A PILOT STUDY TO IDENTIFY SINGLE NUCLEOTIDE VARIANTS IN EXON 19 OF THE *BRCA1* GENE IN FEMALE BREAST CANCER PATIENTS AT MATERO LEVEL ONE HOSPITAL, LUSAKA ZAMBIA

By

Lisa Shakachite

A dissertation submitted to the University of Zambia in partial fulfilment of the requirements for the Master of Science Degree in Molecular Biology

THE UNIVERSITY OF ZAMBIA

LUSAKA

2023

DECLARATION

I, **Lisa Shakachite**, hereby declare that this dissertation represents my work and has not been previously submitted for a degree at this or any other University.

Full name: LISA SHAKACHITE
Signature: 
Date: 04/17/2023

CERTIFICATE OF APPROVAL

This dissertation of Lisa Shakachite has been approved as partially fulfilling the requirements for the award of Master of Science degree in Molecular Biology by the University of Zambia.

Examiner's name	**Signature**	**Date**
…….	…….	…….
Examiner I		
…….	…….	…….
Examiner II		
…….	…….	…….
Examiner III		
…….	…….	…….
Chairperson Board of Examiners		
…….	…….	…….
Supervisor		

ABSTRACT

Breast cancer (BC) patients in Zambia more commonly present with advanced-stage disease and often at young ages (≤50 years). *Breast cancer gene 1* (*BRCA1*) is a tumor suppressor gene in which mutations have been linked to disease presentation at a young age. More than 600 distinct pathogenic (disease causing) single nucleotide variants with the potential for structural and functional effects on BRCA1 have been documented in ClinVar, a major human variation and phenotype database. The *BRCA1* C-terminal (BRCT) region is critical for tumor suppression because it acts in DNA damage repair and cell cycle checkpoint protein control. Mutations in this region have been associated to structural effects and loss of function of BRCA1 in many studies. However, BRCT mutations in young female breast cancer patients from Zambia have not been identified to date. The genetic alterations in exon 19 of the *BRCA1* gene were assessed here for the first time in a cohort from Zambia. Seventy-three patients diagnosed with breast cancer at the age of ≤50 years were recruited irrespective of family history of breast and/or ovarian cancer. Genomic DNA was extracted from ten peripheral blood samples and PCR-amplified with exon 19-targeting primers of the *BRCA1* gene. Amplicons were sequenced using a Sanger sequencing platform and analyzed using bioinformatics tools at various web-based platforms and ClinVar. Overall, eleven single nucleotide mutations were identified in three patients, and all known to play a role in determining the structure of BRCA1. Of the identified mutations, one was pathogenic and seven had uncertain significances. Six of these mutations were heterozygous and classified in the loss of function category in ClinVar. These variants were specifically associated with Breast-ovarian cancer, familial, susceptibility to, 1 (BROVCA1). Four out of the six heterozygous mutations were missense mutations and two were nonsense mutations. Based on an unpaired two-tailed student's *t*-test, the null hypothesis that there was no association between the age of a BC patient and the number of mutations in exon 19 leading to a loss of function of BRCA1, was rejected (P= 0.04). Further studies are clearly necessary to explore how mutations in the entire BRCA1 gene affect disease presentation in the female BC population. The young age (≤50 years) for breast cancer diagnosis in female Zambian patients may be attributed to other factors like other genetic mutation status, epigenetic changes, lifestyle etc. These results will help in decision making with regard to the integration of genomic medicine into breast cancer care in Zambia which has tremendous potential to improve the outcomes and survival of diagnosed patients.

DEDICATION

This dissertation is dedicated to my father and mother for their persistent and tireless support that availed me a chance to complete my master's degree.

ACKNOWLEDGEMENTS

Firstly, I would like to thank our Almighty God, for granting me his grace to pursue this master's degree in Molecular Biology.

Secondly, I would like to thank my supervisors, Dr. Evans Kaimoyo and Dr. Mutumba Songiso, for their guidance in the formulation and structuring of this dissertation. Thirdly, the nurses at Matero level one's breast cancer clinic, Sister Mavis and Sister Mwaba, for their support during the study.

I would also like to appreciate and commend the Centre for Infectious Disease Research in Zambia (CIDRZ) for their provision of the necessary laboratory space and equipment utilized for this study. I would like to specifically commend Kennedy Chibesa and Suwilanji Silwamba from the CIDRZ laboratory for their aid and persistence during the data acquisition processes, analysis of electrophoresis, PCR amplification and purification processes for this study.

Lastly, I extend my gratitude to the University of Zambia's School of Veterinary Medicine, particularly Professor Hang'ombe and Mr. Kabwali, as well as the KAbwe Mine Pollution Amelioration Initiative (KAMPAI) project coordinator, for their provision of equipment used in this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ACRONYMS AND ABBREVIATIONS

BC	Breast Cancer
BRCA1	*Breast cancer gene 1*
BRCT	*BRCA1* C-terminal
BROVCA1	Breast-ovarian cancer, familial, susceptibility to, 1
CDH	Cancer Diseases Hospital
UNZA	University of Zambia
NHRA	National Health Research Authority
DNA	Deoxyribonucleic acid
DSB	DNA double-strand break
HR	Homologous Recombination
DDR	DNA Damage Response
gDNA	Genomic DNA
µL	microliter
mL	milliliter
PCR	Polymerase Chain Reaction
BLAST	Basic Local Alignment Search Tool
UCSC	University of California Santa Cruz
OMIM	Online Mendelian Inheritance in Man
HBOC	Hereditary Breast and Ovarian Cancer
SGE	Saturation Genome Editing
LMICs	Low- and Middle-Income Countries
NGS	Next-Generation Sequencing

CHAPTER ONE

INTRODUCTION

1.0 OVERVIEW

This chapter establishes the background to the study and outlines a statement of the problem. It states the aim of the study as well as the research objectives and questions. The significance of the study as well as the study hypothesis are explained.

1.1 BACKGROUND

Various types of cancer are among the leading causes of deaths worldwide. In 2020, breast cancer became the most commonly diagnosed cancer type in the world with more than 2.26 million new cases and almost 685 000 deaths worldwide (National Cancer Institute, 2020). According to GLOBOCAN, an online platform that provides statistics on cancer trends, 186,598 breast cancer cases were reported in Africa with 85,787 related deaths (Anyigba, Awandare and Paemka, 2021; Sharma *et al.*, 2022). Zambia, a country where breast cancer constitutes the second leading cause of cancer mortality after cervical cancer has not been spared from the impact of the disease (Cabanes *et al.*, 2019). A retrospective observational study using the Zambia National Cancer Registry (ZNCR) population-based data from 2007 to 2014 revealed that BC in Zambia had an age adjusted standardized incidence rate of 38.08 per 100,000 females (Kalubula *et al*., 2021). Data from the Cancer Diseases Hospital (CDH) in the country also illustrates that BC cases make up 9% of all cases seen at CDH and yet the 2-year survival rate among women treated at CDH remains below 50% due to the late stage at diagnosis (MHRZ, 2016). Breast cancer patients in Zambia more commonly present with advanced-stage disease (Stages III- IV) and often at young ages (Songiso *et al.,* 2020). Particularly in women younger than 50 years old, breasts tend to be denser and more prone to develop aggressive disease due to certain risk factors like having fewer childbirths, obesity, and socioeconomic status (Francies *et al.*, 2015; Songiso *et al.,* 2020). To date, genetic data from young female BC patients in Zambia has not been evaluated.

In order to improve the outcome of certain BC treatment strategies, personalized medicine could be the answer. This can be implemented by utilizing techniques like gene sequencing to identify the mutations in cancer-critical genes that have been implicated in BCs. Medical practitioners can then be better positioned to personalize treatment strategies for the benefit of patients like in developed countries e.g., the United States of America (USA) and the United Kingdom (UK). Some of the major human cancers like ovarian, lung, brain, cervical and esophageal cancers whose underlying genetic bases have been elucidated have been shown to be due to specific changes at nucleotide level which may cause changes in protein structure particularly of tumor suppressor genes like *BRCA1, BRCA2 (BReast CAncer gene 2), TP53 (Tumor Protein p53), PTEN (Phosphatase and TENsin homolog), STK11 (Serine/Threonine Kinase 11),* and *CDH1 (CaDHerin-1),* respectively (Siddig *et al.*, 2021; Freedman and Partridge, 2013; Melhem-Bertrandt *et al.*, 2012). These are all considered as high penetrance mutations of breast cancer such that they account for 20% of hereditary risk (Siddig *et al.*, 2021). The *breast cancer susceptibility gene types 1 and 2*, commonly designated as *BRCA1* and *BRCA2*, are some of the cancer critical genes in which inherited mutations have been shown to be associated with high risk of obtaining breast cancer in women (Clark *et al.*, 2012; Mylavarapu *et al.*, 2018). As depicted in figure 2.1, BRCA1 is made up of three major domains one of which is a RING domain (exons 2–7), and the other are a region encoded by exons 11–13, and the Breast Cancer gene C-terminal domain (BRCT) encoded for by exons 16–24 (Clark *et al.*, 2012). Through these domains, BRCA1 forms complexes with several proteins, implicating BRCA1 in multiple cellular functions, such as transcription regulation, cell-cycle checkpoint activation, and DNA repair (Venkitaraman, 2014; Dacheux *et al.*, 2013; Prakash *et. al*, 2015; Deng, 2006).

Mutations in various domains of *BRCA1*, including the N and C-terminal domains, have been linked to hereditary breast and ovarian cancers and have been hypothesized to increase the risk of developing several additional cancers like pancreatic, fallopian tube and peritoneal cancer (Clark *et al.*, 2012; Chen, Roberts and Klein, 2017; Kim *et al.*, 2020). A number of disease-causing mutations, associated with breast cancer, scattered across the entirety of the *BRCA1* gene have been identified in numerous studies elsewhere (ElBiad *et al.*, 2022; Abbad *et al.*, 2018; Francies *et al.*, 2015). In North Africa, *BRCA1* founder variants were found in 36.4% of BC cases, thereby outnumbering *BRCA2* founder variants by a ratio of ≈3:1 (ElBiad *et al.*, 2022). To date, little or

no records exist on mutations in cancer-critical genes and their effects on Zambian patients. This study was thus undertaken as a first step to explore the mutations present in the BRCT region of *BRCA1* in a population of female breast cancer patients and, to contribute to the understanding of the potential correlation between mutations in a cancer critical genes and pathogenesis. Findings from this study will possibly contribute to efforts currently underway to elucidate how pathogenic variants of *BRCA1* are associated with BC in a selected population of patients evaluated at Matero level one hospital.

1.2 STATEMENT OF THE PROBLEM

Breast cancer patients in Zambia and Africa, in general, more commonly present with advanced-stage disease and often at young ages (Francies *et al.*, 2015; Cabanes *et al.*, 2019; Songiso *et al.,* 2020). While mutations in cancer-critical genes such as *BRCA1* have been associated with various types of cancer in women, the underlying genetic basis and specific types of mutations in breast cancer-critical genes in Zambia remains an area that has hardly been explored. Diagnosis typically only occurs at advanced stage due to low access to available diagnostic centers and limited screening (Francies *et al.*, 2015; Sharma *et al.*, 2022) particularly in remote parts of the country. In developed countries, when breast cancer is detected early and is in the localized stage (where there is no sign that the cancer has spread outside of the breast), the 5-year relative survival rate is 100% (National Breast Cancer Foundation Inc., 2020). It is therefore crucially important to couple screening interventions with early diagnosis activities to reduce the late presentation of cases of cancer in our health institutions (MHRZ, 2016).

1.3 SIGNIFICANCE OF THE STUDY

This study is to the best of my knowledge the first to characterize the BRCT variants in *BRCA1* present in BC patients from Zambia. A similar study in South Africa identified three patients with pathogenic gene mutations in *BRCA1* out of eighty-five (3.5%) that influenced its structure and function allowing for their progeny to be genetically screened for the possible inheritance of these variants (Francies *et al.*, 2015). Another study in Morocco identified one, BRCA1:c.116G > A (p.Cys39Tyr), out of three germ-line mutations in *BRCA1* which was found to be pathogenic or

associated with disease (Bakkach *et al.*, 2020). Clearly, these studies highlight how genetic diversity as a result of gene mutations can contribute to pathogenicity and inheritance thereby demonstrating how important it is to consider their influence on BC in Zambia in order to drive targeted research and screening. Breast cancer can be detected early using 2 strategies: screening and early diagnosis. Breast cancer screening is performed using mammogram, clinical breast exam (CBE), and MRI (magnetic resonance imaging) tests (National Cancer Institute, 2022). The mode of early detection supported by the World Health Organization (WHO), i.e., population-based mammography screening programs in asymptomatic women, is not possible presently in Zambia because the facilities and resources required to undertake such a task are limited (Songiso *et al.,* 2020).

Early diagnosis, based on awareness of signs and symptoms associated with cancer, involves the recognition of possible warning signs of cancer, by mammogram or CBE, followed by prompt action (Songiso *et al.,* 2020). In situations where mammography options are limited, like in Zambia, this method of diagnosis can be coupled with genetic information which is a cost-effective method of screening to allow for the proper allocation of resources and treatment strategies. Traditional BC diagnosis journeys that follow diagnosis (mammogram, ultrasound, biopsy) and then treatment (lumpectomy, mastectomy) can cost patients and institutions ~ZMK40-300000, whereas diagnosis coupled with screening (sequencing, panels) and subsequent treatment (possibly non-surgical methods) could cost patients and institutions ~ZMK10-20000 (Yip, 2016; Nicolussi *et al.,* 2019, Foerster *et al.,* 2021). This illustrates the tremendous potential that coupling diagnosis with screening has to save patients and health institutions numerous funds, resources, and valuable time. This proposed research aims at illustrating how genomic medicine can be useful in the implementation of early BC detection and screening.

The young age for development of breast cancer often suggests a genetic predisposition especially germline mutations in the *BRCA1* and *BRCA2* genes (Bakkach *et al.*, 2020). In Morocco, young age (≤40 years) irrespective of family history was proposed as a sufficient criterion for systematic genetic screening for young breast cancer patients, following the western recommendations for *BRCA1/2* genetic testing (Bakkach *et al.*, 2020). This criterion could be utilized to establish and develop genetic screening panels for BC in Zambia according to country specific statistics.

Continued investigation is therefore clearly necessary to determine the presence and precise mechanism(s) by which BRCT mutations in *BRCA1* may contribute to tumorigenesis allowing for the possible generation of *BRCA1* gene screening panels in Zambia, promoting the potential for coupling BC diagnosis with screening.

1.4 AIM OF THE STUDY

The aim of this study was to use molecular methods to characterize single nucleotide gene mutations that may be present in part of the BRCT region (specifically exon 19) in a selected population of patients evaluated at Matero Level One Hospital of Lusaka Zambia, which is a BC referral hospital in Zambia.

1.5 OBJECTIVES OF THE STUDY

In order to achieve the intended purpose, this study was guided by the overall objective and specific objectives.

1.5.1 GENERAL OBJECTIVE

The general objective of the study was to characterize the types of BRCA1 gene mutations that underlie breast cancer pathology in patients presenting breast cancer cases at Matero Level One Hospital in Lusaka.

1.5.2 SPECIFIC OBJECTIVES

In order to achieve the general objective of this study, the four specific objectives formulated for the study were to:

1.5.2.1 Collect blood samples from study participants and characterize the single nucleotide mutations (missense or nonsense) present in a specific part of the BRCT region, i.e. exon 19 of *BRCA1*, in BC patients. Only female BC patients with positive ultrasound-guided core biopsy results were invited to take part in the study.

1.5.2.2 Determine which mutations have a bearing on the structure and function of the mutant gene products.

1.5.2.3 To investigate the association between the age of a BC patient and the number of mutations in exon 19 leading to a loss of function of BRCA1.

1.5.2.4 Identify the pathogenicity associated with identified, potentially deleterious, missense and nonsense mutations of *BRCA1* in Zambian BC patients.

1.6 STUDY HYPOTHESIS

The study tested the null hypothesis that there is no significant association between the age of a BC patient and the number of mutations in exon 19 leading to a loss of function of *BRCA1*.

1.7 ORGANIZATION OF THE DISSERTATION

This dissertation is organized into six chapters. Chapter one presents an introduction to the study. Chapter two presents literature review, while Chapter three describes the sample selection process, materials and methods and discusses the generation of results. The findings of the study are described in Chapter four and the interpretation of, and discussion of the findings are presented in Chapter five. Finally, conclusions and recommendations are provided in Chapter six.

CHAPTER TWO

LITERATURE REVIEW

2.0 OVERVIEW

This chapter reviews studies undertaken by different scholars and other peer-reviewed literature to document the current understanding on *BRCA1* with reference to the structure, major functions, mutations and gaps in knowledge associated with breast cancer.

2.1 STRUCTURE OF *BRCA1*

BRCA1 is a tumor suppressor gene located on chromosome 17, at position 17q21 in humans. (El Khachibi *et al.*, 2015). It has 24 exons (22 of which are protein-coding) and these code for an 1863 amino acid polypeptide (Cunningham *et al*., 2022; Hondrow *et al.*, 2011). The polypeptide is made up of three major domains one of which is a RING (Really Interesting New Gene) domain consisting of a RING finger and two flanking alpha (α) helices encompassing amino acids 1-109 (exons 2–7) (Clark *et al.*, 2012; Lipkowitz and Weissman, 2011). Through seven conserved cysteine residues and one conserved histidine residue, the RING finger coordinates two Zn^{2+} atoms which stabilize the RING structure responsible for the E3-ubiquitin ligase activity of BRCA1 (Lipkowitz and Weissman, 2011). The other domains consist of a region encoded by exons 11–13, and the BRCT domain (exons 16–24), as shown in figure 2.1 (Clark *et al.*, 2012; Di Masi *et al.*, 2011).

