatics lab, providing me with resources and environment necessary to enhance my skills and knowledge

Special thanks goes to PUBReC masters students, whose collaboration and assistance were crucial in helping me navigate the complexities of bioinformatics. Their willingness to share knowledge and provide support has been a source of great inspiration.

Thank you all for your contributions to my academic and personal growth.

ABSTRACT

Gonorrhea, a globally prevalent sexually transmitted infection caused by *Neisseria gonorrhoeae*, has become an urgent public health concern due to the emergence of multidrug-resistant (MDR) strains. These strains threaten the efficacy of standard treatments and underscore the critical need for innovative control strategies and enhanced surveillance mechanisms. In Kenya, resistance characterization has primarily relied on phenotypic profiling through Minimum Inhibitory Concentrations (MICs), offering limited insights into the genetic mechanisms underlying resistance. This knowledge gap hampers efforts to develop targeted therapies and effective preventive measures. This study aims to assess the population genetic structure and genetic diversity of MDR *N. gonorrhoeae* isolates in Kenya using genome-wide Single Nucleotide Polymorphism (SNP) analysis. By integrating advanced bioinformatics approaches, the research seeks to elucidate the genetic underpinnings of antibiotic resistance and their correlation with epidemiological factors. Specifically, the study will: (1) analyze the genetic structure of *N. gonorrhoeae* isolates using Principal Component Analysis (PCA), admixture analysis, and Neighbor-Net analysis to uncover clustering, ancestral mixing, and gene flow patterns; and (2) evaluate genetic diversity and linkage disequilibrium patterns through metrics such as nucleotide diversity and heterozygosity rates. The research will leverage high-quality genomic data from previously sequenced *N. gonorrhoeae* isolates collected across various regions in Kenya. Sophisticated bioinformatics tools will be employed to process SNP datasets, ensuring robust and reliable analyses. The expected outcomes include identifying distinct genetic clusters of *N. gonorrhoeae* isolates with unique resistance profiles and revealing significant genetic differentiation among isolates from different geographic regions. These findings will enhance understanding of the evolutionary dynamics of antibiotic resistance and inform public health strategies for gonorrhea management. This study's significance lies in its potential to advance the understanding of the genetic factors driving antibiotic resistance, contributing to the formulation of evidence-based therapeutic strategies and public health interventions. By addressing a critical research gap, this work will support global efforts to combat MDR *N. gonorrhoeae* and improve disease control and prevention in Kenya and beyond.

The proposed study seeks to fill this critical research gap by employing genome-wide Single Nucleotide Polymorphism (SNP) analysis to thoroughly assess the genetic diversity and population structure of multidrug-resistant *Neisseria gonorrhoeae* isolates in Kenya. By focusing on the genomic aspects, this research aims to elucidate the genetic foundations that contribute to antibiotic resistance and how these genetic variations correlate with epidemiological factors.The objectives of this research are twofold: First, to analyze the genetic structure of these isolates through advanced bioinformatics tools such as Principal Component Analysis (PCA), admixture analysis, and Neighbor-Net analysis, which will reveal the clustering, ancestral mixing, and gene flow patterns among the populations. Second, to evaluate the genetic diversity and patterns of linkage disequilibrium among the isolates using metrics such as nucleotide diversity and heterozygosity rates. This dual approach will provide a comprehensive overview of the genetic landscape of *Neisseria gonorrhoeae* in Kenya, highlighting potential regional differences in genetic makeup and identifying evolutionary pressures that may influence resistance patterns.The methodology for this study involves the bioinformatics analysis of previously sequenced genomic data. The SNP datasets will be derived from high-quality genomic sequences obtained from isolates collected from various regions in Kenya. These sequences have

already been processed and quality-assured, ensuring reliable input data for subsequent analyses. The bioinformatics phase will utilize sophisticated software tools for data processing and analysis, ensuring rigorous evaluation and interpretation of the genetic data.Expected results from this study include the identification of distinct genetic clusters of *Neisseria gonorrhoeae* isolates in Kenya, which may exhibit unique resistance profiles. Additionally, this research is anticipated to reveal significant genetic differentiation among isolates from different geographic regions, which could be instrumental in shaping future public health policies and treatment protocols. The findings could also provide insights into the evolutionary dynamics of antibiotic resistance in *Neisseria gonorrhoeae*, contributing to global efforts in combating this resilient pathogen.The significance of this study lies in its potential to advance the understanding of genetic factors in antibiotic resistance, which will aid in the formulation of more effective therapeutic strategies and public health interventions. By providing detailed genetic insights, this research will contribute to the global body of knowledge on gonorrhea management and resistance mitigation, ultimately leading to improved disease control and prevention measures.

TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

BUSCO	Benchmarking Universal Single-Copy Orthologs
DGI	Disseminated Gonococcal Infection
DNA	Deoxyrinonucleic Acid
GATK	Genome Analysis Tool-Kit
GW-SNP	Genome-Wide SNP
MDR-NG	Multidrug-resistant *Neisseria gonorrhoeae*
MICs	Minimum Inhibitory Concentrations
NCBI	National Center for Biotechnology Information
PCA	Principle Component Analysis
PID	Pelvic Inflammatory Disease
PROKKA	Rapid Prokaryotic genome Annotation
QUAST	Quality Assessment Tool for Genome Assemblies
SNPs	Single Nucleotide polymorphisms
SPAdes	St. Petersburg genome Assembler (de novo)
STIs	Sexually Transmitted Infections
USAMRD-A	US Army Medical Research Directorate-Africa

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background

Neisseria gonorrhoeae is a bacterium responsible for gonorrhea, a common sexually transmitted infection that affects both men and women. The bacterium primarily targets the genitals, where it thrives in the favorable environment of the urethra and endocervix. If untreated, gonorrhea can lead to severe health complications, including disseminated gonococcal infection (DGI), pelvic inflammatory disease (PID), and epididymis in men, as well as complications during childbirth (Quillin & Seifert, 2018). The bacteria has been building up resistance to many of the antibiotics which were originally used to combat it, in the last 70 years. The medications which were used, including penicillin, aminoglycosides, macrolides and tetracycline could cure the gonorrhoea disease in relatively small amount, which could have high efficacy. Mutations occurred drastically in the 1980s, where the bacteria was reported to have developed resistance against tetracycline medications, which resulted to the removal of tetracycline as treatment options. The medication which were now being used after the withdrawal of the tetracycline was ciprofloxacin and fluoroquinolones, which were effective enough to cure gonorrhoea (Belland et al, 1994). It was not until 2007, when the fluoroquinolones could not be an effective treatment option for the disease, due to the development of resistance by the bacteria, with also high cases of the bacteria being resistant to the ciprofloxacin (Del et al, 2007).

The resistance to ciprofloxacin and fluoroquinolones led to the use of the third-generation cephalosporins (cifixime and cefriaxone), as they could treat the disease with lower concentrations. They were then proposed to be used as the first line of defense, as these drugs could confidently be used to effectively combat this bacteria. Despite being effectively used in this case, the bacteria has still mounted resistance towards these drugs, and the resistance strains of this bacteria have been highly transmitted across the world. In Kenya, the fear of this resistance strains to most of the antibiotics available for its treatment is justified by a study done on gyrA and parC mutations in fluoroquinolone-resistant Neisseria gonorrhoeae isolates indicated increased MICs in ciprofloxacin, as a number of samples (20) were showing resistance of ciprofaoxacin with only 2 being susceptible, all being resistant to penicillin and 16 being resistant to tetracycline (Kivata et al, 2019). This is a clear indication

of the global spread of the antimicrobial resistance strains, and it posses a challenge in the health sector given the history of how fast the bacteria is capable of mutating.

This emergence of antibiotic-resistant Neisseria gonorrhoeae strains, particularly those resistant to penicillin and tetracycline, presents a significant challenge to public health (Ndowa & Lusti-Narasimhan, 2012). The prevalence of multidrug-resistant Neisseria gonorrhoeae is particularly high among men who have sex with men, young adults aged 15-24 years, and other populations with high-risk behaviors.

In this study, we will analyze genome sequences of 36 *Neisseria gonorrhoeae* isolates from Kenya, which exhibit plasmid-mediated resistance to penicillin and tetracycline. This study aims to provide insights into the genetic diversity, and evolutionary pathways, and determine whether geolocation plays a role in the development of resistance of this bacterium, thereby contributing to the development of targeted treatment strategies.

1.2 Problem statement

There has been limited research which have been conducted in Kenya in the surveillance of multi-drug-resistant *Neisseria gonorrhoeae,* with most researchers focusing on the phenotypic characterization through determination of minimum inhibitory concentration (MICs) and few kinds of research carried out on genome-wide sequencing techniques to determine the genes and antimicrobial determinants in the resistant strains of *Neisseria gonorrhoeae*. There is still limited knowledge on the genetic population structure and transmissibility of the Kenyan gonococcal strains. STI surveillance at the US Army Medical Research Directorate-Africa (USAMRD-A) which is hosted by Kenya Medical Research Institute has shown the reduction in the susceptibility of some Neisseria gonorrhoeae to locally available medications such as penicillin, tetracycline, ciprofloxacin, and azithromycin, using Minimum Inhibitory Concentrations (MICs)(Nacht et al., 2020). The determination of the genetic population structure and the transimissibility of these Kenyan strains is crucial in understanding transmission nature of these strains and also providing complete understanding on the genetic population structure of the Kenyan strains and how they differ from other strains in other geographic regions, for enhancement of target mitigation strategies.

