**UNIVERSITY OF GHANA**

**COLLEGE OF HEALTH SCIENCES**

# MUT L HOMOLOG 1 (MLH1) EXON 16 MUTATIONAL ANALYSIS IN GHANAIAN COLORECTAL CANCER PATIENTS ON TREATMENT AT THE KORLE- BU TEACHING HOSPITAL.

**BY**

**AGNES SITSOFE DORBU**

**(10805179)**

**THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF MPHIL IN MEDICAL BIOCHEMISTRY DEGREE**

**SEPTEMBER, 2023**

# DECLARATION

## DECLARATION

I, Agnes Sitsofe Dorbu, of MPhil Medical Biochemistry Department hereby declare that this project, aside other sited articles and references, is the result of a research I duly conducted; which was supervised by Dr. Bartholomew Dzudzor, a senior lecturer at the Department of Medical Biochemistry (UGMS-Korle-Bu), and Dr. Antoinette Bediako-Bowan, a surgeon at the Department of Surgery, College of Health Sciences and Dr. Lucas Amenga- Etego, also a senior lecturer at West African Center for Cell Biology of Infectious Pathogens (WACCBIP)- University of Ghana, Legon.

Signature ..... Date 29th July, 2022

Agnes Sitsofe Dorbu (Student)

Signature ..... Date 29th July, 2022

Dr. Bartholomew Dzudzor (Supervisor)

Signature ..... Date 29th July 2022

Dr. Lucas Amenga- Etego (Co-supervisor)

Signature ..... Date 29th July 2022

Dr. Antoinette Bediako-Bowan (Co-supervisor)





ii

## DEDICATION

For their unwavering support, patience, prayers, and love during my studies, I loving dedicate this thesis to my wonderful husband, Daniel Vivor, and two children, Kevin Selator Vivor and Daniella Nalana Kafui Vivor.

I also dedicate this book to my mother-in-law (Madam Annie Aku Agbleze), who has beena huge support to me and her grandchildren throughout the entire time. Thank you very much! And know that with you this was made possible.

Lastly, to my late father (Mr. Dickson Fiawona) and Mother (Madam Saviour Amediku) for their endless support.

**ACKNOWLEDGEMENT**

Throughout this research, the Lord Almighty has shown me his abundant grace, mercies, and protection. I thank HIM for the tremendous strength He has bestowed upon me to complete this course successfully.

I am very grateful to Dr. Bartholomew Dzudzor, my supervisor, for his invaluable advice, guidance, support, and contributions in all aspects during this work. I acknowledge that without his counsel and direction, this book and qualification wouldn't be realized. God bless you, I say.

I show appreciation to my co- supervisors: Dr. Lucas Amenga-Etego of WACCBIP, University of Ghana, Legon, and Dr. Antoinette Bediako-Bowan of Dept. of Surgery for their enormous contributions and supports, towards a successful completion of this work.I say God bless you both! I thank Dr. Tagoe also, for his constant checks and concern for me to finish the work on time.

I thank Mr. Elvis Twumasi and Mr. Dominic Amuzu, all of (WACCBIP), University of Ghana, Legon, for their time and patience to always help me out with explanations and directions. A special thanks goes to Mr. Bright Yemi, also of WACCIB, for his constant support and services rendered during the bench work. The brain behind the bioinformatics analysis of this work, Mr. Collins M'oranga Misita -I say thank you.

Another thanks goes to Mr. Hope Azasu and Mr. Emmanuel Bannerman, all of UGMS, for helping greatly in the sample collection, processing and storage. I really do appreciate and say thank you gentlemen. To all my friends and senior colleagues within the university,

especially Mrs. Dorothy Bendah and Mrs. Gifty Owusu Antwi, I say thank you for your encouragement and support.

And lastly, my backbone Mr. Daniel Vivor, my husband and two wonderful children Kevin and Daniella, have been my strength and aspirations. I say I love and appreciate you guys. Not forgetting my mom- Madam Saviour Amediku and wonderful mother in law, Madam Annie Aku Agbleze, and entire family, I ask God to bless you all for your enormous support. I am very grateful!