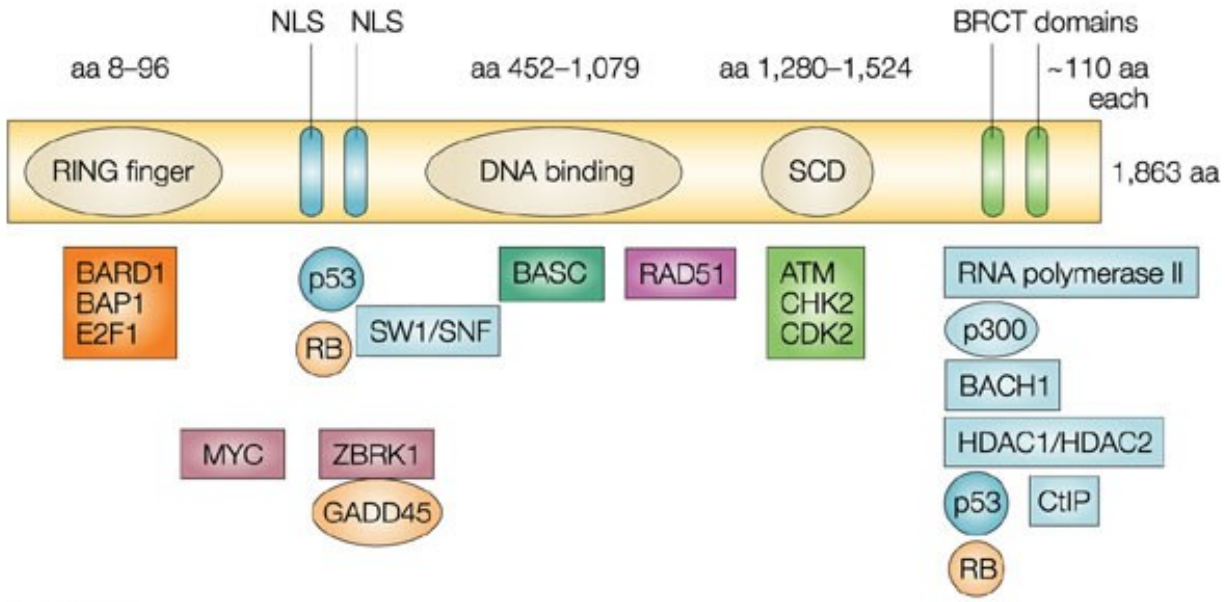


Figure 2.1 Schematic of major BRCA1 regions described. Narod & Foulkes, 2004.

The C-terminal domain of BRCA1 (BRCT) is made up of two domains: BRCT1 (amino acids 1642-1736) at the N-terminus and BRCT2 (amino acids 1756-1855) at the C-terminus (UniProt Consortium, 2021; Di Masi *et al*., 2011; Clark *et al.*, 2012). Each BRCT domain is composed of 85-95 amino acids connected by a 22-amino acid linker with each domain consisting of three α helices packed around a four-strand beta (β) -sheet, Figure 2.1 (Di Masi *et al*., 2011; Clark *et al.*, 2012). Exon 19 codes for amino acids 1732-1759 and encompasses the 22-amino acid linker and a small portion of the BRCT1 and BRCT2 domains (Cunningham *et al*., 2022; Clark *et al.*, 2012). Cancer-causing missense mutations occur at the interface between the two BRCT repeats, and these have been hypothesized to destabilize the structure of BRCA1, thereby affecting its role as a tumor suppressor (Williams, Green and Glover, 2001; Glover, Williams and Lee, 2004).

Through its three major domains, BRCA1 forms complexes with and interacts with several proteins, as shown in figure 2.1 and 2.2, implicating it in multiple cellular functions, such as transcription regulation, cell-cycle checkpoint activation, and DNA repair (Bunch *et al.*, 2021; Prakash *et al.*, 2015; Venkitaraman, 2014). A study in Burkina Faso analyzed the prevalence and variants of various putative genes involved in cancer regulation in over 130 participants and found

germline variants in exon 11 of *BRCA1* which were linked to pathogenesis (Kiendrebeogo *et al.*, 2022).

2.2 MAJOR FUNCTIONS OF BRCA1

BRCA1 has been reported to function in various molecular pathways and processes including transcription regulation, the ordering, timing and progression of the cell cycle based on its cellular localization and DNA repair (Dacheux *et al.*, 2013; Venkitaraman, 2002; Deng, 2006). Some of the various functions of BRCA1 are summarized under the respective subheadings below.

2.2.1 TRANSCRIPTION REGULATION

Depletion of BRCA1 expression has been shown to lead to transcriptional changes in about 1294 genes, out of which only 44 have been demonstrated to have promoter elements bound by BRCA1 (Gorski *et al.*, 2011). It is clear that although dysregulation of BRCA1 has a profound impact on transcriptional regulation, much of this is indirect (Gorski *et al.*, 2011). It is possible that BRCA1 facilitates the fine-tuning of transcriptional activation or repression in response to different external stimuli, through phosphorylation-mediated assembly of various transcriptional complexes (Gorski *et al.*, 2011) as illustrated in Figure 2.2. BRCA1 may clearly play a role in regulating the transcription of several genes through its multitude of interactions as illustrated in Figure 2.2.

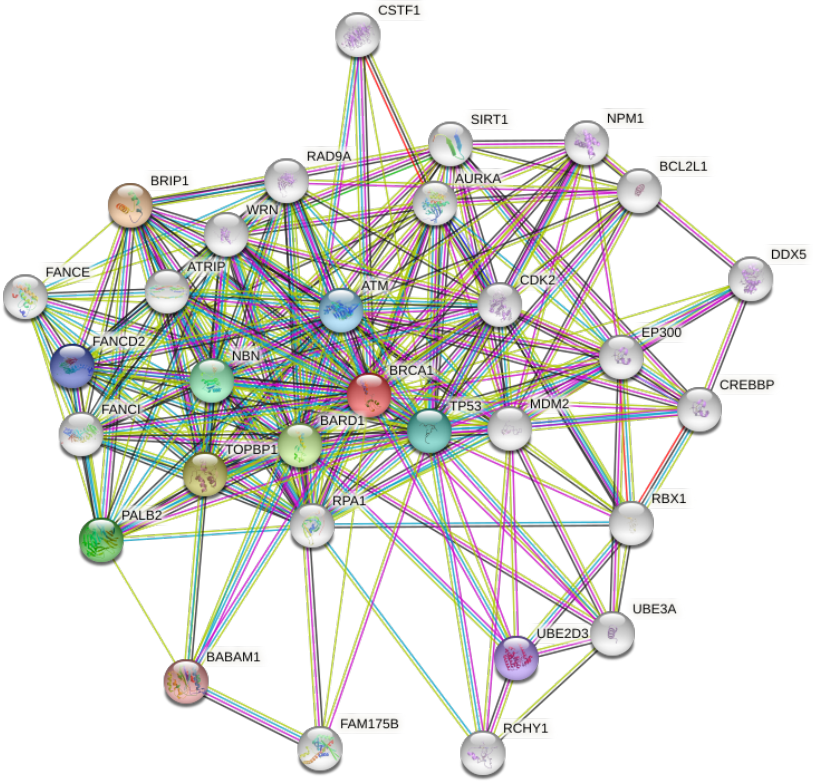


Figure 2.2 Network of selected proteins that BRCA1 interacts with. Obtained through STRING database (https://string-db.org) Accessed on 26th May 2022.

2.2.2 CELL-CYCLE CHECKPOINT ACTIVATION

BRCA1 through its association with specific proteins and multi-protein complexes has also been proposed to act as a sentinel of the normal cell cycle control (Christou and Kyriacou, 2013). One molecular mechanism by which genomic integrity and stability are maintained is mediated by a cellular network of signaling events and the DNA Damage Response (DDR), triggered in response to genotoxic stress (Roy, Chun and Powell, 2016). The DDR includes the activation of checkpoints that delay the cell cycle before or during replication (G1/S or intra S-phase checkpoints) or before cell division (G2/M checkpoint) to ensure that genetic errors are not transmitted to subsequent generations by allowing time for DNA repair (Roy, Chun and Powell, 2016). BRCA1 has thus been shown to play a major role in cell-cycle checkpoint activation by forming complexes and phosphorylating or ubiquitinating other proteins involved in checkpoint control.

The association of BRCA1 with ATM/ATR, CHK2 and Aurora A protein kinases regulates cell cycle progression by affecting its post-translational modification (i.e., phosphorylation by ATM) and localization within a cell (Christou and Kyriacou, 2013). This allows BRCA1 to mediate the ordering, timing, and progression of the cell cycle based on its cellular localization. The G1/S-checkpoint requires phosphorylation of BRCA1 by ATM or ATR, which facilitates phosphorylation of p53 on S15 and BRCA1 regulates the intra S-phase checkpoint by modulating CHK1 phosphorylation (Roy, Chun and Powell, 2016; Masuda *et al.*, 2016). The exact mechanism of how phosphorylation of CHK1 is affected by DNA binding of BRCA1 remains unclear (Masuda *et al.*, 2016; Deng, 2006). BRCA1 also displays E3 ubiquitin ligase activity which has been implicated in the regulation of the G2/M cell cycle checkpoint (Shakya *et al.*, 2011). Through this activity the polypeptide has been hypothesized to contribute to maintenance of genomic stability by regulating the presence, activity and degradation of several proteins involved in regulating the cell-cycle (Shabbeer *et al.*, 2013).

2.2.3 DNA REPAIR

Detection and faithful repair of damaged DNA is essential for genome integrity (Ceccaldi, Rondinelli and Andrea, 2016). DSBs are lesions formed when both strands of the DNA duplex are broken and two major DSB repair pathways that have been studied extensively are homologous recombination (HR) and nonhomologous end-joining (NHEJ) (Chapman, Taylor and Boulton, 2012). BRCA1 is a versatile protein that links DNA damage sensing and DDR effectors (Roy, Chun and Powell, 2016). Among its various DDR functions, the polypeptide normally promotes error-free, homologous recombination-type, DNA damage repair (HRR) (Hu *et al.*, 2014). HRR or HR is activated by DNA damage, and it is a vital DNA repair mechanism that uses the undamaged sister chromatid to carry out high-fidelity repair of predominantly replication associated DSBs (Roy, Chun and Powell, 2016).

The association of BRCA1 with RAD51 has a direct impact on the repair of DNA double-strand breaks (DSBs) by homologous recombination (HR) (Christou and Kyriacou, 2013). The main role of RAD51 is catalyzing the core reactions of HR, including strand invasion into duplex DNA and enabling strand exchange through the pairing of homologous DNA strands. BRCA1 is required for

RAD51 recruitment to sites of DNA damage through its interactions with PALB2 and BRCA2 (Roy, Chun and Powell, 2016). The defining step of HR is homologous strand exchange, including strand invasion into duplex DNA and the pairing of homologous DNA strands enabling strand exchange, mediated by the RAD51 protein (Prakash *et al.*, 2015; Roy, Chun and Powell, 2016; Bhattacharya *et al.*, 2017). PALB2 is a binding partner of BRCA2 that is responsible for its localization to the nucleus during HR (Roy, Chun and Powell, 2016; Nepomuceno *et al.*, 2017). BRCA1 is recruited to foci independently of BRCA2 and PALB2, yet BRCA1 influences PALB2 accumulation, which in turn promotes BRCA2 recruitment to DSBs (Chapman, Taylor and Boulton, 2012). The recruitment of the RAD51 strand exchange protein to repair foci requires all three proteins, and consequently loss of any of these proteins impairs HR (Chapman, Taylor and Boulton, 2012).

The role of BRCA1 in DNA repair, in particular in HR repair, has thus far been one of its most well-recognized functions thought to be critical for its tumor suppressor activity (Prakash *et. al*, 2015). Compelling evidence suggests that BRCA1 functions in two distinct steps: (i) 5′ to 3′ resection of DSBs to generate 3′ single stranded DNA (ssDNA) overhangs, and (ii) loading of the RAD51 recombinase onto the ssDNA (Prakash *et al.*, 2015). BRCA1 interactions with the MRN complex or proteins, with the BRCC complex of proteins that exhibit E3 ligase activity and with the phosphor proteins CtIP, BACH1 (BRIP1) and Abraxas (CCDC98) have also been implicated in DNA repair mechanisms (Christou and Kyriacou, 2013).

2.3 *BRCA1* MUTATIONS

BRCA1 gene mutations have been associated with increased risk of certain types of cancers (Clark *et al.*, 2012; Chen, Roberts and Klein, 2017; Kim *et al.*, 2020). A harmful mutation in BRCA1 can result from errors in DNA replication during cell division, exposure to mutagens or a viral infection (Clark *et al.*, 2012; Roberts and Klein, 2017; Kim *et al.*, 2020). They can be inherited from either parent and each child of a parent who carries any mutation in one of these genes has a 50% chance (or 1 in 2 chance) of inheriting the mutation. In the presence of a *BRCA1* mutation, women have a 70-80% lifetime risk of developing breast cancer and a 50% risk of developing ovarian cancer (Petrucelli, Daly and Pal, 1998; Roy, Chun and Powell, 2016; Chehade *et al.*, 2016). Of the 1700 unique *BRCA1* mutations reported to the Breast Cancer Information Core Database, 858 have been

confirmed as being "Clinically significant" causing an increased risk of cancer and resulting in a protein with reduced function or no protein product (Clark *et al.*, 2012). From a medical standpoint, the key function of the BRCA1 protein is its ability to suppress tumor formation in breast and ovarian tissues (Shakya *et al.*, 2011). *BRCA1* is most often mutated in three domains or regions: the N-terminal RING domain, exons 11-13, and the BRCT domain (Chehade *et al.*, 2016; Clark *et al.*, 2012). Several key BRCA1 functional domains, including the N-terminal RING domain and the C-terminal BRCT tandem repeats, promote HR and are necessary for tumor suppression (Moynahan and Jasin, 2010).

The BRCT region has a functional duplicated globular domain, which helps with DNA damage repair and cell cycle checkpoint protein control (Chakraborty *et al.*, 2013). Most pathogenic *BRCA1* lesions are frameshift or nonsense mutations that would eliminate or grossly disrupt the BRCT domain (Billing *et al.*, 2018). However, in some families, tumor susceptibility can be attributed to a single amino acid substitution in BRCA1, often involving residues within its BRCT domain (Billing *et al.*, 2018). Because most tumor-associated *BRCA1* alleles have frameshift/nonsense mutations that eliminate one or both BRCT motifs, BRCT phospho-recognition may be critical for tumor suppression (Shakya *et al.*, 2011). Exon 19 codes for amino acids 1732-1759 which encompasses the 22-amino acid linker with a small portion of the BRCT1 and BRCT2 regions (Cunningham *et al.*, 2022; Clark *et al.*, 2012). Cancer-causing missense mutations at this interface between the two repeats have been hypothesized to destabilize the structure of BRCA1, thereby affecting its role as a tumor suppressor (Williams, Green and Glover, 2001; Glover, Williams and Lee, 2004). ClinVar lists various mutations in exon 19 such as V1736G, R1737Ter, G1748D etc., which have all been identified as pathogenic and shown to influence the structure and function of BRCA1 (https://www.ncbi.nlm.nih.gov/clinvar/). Therefore, it is necessary to understand how genome alteration is related to the development of cancer (Khodadadian *et al.*, 2020).

2.4 RESEARCH GAPS

More than 600 distinct pathogenic single nucleotide variants with the potential for structural and functional effects on BRCA1 have been documented in ClinVar (Tsai *et al*., 2019; Findlay *et al.*,

2014; Lee *et al*., 2010). Some of these mutations have been shown to affect the BRCT linker region of BRCA1 which is critical for its role in tumor suppression because it helps with DNA damage repair and cell cycle checkpoint protein control. It is clear that late diagnosis in young patients has been linked to high death rates from BC in resource-poor countries and to date, genetic data of *BRCA1*, i.e., a gene in which mutations have been linked with high risk of BC susceptibility, from young female BC patients in Zambia has not been evaluated. This work was therefore undertaken to assess genetic alterations within a selected domain of *BRCA1* for the first time in a cohort from Zambia.

CHAPTER THREE

MATERIALS AND METHODS

3.0 OVERVIEW

This chapter describes the different methods and approaches used to execute this study. The methods of this study were based on the set objectives of the research presented in Chapter One of this dissertation and are presented under the following headings: (1) Location (2) Inclusion/exclusion criteria, assumptions and sample size (3) Ethical considerations (4) Data Requirements (5) Data Collection (6) Data Analysis Methods (7) Sample demographics and Statistical analysis and (8) Limitations of the Study.

3.1 LOCATION

This study was conducted at Matero Level One Hospital in Lusaka, which has a referral breast cancer clinic. Based on positive ultrasound-guided core biopsy results, female breast cancer patients were invited to take part in the study.

3.2 INCLUSION/EXCLUSION CRITERIA, ASSUMPTIONS AND SAMPLE SIZE

3.2.1 Inclusion Criteria and Assumptions

i) Female patients that have been clinically diagnosed with breast cancer at a young age (≤ 50 years) irrespective of family history of breast cancer (Appendix IV). The selected age range was based on the assumption that risk for breast cancer increases with age and reproductive history (Center for Disease Control and Prevention, 2020). Previous data also indicates that approximately 60% of breast cancer cases detected at Matero Level One Hospital are in women younger than 50 years. (Songiso *et al.,* 2020).

3.2.2 Exclusion Criteria

i) According to known clinical records, patients in the selected cohort that presented with underlying conditions like HIV, cardiovascular disorders, diabetes mellitus and chronic obstructive pulmonary disease were excluded from this study to prevent potential confounders that may corrupt the causal relationship of interest.

3.2.3 Sample Size

i) From September 2021 to January 2022, a total of 75 young patients (≤ 50 years) with invasive breast cancer were referred to the breast cancer clinic at Matero Level One Hospital. Seventy-three participants agreed to take part in the study.

ii) Ten participants were selected randomly and representative of the following age groups: 21-30 years, 31-40 years and 41-50 years (Table 4.3). Because this was a pilot study, the selected sample size was justified by its nature and scope (Martínez-Mesa *et al.*, 2014).

3.3 ETHICAL CONSIDERATIONS

Ethical clearance was obtained from the relevant organizations and informed consent was obtained from all participants as described below.