1.3 Justification

Due to the urgent need in the understanding of the genetic makeup of the multidrug-resistant N. gonorrhoeae in Kenyan isolates to combat its spread, the characterization of the genetic population structure and the genetic diversity of this bacteria can lead to a better understanding on how this bacterium is transmitted or spread across the population. Understanding the SNPs will give a guiding criterion on the effective treatment protocol and control of these resistant strains. This study will also help in understanding the genetic mechanisms driving antibiotic resistance in these strains, which will also be helpful in the formulation of a more targeted treatment therapy to combat the threat.

Most of the studies which has been carried out concerning the Kenyan Mulitdrug-resistant N. gonorrhoeae have focused on the determination via the phenotypic methods, such as determination of minimum inhibitory concentrations (MICs) and also few studies focusing on the genetic antimicrobial determinants and characterization of these these antimicrobial determinants. There is still few data on the population structure, diversity and biogeography of the Kenyan MDR-NG. This study therefore seeks to provide a more comprehensive data on the population genetic structure, diversity and biogeography to understand better on the transimissibility of this bacteria. The analysis of genome-wide SNPs will also offer insights into the evolutionary processes and the selection pressures which contribute to the spread of this bacteria in Kenya.

1.4 Research Question

i. What are the genetic population structures of multidrug-resistant Neisseria gonorrhoeae isolates in different geographic regions of Kenya?

ii. To what extent do regional factors influence the genetic diversity and differentiation of multidrug-resistant Neisseria gonorrhoeae isolates in Kenya?

iii. How do recombination and linkage disequilibrium contribute to the spread of resistant genes among Neisseria gonorrhoeae populations in Kenya?

iv. Is there a significant difference in the genetic population structures of multidrug-resistant Neisseria gonorrhoeae isolates between various geographic regions in Kenya?

v. What role do evolutionary factors play in shaping the genetic diversity of multidrug-resistant Neisseria gonorrhoeae isolates in Kenya?

1.5 Objectives

1.5.1 General Objectives

To assess the population genetic structure and genetic diversity of multidrug-resistant *Neisseria gonorrhoeae* isolates in Kenya through genome-wide SNP analysis.

1.5.2 Specific Objectives

i. To analyze the population genetic structure of Kenyan multidrug-resistant Neisseria gonorrhoeae isolates.
ii. To determine genetic diversity and patterns of linkage disequilibrium in the Kenyan multidrug-resistant *N. gonorrhoeae* isolates.
iii. To analyze the selective pressures driving resistance development in Kenyan Neisseria gonorrhoeae isolates using genome-wide SNP data.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 *Neisseria gonorrhoeae*

Neisseria gonorrhoeae is the causative agent of disease gonorrhoeae which is a sexually transmitted disease. It is a delicate pathogen which readily dies outside the body. The disease leads to the dissemination of the genitalia, where it causes urethritis in men and cervicitis in women (Unemo & Shafer, 2014). The bacteria not only exudates in the genitalia but also proliferates within the throat and rectum. The most difficulty to treat is the throat or pharyngeal gonococci because of the nature of the tissue in the region, where there is a outer covering of mucus membrane preventing adequate penetration of drugs. This type of gonorrhoeae infection is prevalent in men who have sex with men, and has less or no symptoms, leading to enhanced growth and transmission(Takahashi et al., 2008).Due to this, the *N. gonorrhoeae* bacteria can lead to a significant development of resistance to drugs because the concentration of drugs reaching its targets is very minimal, leading to buildup of resistance due to adaptation over time.

2.2 *N. gonorrhoeae* antimicrobial resistance.

Reports by WHO on the surveillance of antimicrobial resistance in *N. gonorrhoea* bacteria have indicated that there is a significant increase in the resistance of the bacteria to most antibiotics across different parts of the world like Asia, Europe, America and Africa (Unemo et al., 2021). Kenya has also been on the list of the affected countries in the pandemic, and research has been carried in efforts to combat this bacterial resistance. A study carried out in Kenya has shown an increase in the minimum inhibitory concentrations (MICs) on isolates collected in different clinics in Kenya under US Army Medical Research Directorate – Africa, on surveillance in the STIs (Kivata et al., 2020). The characterization of the genes coding for different resistance has been shown, where the most resistant gene are the blaTEM-1 gene, which is a gene encoding for resistance against β-lactams.