3.3.1 ETHICAL CLEARANCE

This study was conducted after seeking ethical clearance and approval from the following ethical bodies:

i. The University of Zambia (UNZA) Natural and Applied Sciences Research Ethics Committee (NASREC).
ii. The Zambia National Health Research Authority (NHRA).

Ethical clearance certificates and letters are presented in Appendix I and II, respectively.

3.3.2 INFORMED CONSENT PROCEDURES

An informed consent form was used to obtain permission from all participants in this study according to guidelines set by UNZA NASREC. Results were treated with confidentiality and were not included in participants' medical records.

3.4 DATA REQUIREMENTS

To achieve the objectives set out, work was conducted to generate information and results reported herein, via extraction of genomic DNA (gDNA) from human whole blood samples and authentication of its quality by NanoDrop spectrophotometry and agarose gel electrophoresis. Amplicons of a target region within the C-terminal domain of *BRCA1* were generated by the polymerase chain reaction (PCR) and sequencing of *BRCA1* C-terminal domain of PCR products was conducted. Nucleotide sequence data was analyzed using bioinformatics tools available at different web-based platforms.

3.5 DATA COLLECTION

The materials and methods used in the collection of the data listed in section 3.4 are described below.

3.5.1 GENOMIC DNA EXTRACTION

Blood for genomic DNA extraction was collected in 4 mL purple top Ethylene diaminetetraacetic acid (EDTA) Vacutainer® tubes from participants who agreed to take part in the study by way of informed consent. It was transported to the laboratory at the School of Veterinary Medicine of the University of Zambia for storage in a -20ºC refrigerator until use. Demographic information and date of collection were noted for each participant. Genomic DNA was extracted from blood samples by the Zymogen Research Quick-DNA™ Miniprep kit (Catalog # D3024) according to Zymogen Research instructions. Briefly, 100µL of whole blood from each sample collected from participants was mixed with 400µL of genomic lysis buffer from the genomic DNA extraction kit.

The mixture was vortexed and transferred to spin columns placed into labeled 1.5 mL collection tubes. These were spun at 12,000 revolutions per minute (rpm) in a Qiagen microcentrifuge at room temperature (rtp) for one minute. Spin columns were then transferred to new labeled 1.5 mL Eppendorf tubes and 200µL of DNA pre-wash buffer was added to each column and spun at 12,000 rpm in a Qiagen microcentrifuge at rtp for one minute. 500µL of g-DNA wash buffer was added to each spin column and spun at 12,000 rpm at rtp for another minute as described. The wash columns were then transferred to new labeled 1.5 mL microcentrifuge tubes and 60µL of DNA elution buffer was added to each column and these were again spun at maximum speed for 30 seconds. Genomic DNA was stored in the DNA elution buffer and stored at -20ºC until further analysis.

3.5.2 DNA CONCENTRATION DATA

The quality and concentration of the extracted gDNA was characterized using NanoDrop spectrophotometer (DeNovix, Wilmington, DE, United States) and details of the DNA quality were recorded in table 4.1 and 4.3 in Chapter 4. Sample trace data obtained using the NanoDrop spectrophotometer is shown below.

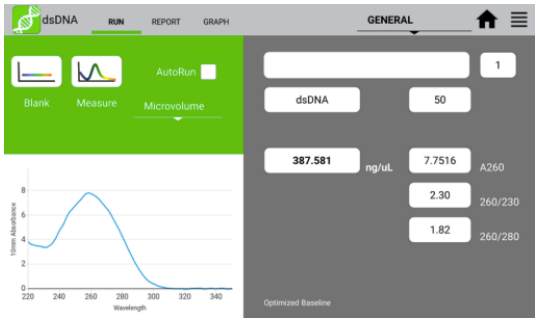


Figure 3.1 Sample trace using the NanoDrop spectrophotometer.

3.5.3 PCR AMPLIFICATION OF GENOMIC DNA

The protocol by Chakraborty *et al*. (2013), was used to PCR-amplify exon 19 in the *BRCA1* gene using oligonucleotides 5'-ATATGACGTGTCTGCTCCAC-3' (forward) and 5'-

GGGAATCCAAATTACACAGC-3' (reverse) (Chakraborty *et al.*, 2013). These primers were selected because they were previously shown to be effective in the amplification of a 401bp region of DNA containing exon 19. The expected amplicon size yielded by the specific primers was 401bp as shown in figure 3.1. The PCR reaction was set up using One Taq® Quick-Load® 2X Master Mix with Standard Buffer (Cat # M0486S). The thermal cycle profile included an initial denaturation step at 94°C for 2 minutes and, thirty amplification cycles that included a denaturation step of 30 seconds at 94°C, primer annealing at 45°C for 30 seconds and primer extension at 72°C for 30 seconds. A final primer extension step at 68°C for 5 minutes was included. PCR amplicon size was determined on a 2% agarose gel stained with EZ-vision® Bluelight DNA Dye and 0.05 µg/mL ethidium bromide final concentration. The GeneRuler 1000bp DNA Ladder (Cat# SM0312) was used on all gels as size standard. Passing criteria included a clear gel image with no signs of genomic DNA degradation.

Exon 19 of BRCA1

(RefSeq: NM_007298.3)
401 bp fragment
(i.e. position 1882-1965; aa 628-655 Refer to link below)

TGTCTGCTCCACTTCCATTGAAGGAAGCTTCTCTTTCTCTTATCCTGATGGGTTGTGTTTGGTTTCTTTCAGCA
TGATTTTGAAGTCAGAGGAGATGTGGTCAATGGAAGAAACCACCAAGGTCCAAAGCGAGCAAGAGAATCCCAGG
ACAGAAAGGTAAAGCTCCCTCCCTCAAGTTGACAAAAATCTCACCCCACCACTCTGTATTCCACTCCCCTTTGC
AGAGATGGGCCGCTTCATTTTGTAAGACTTATTACATACATACACAGTGCTAGATACTTTCACACAGGTTCTTT
TTTCACTCTTCCATCCCAACCACATAAATAAGTATTGTCTCTACTTTATGAATGATAAAACTAAGAGATTTAGA
GAGGCTGTGTAATTTGGATTCCCGTCTCGGG

Figure 3.2 401 nucleotide fragment containing exon 19 (highlighted in purple) using specific primers. Sequence obtained from the University of California Santa Cruz's (UCSC) genome browser. Accessed on 10th June 2022.

Of the genomic DNA obtained from samples submitted by the seventy-three participants, ten were randomly selected for sequencing of the exon 19 *BRCA1* region. These were ratio representatives of the following age groups: 21-30 years (1), 31-40 years (4) and 41-50 years (5) as shown in table 4.3. The total number of samples for each group was four, twenty-eight and forty-two, respectively. The amplicons obtained from PCR amplification of these participants genomic DNA was extracted from the agarose gel slab and purified using the QIAquick® Gel Extraction Kit (Cat # 28706X4) in preparation for sequencing. Briefly, gel slices containing DNA bands from each of the ten

participants were excised with a clean, sharp scalpel. Gel slices were weighed in labeled Eppendorf tubes and 2 volumes of diffusion buffer were added to 1 volume of gel i.e., 200µL diffusion buffer for each 100mg of gel. Labeled Eppendorf tubes were incubated on a heating block at 50°C for 30 minutes to allow agarose to melt and later span at 12,000 rpm for 1 minute. The supernatant was carefully transferred to labeled spin columns placed in collection tubes using a disposable syringe containing a Whatman GF/C filter to remove any residual agarose. The volume of recovered supernatant, containing amplicon DNA, was determined. Three volumes of Buffer QG were added to each spin column to one volume of supernatant. The color of the mixture was determined to be yellow showing no residual agarose was present. The labeled spin columns were span at 12,000 rpm at rtp for 1 minute and the flow through was discarded. The column was washed with 750µL of Buffer PE and span again as described. The QIAquick spin columns were transferred to new labeled Eppendorf tubes and 50µL of Buffer EB (10mM Tris-Cl, pH 8.5) was added to the center of the spin column. To elute DNA, after 1 minute, the spin columns were centrifuged at maximum speed. The quality and concentration of the purified amplicons was characterized using NanoDrop spectrophotometer (DeNovix, Wilmington, DE, United States) and details of the DNA quality were recorded in table 4.3 in Chapter 4. A 260/280 ratio of ~1.8 was generally accepted as "pure" for DNA i.e., above the quality threshold and amplicons from the ten participants were then considered ready for sequencing.

3.5.4 SEQUENCING DATA OF PCR AMPLIFIED PRODUCTS

The amplicons were submitted to Inqaba Biotechnical Services Pvt in the Republic of South Africa where they were sequenced in the forward and reverse direction using the ABI 3730xl Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific).

3.6 DATA ANALYSIS METHODS

Sequences were trimmed and contiguous sequences were generated in BioEdit using the forward and reverse sequences. The data that was obtained from BioEdit contained nucleotide sequences for each sequenced sample. GenScript was used to obtain reverse complements of sequences if needed.

3.6.1 MUTATION IDENTIFICATION

Exon 19 of *BRCA1* for each sample was compared to the wild type nucleotide consensus sequence (NM_007298.3) obtained from UCSC's genome browser. A nucleotide Basic Local Alignment Search Tool, commonly known as BLAST (Altschul *et al.*, 1990) was utilized to align sequences from each participant with the wild type consensus nucleotide sequence. This was to confirm whether or not the amplified region was exon 19 from *BRCA1*.

3.6.2 SWISS-MODEL

DNA sequences with putative mutations were translated into amino acid (protein) sequences using Ensembl (https://www.ensembl.org) and these were then compared to the wild type consensus protein sequence (NP_009229.2) obtained from Ensembl. SWISS-MODEL (https://swissmodel.expasy.org/) was utilized to generate reliable three-dimensional protein structure models for protein sequences that contained mutations in the exon of interest. This was achieved by inputting the protein sequences into SWISS-MODEL and selecting the most likely three-dimensional model based on statistical data e.g., Qualitative Model Energy ANalysis (QMEAN) z-score, provided by SWISS-MODEL (Biasini *et al.*, 2014). Briefly, the QMEAN scoring function allowed for the identification of the native structure of a protein and allowed for the discrimination of good models from bad ones based on multiple databases (Biasini *et al.*, 2014). QMEAN z-scores around zero indicated a good agreement between the model structure and experimental structures of similar size (Swiss model, 2023). The use of SWISS-MODEL over other modeling software's is justified by the fact that the expected accuracy of each specific model was communicated to the user and the overall accuracy of SWISS-MODEL was continuously monitored over time (Biasini *et al.*, 2014). Finally, generated models were compared to the wild-type three-dimensional protein structure of BRCA1 with the assumption that all the other exons of mutant *BRCA1*, if translated, remained the same. Results were used to infer structural changes conferred by mutations, if any.

3.6.3 VARIANT ANALYSIS

The clinical significance of identified mutations or variants was assessed based on information from a major BRCA1 database, ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). Results were used to infer functional changes, if any.

3.7 SAMPLE DEMOGRAPHICS AND STATISTICAL ANALYSIS

Sample demographics including age, sex, ethnicity, and geographic location of participants were collected. Statistical analysis was performed using Microsoft Excel (Version 16.27) to test the significance of the null hypothesis previously described in chapter one, section 1.6. Briefly, the unpaired two-tailed student's *t*-test equation (=TTEST(array1,array2,1,3)) was utilized for this study to test the null hypothesis that there was no significant association between the age of a BC patient and the number of mutations in exon 19 leading to a loss of function of BRCA1. For this study, the use of the student's t-test is justified because it is a commonly used statistical test for reasonably small sample sizes (less than approximately 30) (Science Direct, 2023). P-values of <0.05 yielded by the unpaired two-tailed student's *t*-test were considered significant.

3.8 LIMITATIONS OF THE STUDY

The researcher faced some challenges during the data collection process due to the following reasons:

i. Limited time and resources to collect a representative sample size of the young female breast cancer patient population evaluated at Matero Level One Hospital. Nurses and doctors were only available for the indicated period of time (September 2021 to January 2022).
ii. Limited funding to sequence all the collected samples.

CHAPTER FOUR

RESULTS

4.0 OVERVIEW

This chapter presents the findings of this study and are presented under the following headings: (1) DNA concentration data (2) PCR amplified products (3) Sequencing data of PCR amplified products (4) Mutation identification and SWISS-MODEL data (5) Variant analysis and (6) Sample demographics and Statistical analysis.

4.1 DNA CONCENTRATION DATA

Genomic DNA was extracted from peripheral whole blood samples collected from study participants and its concentration was characterized by NanoDrop spectrophotometry. Table 4.1 shows results obtained for each sample i.e., gDNA volume (µL), concentration (ng/µL), mass (ng) and the ratio of absorbance at 260 nm and 280 nm (A260/A280). The concentration of the extracted genomic DNA ranged from 1.556ng/µL to 40.681ng/µL and the quality based on A260/A280 ratio was deemed acceptable for the majority of the extracted samples. A 260/280 ratio of ~1.8 was generally accepted as "pure" for DNA and subsequent purification was conducted if DNA was considered contaminated for sequencing purposes. DNA that contained contaminants was purified as detailed in section 3.5.3 and the absorbance measured again by NanoDrop spectrophotometry as shown in table 4.3.

Table 4.1 gDNA concentration data determined by NanoDrop spectrophotometry.

Sample No	DNA Volume (µL)	DNA Concentration (ng/ µL)	DNA Mass (ng)	A260/A280
1	60	11.924	715.44	1.55
2	60	6.489	389.34	1.78
3	60	6.065	363.90	1.37
4	60	7.433	445.98	1.86
5	60	8.843	530.58	1.52

6	60	18.139	1088.34	1.69
7	60	16.013	960.78	1.72
8	60	10.404	624.24	1.51
9	60	10.747	644.82	1.51
10	60	12.265	735.90	1.50
11	60	12.791	767.46	1.64
12	60	31.923	1915.38	1.79
13	60	7.345	440.70	1.63
14	60	2.391	143.46	1.09
15	60	8.045	482.70	1.67
16	60	11.428	685.68	1.52
17	60	1.556	93.36	0.58
18	60	5.648	338.88	1.48
19	60	14.868	892.08	1.28
20	60	20.333	1219.98	1.71
21	60	7.852	471.12	1.62
22	60	40.681	2440.86	1.40
23	60	9.110	546.60	1.57
24	60	10.814	648.84	1.73
25	60	10.105	606.30	1.62
26	60	11.391	683.46	1.68
27	60	5.026	301.56	1.53
28	60	3.374	202.44	1.40
29	60	32.575	1954.50	1.82
30	60	12.292	737.52	1.64
31	60	22.920	1375.20	1.70
32	60	23.000	1380.00	1.65
33	60	19.965	1197.90	1.66
34	60	7.961	477.66	1.39
35	60	9.051	543.06	1.57
36	60	4.931	295.86	1.58
37	60	17.244	1034.64	1.02
38	60	5.642	338.52	1.30
39	60	5.102	306.12	1.37
40	60	6.133	367.98	1.19
41	60	10.603	636.18	1.19
42	60	7.363	441.78	1.55
43	60	15.866	951.96	1.63
44	60	29.867	1792.02	1.75

45	60	13.377	802.62	1.76
46	60	25.584	1535.04	1.82
47	60	10.773	646.38	1.80
48	60	16.755	1005.30	1.65
49	60	22.492	1349.52	1.88
50	60	15.647	938.82	1.05
51	60	6.168	370.08	1.39
52	60	16.309	978.54	1.70
53	60	9.636	578.16	1.83
54	60	12.400	744.00	1.52
55	60	10.978	658.68	1.65
56	60	10.364	621.84	1.90
57	60	8.077	484.62	1.49
58	60	2.987	179.22	1.22
59	60	4.758	285.48	1.59
60	60	11.078	664.68	1.64
61	60	30.085	1805.10	1.73
62	60	14.195	851.70	1.67
63	60	6.321	379.26	1.60
64	60	10.971	658.26	1.60
65	60	3.817	229.02	1.72
66	60	2.757	165.42	1.39
67	60	4.214	252.84	1.48
68	60	4.794	287.64	1.50
69	60	8.935	536.10	2.05
70	60	4.815	288.90	1.56
71	60	15.435	926.10	1.81
72	60	23.944	1436.64	1.71
73	60	22.559	1353.54	1.64

4.2 PCR AMPLIFIED PRODUCTS

Exon 19, which codes for part of the BRCT domain of BRCA1, was amplified using a polymerase chain reaction platform. The reaction volume for each sample was 25µL and the annealing temperature was set at 45℃ for 30 cycles. A 401bp fragment was obtained for each sample and the amplicons were visualized on a 2% agarose gel stained with EZ-vision® Bluelight DNA Dye and ethidium bromide at final concentration of 0.05 µg/mL. The results of the amplification

reactions were captured on a gel documentation system (UVP 2UV Transilluminator LM-20) and are presented in Figure 4.1A to 4.1C.

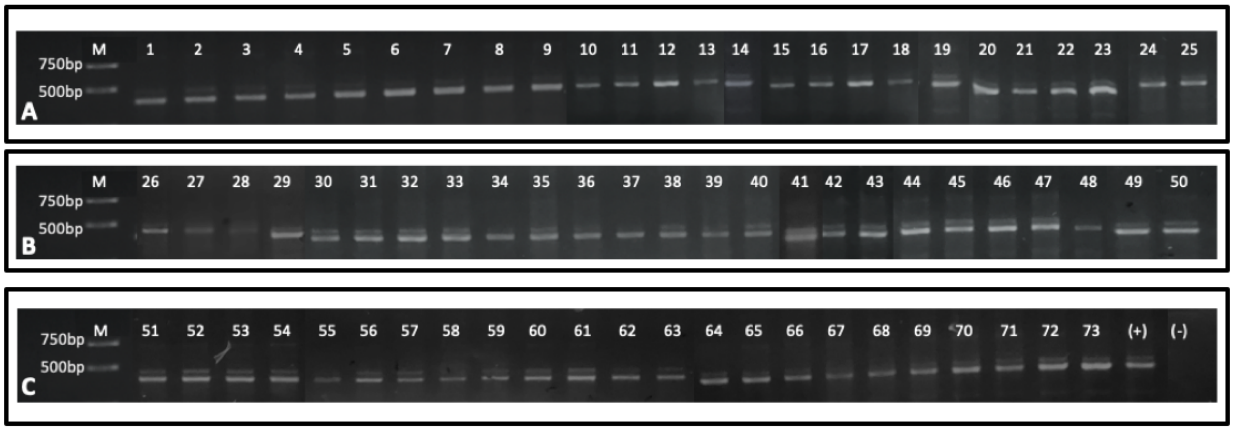


Figure 4.1 Agarose gel electrophoresis analysis of amplicon DNA samples 1-25 (A), 26-50 (B) and 51-73 (C) using exon-19-specific primers (Chakraborty *et al.*, 2013) on ethidium bromide. M= 1000bp molecular weight marker; (+) = positive control gDNA; (-) = negative control in which all reagents were added to the PCR tube except template DNA which was replaced with molecular grade nuclease-free water; All seventy-three isolates included in this study showed similar band migration profiles. Samples 28, 41, and 55 had limited amounts of amplified DNA as denoted by the faint bands but this was considered sufficient for sequencing by Inqaba Biotechnical Services Pvt.

4.3 SEQUENCING DATA OF PCR AMPLIFIED PRODUCTS

PCR products from ten of the seventy-three participants were sequenced at Inqaba Biotechnical Services Pvt., in the forward and reverse direction using the ABI 3730xl Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific). The contig-generating tool in BioEdit was used to generate the contiguous sequence for each sample, using the forward and reverse sequences, and the following results were obtained (Figure 4.2):

DS16
>Contig-1ds16
GTTACTGATACTATATCGACGCGGCTCTGCATAATGTAGGTGGGACGCTCCACACAAATGGAAC
TTGTGACAACGATCTAGCAGTGAGTATGTTGAAATAAGTGGAACAATATGAAGCGGCCTTCTCT
GCCTGCCTTAGGGAATATTGCCCTTGGGTCGACATTTTATTTATCATGAGGGAGAAGGCCTTAC
CTGTCTGATCTAGCTCTGTCACGCTCTGCTTGTGCCTTGGGTGCTGCTTACACTATGCGATCTC
CTCCGAATCACACTTCATGACGAAACAAACCAATCGCAACCCTTCATTGTGAGAGAAGGACAAG
CTCCCCTTCCGGGAAGGAGAACAGACAGGGATTACAAACTCACATGTCTCTGGGATGGAGCACC
CAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATA

DS19
>Contig-1DS19
AAGGATCTCTGCTACTTGCATGAGGCTCTTCTGATATATGTGTGTGTGATGGAACCCTCAACAA
TGGTTAGTGGGCTGTATCTATCACAGGGTATGTGGACATAAGTTTTGGAAATGAGAGGCCTTTC
TTGTCTCGCTAGGGGAGTACTGACCTTGGTTCGACATATAAGATATCTCGGAGCCGGAAGCCCT
ACATTTCTTATAAGCGACTCCCTTGCTCTGTATGTGACTCGGGGACAACTTCCTTCGTCTCCCT
CTGGTCCCGCACCAACTCCATGACGAATCAACCCAAAAGCAACCCTTCAGGGTAACAGAAAGAC
AAACCCTCCTTCCTGGAACGGTAACACAGACTGCTTGTAAACTCACATGTCTCTGGGATGGAGC
ACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS26
>Contig-1DS26
GATCTAGGTACACTGCAGCGGCACTACCTCCTGATCTAGGGTCTCGATGCAACCCACCACTGGA
GTCTGTGGGATAGTGCTCACAGAGTGTACGCGGTCATAAGTGCAGGAAAAGTGAAGTCCCGTCT
TGACAGCTATAGGGAGATCTGCCCTTGCTATCGTCATAATAGTTAACTTGAGGGCGGAGACTTT
ACCTGTCAGATCAGGACCTCCCATGCTCAGTTTGTGACCCAAGGAACAACCCCATTCGTCTAAC
TCACGTACCACATCCACTCCACTCCTCAACACCCCAAACGCAACCCTGGATTGTAACAGAAACA
GAAACCCTCCTTCCTGGAACGGTAACACAGACTGATTGCAAACTCACATGTCTCTGGGATGGAG
CACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATA

DS41
>Contig-1DS41
GAACTCTGATACTTGATGAGGCAGATCCGACTAGACTGTGTGTCTAGCATCCCACCACTGGAGC
TATATGGCAATAATCATCGACTGTGTCTGTGGATATAACTTTAGGTTGTATGAAATCCTTTCAG
CCACTGTCATTGGGTGGCACTGCCCTTTCTATCCTAGCAAAAGACAACACGTATGCAAAAGCTG
GACGTGTCTTATCTACACCTCCCTAAGGTGTTGGGACTTAAAGAACAAGTCACTCCATGTGACG
CTGGTACCGCTTAAACTCCACTGCTCATGACACCTCCGAGAACTATGCCTTGAGACAACAGCAC
AGACCCTCCATCCTAGGCCGGTAGCACTGACTGATTGTAAACTCACATGTCTCTGGGATGGAGC
ACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS55
>Contig-1DS55
AGGGGGGCATATCATTCATAATTGCATAGACAATACTTATTTATGTGGTACTGATGGAAGAGAG
CAAAGAGAACCTGAGTGAAAGTATCTACCACAAGTTATGTAGGAAATCGCTCTAACGAACTGAA
TAGGCCCGTCTAGACCACGCGGAGTGGGATACATAGTGTTGCTTCCAGACTTTAGTCAACTTGA
GGGAGGGAGCTTTACCTTTCTGACCTGGGATTGCCATGCTCTGTATGTCCCTTGGGGGCTTCTC
CCATTGTCCAGATCTCCTCTGACTTCCAAATCATGCTGAAAGAAACCTAACGCAACCCTGCATG

ATGCGAGAAAGACAAACTTCCTTCAATGGAAGTGGAACACACACGGAATATAAACTCACATGTC
TCTGGGATGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS56
>Contig-1DS56
TTATCTTTTGCTACTTGTTTGAGACAGTACCTACTGATGTGGTACTGATGCAAGAGTGAAAAGA
GAGTTAGTGTGCAAGTATCTATCACTAGGTACTTTGTCATAAGTCTAACAAACAGAAGCAGCCC
ATCACTGCAACGGGTTGTGGAATATATAGCGGTGGATCGAGACTTTAGTCATCTTGAGGGAGAA
AGCTTTACCTTTCTGTCCTGGGATTCTCTTGATCTCTTTGGACCTTGGGGGCTACTCACATTGT
CCACACCACCTCTGACTTAAAAATCATGCTGAAAGAAACCAATCGCAACCCATCATGGTAACAG
AAAGACAAACTTCCTTCAATGGAACTGGAACACACACGGAATACAAACTCACATGTCTCTGGGA
TGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATAC

DS60
>Contig-0DS60
AGTCCGTGGGAATCCAAATTACACAGCAGCCCTTCAGAAAAGTGGCCAGACTCTGTACTTGATG
GGCAGATCCTCCTTGGGAATCCAAATTACACAGCCTCTCTAAATCTCGTAGTCTTATCATTGCA
TAAAGTAGAGACAATACTTATTTATGTGGTTGGGATGGAAGAGTCAAAAAAGAACCTGTGTGAA
AGTATCTAGCACTGTGTATGTATGAAATAAGTCTTACAAAATGAAGCGGCCCATCTCTGCAAAG
GGGAGTGGAATACAGACTGGTGGGGCGAGATTTTAGTCAACTTGAGGGAGGGAGCCCTACATTT
CATGTCCTGGGATTCTCTTGCTCACTATAGGACCGTCGGGGTTGCTTACATCGACCACATCTCC
TCTGACTTCAAAATCATGATGAAACAAACCAAACAGCAACCCATCAGGAGAAGAGAACAGGTAC
AAGCTTCCCTCAATGGGAGTGGAGCACAGACGTCATATAAACTCACATGTCTCTGGGATGGAGC
ACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATAAAGCTGTTC

DS62
>Contig-1DS62
GTGGGGCATACTCTGATACTTGATGAGACAGTACTTACTGGTCTAGGTGTTATGGAAGCCTCAG
ATGAGAACCAGTGTGAAGGTTCTATCACTGTGTATGTTGTAATAGCTCTTACGAAATGAAGCGG
CCCATCTCTGCCTAGGGGTATGGAATACATGGCCGTGGGGCGACATATGTGTCATCTTGAGGGA
GGGAACTTTACGTTCCTGTCCAGGACTTCTCTTGAGCGCTTTGGACCTCGGTGGCTACTCCCAT
TGTCCACACCACCTATGACTTACAAATCATGATGAATCAAACCACACGGAACCCAGCATTGTAA
CAGAAAGACAAACTCCCTTCCCTGGAACTGGAACACACACGGATTACAAACTCACATGTCTCTG
GGATGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATAA

DS63
>Contig-1DS63
TATAACTCTGATACTTGATGGGCAGTATCCTACTGACCTATTTGTCTAGGTGCCCAGCACAGGA
GCTATATATGAAAGATCTCTCCTGTGTACGTGGAATAGCTTTTACGAAATGAGAAAGCCTTTCC
TGTCTGCTAGATGGAATACTTGCCTAGGAGGCACATAATAGTTAACTCGGGGCCGAAAACTCTA
CCTGTCTGATCTGCGCCTCCATTGATGTCTTTCTGACCAAAGGGGCTTCTCAATTCATCTACCT
CTGGGCCCACACCCACTCCGTTACGAATCACCCCTCAAGCAACCCTGGATGATGACAAAAGCAC
AAACCCTCCTTCCTGGAACGGTAACACAGACTGCTTGTTAACTCACATGTCTCTGGGATGGAGC
ACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS65
>Contig-1DS65

AAATCATCTGCTACTATCACATGCAATTCTTACTGTCCTGTGTGTCTAGGCAATACACCAATGG
AGCTTTCTGAGATTGGTATACTCAGTGGGTACGTGTACCTAATTTTTGCAATGTGACGAAGGCC
CTTTAGGAGAGTGTATGGGGGATAATATCCCTTGCTATCCACATATTATGTATCATGGGGGCCA
GAGCTCTAGACGTCAGATCTGCACGTCCCATGGTCTGTTGGTAACTCAGGGACTGGTCACGTCG
TCTCCCTTCGTCCGATCCCCACTCCGTTACGCATCACCCCTAAAGCAACCCTGCATGATAACCA
AAACACAGACTCTCCTTCCTGGAACGATAACACAGACTGATTGTAAACTCACATGTCTCTGGGA
TGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

Figure 4.2 Contiguous sequences of region of interest.

4.4 MUTATION IDENTIFICATION AND SWISS-MODEL DATA

Exon 19 of *BRCA1* for each sample was compared to the consensus wild type nucleotide sequence obtained from UCSC's genome browser. Nucleotide BLAST was utilized to align sequences from each participant with the wild type nucleotide sequence. This was to confirm the amplified region was exon 19 from *BRCA1*, as shown in Figure 4.3 below.

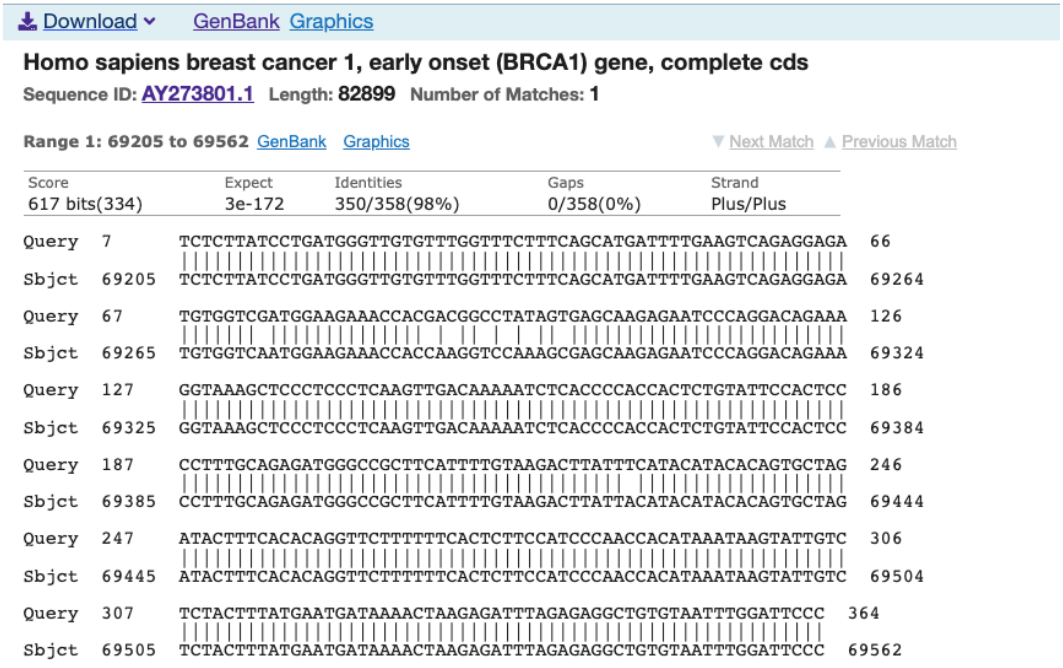


Figure 4.3 Nucleotide BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) results confirming exon 19 from a 24-year-old patient (sample 55). Query is the wild-type sequence.

Three of the samples of sequenced amplicon DNA from one 24-year-old (sample 55), one 32-year-old (sample 60) and one 50-year-old (sample 56) participant were found to carry mutations of biological significance. Nucleotide sequences from these were converted into amino acid sequences and compared to the consensus wild type protein sequence as shown below in Figure 4.4:

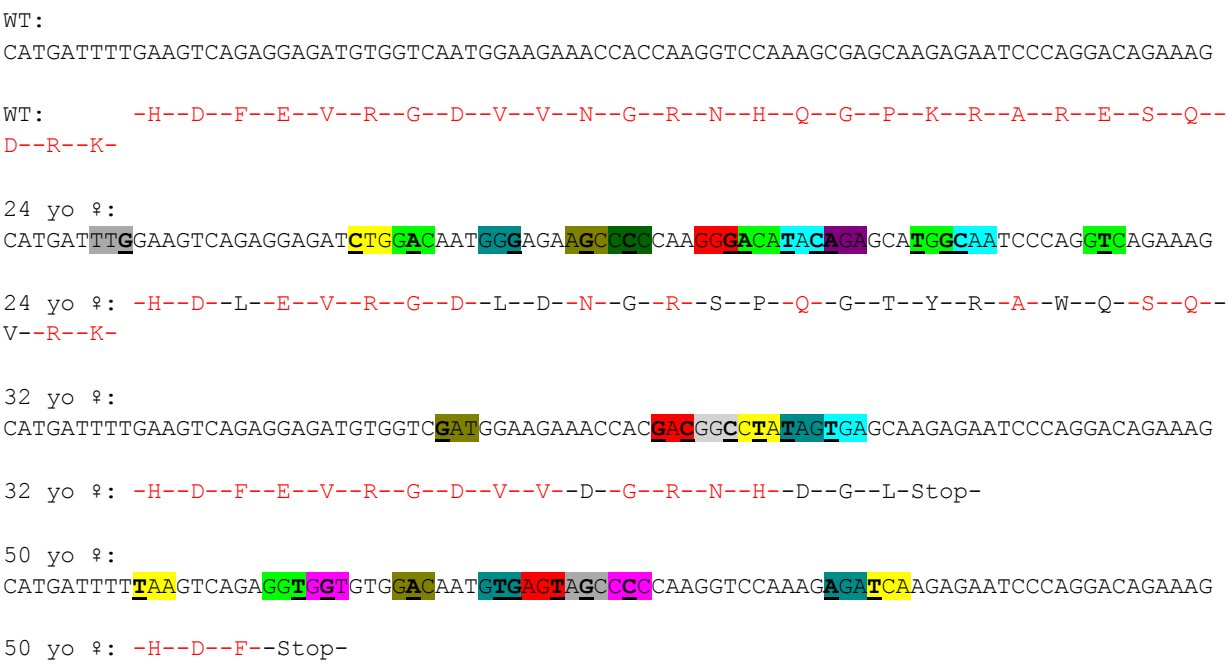


Figure 4.4 Wild type nucleotide and amino acid sequences of exon 19 of *BRCA1* and mutated nucleotide and amino acid sequences of a 24-year-old patient (sample 55), a 32-year-old patient (sample 60) and a 50-year-old patient (sample 56). Codons with mutations are highlighted and nucleotide changes are bold and underlined. Amino acid substitutions are black in color. Only the region of interest (exon 19) is shown in the image.

Three-dimensional protein structure models were then obtained using SWISS-MODEL (figure 4.5). Results were used to infer structural changes that arose from the putative changes at nucleotide or gross protein level.