2.3 Mechanisms of antibiotic resistance

There are different mechanisms in which this bacterium develops resistance towards the administered antibiotic. As discussed above, the pharyngeal gonococci are complicated to cure, and effective in building up of resistance due to the adaptation of the bacteria from the low concentrations reaching the targets (Peterson & Kaur, 2018). This is one of the major

mechanisms that is leading to the complications in the healthcare system, as it is difficult to know when an individual suffers from this kind of gonorrhoea, therefore posing a challenge in the control of the disease. Despite using the site, that is the pharynx to evade drug action, bacteria also produce enzymes that target to metabolize the drug, after significant exposure to the drug. An example of this is the transcription of the β-lactamase enzyme in *N. gonorrhoeae,* which targets to degrade the β-lactam ring in the β-lactams, therefore destroying the active component of the drug, reducing its effectiveness in the treatment (Kapoor et al., 2017). This is also a challenge, as the traits are transmitted to the offspring genetically, which leads to proliferation of the antibiotic-resistant strains.

2.4 Treatment of *N gonorrhoeae*

Fluoroquinolones and cefriaxones were the first line antibiotics which were recently used to treat gonorrhoea disease, but they are currently not effective in the treatment due to the development of the resistance in the bacteria, towards these drugs. The threat that is currently a public and health concern is in the AMR in this bacterium, as most of the available drugs are not effective in the treatment of the disease (Sarenje et al., 2024). Due to the development of the resistance in this bacterium, most of the used classes of medication such as penicillin and β-lactams, the solution that can currently be used is tracing this problem back to basic molecular biology, where the genes are now studied and analyzed in efforts to understand and combat this bacteria (World Health Organization, 2016). With a better understanding of how this bacterium develops resistance, more targeted antibiotics or gene modification to reduce or mitigate the spread of the resistant strains. The study of the genetic population structure using the genome-wide SNPs data can be very important in the determination of the transmissibility of gonorrhoea. There is urgent need of addressing the treatment option for this highly mutating bacteria, because as it is already posing a threat in health sector, it can be very difficulty to treat in future, if the resistance development of the bacteria is not effectively determined.

2.5 Bacteria growth and transmission.

Neisseria gonorrhoeae bacteria has been reported to mainly colonize genital mucosa alongside nasopharyngeal, ocular and anal mucosa. This regions provides a conducive environment for the bacteria to grow and also be easily transmitted through sexual intercourse. This means that individuals with multiple sex partners are at higher risk of this resistant disease. Because of this distinction in the regions where the bacteria can thrive well, two major sources are taken into consideration when the bacteria is isolated from patients,

which are urethra and endocervix. These two sources of the bacteria are key to the study because the prevalence of the infection is not the same. Female sex workers are at higher risk of contracting the disease, and they are also considered to be a major source of spread of this disease due to their sexual behaviors (Abdullahi et al, 2022). This is associated to their sexual behavior, where they are capable of contracting this disease to a larger population from just a single individual. In Nairobi, It was found out that the female sex workers in the age of between 38-49 years were most likely to get the disease, and also the ones who had only the primary education (Abdullahi et al, 2022). Due to horizontal gene transfer, this behavior can lead to acquired resistance, where the strains being introduced from other individual having gonorrhoea disease may modify the existing strains (Reygaert, 2018). This interchange of the strains by this commercial sex workers lead to elevation of resistance, making this bacteria to be more resistant to medication.

2.5 Single nucleotide polymorphisms

Single nucleotide polymorphisms (SNPs) in genomic data is the assessment of the change and variation of a genome at a nucleotide or base level. SNPs occur when nucleotide bases (Adenine, Guanine, Cytosine and Thymine) differ between individuals. One of the DNA nucleotide base may be replaced or substituted with another, for example, Adenine can be replaced with Guanine, which may bring variations amongst individuals (Leaché & Oaks, 2017). These SNPs have a significant contribution in the gene function and may bring about issues with disease susceptibility and antimicrobial resistance among other effects (Chander et al., 2021). It is therefore very useful in the study of genetic diversity, population structure and evolution, and also determination of transmission of a bacteria or pathogen in study. SNPs are identified using bioinformatics tools such as GATK, BCFtools and SAMtools, and then annotation can be done to determine coding and non-coding regions. GW-SNP is also used as a tool in the determination disease risk factors,by identifying common genetic variability underlying a disease.