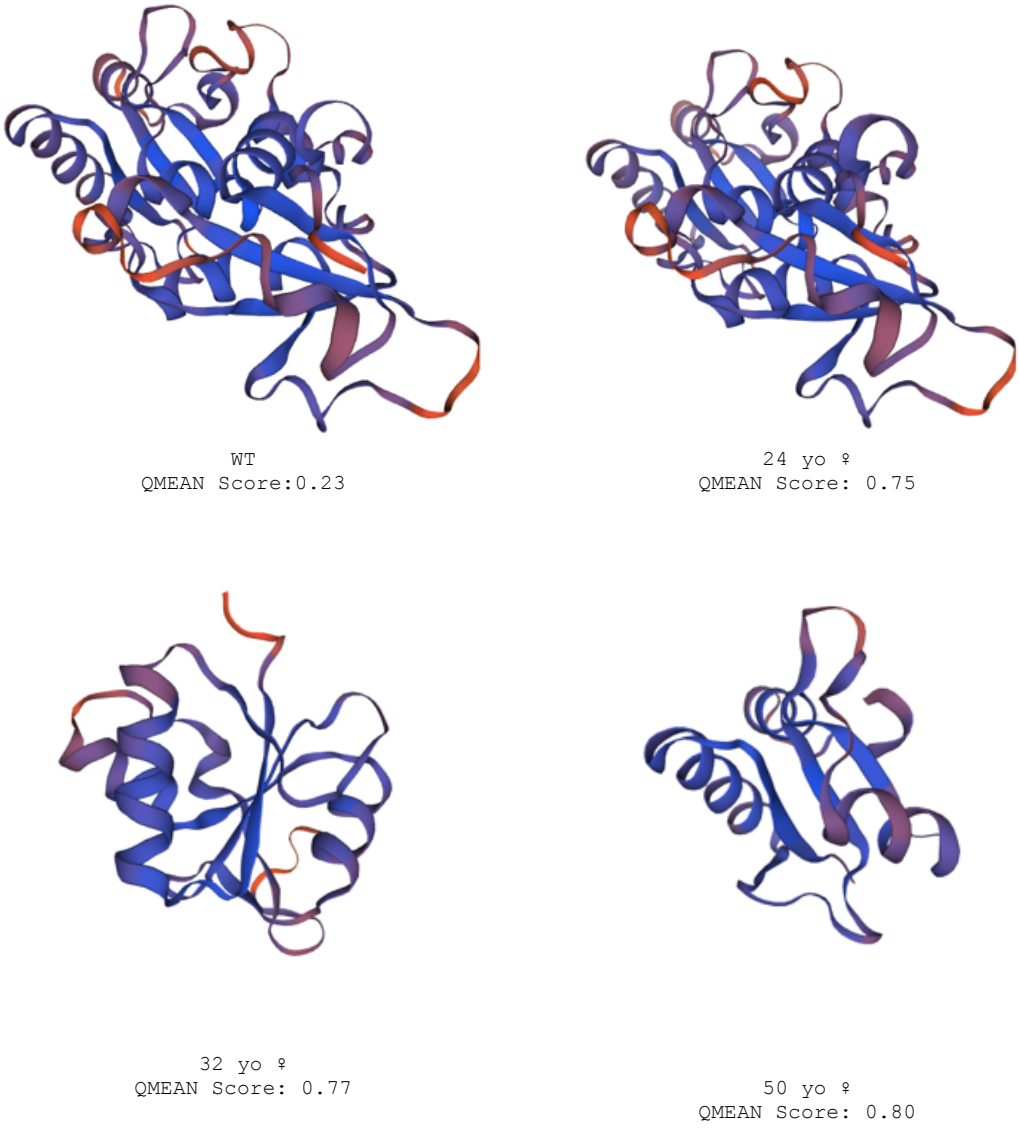


Figure 4.5 Three-dimensional protein structure models in the same orientation, obtained by the author using SWISS-MODEL (as detailed in section 3.6.2), of the wild type BRCA1 protein and mutated proteins of a 24-year-old patient (sample 55), a 32-year-old patient (sample 60) and a 50-year-old patient (sample 56), respectively. The assumption was that all other exons of mutant *BRCA1*, if translated, remained the same. QMEAN z-scores around zero indicated good agreement between the model structure and experimental structures of similar size.

4.5 VARIANT ANALYSIS

The clinical significance of identified mutations was assessed based on information from a major BRCA1 database, ClinVar. Results were used to infer functional changes, if any and a summary of all mutations leading to amino acid substitutions were recorded in table 4.2.

Table 4.2 Summary of eleven identified mutations from three out of ten analyzed samples, their associated pathogenicity and functional classification based on reference data from ClinVar. Rows containing mutations with a functional classification of 'loss of function' are highlighted in grey. Unc. Sig.= Uncertain Significance, NP= Not Provided (based on literature), Het.= Heterozygous, Homo.= Homozygous, F- effect= Functional effect, Funct.= Functional, LOF= Loss of function, Inter.= Intermediate, F- score= Functional score. SGE function score ranges for classification are as follows: functional = score > -0.748; intermediate = -0.748 > score > -1.328 and non-functional = score < -1.328 (National Center for Biotechnology Information. ClinVar archive; [VCV000865201.5], 2022)

Age	Nucleotide	Protein	Consequen-ce	ClinVar Interpr-etation	Allele Status	F-effect	F-score
24 yo ♀ (Sample 55)	NM_007298.3: c.1890T>G	NP_009229.2:p.Phe630Leu	missense	Unc. Sig.	Het.	LOF	-1.87
24 yo ♀	NM_007298.3: c.1906G>C	NP_009229.2:p.Val636Leu	missense	NP	Het.	Funct.	-0.06
24 yo ♀	NM_007298.3: c.1910T>A	NP_009229.2:p.Val637Asp	missense	NP	Het.	LOF	-1.91
24 yo ♀	NM_007298.3: c.1931A>G	NP_009229.2:p.Asn641Ser	missense	Unc. Sig.	Het.	Funct.	0.2

24 yo ♀	NM_007298.3: c.1925A>C	NP_009229.2:p. His642Pro	missense	Unc. Sig.	Het.	LOF	-2.57
24 yo ♀	NM_007298.3: c.1933C>A	NP_009229.2:p. Pro645Thr	missense	Unc. Sig.	Het.	Funct.	-0.65
24 yo ♀	NM_007298.3: c.1948G>C	NP_009229.2:p. Glu650Gln	missense	Unc. Sig.	Het.	Funct.	0.02
32 yo ♀ (Sample 60)	NM_007298.3: c.1912A>G	NP_009229.2:p. Asn638Asp	missense	Unc. Sig.	Homo.	Inter.	-0.75
32 yo ♀	NM_007298.3: c.1934C>T	NP_009229.2:p. Pro645Leu	missense	Unc. Sig.	Het.	LOF	-2.09
32 yo ♀	NM_007298.3: c.1936A>T	NP_009229.2:p. Lys646Ter	nonsense	Pathogenic	Het.	LOF	-2.37
50 yo ♀ (Sample 56)	NM_007298.3: c.1891G>T	NP_009229.2:p. Glu631Ter	nonsense	NP	Het.	LOF	-2.51

4.6 SAMPLE DEMOGRAPHICS AND STATISTICAL ANALYSIS

Sample demographics for this study consisted of patients diagnosed with breast cancer at the age of ≤50 years irrespective of family history of breast and/or ovarian cancer. All recruited participants were ethnically black females that were located in Lusaka. Ten participants were selected randomly for sequencing and were representative of the following age groups: 21-30 years, 31-40 years and 41-50 years. Table 4.3 shows sample demographics like age, sex, ethnicity and demographic location for the sequenced participants. The age range was from 20 to 50 years old.

Table 4.3 Details of participants from whom genomic DNA samples were sequenced.

Sample No	A260/ A280	Sex	Ethnicity	Geographic Location	Age (years)	Age Group (years)	Age Group Total
55	1.83	Female	Black	Lusaka	24	21-30	1
63	1.81	Female	Black	Lusaka	31	31-40	4
60	1.86	Female	Black	Lusaka	32		
26	1.82	Female	Black	Lusaka	35		
19	1.81	Female	Black	Lusaka	40		
16	1.89	Female	Black	Lusaka	41	41-50	5
62	1.86	Female	Black	Lusaka	43		
41	1.80	Female	Black	Lusaka	48		
65	1.88	Female	Black	Lusaka	49		
56	1.85	Female	Black	Lusaka	50		

CHAPTER FIVE

DISCUSSION

5.0 OVERVIEW

This chapter discusses the research findings of this study in accordance with research objectives.

5.1 CLINICAL SIGNIFICANCE OF IDENTIFIED MUTATIONS

PCR amplicons of genomic DNA samples were sequenced with both forward and reverse primers and in the total sequenced study population (n= 10), Eleven (11) mutations were identified in 3 patients: a 24-year-old patient (sample 55), a 32-year-old patient (sample 60) and a 50-year-old patient (sample 56) (Table 4.2). Six heterozygous mutations leading to a loss of function of BRCA1 were identified in ClinVar, a web-based archive (https://www.ncbi.nlm.nih.gov/clinvar) (Table 4.2). Of the six loss of function mutations, four were missense mutations and two were nonsense mutations.

The 24-year-old female breast cancer patient had three of the four heterozygous missense mutations. The mutations in this particular case led to a substitution of thymine (T) for guanine (G) at position 1890. This changed the codon at 630 from TTT to TTG (TTT→TTG) and resulted in the replacement of phenylalanine by leucine (F630L) and substitution of nucleotide adenine (A) with cytosine (C) at position 1925 changed the codon at 642 from CAC→CCC which resulted in the replacement of histidine by proline (H642P). For this mutation, loss of the positive charge of histidine had the potential to contribute to the loss of function of BRCA1. All these missense variants were interpreted as having uncertain significance in ClinVar meaning the criteria for the variant to be labeled as pathogenic, likely pathogenic, benign and likely benign were not met or were contradictory (Richards *et al*., 2015). The third missense variant at position 1910 caused a transversion mutation whereby the nucleotide thymine was changed to adenine resulting in the change of codon 637 from GTC→GAC. This resulted in the replacement of valine by aspartic acid (V637D). The interpretation of this missense variant was not provided in ClinVar meaning

the mutation was only provided in literature, research submissions, and phenotyping only submissions from clinics or physicians which did not lead to the interpterion of the clinical significance of this variant (https://www.ncbi.nlm.nih.gov/clinvar).

The 32-year-old female breast cancer patient had one heterozygous missense mutation in which substitution of nucleotide cytosine with thymine at position 1934 changed the codon at 645 from CCA→CTA. This resulted in the replacement of proline by leucine (P645L). This missense variant was interpreted as having uncertain significance in ClinVar meaning the criteria for the variant to be labeled as pathogenic, likely pathogenic, benign and likely benign were not met or were contradictory (Richards *et al*., 2015). A heterozygous nonsense mutation at position 1936 in which the substitution of nucleotide adenine with thymine changed codon 646 from AAG→TAG resulting in the replacement of lysine with a stop codon (K646Stop/Ter), was identified in this patient. This nonsense variant was interpreted as pathogenic and correlated with cancer development in the ClinVar archive. The 50-year-old female breast cancer patient had one heterozygous nonsense mutation. The substitution of nucleotide guanine with thymine at position 1891 changed codon 631 from GAA→TAA. This resulted in the replacement of glutamic acid by a stop codon (E631Stop/Ter). This nonsense mutation could contribute to loss of function of BRCA1 due to the loss of a huge part of the protein, however, the interpretation of this nonsense variant was not provided in the ClinVar archive.

All the loss of function variants were specifically associated to Breast-ovarian cancer, familial, susceptibility to, 1 i.e., BROVCA1 (Online Mendelian Inheritance in Man (OMIM) #604370). Characteristic clinical features of BROVCA1 are familial breast cancer, younger age at diagnosis and frequent bilateral disease (Online Mendelian Inheritance in Man, 2022). *BRCA1*- and *BRCA2*-associated hereditary breast and ovarian cancer (HBOC) have been characterized by an increased risk for female and male breast cancer, ovarian cancer (including fallopian tube and primary peritoneal cancers), and to a lesser extent other cancers such as prostate cancer, pancreatic cancer, and melanoma primarily in individuals with a *BRCA2* pathogenic variant (Petrucelli, Daly and Pal, 1998). The diagnosis of *BRCA1*- and *BRCA2*-associated HBOC is established in a proband by identification of a heterozygous germline pathogenic variant in *BRCA1* or *BRCA2* on molecular genetic testing (Petrucelli, Daly and Pal, 1998). Genetic testing would therefore provide useful

information on the risk of inheriting *BRCA1*- and *BRCA2*-associated HBOC in female BC patients from Zambia.

5.2 STRUCTURAL AND FUNCTIONAL EFFECTS OF IDENTIFIED MUTATIONS

The particular BRCA1 isoform (P38398-3) of interest to this study was a 759 amino acid protein. Exon 19 codes for part of the BRCT domain for this isoform. As depicted in figure 4.5, mutations in this region have the potential to affect the three-dimensional folding of the entire protein structure. The presence of a nonsense mutation in the 32-year-old patient and the 50-year-old patient greatly affected the structure of the BRCA1 protein because the presence of a premature stop codon resulted in an incomplete protein that was shorter than the wild type (Figure 4.5). The loss of a significant portion of the amino acids that maintained the secondary structure of the BRCA1 protein by forming hydrogen bonds, made it difficult for the two shorter proteins to fold correctly i.e., forming β sheets and α helices. This is particularly evident in the 32-year-old patient because the α helices are clearly misfolded and incomplete. This correlates with ClinVar's functional consequence for the pathogenic variant of loss of function or functionally abnormal i.e., a sequence variant in which the function of a gene product is altered with respect to a reference (Eilbeck *et al*., 2005).

Saturation genome editing (SGE) is a CRISPR/Cas9-based method to functionally test the effects of large numbers of variants in their native genomic context (Findlay *et al*., 2014). SGE function score ranges for classification are as follows: functional = score > -0.748; intermediate = -0.748 > score > -1.328 and non-functional = score < -1.328 (National Center for Biotechnology Information. ClinVar archive; [VCV000865201.5], 2022). SGE function score ranges for all identified mutations, with a corresponding functional consequence of loss of function, were < -1.328 (Table 4.2). In particular, the SGE assay for the nonsense variant (c.1936A→T) for the 32-year-old patient produced a function score of -2.37 (National Center for Biotechnology Information. ClinVar archive; [VCV000865201.5], 2022). This sequence change apparently created a premature translational stop signal (K646Stop/Ter) in the *BRCA1* gene, and a disrupted protein product was produced as shown in figure 4.5. When criteria for pathogenicity is met, loss of function variants in BRCA1 are known to be pathogenic (PMID: 20104584) and for these

reasons, this variant was classified as pathogenic or disease-causing (National Center for Biotechnology Information. ClinVar archive; [VCV000865201.5], 2022). ClinVar utilizes information curated by clinicians, statisticians and other major world databases to assign pathogenicity to a variant. Though most nonsense mutations are deleterious, they are not all associated with disease. According to ClinVar, it is important to consider the differences between implicating a variant as pathogenic (i.e., causative) for a disease, and a variant that may be predicted to be disruptive/damaging to the protein for which it codes but is not necessarily implicated in a disease (Richards *et al*., 2015). Therefore, variants could only be interpreted as pathogenic if all criteria for pathogenicity were met (Richards *et al*., 2015).

ClinVar lists various mutations in exon 19 such as V1736G, R1737Stop/Ter, G1748D etc., which have all been identified as pathogenic and shown to influence the structure and function of BRCA1 based on SGE scores (https://www.ncbi.nlm.nih.gov/clinvar/). Disruptions to the BRCA1 linker region, i.e., exon 19, caused by these mutations and those identified in this study greatly impact: (i) transcriptional regulation due to changes in affinity of BRCA1 for its interacting partners (ii) cell cycle checkpoint activation through the loss of amino acids that are essential for the formation of hydrogen bonds that lead to protein complex formation and ubiquitination of cell-cycle proteins tagging them for proteasomal degradation and, (iii) DNA repair through the gain or loss of charges that play a crucial role in BRCA1 folding thereby affecting its ability to recruit and specifically interact with proteins like RAD51 and PALB2 (Cunningham *et al.*, 2022; Billing *et al.*, 2018; Coquelle, Green and Glover, 2011).

5.3 STATISTICAL FINDINGS

Based on an unpaired two-tailed student's *t*-test and the specific sample size evaluated in this study, the null hypothesis, previously described in Chapter one section 1.6 for this study, was rejected meaning that there was a significant association between the age of a BC patient and the number of mutations in exon 19 leading to a loss of function of BRCA1 (P= 0.04). However, further studies are clearly necessary to explore how mutations in the entire *BRCA1* gene affect disease presentation in the female BC population from Zambia. Data in this study is not enough to suggest a significant association between age of a BC patient and the number of mutations in

exon 19 leading to a loss of function of BRCA1 in the entire population of BC patients evaluated at Matero level one hospital. The limited size of this study makes it difficult to draw firm conclusions about the contribution of single nucleotide mutations to disease presentation in the entire Zambian female BC population too. Therefore, a study with a larger sample size that appropriately represents this population is highly recommended.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.0 OVERVIEW

This chapter reports on the conclusions drawn from the study and recommendations based on results obtained.

6.1 CONCLUSIONS

This pilot study was the first of its kind to experimentally investigate and generate genetic data from samples of female breast cancer patients from Matero level one hospital. The goal was to investigate the use of molecular methods to characterize single nucleotide variants that may be present in part of the BRCT region (i.e., exon 19) in a selected population of patients evaluated at Matero level one hospital. The results presented herein are particularly relevant for the potential incorporation of molecular testing in the population of female BC patients in Zambia, through the coupling of diagnosis and screening in a previously under-researched group. The integration of genomic medicine into BC care has tremendous potential to improve the outcomes and survival of diagnosed patients in low- and middle-income countries (LMICs) like Zambia because resources can be better allocated, and treatment options can be assigned to patients according to necessity. The finding of missense and nonsense mutations in the sample size tested has the potential to alter treatment decisions; for example, women with pathogenic mutations in BRCA1 associated with HBOC could consider obtaining a family history thereby increasing the likelihood of molecular testing for their entire family. Relatives of an individual who has a BRCA1 pathogenic variant could be counseled regarding their risk of having inherited the same variant, their possible cancer risk, recommendations for screening and their options for molecular genetic testing.

In situations where chemotherapy or mammography options are limited, personalized medicine through the use of genomic information can have the greatest financial impact. Our results have highlighted the possible contributions of *BRCA1* single nucleotide variants to the structure and

function of the protein as a tumor suppressor in Zambian female BC patients and it has also highlighted how results from genomic tests can be incorporated into the decisions physicians make concerning BC treatment strategies. The Sanger sequencing approach utilized in this study shows great promise for *BRCA1* screening in developing countries like Zambia. To the best of our knowledge, the current study is the first study for molecular *BRCA1* testing in BC patients irrespective of family history of breast and/or ovarian cancer.

6.2 RECOMMENDATIONS

The findings from this study will contribute to efforts currently underway to evaluate genetic data from young female BC patients in Zambia. Based on study findings, the following recommendations are hereby being made:

i. A study focusing on exon 19 with a larger sample size that appropriately represents this population is highly recommended. The limited size of this study made it difficult to draw firm conclusions and to extrapolate on the impact of the contributions of single nucleotide mutations to HBOC in the entire Zambian female BC population.

ii. It is crucial to further explore the other possible *BRCA1* gene mutations (insertions, deletions, [indels], duplications) in exon 19 and other regions of the gene which could adversely impact BRCA1. This might lead to new insights that could not be gained from this study which mainly focused on characterizing single nucleotide variants present in exon 19 of the gene.

iii. In Zambia, the prevalence of *BRCA1* mutations is yet to be well defined, which adds further challenge in tailoring a gene screening panel and international guidelines to the national context. The *BRCA1* gene is very large, therefore, looking at next-generation sequencing (NGS) data will allow its complete evaluation which may provide useful information that may aid in the implementation of genetic screening panels, genetic testing, and subsequent genetic counseling of BC patients from Zambia.

iv. Further, a study on the impact of implementing molecular genetic testing of young female BC patients in Zambia, following the western recommendations, is highly encouraged. In order to successfully utilize this powerful technology in LMICs, this will require well-trained teams for genetic testing and counseling experts, proper implementation and community education.

v. Until recently, scientists thought BCs were mainly caused by changes in DNA sequence, which was a major focus of this research project. However, epigenetic processes are increasingly being implicated in gene dysregulation in BCs thereby contributing to the disease phenotype. A study exploring the possible epigenetic changes in female BC patients from Zambia is also being recommended.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), pp. 403–410. [Online]: 10.1016/S0022-2836(05)80360-2 (Accessed: 15 March 2022).

Abbad, A., Baba, H., Dehbi, H., Elmessaoudi-Idrissi, M., Elyazghi, Z., Abidi, O., & Radouani, F. (2018). Genetics of breast cancer in African populations: a literature review. *Global health, epidemiology and genomics*, 3, e8. [Online]: 10.1017/gheg.2018.8 (Accessed: 16 July 2022).

Anyigba, C. A., Awandare, G. A., & Paemka, L. (2021). Breast cancer in sub-Saharan Africa: The current state and uncertain future. Experimental biology and medicine (Maywood, N.J.), 246(12), 1377–1387. https://doi.org/10.1177/15353702211006047 (Accessed: 15 January 2023).

Bakkach, J., Mansouri, M., Derkaoui, T., Loudiyi, A., El Fahime, E., Barakat, A., Nourouti, N.G., De Villarreal, J.M., Bringas, C.C., Mechita, M.B. (2020). Contribution of *BRCA1* and *BRCA2* germline mutations to early onset breast cancer: a series from north of Morocco. *BMC Cancer*, 20(1), pp. 859. [Online]: 10.1186/s12885-020-07352-9 (Accessed: 16 September 2020).

Bhattacharya, S., Srinivasan, K., Abdisalaam, S., Su, F., Raj, P., Dozmorov, I., Mishra, R., Wakeland, E. K., Ghose, S., Mukherjee, S., & Asaithamby, A. (2017). RAD51 interconnects between DNA replication, DNA repair and immunity. Nucleic acids research, 45(8), 4590–4605. https://doi.org/10.1093/nar/gkx126 (Accessed: 15 January 2023).

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., & Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic

acids research, 42(Web Server issue), W252–W258. https://doi.org/10.1093/nar/gku340 (Accessed: 15 January 2023).

Billing, D., Horiguchi, M., Wu-Baer, F., Taglialatela, A., Leuzzi, G., Nanez, S.A., Jiang, W., Zha, S., Szabolcs, M., Lin, C.S., Ciccia, A., Baer, R. (2018). The BRCT Domains of the BRCA1 and BARD1 Tumor Suppressors Differentially Regulate Homology-Directed Repair and Stalled Fork Protection. *Molecular Cell*, 72(1), pp. 127-139. [Online]: 10.1016/j.molcel.2018.08.016 (Accessed: 18 October 2020).

Bunch, H., Jeong, J., Kang, K., Sin Jo, D., Cong, A., Kim, D., Kim, D., Cho, D. H., Lee, Y. M., Chen, B., Schellenberg, M. J., Calderwood, S. K. (2021). BRCA1-BARD1 regulates transcription through modulating topoisomerase II*β*. *Open Biology*, 11(10). [Online]: 10.1098/rsob.210221 (Accessed: 23 May 2021).

Cabanes, A., Kapambwe, S., Citonje-Msadabwe, S., Parham, G.P., Lishimpi, K., Cruz, T.A., Shastri, S. (2019). Challenges, Opportunities, and Priorities for Advancing Breast Cancer Control in Zambia: A Consultative Meeting on Breast Cancer Control. *Journal of Global Oncology*, 5, pp. 1-7. [Online] DOI: 10.1200/JGO.18.00222 (Accessed: 3 October 2020).

Ceccaldi, R., Rondinelli, B., D'Andrea, A.D. (2016). Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol*, 26(1), pp. 52–64. [Online]: 10.1016/j.tcb.2015.07.009 (Accessed: 21 September 2020).

Center for Disease Control and Prevention. (2020). What Are the Risk Factors for Breast Cancer? Retrieved16 November 2020 fromhttps://www.cdc.gov/cancer/breast/basic_info/risk_factors.htm

Chakraborty, A., Katarkar, A., Chaudhuri, K., Mukhopadhyay, A., Basak, K. (2013). Detection of a novel mutation in exon 20 of the *BRCA1* gene. *Cellular and Molecular Biology letters*, 18 (4), pp. 631-638. [Online]: 10.2478/s11658-013-0110-3 (Accessed: 1 December 2020).

Chapman, J. R., Taylor, M. R., Boulton, S. J. (2012). Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Molecular Cell*, 47(4), pp.497–510. [Online]: 10.1016/j.molcel.2012.07.029 (Accessed: 23 May 2021).

Chehade, R., Pettapiece-Phillips, R., Salmena, L., Kotlyar, M., Jurisica, I., Narod, S. A., Akbari, M. R., & Kotsopoulos, J. (2016). Reduced BRCA1 transcript levels in freshly isolated blood leukocytes from *BRCA1* mutation carriers is mutation specific. *Breast cancer research: BCR*, 18(1), pp. 87. [Online]: 10.1186/s13058-016-0739-8 (Accessed: 16 July 2022).

Chen, F., Roberts, N. J., & Klein, A. P. (2017). Inherited pancreatic cancer. *Chinese clinical oncology*, 6(6), pp. 58. [Online]: 10.21037/cco.2017.12.04 (Accessed: 16 July 2022).

Christou, C. M., Kyriacou, K. (2013). BRCA1 and Its Network of Interacting Partners. *Biology*, 2(1), pp. 40–63. [Online]: 10.3390/biology2010040 (Accessed: 23 May 2021).

Clark, S. L., Rodriguez, A. M., Snyder, R. R., Hankins, G. D., & Boehning, D. (2012). Structure-Function of The Tumor Suppressor BRCA1. *Computational and structural biotechnology journal*, *1*(1), e201204005. [Online]: 10.5936/csbj.201204005 (Accessed: 16 November 2020).

Clark, S. L., Rodriguez, A. M., Snyder, R. R., Hankins, G. D., & Boehning, D. (2012). Structure-Function of The Tumor Suppressor BRCA1. *Computational and structural biotechnology journal*, *1*(1), e201204005. [Online]: 10.5936/csbj.201204005 (Accessed: 16 November 2020).

Coquelle, N., Green, R., & Glover, J. N. (2011). Impact of BRCA1 BRCT domain missense substitutions on phosphopeptide recognition. Biochemistry, 50(21), 4579–4589. https://doi.org/10.1021/bi2003795 (Accessed: 11 January 2023).

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., … Flicek, P. (2022). Ensembl 2022. *Nucleic acids research*, 50(D1), D988–D995, Version 107. Transcript: ENST00000357654.9 BRCA1-203. [Online]: 10.1093/nar/gkab1049 (Accessed: 16 July 2022).

Dacheux, E., Vincent, A., Nazaret, N., Combet, C., Wierinckx, A., Mazoyer, S., Diaz, J. J., Lachuer, J., & Venezia, N. D. (2013). BRCA1-Dependent Translational Regulation in Breast Cancer Cells. *PloS one*, 8(6), e67313. [Online]: 10.1371/journal.pone.0067313 (Accessed: 16 July 2022).

Deng C. X. (2006). BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic acids research*, 34(5), pp. 1416–1426. [Online]: 10.1093/nar/gkl010 (Accessed: 16 July 2022).

Di Masi, A., Gullotta, F., Cappadonna, V., Leboffe, L., Ascenzi, P. (2011). Cancer Predisposing Mutations in BRCT Domains. *IUBMB Life*, 63(7), pp. 503–512. [Online]: 10.1002/iub.472 (Accessed: 16 November 2020).

El Khachibi, M., Diakite, B., Hamzi, K., Badou, A., Senhaji, M.A., Bakhchane, A., Jouhadi, H., Barakat, A., Benider, A., Nadifi, A. (2015). Screening of exon 11 of *BRCA1* gene using the high-resolution melting approach for diagnosis in Moroccan breast cancer patients. *BMC Cancer*, 15 (81), pp. 1-5. [Online]: 10.1186/s12885-015-1040-4 (Accessed: 21 October 2020).

ElBiad, O., Laraqui, A., El Boukhrissi, F., Mounjid, C., Lamsisi, M., Bajjou, T., Elannaz, H., Lahlou, A. I., Kouach, J., Benchekroune, K., Oukabli, M., Chahdi, H., Ennaji, M. M., Tanz, R., Sbitti, Y., Ichou, M., Ennibi, K., Badaoui, B., & Sekhsokh, Y. (2022). Prevalence of specific and recurrent/founder pathogenic variants in BRCA genes in breast and ovarian

cancer in North Africa. *BMC Cancer,* 22(1), pp. 208. [Online]: 10.1186/s12885-022-09181-4 (Accessed: 08 July 2022).

Foerster, M., McKenzie, F., Zietsman, A., Galukande, M., Anele, A., Adisa, C., Parham, G., Pinder, L., Schüz, J., McCormack, V., & Dos-Santos-Silva, I. (2021). Dissecting the journey to breast cancer diagnosis in sub-Saharan Africa: Findings from the multicountry ABC-DO cohort study. International journal of cancer, 148(2), 340–351. https://doi.org/10.1002/ijc.33209 (Accessed: 28 January 2023).

Francies, F.Z., Wainstein, T., De Leeneer, K., Cairns, A., Murdoch, M., Nietz, S., Cubasch, H., Poppe, B., Van Maerken, T., Crombez, B., Coene, I., Kerr, R., Slabbert, J.P., Vral, A., Baeyens, A., Claes, K.B.M. (2015). BRCA1, BRCA2 and PALB2 mutations and CHEK2 c.1100delC in different South African ethnic groups diagnosed with premenopausal and/or triple negative breast cancer. *BMC Cancer*, 15(912), pp. 1-10. [Online]: 10.1186/s12885-015-1913-6 (Accessed: 10 October 2020).

Freedman, R. A., & Partridge, A. H. (2013). Management of breast cancer in very young women. *Breast (Edinburgh, Scotland)*, 22 Suppl 2, S176–S179. [Online]: 10.1016/j.breast.2013.07.034 (Accessed: 16 July 2022).

Glover, J. N., Williams, R. S., & Lee, M. S. (2004). Interactions between BRCT repeats and phosphoproteins: tangled up in two. *Trends in biochemical sciences*, 29(11), pp. 579–585. [Online]: 10.1016/j.tibs.2004.09.010 (Accessed: 12 July 2022).

Gorski, J. J., Savage, K. I., Mulligan, J. M., McDade, S. S., Blayney, J. K., Ge, Z., & Harkin, D. P. (2011). Profiling of the BRCA1 transcriptome through microarray and ChIP-analysis. *Nucleic acids research*, *39*(22), pp. 9536–9548. [Online]: 10.1093/nar/gkr679 (Accessed: 21 October 2020).

Hondrow, H. L., Fox, S.B., Mitchell, G., Scott, R.J., Beshay, V., Wong, S.Q., konFab Investigators, Dobrovic, A. (2011). A high-throughput protocol for mutation scanning of

the *BRCA1* and *BRCA2* genes. *BMC Cancer*, 11 (265), pp. 1-11. [Online]: 10.1186/1471-2407-11-265 (Accessed: 21 October 2020).

Hu, Y., Petit, S. A., Ficarro, S. B., Toomire, K. J., Xie, A., Lim, E., Cao, S. A., Park, E., Eck, M. J., Scully, R., Brown, M., Marto, J. A., & Livingston, D. M. (2014). PARP1-driven poly-ADP-ribosylation regulates BRCA1 function in homologous recombination-mediated DNA repair. *Cancer discovery*, *4*(12), pp. 1430–1447. [Online]: 10.1158/2159-8290.CD-13-0891 (Accessed: 1 December 2020).

Kalubula, M., Shen, H., Makasa, M., & Liu, L. (2021). Epidemiology of Cancers in Zambia: A significant variation in Cancer incidence and prevalence across the nation. Malawi medical journal: the journal of Medical Association of Malawi, 33(3), 186–195. https://doi.org/10.4314/mmj.v33i3.6

Khodadadian, A., Darzi, S., Haghi-Daredeh, S., Eshaghi, F.S., Babakhanzadeh, E., Mirabutalebi, S.H., Nazari, M. (2020). Genomics and Transcriptomics: The Powerful Technologies in Precision Medicine. *Int J Gen Med*, 13, pp. 627-640. [Online]: 10.2147/IJGM.S249970 (Accessed: 21 September 2020).

Kiendrebeogo, I. T., Zoure, A. A., Zongo, F. I., Ouattara, A. K., Ouedraogo, M., Amegnona, J., Yonli, A. T., Bayala, B., Zongo, N., Bambara, A. H., Sawadogo, A. Y., Zohoncon, T. M., Obiri-Yeboah, D., & Simpore, J. (2022). Mutations in exon 11 (11.1 and 11.2) of the *BRCA1* gene and risk factors for breast cancer in Burkina Faso. *Journal of public health in Africa*, 13(1), pp. 1921. [Online]: 10.4081/jphia.2022.1921 (Accessed: 16 July 2022).

Kim, S. I., Lee, M., Kim, H. S., Chung, H. H., Kim, J. W., Park, N. H., & Song, Y. S. (2020). Germline and Somatic *BRCA1/2* Gene Mutational Status and Clinical Outcomes in Epithelial Peritoneal, Ovarian, and Fallopian Tube Cancer: Over a Decade of Experience in a Single Institution in Korea. *Cancer research and treatment*, 52(4), pp. 1229–1241. [Online]: 10.4143/crt.2020.557 (Accessed: 16 July 2022).

Lee, M. S., Green, R., Marsillac, S. M., Coquelle, N., Williams, R. S., Yeung, T., Foo, D., Hau, D. D., Hui, B., Monteiro, A. N., & Glover, J. N. (2010). Comprehensive analysis of missense variations in the BRCT domain of BRCA1 by structural and functional assays. *Cancer research*, 70(12), pp. 4880–4890. [Online]: 10.1158/0008-5472.CAN-09-4563 (Accessed: 08 July 2022).

Lipkowitz, S., Weissman, A. (2011). RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat Rev Cancer*, 11, pp. 629–643. [Online]: 10.1038/nrc3120 (Accessed: 10 July 2022).

Martínez-Mesa, J., González-Chica, D. A., Bastos, J. L., Bonamigo, R. R., & Duquia, R. P. (2014). Sample size: how many participants do I need in my research? *Anais brasileiros de dermatologia*, *89*(4), 609–615. https://doi.org/10.1590/abd1806-4841.20143705

Masuda, T., Xu, X., Dimitriadis, E. K., Lahusen, T., Deng, C. X. (2016). "DNA Binding Region" of BRCA1 Affects Genetic Statbility through modulating the Intra-S-Phase Checkpoint. *International Journal of Biological Sciences*, 12(2), pp. 133–143. [Online]: 10.7150/ijbs.14242 (Accessed: 23 May 2021).

Melhem-Bertrandt, A., Bojadzieva, J., Ready, K. J., Obeid, E., Liu, D. D., Gutierrez-Barrera, A. M., Litton, J. K., Olopade, O. I., Hortobagyi, G. N., Strong, L. C., & Arun, B. K. (2012). Young onset HER2-positive breast cancer is associated with germline TP53 mutations. *Cancer*, 118(4), pp. 908–913. [Online]: 10.1002/cncr.26377 (Accessed: 08 July 2022).

Ministry of Health, Republic of Zambia. (2016). National Cancer Control Strategic Plan (NCCSP) 2016-2021. Retrieved 3 October 2020 from https://www.iccp-portal.org/system/files/plans/NCCSP%20Final%20Version%20Zambia.pdf

Moynahan M.E., Jasin, M. (2010). Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol*, 11(3), pp. 196-207. [Online]: 10.1038/nrm2851 (Accessed: 06 September 2020).

Mylavarapu, S., Das, A., & Roy, M. (2018). Role of BRCA Mutations in the Modulation of Response to Platinum Therapy. *Frontiers in oncology*, 8, pp. 16. [Online]: 10.3389/fonc.2018.00016 (Accessed: 08 July 2022).

Narod, S. A., & Foulkes, W. D. (2004). BRCA1 and BRCA2: 1994 and beyond. Nature reviews. Cancer, 4(9), 665–676. https://doi.org/10.1038/nrc1431

National Breast Cancer Foundation Inc. (2020). About Breast Cancer- Stages 0 & 1. Retrieved 16 November 2020 from https://www.nationalbreastcancer.org/breast-cancer-stage-0-and-stage-1

National Cancer Institute. (2020). Cancer Statistics. Retrieved 27 September 2020 from https://www.cancer.gov/about-cancer/understanding/statistics#:~:text=The%20rate%20of%20new%20cases,on%202013%E2%80%932017%20deaths

National Cancer Institute. (2022). Breast Cancer Screening (PDQ®)–Patient Version. Retrieved 15 January 2023 from https://www.cancer.gov/types/breast/patient/breast-screening-pdq

National Center for Biotechnology Information. ClinVar archive; [VCV000865201.5], https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000865201.5 (Accessed July 08, 2022).