2.6 Principal component analysis

Principle component analysis (PCA) is an analysis that reduces the dimensionalities of the genomic data, while it preserves the covariance within the samples. It is widely applied in the determination of the genetic population structure, by reducing the otherwise multi-component data into small components which can therefore be visualized by a plot (scatter-plot). The scatter plot shows how close-relatedness samples are to each other, which can provide inferences to how many subpopulations are found in a population (Kurita, 2019). PCA has

wide application in genomics, but in this case, we will be focusing on the examination of the population structure in Kenyan samples to determine ancestry, and admixture and also determine the demographic history of the samples in the Kenyan population. It is used to quantify and detect genetic difference within the population by clustering individuals based on the relatedness.

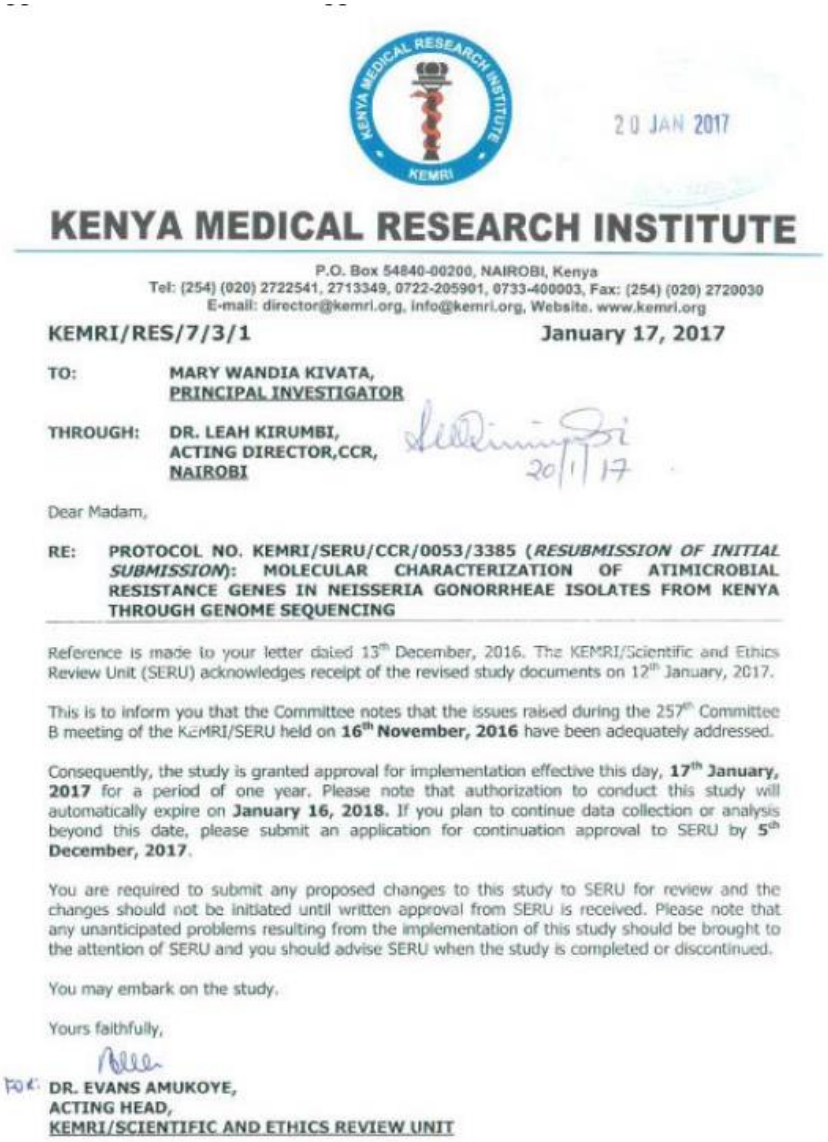
CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Sample Collection

The bacterial isolates used in this study were obtained as part of ongoing research on sexually transmitted infections under the Armed Forces Health Surveillance program at the US Army Medical Research Directorate-Africa. The samples were collected from patients seeking treatment for gonorrhea at various healthcare facilities across Kenya, including Nairobi, Rift Valley, Coastal region, Kombewa, and Kisumu. The samples, primarily from urethral and endocervical sources, were collected between 2013 and 2018.