Nicolussi, A., Belardinilli, F., Mahdavian, Y., Colicchia, V., D'Inzeo, S., Petroni, M., Zani, M., Ferraro, S., Valentini, V., Ottini, L., Giannini, G., Capalbo, C., & Coppa, A. (2019). Next-generation sequencing of BRCA1 and BRCA2 genes for rapid detection of germline mutations in hereditary breast/ovarian cancer. PeerJ, 7, e6661. https://doi.org/10.7717/peerj.6661 (Accessed January 28, 2023).

Petrucelli N, Daly MB, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. 1998 Sep 4 [Updated 2022 May 26]. In: Adam MP, Everman DB, Mirzaa GM, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1247/ (Accessed January 28, 2023).

Prakash, R., Zhang, Y., Feng, W., Jasin, W. (2015). Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. *Cold Spring HarbPerspect Biol*, 7(4), pp. 1-27. [Online]: 10.1101/cshperspect.a016600 (Accessed: 1 December 2020).

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H. L., & ACMG Laboratory Quality Assurance Committee (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine: official journal of the American College of Medical Genetics, 17(5), 405–424. https://doi.org/10.1038/gim.2015.30 (Accessed: 06 January 2023).

Roy, R., Chun, J., Powell, S. N. (2016). BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat Rev Cancer*, 12(1), pp. 68–78. [Online]: 10.1038/nrc3181 (Accessed: 06 September 2020).

ScienceDirect. (2023). Student's t-Test. Retrieved 15 January 2023 from https://www.sciencedirect.com/topics/mathematics/students-t-test

Shabbeer, S., Omer, D., Berneman, D., Weitzman, O., Alpaugh, A., Pietraszkiewicz, A., Metsuyanim, S., Shainskaya, A., Papa, M. Z., &Yarden, R. I. (2013). BRCA1 targets G2/M cell cycle proteins for ubiquitination and proteasomal degradation. *Oncogene*, *32*(42), pp. 5005–5016. [Online]: 10.1038/onc.2012.522 (Accessed: 8 December 2020).

Shakya, R., Reid, L.J., Reczek, C.R., Cole, F., Egli, D., Lin, C.S., deRooij, D.G., Hirsch, S., Ravi, K., Hicks, J.B., Szabolcs, M., Jasin, M., Baer, R., Ludwig, T. (2011). BRCA1 tumor suppression depends on BRCT phosphoprotein binding, but not its E3 ligase activity. *Science*, *334*(6055), pp. 525–528. [Online]: 10.1126/science.1209909 (Accessed: 06 September 2020).

Sharma, R., Aashima, Nanda, M., Fronterre, C., Sewagudde, P., Ssentongo, A. E., Yenney, K., Arhin, N. D., Oh, J., Amponsah-Manu, F., & Ssentongo, P. (2022). Mapping Cancer in Africa: A Comprehensive and Comparable Characterization of 34 Cancer Types Using Estimates from GLOBOCAN 2020. Frontiers in public health, 10, 839835. https://doi.org/10.3389/fpubh.2022.839835 (Accessed: 15 January 2023).

Siddig, A., Tengku Din, T., Mohd Nafi, S. N., Yahya, M. M., Sulong, S., & Wan Abdul Rahman, W. F. (2021). The Unique Biology behind the Early Onset of Breast Cancer. *Genes*, 12(3), pp. 372. [Online]: 10.3390/genes12030372 (Accessed: 08 July 2022).

Songiso, M., Pinder, L. F., Munalula, J., Cabanes, A., Rayne, S., Kapambwe, S., Shibemba, A., & Parham, G. P. (2020). Minimizing Delays in the Breast Cancer Pathway by Integrating Breast Specialty Care Services at the Primary Health Care Level in Zambia. JCO global oncology, 6, 859–865. https://doi.org/10.1200/GO.20.00083 (Accessed: 15 January 2023).

Swiss model (2023). QMEAN: Quantitative Model energy ANalysis. Retrieved 28 January 2023 from https://swissmodel.expasy.org/.

Tsai, G. J., Rañola, J., Smith, C., Garrett, L. T., Bergquist, T., Casadei, S., Bowen, D. J., & Shirts, B. H. (2019). Outcomes of 92 patient-driven family studies for reclassification of variants of uncertain significance. *Genetics in medicine: official journal of the American College of Medical Genetics*, 21(6), pp. 1435–1442. [Online]: 10.1038/s41436-018-0335-7 (Accessed: 08 July 2022).

UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic acids research*, 49(D1), D480–D489. P38398, BRCA1_HUMAN. [Online]: 10.1093/nar/gkaa1100 (Accessed: 08 July 2022).

Venkitaraman A. R. (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*, 108(2), pp. 171–182. [Online]: 10.1016/s0092-8674(02)00615-3 (Accessed: 12 July 2022).

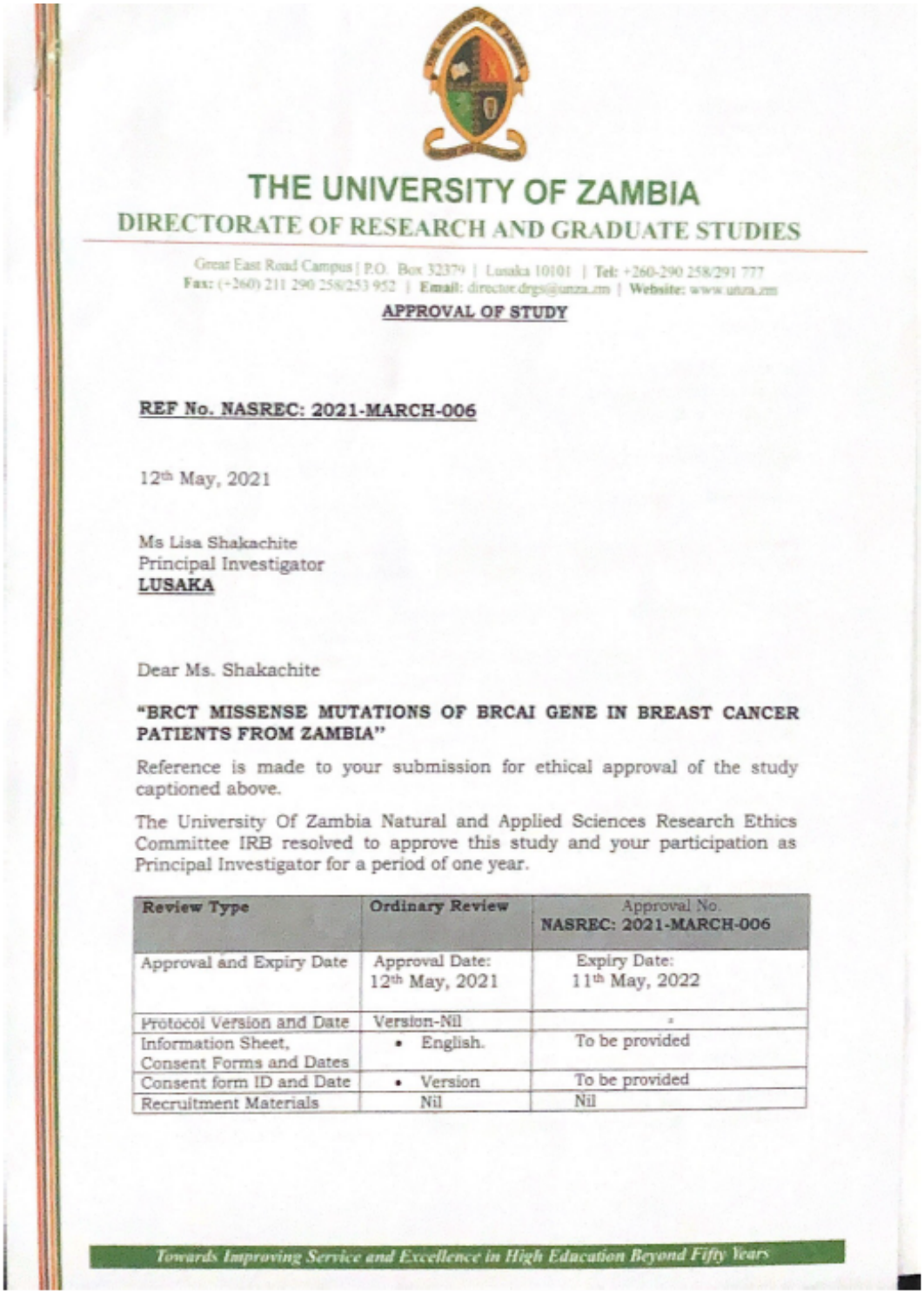
Venkitaraman, A. R. (2014). Cancer suppression by the chromosome custodians, BRCA1 and BRCA2. *Science (New York, N.Y.)*, 343(6178), pp. 1470–1475. [Online]: 10.1126/science.1252230 (Accessed: 16 November 2020).

Williams, R. S., Green, R., & Glover, J. N. (2001). Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nature structural biology*, *8*(10), pp. 838–842. [Online]: 10.1038/nsb1001-838 (Accessed: 12 July 2022).

Yip, C.H. (2016) Challenges in the early detection of breast cancer in resource-poor settings. *Breast Cancer Management*. Retrieved 28 January 2023 from https://www.futuremedicine.com/doi/abs/10.2217/bmt-2016-0026.

APPENDICES

APPENDIX I: UNZA ETHICAL CLEARANCE CERTIFICATE



THE UNIVERSITY OF ZAMBIA

DIRECTORATE OF RESEARCH AND GRADUATE STUDIES

Great East Road Campus | P.O. Box 32379 | Lusaka 10101 | Tel: +260-290 258/291 777
Fax: (+260) 211 290 258/253 952 | Email: director.drgs@unza.zm | Website: www.unza.zm

APPROVAL OF STUDY

REF No. NASREC: 2021-MARCH-006

12th May, 2021

Ms Lisa Shakachite
Principal Investigator
LUSAKA

Dear Ms. Shakachite

"BRCT MISSENSE MUTATIONS OF BRCAI GENE IN BREAST CANCER PATIENTS FROM ZAMBIA"

Reference is made to your submission for ethical approval of the study captioned above.

The University Of Zambia Natural and Applied Sciences Research Ethics Committee IRB resolved to approve this study and your participation as Principal Investigator for a period of one year.

Review Type	Ordinary Review	Approval No. NASREC: 2021-MARCH-006
Approval and Expiry Date	Approval Date: 12th May, 2021	Expiry Date: 11th May, 2022
Protocol Version and Date	Version-Nil	
Information Sheet, Consent Forms and Dates	• English.	To be provided
Consent form ID and Date	• Version	To be provided
Recruitment Materials	Nil	Nil

Towards Improving Service and Excellence in High Education Beyond Fifty Years

Specific conditions will apply to this approval. As Principal Investigator it is your responsibility to ensure that the contents of this letter are adhered to. If these are not adhered to, the approval may be suspended. Should the study be suspended, study sponsors and other regulatory authorities will be informed.

Conditions of Approval

- No participant may be involved in any study procedure prior to the study approval or after the expiration date.
- All unanticipated or Serious Adverse Events (SAEs) must be reported to NASREC within 5 days.
- All protocol modifications must be approved by NASREC prior to implementation unless they are intended to reduce risk (but must still be reported for approval). Modifications will include any change of investigator/s or site address.
- All protocol deviations must be reported to NASREC within 5 working days.
- All recruitment materials must be approved by NASREC prior to being used.
- Principal investigators are responsible for initiating Continuing Review proceedings. HSSRECwill only approve a study for a period of 12 months.
- It is the responsibility of the PI to renew his/her ethics approval through a renewal application to NASREC.
- Where the PI desires to extend the study after expiry of the study period, documents for study extension must be received by NASREC at least 30 days before the expiry date. This is for the purpose of facilitating the review process. Documents received within 30 days after expiry will be labelled "late submissions" and will incur a penaltyfee of K500.00. No study shall be renewed whose documents are submitted for renewal 30 days after expiry of the certificate.
- Every 6 (six) months a progress report form supplied by The University of Zambia Humanities and Social Sciences Research Ethics Committee as an IRB must be filled in and submitted to us. There is a penalty of K500.00 for failure to submit the report.
- When closing a project, the PI is responsible for notifying, in writing or using the Research Ethics and Management Online (REMO),both NASREC and the National Health Research Authority (NHRA) when ethics certification is no longer required for a project.
- In order to close an approved study, a Closing Report must be submitted in writing or through the REMO system. A Closing Report should be filed when data collection has ended and the study team will no longer be using human participants or animals or secondary data or have any direct or indirect contact with the research participants or animals for the study.
- Filing a closing report (rather than just letting your approval lapse) is important as it assists NASREC in efficiently tracking and reporting on

projects. Note that some funding agencies and sponsors require a notice of closure from the IRB which had approved the study and can only be generated after the Closing Report has been filed.

- A reprint of this letter shall be done at a fee.
- All protocol modifications must be approved by NASREC by way of an application for an amendment prior to implementation unless they are intended to reduce risk (but must still be reported for approval). Modifications will include any change of investigator/s or site address or methodology and methods. Many modifications entail minimal risk adjustments to a protocol and/or consent form and can be made on an Expedited basis (via the IRB Chair). Some examples are: format changes, correcting spelling errors, adding key personnel, minor changes to questionnaires, recruiting and changes, and so forth. Other, more substantive changes, especially those that may alter the risk-benefit ratio, may require Full Board review. In all cases, except where noted above regarding subject safety, any changes to any protocol document or procedure must first be approved by NASREC before they can be implemented.

Should you have any questions regarding anything indicated in this letter, please do not hesitate to get in touch with us at the above indicated address.

On behalf of NASREC, we would like to wish you all the success as you carry out your study.

Yours faithfully,

Dr. E. Mwanaumo

CHAIRPERSON
THE UNIVERSITY OF ZAMBIA NATURAL AND APPLIED SCIENCES
RESEARCH ETHICS COMMITTEE IRB

cc Director, Directorate of Research and Graduate Studies
Assistant Registrar (Research), Directorate of Research and Graduate Studies
Acting Senior Administration (R), Directorate of Research and Graduate Studies

APPENDIX II: NHRA ETHICAL CLEARANCE CERTIFICATE

NATIONAL HEALTH RESEARCH AUTHORITY
Paediatric Centre of Excellence, University Teaching Hospital, P.O. Box 30075, LUSAKA
Tell: +260211 250309 | Email: znhrasec@gmail.com | www.nhra.org.zm

Ref No:……. **Date: 9th August 2021**

The Principal Investigator
Lisa Shakachite,
School of Natural Sciences
University of Zambia
Lusaka.

Dear Ms Shakachite,

Re: Request for Authority to Conduct Research

The National Health Research Ethics Board (NHREB) is in receipt of your request for authority to conduct research titled **"BRCT Missense mutations of BRCA1 gene in Breast Cancer patients from Zambia."**

I wish to inform you that following submission of your request to the Board, its review of the same and in view of the ethical clearance, this study has been **approved** on condition that:

1. **A Material Transfer Agreement is obtained and cleared by the National Health Research Ethics Board should there be any need for samples to be sent outside the country for analysis.**
2. The relevant Provincial and District Medical Officers where the study is being conducted are fully appraised;
3. Progress updates are provided to NHRA quarterly from the date of commencement of the study;
4. The final study report is cleared by the NHRA before any publication or dissemination within or outside the country;
5. After clearance for publication or dissemination by the NHRA, the final study report is shared with all relevant Provincial and District Directors of Health where the study was being conducted, and all key respondents.

Yours sincerely,

Prof. Victor Chalwe
For Chairperson
National Health Research Ethics Board

All correspondences should be addressed to the Director/CEO National Health Research Authority

APPENDIX III: FORWARD AND REVERSE SEQUENCES FOR SAMPLES ANALYZED.

DS16

>DS16_Forward
CGGAATTGGGGTTCACAAAGGGTGGTGGGATTTCACCTGGGGGCTTGTGTGGAGGAAATGATGT
GACTTTGTGCCCCCGCCATGGAGGCCTGGTTTAGAAGATATCGCATGGCTGAAAGGTCCATGTC
CTTACGTCTTGTGACGAACATCTCACCACATGTCTCTGTGTTCCACTCCCCTTTTCAGAGATGG
GCCGCTTTGTTTTGTAGAAATTATTTCTTACAAAAACAGGGCTAGATACTTTCACACAGAGACG
TTTATCACTTTTCCATCCCTGCTGCAAAAATTAGCAATGGATCTGCTGTATTGCTGATGGAACT
AAGAGACCTAGAGAGGCTGGGTAATATGGACTCCCAAGGAGGATCTGCCCATCAAGTACAGAGT
CTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCAACAGA

>DS16_Reverse
GTTACTGATACTATATCGACGCGGCTCTGCATAATGTAGGTGGGACGCTCCACACAAATGGAAC
TTGTGACAACGATCTAGCAGTGAGTATGTTGAAATAAGTGGAACAATATGAAGCGGCCTTCTCT
GCCTGCCTTAGGGAATATTGCCCTTGGGTCGACATTTTATTTATCATGAGGGAGAAGGCCTTAC
CTGTCTGATCTAGCTCTGTCACGCTCTGCTTGTGCCTTGGGTGCTGCTTACACTATGCGATCTC
CTCCGAATCACACTTCATGACGAAACAAACCAATCGCAACCCTTCATTGTGAGAGAAGGACAAG
CTCCCCTTCCGGGAAGGAGAACAGACAGGGATTACAAACTCACATGTCTCTGGGATGGAGCACC
CAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATA

DS19

>DS19_Forward
CCATATATCACTAGACATGATTGTGTATTGGTCTCACTGCGCAATCTTGTCCAAAAAAGAGGGT
TTTGACCATGGGACCAAGCCACGGTTCTCTGACTTGCGATGACTATTGCCGGATGGGAAGACCC
TGCGCCCTATGGCTTTGCTCAATATTCTCACCACCTCACTCTGTTTGCCCTACACTTATGGACA
TGAGGCGCTTTGCTCTGTAGAAATTTTAACATACGAAAACACGGCTAGATCCCATCACACGGGG
TCTTTTATCGGATTTCCTTCCCTCCTACAAGATTTGGCCATAGAATTGCTGAATGGATACTGAA
TTATGACACCTAGGGTGGGCTTGTAATATACACTCCCAAGGAGGATCTGCCCATCAAGTACAGA
GTCTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCA

>DS19_Reverse
AAGGATCTCTGCTACTTGCATGAGGCTCTTCTGATATATGTGTGTGTGATGGAACCCTCAACAA
TGGTTAGTGGGCTGTATCTATCACAGGGTATGTGGACATAAGTTTTGGAAATGAGAGGCCTTTC
TTGTCTCGCTAGGGGAGTACTGACCTTGGTTCGACATATAAGATATCTCGGAGCCGGAAGCCCT
ACATTTCTTATAAGCGACTCCCTTGCTCTGTATGTGACTCGGGGACAACTTCCTTCGTCTCCCT
CTGGTCCCGCACCAACTCCATGACGAATCAACCCAAAAGCAACCCTTCAGGGTAACAGAAAGAC
AAACCCTCCTTCCTGGAACGGTAACACAGACTGCTTGTAAACTCACATGTCTCTGGGATGGAGC
ACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS26

>DS26_Forward
CGGGGCTCACCGACATCGTTAGGTACGGATCACTCCTTGCAATCGTGCCCAGGATGAATGGTTG
TGGCTGAGGGACCAAACCACGGTTGCCTGTCGGGGATGAGCCTCGCAGGGCGGGAGGAACCTGC
TCCCTATCTATTGTGACGATCTTCTCACCACCTGAATCTGATCTCCTCTCCACTTATGGACGGA
CCCTCTTTGTTCTCTAGAAATATCTACTTACGAAAACAAGGCTAGATCCCTGCACACGGAGTGT

TTTATAGGATTTCCTTCCCTCCTGCAAGCTCAGTTCAAGGATCAGCTTTATGACTGCTGAATTG
ATAACACCAATGGTGGGCTTGCAATATACACTCGCAAGGAGGATCTGCCCATCAAGTACAGAGT
CTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCA

>DS26_Reverse
GATCTAGGTACACTGCAGCGGCACTACCTCCTGATCTAGGGTCTCGATGCAACCCACCACTGGA
GTCTGTGGGATAGTGCTCACAGAGTGTACGCGGTCATAAGTGCAGGAAAAGTGAAGTCCCGTCT
TGACAGCTATAGGGAGATCTGCCCTTGCTATCGTCATAATAGTTAACTTGAGGGCGGAGACTTT
ACCTGTCAGATCAGGACCTCCCATGCTCAGTTTGTGACCCAAGGAACAACCCCATTCGTCTAAC
TCACGTACCACATCCACTCCACTCCTCAACACCCCAAACGCAACCCTGGATTGTAACAGAAACA
GAAACCCTCCTTCCTGGAACGGTAACACAGACTGATTGCAAACTCACATGTCTCTGGGATGGAG
CACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATA

DS41
>DS41_Forward
ATGGGGTTGATCCCAGACATCGTGTACATGCCGATTAGTTAGTGTAGTCCGCCCACGAGAGAGT
GCGGTGCTATCCAACGAAAGAGGGCTACCCTGTCTGGTGATCTGTATAGCCGGACGGGTGTAAC
GTGTACGAGCACGATTGTCTAGTTCTTCGGATCCCCCCCATCTTATTGTACGTGCCGATATGGC
ACTAAAGGACATCGCTCCCTCGATATATATACTTACGTACATAAGGGAATACAACATCACACTC
ACTGTATTTAAGGACTTCCATTGCTCCTGCTAAATCTGTACAAGCATTAGCTTAGTGACTTCTG
GCTTGATAATACCAATGGAGGCTTGGCTATAGACATAGCCAAGGAGGATCTGCCCATCAAGTAC
AGAGTCTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCA

>Ds41_Reverse
GAACTCTGATACTTGATGAGGCAGATCCGACTAGACTGTGTGTCTAGCATCCCACCACTGGAGC
TATATGGCAATAATCATCGACTGTGTCTGTGGATATAACTTTAGGTTGTATGAAATCCTTTCAG
CCACTGTCATTGGGTGGCACTGCCCTTTCTATCCTAGCAAAAGACAACACGTATGCAAAAGCTG
GACGTGTCTTATCTACACCTCCCTAAGGTGTTGGGACTTAAAGAACAAGTCACTCCATGTGACG
CTGGTACCGCTTAAACTCCACTGCTCATGACACCTCCGAGAACTATGCCTTGAGACAACAGCAC
AGACCCTCCATCCTAGGCCGGTAGCACTGACTGATTGTAAACTCACATGTCTCTGGGATGGAGC
ACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS55
>DS55_Forward

ATCGTGGTTCTTATCCTGATGGGTGGCGTTTGGTCTCTTACACTATGCTTTTGCGCTAAGAGGA
GATGGTGTTGATCGAAGAAACCACCAAGGTCCGAAGTGAGCAAGAGAATCCCAGGATGGAAAGG
TAATGCTCCCTCTCGCATGTTGACGAAAATTTGACCCACTCACTCTGTGTTCCTCTCCACTTTG
CAGAGATGGGGCGCTTCGTTTTGACAAACTTATTACTTACGTTCACAAGGCTAGATACTTTCAC
ACGCGTACTTAGATCGCATTTCCATCCCTACCGGATAAATAAGTATTGGCTCTGCTGTATGAAT
GATGAAACTAAGACATCTAGAGAGGGCTTGTAATATGGACTCCCAAGGAGGATCTGCCCATCAA
GTACAGAGTCTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCA

>DS55_Reverse
AGGGGGGCATATCATTCATAATTGCATAGACAATACTTATTTATGTGGTACTGATGGAAGAGAG
CAAAGAGAACCTGAGTGAAAGTATCTACCACAAGTTATGTAGGAAATCGCTCTAACGAACTGAA
TAGGCCCGTCTAGACCACGCGGAGTGGGATACATAGTGTTGCTTCCAGACTTTAGTCAACTTGA

GGGAGGGAGCTTTACCTTTCTGACCTGGGATTGCCATGCTCTGTATGTCCCTTGGGGGCTTCTC
CCATTGTCCAGATCTCCTCTGACTTCCAAATCATGCTGAAAGAAACCTAACGCAACCCTGCATG
ATGCGAGAAAGACAAACTTCCTTCAATGGAAGTGGAACACACACGGAATATAAACTCACATGTC
TCTGGGATGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS56

>DS56_Forward
CCCTATAATACCCAGATATGGTTGTGTTTGGTCTCTCTCAGCATGATTTTGAAGTAAGAGGAGA
TTGGATTAAGCAAAGGCACCACCAAGGTTCAAGACGAGAAGAGAATCGCCTGGCGTAAAGGTCA
AGCTCCCGCCGTCATAATGTTGATCAACTCACCCCATGTATCTGAATTCCACTCCACTTTGCAG
CGATGGGCGACATCATTCTGAAGAAATTATAAATGAGGTACACAGGGCTAGATACTTGCACAGA
GATTGGATTTACACTCTTCCTGCGCAACCAGATACATTATCAATGGAACTGCTTTGTGAATGAT
AACTCTAAAAGATTTAGAGAGGGTGGGCTATTGGGATTCCCCAGGAGGATCTGCCCATCAAGTA
CAGAGTCTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCA

>DS56_Reverse
TTATCTTTTGCTACTTGTTTGAGACAGTACCTACTGATGTGGTACTGATGCAAGAGTGAAAAGA
GAGTTAGTGTGCAAGTATCTATCACTAGGTACTTTGTCATAAGTCTAACAAACAGAAGCAGCCC
ATCACTGCAACGGGTTGTGGAATATATAGCGGTGGATCGAGACTTTAGTCATCTTGAGGGAGAA
AGCTTTACCTTTCTGTCCTGGGATTCTCTTGATCTCTTTGGACCTTGGGGGCTACTCACATTGT
CCACACCACCTCTGACTTAAAAATCATGCTGAAAGAAACCAATCGCAACCCATCATGGTAACAG
AAAGACAAACTTCCTTCAATGGAACTGGAACACACACGGAATACAAACTCACATGTCTCTGGGA
TGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATAC

DS60

>DS60_Forward
ACCTGGTCTCTTATCCTGATGGGTTGTGTTTGGTTTCTTTCAGCATGATTTTGAAGTCAGAGGA
GATGTGGTCGATGGAAGAAACCACGACGGCCTATAGTGAGCAAGAGAATCCCAGGACAGAAAGG
TAAAGCTCCCTCCCTCAAGTTGACAAAAATCTCACCCCACCACTCTGTATTCCACTCCCCTTTG
CAGAGATGGGCCGCTTCATTTTGTAAGACTTATTTCATACATACACAGTGCTAGATACTTTCAC
ACAGGTTCTTTTTTCACTCTTCCATCCCAACCACATAAATAAGTATTGTCTCTACTTTATGAAT
GATAAAACTAAGAGATTTAGAGAGGCTGTGTAATTTGGATTCCCAAGGAGGATCTGCCCATCAA
GTACAGAGTCTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCACGGACT

>DS60_Reverse
TCGTGTCTTATCATTGCTAAAGTAGAGACGATACTTATTTATGTGGTTGGGATGGAATAGTCAG
CAATGAACCTGTGTGCAAGTATCTAGGACTGTGTATGTATGAAGTAATTCTTACAAAATGAAGC
GGCCCGTCTCTGCAAATGGGAGTGGAATACAGACTGGTGGGGCGTGATTTTAGTTAACTTGAGG
GAGGGAGCCCTACATTTCATTCCTGGGATTCTCTTGCTCGGTTTGTACCTTGGGGGTTGGTTAC
ATCGACCACATCTCCTCTGACTTCAAAATCATGATGAAACAAACCAAAAGCAACCCTTGTGGAG
AAGAGAAAGACAAGCTTCCCTCAATGGGAGTGGAGCACAGACGTCATATAAACTCACATGTCTC
TGGGATGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATAAAGCTGTTC

DS62

>DS62_Forward
TCTTCGTTATACCTAGATATGGTTGTGTTTCGGTTTCATAGTGCATGCTGTTCGAAGAAGAGGG
GATGGGGTCAAGGGACCAAACCACCGAGCCCCGATCTGGCCATGAGAATCGCAGGATGGGAAGG

AAAAGTTGCCTACCTCTTTTCTCATTCATCGCACCCCACGACTCTGAATGGCAGTCCCTTTTGG
AGAGATTAGGGGCTTCGTTTTGTAAAACTTATTACTTACAATCACAGGGCTAGATACTTTCACA
CAGATTGTTTTATCACTCTTCCTTCCCAACCGCATAAATTATCAATGGAACTGCTTTATGAATG
ATAACTCTAAGAGATTTAGAGTGGGTGGGCTATATAGATTCCCCAGGAGGATCTGCCCATCAAG
TACAGAGTCTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCA

>DS62_Reverse
GTGGGGCATACTCTGATACTTGATGAGACAGTACTTACTGGTCTAGGTGTTATGGAAGCCTCAG
ATGAGAACCAGTGTGAAGGTTCTATCACTGTGTATGTTGTAATAGCTCTTACGAAATGAAGCGG
CCCATCTCTGCCTAGGGGTATGGAATACATGGCCGTGGGGCGACATATGTGTCATCTTGAGGGA
GGGAACTTTACGTTCCTGTCCAGGACTTCTCTTGAGCGCTTTGGACCTCGGTGGCTACTCCCAT
TGTCCACACCACCTATGACTTACAAATCATGATGAATCAAACCACACGGAACCCAGCATTGTAA
CAGAAAGACAAACTCCCTTCCCTGGAACTGGAACACACACGGATTACAAACTCACATGTCTCTG
GGATGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATATAA

DS63
>DS63_Forward
TTCTCTTCCACTACACAGGGTGGTTTATCGATCACTCTCTACAATCTTGCCCAAGATGAATAGA
TTGGGCTGTTCGACCAAACCAGGGAGCTCCGAACGAGAGGCGGTCCCATTGCGTATGTTACATG
TTCGCGCCGGATTGGGACGAACTTCTCACCACCTGACTACGAATGCACTCCCCTTATGGACTGG
GCGTCACTACTTTCGAGAAATATTTACTTACATTAGCAAGGGCAGATACCATCACACAGGGAGT
TTTTACGGTTTTCCTTCCCTCCTGCAAGATCTGTCCAAGGATCTACTGAATGACTGATGAATTG
ATAACACCTATGGTGGGCTTGTAATATACACTCGCAAGGAGGATCTGCCCATCAAGTACAGAGT
CTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCG

>DS63_Reverse
TATAACTCTGATACTTGATGGGCAGTATCCTACTGACCTATTTGTCTAGGTGCCCAGCACAGGA
GCTATATATGAAAGATCTCTCCTGTGTACGTGGAATAGCTTTTACGAAATGAGAAAGCCTTTCC
TGTCTGCTAGATGGAATACTTGCCTAGGAGGCACATAATAGTTAACTCGGGGCCGAAAACTCTA
CCTGTCTGATCTGCGCCTCCATTGATGTCTTTCTGACCAAAGGGGCTTCTCAATTCATCTACCT
CTGGGCCCACACCCACTCCGTTACGAATCACCCCTCAAGCAACCCTGGATGATGACAAAAGCAC
AAACCCTCCTTCCTGGAACGGTAACACAGACTGCTTGTTAACTCACATGTCTCTGGGATGGAGC
ACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

DS65
>DS65_Forward
CCTTATACACGCAGGAATGGTGTACGTGGTGATCTCATTCCGCAAGCTTTTCAAGCAGACGAGT
CGGGGCTGTTGGACCACACAACCGTACCCTGTATGAGATGACTATCCTAGGGCGGGAATGACCT
GTGCCCTACCTATAATGACGCTCATCTCACTACCCCCATACGTGTGTCGGCCCTATTTGGAGAA
AAGCACTTTGCTCTCGCGAAAATTTAAATTACATAAACACGGCTATATCCTATCACACGAACTG
TATTTATCCACTTCAATCCCTCCTGCTAGATTGGTACAAAGATTTACTTTGTGACTGCTGAATT
GATAATACCAATAGTGGGCTTGTAATATACACTCCCAAGGAGGATCTGCCCATCAAGTACAGAG
TCTGGCCACTTTTCTGAAGGGCTGCTGTGTAATTTGGATTCCCA

>DS65_Reverse
AAATCATCTGCTACTATCACATGCAATTCTTACTGTCCTGTGTGTCTAGGCAATACACCAATGG
AGCTTTCTGAGATTGGTATACTCAGTGGGTACGTGTACCTAATTTTTGCAATGTGACGAAGGCC

CTTTAGGAGAGTGTATGGGGGATAATATCCCTTGCTATCCACATATTATGTATCATGGGGGCCA
GAGCTCTAGACGTCAGATCTGCACGTCCCATGGTCTGTTGGTAACTCAGGGACTGGTCACGTCG
TCTCCCTTCGTCCGATCCCCACTCCGTTACGCATCACCCCTAAAGCAACCCTGCATGATAACCA
AAACACAGACTCTCCTTCCTGGAACGATAACACAGACTGATTGTAAACTCACATGTCTCTGGGA
TGGAGCACCCAGGAGACAAGCAAAGTGGTGGAGCAGACACGTCATAA

APPENDIX IV: SAMPLE DEMOGRAPHICS.

Sample No	Age	Sex	Ethnicity	Geographic Location
1	36	Female	Black	Lusaka
2	31	Female	Black	Lusaka
3	46	Female	Black	Lusaka
4	28	Female	Black	Lusaka
5	44	Female	Black	Lusaka
6	45	Female	Black	Lusaka
7	36	Female	Black	Lusaka
8	50	Female	Black	Lusaka
9	50	Female	Black	Lusaka
10	37	Female	Black	Lusaka
11	47	Female	Black	Lusaka
12	42	Female	Black	Lusaka
13	37	Female	Black	Lusaka
14	41	Female	Black	Lusaka
15	43	Female	Black	Lusaka
16	41	Female	Black	Lusaka
17	50	Female	Black	Lusaka
18	33	Female	Black	Lusaka
19	40	Female	Black	Lusaka
20	38	Female	Black	Lusaka
21	38	Female	Black	Lusaka
22	35	Female	Black	Lusaka
23	47	Female	Black	Lusaka
24	45	Female	Black	Lusaka
25	37	Female	Black	Lusaka
26	35	Female	Black	Lusaka
27	39	Female	Black	Lusaka
28	37	Female	Black	Lusaka
29	43	Female	Black	Lusaka
30	48	Female	Black	Lusaka
31	50	Female	Black	Lusaka
32	30	Female	Black	Lusaka
33	41	Female	Black	Lusaka
34	43	Female	Black	Lusaka
35	37	Female	Black	Lusaka

36	44	Female	Black	Lusaka
37	50	Female	Black	Lusaka
38	35	Female	Black	Lusaka
39	38	Female	Black	Lusaka
40	46	Female	Black	Lusaka
41	48	Female	Black	Lusaka
42	45	Female	Black	Lusaka
43	45	Female	Black	Lusaka
44	35	Female	Black	Lusaka
45	49	Female	Black	Lusaka
46	49	Female	Black	Lusaka
47	45	Female	Black	Lusaka
48	49	Female	Black	Lusaka
49	38	Female	Black	Lusaka
50	44	Female	Black	Lusaka
51	39	Female	Black	Lusaka
52	39	Female	Black	Lusaka
53	44	Female	Black	Lusaka
54	37	Female	Black	Lusaka
55	24	Female	Black	Lusaka
56	50	Female	Black	Lusaka
57	42	Female	Black	Lusaka
58	48	Female	Black	Lusaka
59	49	Female	Black	Lusaka
60	32	Female	Black	Lusaka
61	46	Female	Black	Lusaka
62	43	Female	Black	Lusaka
63	31	Female	Black	Lusaka
64	41	Female	Black	Lusaka
65	49	Female	Black	Lusaka
66	43	Female	Black	Lusaka
67	35	Female	Black	Lusaka
68	37	Female	Black	Lusaka
69	36	Female	Black	Lusaka
70	43	Female	Black	Lusaka
71	41	Female	Black	Lusaka
72	32	Female	Black	Lusaka
73	45	Female	Black	Lusaka