Ethical clearance



KENYA MEDICAL RESEARCH INSTITUTE

20 JAN 2017

P.O. Box 54840-00200, NAIROBI, Kenya
Tel: (254) (020) 2722541, 2713349, 0722-205901, 0733-400003, Fax: (254) (020) 2720030
E-mail: director@kemri.org, info@kemri.org, Website. www.kemri.org

KEMRI/RES/7/3/1 January 17, 2017

TO: MARY WANDIA KIVATA,
PRINCIPAL INVESTIGATOR

THROUGH: DR. LEAH KIRUMBI,
ACTING DIRECTOR,CCR,
NAIROBI

20/1/17

Dear Madam,

RE: PROTOCOL NO. KEMRI/SERU/CCR/0053/3385 (*RESUBMISSION OF INITIAL SUBMISSION*): MOLECULAR CHARACTERIZATION OF ATIMICROBIAL RESISTANCE GENES IN NEISSERIA GONORRHEAE ISOLATES FROM KENYA THROUGH GENOME SEQUENCING

Reference is made to your letter dated 13th December, 2016. The KEMRI/Scientific and Ethics Review Unit (SERU) acknowledges receipt of the revised study documents on 12th January, 2017.

This is to inform you that the Committee notes that the issues raised during the 257th Committee B meeting of the KEMRI/SERU held on **16th November, 2016** have been adequately addressed.

Consequently, the study is granted approval for implementation effective this day, **17th January, 2017** for a period of one year. Please note that authorization to conduct this study will automatically expire on **January 16, 2018.** If you plan to continue data collection or analysis beyond this date, please submit an application for continuation approval to SERU by **5th December, 2017**.

You are required to submit any proposed changes to this study to SERU for review and the changes should not be initiated until written approval from SERU is received. Please note that any unanticipated problems resulting from the implementation of this study should be brought to the attention of SERU and you should advise SERU when the study is completed or discontinued.

You may embark on the study.

Yours faithfully,

FOR: DR. EVANS AMUKOYE,
ACTING HEAD,
KEMRI/SCIENTIFIC AND ETHICS REVIEW UNIT

DEPARTMENT OF THE ARMY
WALTER REED ARMY INSTITUTE OF RESEARCH
503 ROBERT GRANT AVENUE
SILVER SPRING, MD 20910-7500

REPLY TO ATTENTION OF

MCMR-UWZ-C

27 July 2017

MEMORANDUM FOR Mary Wandia Kivata, Msc, Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya

SUBJECT: Project Qualifies as Research Not Involving Human Subjects, **WRAIR #1743A**

1. A determination was made that the project **WRAIR #1743A**, entitled "Molecular Characterization of Antimicrobial Resistance Genes in *Neisseria gonorrhoeae* Isolates from Kenya through Whole Genome Sequencing," (Version 1.2, dated 9 June 2017) does not require review by the Walter Reed Army Institute of Research (WRAIR) Institutional Review Board (IRB) in accordance with WRAIR Policy Letter #12-09, as the project involves the analysis of coded isolates and a limited patient data set to which the sub-project investigator does not have access to identifiable information; therefore, this research activity does not meet the definition of research involving human subjects and 32 CFR 219 does not apply.

2. This is a retrospective laboratory based study nested in an ongoing sexually transmitted illness (STI) surveillance program under the Walter Reed Project at the Kenya Medical Research Institute (KEMRI). Archived *Neisseria gonorrhoeae* isolates obtained from the ongoing study WRAIR# 1743, entitled "A Surveillance Study of Antimicrobial Susceptibility Profiles of *Neisseria gonorrhoeae* Isolates from Patients Seeking Treatment in Selected Military and Civilian Clinics in Kenya," and antimicrobial susceptibility information from the archived isolates will be used in this project.

The general objective of this project is to characterize antimicrobial resistance (AMR) in *Neisseria gonorrhoeae* isolates from Kenya through whole genome sequencing. The specific objectives are to:

a) Determine and characterize the full genome of antimicrobial resistant Kenyan *Neisseria gonorrhoeae* isolates;
b) Geophylogenically characterize the origin and genetic relatedness of Kenya *Neisseria gonorrhoeae* isolates;
c) Characterize chromosomal and plasmid gene determinants of resistance expressed by the Kenyan *Neisseria gonorrhoeae* isolates;
d) Evaluate the differences between identified AMR determinants and existing known AMR determinants; and
e) Characterize the molecular mechanisms involved in AMR development in the Kenyan *Neisseria gonorrhoeae* isolates.

3.2 DNA extraction and sequencing

Both Genomic and plasmid DNA were extracted using QIAamp DNA Mini Kit and QIAprep Spin Miniprep Kit (QIAGEN, Hilden, Germany) "respectively" according to the manufacturer's instructions. Qubit dsDNA HS Assay was used to quantitate DNA using Qubit 3.0 fluorometer, (Thermo Fisher Scientific Inc. Wilmington, Delaware USA) according to the manufacturer's instructions, and DNA stored at − 20 °C prior to sequencing.

3.3 Whole-genome sequencing and sequence analysis

Illumina Nextera XT kit (Illumina Inc. San Diego, CA, USA) was used to prepare libraries from 1 ng of genomic DNA of each sample as per manufacturer's instructions. Sequence reads were generated on Illumina MiSeq platform (Illumina, San Diego, CA, USA) using a paired-end 2 × 300 bp protocol [49]. The generated reads are linked to NCBI BioProjects: PRJNA481622 and PRJNA590515.

3.4 Data Processing

The quality of the sequencing reads will be assessed using FastQC v0.11.9, and low-quality bases (Phred score <20) will be trimmed using Trimmomatic PE v0.39. The cleaned reads were assembled de novo using SPAdes v4.0.0, with assembly parameters k: [21, 33, 55, 77]. The quality of the genome assemblies will be evaluated using QUAST v5.2, and the completeness of each assembly will be assessed using the BUSCO tool with the neisseriales_odb10 database.

3.5 Gene prediction and annotation

Gene prediction will be performed using Prodigal v2.6.3, while gene annotation will be carried out using Prokka v1.14.6. The number of predicted genes and their functions will be determined to identify genes associated with antibiotic resistance.

3.6 Variant calling and phylogenetic analysis

Variants, including SNPs and Indels, will be identified using Bcftools v1.10.2. Phylogenetic analysis will be conducted based on gene clusters across all samples using the roary tool, and a phylogenetic tree will be constructed using the maximum likelihood algorithm in IQ-TREE v2.2.2.7. The SNPs data obtained here will then be used to examine the population genetic structure, and also provide insights on the transmission of the *N. gonorrhoea*.

REFERENCES

1. Chander, S., Garcia-Oliveira, A. L., Gedil, M., Shah, T., Otusanya, G. O., Asiedu, R., & Chigeza, G. (2021). Genetic Diversity and Population Structure of Soybean Lines Adapted to Sub-Saharan Africa Using Single Nucleotide Polymorphism (SNP) Markers. *Agronomy*, *11*(3), 604. https://doi.org/10.3390/agronomy11030604
2. Kapoor, G., Saigal, S., & Elongavan, A. (2017). Action and resistance mechanisms of antibiotics: A guide for clinicians. *Journal of Anaesthesiology Clinical Pharmacology*, *33*(3), 300. https://doi.org/10.4103/joacp.JOACP_349_15
3. Kivata, M. W., Mbuchi, M., Eyase, F., Bulimo, W. D., Kyanya, C. K., Oundo, V., Mbinda, W. M., Sang, W., Andagalu, B., Soge, O. O., McClelland, R. S., & Distelhorst, J. (2020). Plasmid mediated penicillin and tetracycline resistance among Neisseria gonorrhoeae isolates from Kenya. *BMC Infectious Diseases*, *20*(1), 703. https://doi.org/10.1186/s12879-020-05398-5
4. Kurita, T. (2019). Principal component analysis (PCA). *Computer vision: a reference guide*, 1-4.
5. Leaché, A. D., & Oaks, J. R. (2017). The Utility of Single Nucleotide Polymorphism (SNP) Data in Phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 69–84. https://doi.org/10.1146/annurev-ecolsys-110316-0226
6. Nacht, C., Agingu, W., Otieno, F., Odhiambo, F., & Mehta, S. D. (2020). Antimicrobial resistance patterns in *Neisseria gonorrhoeae* among male clients of a sexually transmitted infections clinic in Kisumu, Kenya. *International Journal of STD & AIDS*, *31*(1), 46–52. https://doi.org/10.1177/0956462419881087
7. Ndowa, F., & Lusti-Narasimhan, M. (2012). The threat of untreatable gonorrhoea: Implications and consequences for reproductive and sexual morbidity. *Reproductive Health Matters*, *20*(40), 76–82. https://doi.org/10.1016/S0968-8080(12)40653-X
8. Peterson, E., & Kaur, P. (2018). Antibiotic Resistance Mechanisms in Bacteria: Relationships Between Resistance Determinants of Antibiotic Producers, Environmental Bacteria, and Clinical Pathogens. *Frontiers in Microbiology*, *9*, 2928. https://doi.org/10.3389/fmicb.2018.02928
9. Quillin, S. J., & Seifert, H. S. (2018). Neisseria gonorrhoeae host adaptation and pathogenesis. *Nature Reviews Microbiology*, *16*(4), 226–240. https://doi.org/10.1038/nrmicro.2017.169
10. Sarenje, K. L., Van Zwetselaar, M., Kumburu, H., Sonda, T., Mmbaga, B., Ngalamika, O., Maimbolwa, M. C., Siame, A., Munsaka, S., & Kwenda, G. (2024).

Antimicrobial resistance and heterogeneity of Neisseria gonorrhoeae isolated from patients attending sexually transmitted infection clinics in Lusaka, Zambia. *BMC Genomics*, *25*(1), 290. https://doi.org/10.1186/s12864-024-10155-y

11. Takahashi, S., Kurimura, Y., Hashimoto, J., Takeyama, K., Tsukamoto, T., Koroku, M., Tanda, H., & Nishimura, M. (2008). Pharyngeal Neisseria gonorrhoeae detection in oral-throat wash specimens of male patients with urethritis. *Journal of Infection and Chemotherapy*, *14*(6), 442–444. https://doi.org/10.1007/s10156-008-0649-5
12. Unemo, M., & Shafer, W. M. (2014). Antimicrobial Resistance in Neisseria gonorrhoeae in the 21st Century: Past, Evolution, and Future. *Clinical Microbiology Reviews*, *27*(3), 587–613. https://doi.org/10.1128/CMR.00010-14
13. Unemo, M., Lahra, M. M., Escher, M., Eremin, S., Cole, M. J., Galarza, P., Ndowa, F., Martin, I., Dillon, J.-A. R., Galas, M., Ramon-Pardo, P., Weinstock, H., & Wi, T. (2021). WHO global antimicrobial resistance surveillance for Neisseria gonorrhoeae 2017–18: A retrospective observational study. *The Lancet Microbe*, *2*(11), e627–e636. https://doi.org/10.1016/S2666-5247(21)00171-3
14. World Health Organization. (2016). *WHO guidelines for the treatment of Neisseria gonorrhoeae*. World Health Organization. https://iris.who.int/handle/10665/246114
15. *The Lancet Microbe, 2(11), e627–e636. Clinical Microbiology Reviews*, *27*(3), 587–613. *WHO guidelines for the treatment of Neisseria gonorrhoeae*. World Health Organization. Chander, S., Garcia-Oliveira, A. L., Gedil, M., Shah, T., Otusanya, G. O., Asiedu, R., & Chigeza, G. (2021). Genetic Diversity and Population Structure of Soybean Lines Adapted to Sub-Saharan Africa Using Single Nucleotide Polymorphism (SNP) Markers. *Agronomy*, *11*(3), 604.
16. Belland, R. J., Morrison, S. G., Ison, C., & Huang, W. M. (1994). Neisseria gonorrhoeae acquires mutations in analogous regions of gyrA and parC in fluoroquinolone-resistant isolates. *Molecular microbiology*, *14*(2), 371-380.
17. Del Rio, C., Hall, G., Hook Iii, E. W., Holmes, K. K., Whittington, W. L. H., Judson, F. N., ... & Weinstock, H. S. (2007). Update to CDC's sexually transmitted diseases treatment guidelines, 2006: fluoroquinolones no longer recommended for treatment of gonococcal infections. *JAMA: Journal of the American Medical Association*, *297*(22).
18. Kivata, M. W., Mbuchi, M., Eyase, F. L., Bulimo, W. D., Kyanya, C. K., Oundo, V., ... & Mancuso, J. D. (2019). gyrA and parC mutations in fluoroquinolone-resistant Neisseria gonorrhoeae isolates from Kenya. *BMC microbiology*, *19*, 1-9.

19. Abdullahi, A., Nzou, S. M., Kikuvi, G., & Mwau, M. (2022). Neisseria gonorrhoeae infection in female sex workers in an STI clinic in Nairobi, Kenya. *Plos one*, *17*(2), e0263531.
20. Reygaert, W. C. (2018). An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS microbiology*, *4*(3), 482.

APPENDICES

Work-plan

Table 1: Timeline

Month / Activity	Oct'	Nov'	Dec'	Jan'	Feb	Mar'	Apr'	May
Topic Refinement								
Literature Review								
Proposal Writing								
Proposal Defense			BREAK					
Data Preparation Analysis								
Writing and Review								
Project Presentation and defense								

Budget

Table 2: budget

Item	Description	Cost (Ksh)
1. Software/Tools		
Cloud computing resources	For high-throughput analysis	Ksh 6,500
External storage drives (1TB)	For secure data backup	Ksh 6,500
2. Publication and Printing		
Printing of Report	Hard copies of thesis for submission	Ksh 1000
5. Miscellaneous	Internet Access	Ksh 2,000
Total Estimated Budget		**Ksh 16,000**