# TABLE OF CONTENTS

Table of Contents

## LIST OF TABLES



## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AFAP: | Attenuated Familial Adenomatous Polyposis |
| APC: | Adenomatous polyposis coli |
| CD: | Crohn's disease |
| CRC: | Colorectal Cancer |
| DNA: | Deoxyribose Nucleic Acid |
| EDTA: | Ethylenediaminetetraacetic acid |
| FAP: | Familial Adenomatous Polyposis |
| hMSH2: | human Mut S Homolog 2 |
| hMSH6: | human Mut S Homolog 6 |
| hMLH1: | human Mut L Homolog 1 |
| hMLH3: | human Mut L Homolog 3 |
| hMSH3 | human Mut S Homolog 3 |
| HNPCC: | Hereditary Non-Polyposis Colorectal Cancer |
| HPMS1: | Human Post-meiotic Segregation 1 |
| HPMS2: | Human Post-meiotic Segregation 2 |
| KBTH: | Korle Bu Teaching Hospital |
| LS: | Lynch syndrome |
| M | Molecular Weight Ladder |
| MMR: | Mismatch Repair |
| PCNA: | Proliferating Cell Nuclear Antigen |

| | |
|---|---|
| PCR: | Polymerase Chain Reaction |
| QC: | Quality control |
| RFC: | Replicative Factor C |
| RER+: | Positive Replication Error |
| SD: | Standard Deviation |
| SNPS: | Single Nucleotide Polymorphisms |
| SPSS: | Statistical Package for the Social Sciences |
| UC: | Ulcerative colitis |
| WACCBIP: | West African Center for Cell Biology of Infectious Pathogens |

## ABSTRACT

**Background:** Colorectal cancer (CRC), commonly known as cancer of the colon or rectum is a significant cause of morbidity and mortality around the world. In developing nations, notably in Sub-Saharan Africa, colorectal cancer is becoming more common. According to World Cancer Research Fund International Statistics in 2020, colorectal cancer is the third most common cancer worldwide. It is the third most common cancer in men and the second most common cancer in women with approximately 1.9 million new cases and 900,000 expected deaths. In Ghana's capital, Accra, colorectal cancer ranks third among the most frequent cancers, and also ranked tenth among cancer death in the country. Mutations in the hMLH1 gene are responsible for the bulk of known instances of hereditary non-polyposis colorectal cancer despite the presence of a minor number of variation in the other mismatch repair (MMR) genes (such as hMSH2,hMSH3, hMSH6, hPMS1, and hPMS2). In fact, over 300 germ line mutations have been identified in the MLH1 gene which can be diagnosed easily by automated sequencing. Accordingly, exons 16 and 18 of the MLH1 gene are designated as genetic hotspots for mutations, which agrees with other populations findings. Several studies have also discovered alterations in this gene and exon, and depending on where you live, its presence varies.

**Aim:** Study therefore looked at mutations in exon 16 of the MLH1 gene associated with colorectal cancer in Ghanaian colorectal cancer patients.

**Methodology:** This was a case-control study in which 71 Ghanaian confirmed colorectal cancer patients (cases), who had been clinically and histologically diagnosed by doctors, compared to 68 non-CRC healthy Ghanaians (controls). A questionnaire was used to obtain patients information, and 5mls of whole blood was taken into an EDTA tube. The buffy

coat (white blood cells) obtained from the whole blood was used to extract DNA. Primers targeting exon 16 of the MLH1 gene were designed and used to amplify exon 16 from the extracted DNA using polymerase chain reaction (PCR). Amplified DNA (amplicons) were afterwards sequenced and the sequenced products subjected to bioinformatics analysis to look for any genetic alterations.

**Results:** Out of the 71 case samples, 28(39.4%) samples DNA amplification was detected on the agarose gel to the expected band size of 195 bps. All the 68 control samples were amplified to the expected band size of 195 bps.

Out of the 139 study samples collected, 96 samples (71 cases, 25 controls) were sequenced. For the case samples; 14(19.7%) out of the 71 case samples were without MLH1 exon 16 region; 30(42.3%) samples had good nucleotide sequence of which 2 had a single guanine (G) nucleotide insertion when aligned with healthy control and reference sequence. Twenty-five (25) case samples (35.2%) had sequences that were not clear.

For the control samples, all (25) had fairly good nucleotide sequences that had 100% alignment with the reference sequence.

**Conclusion:** This study is the first compiled molecular information on the Ghanaian CRC population. The study reported two types of mutations; whole exon (16) deletion and a single guanine nucleotide insertion (frameshift) mutation at codon 596 of the coding sequence of exon 16 MLH1 protein. Whole exon (16) deletion in this study accounted for 19.7% (14/71) of the total mutations found within the MLH1 exon 16 gene while insertion frameshift mutation was observed in 2.8% (2/71) of the cases.

# CHAPTER ONE

## 1.0 Introduction

## 1.1 Research Background

Cancer is a type of abnormal cell growth that can assault or spread throughout the body. It usually begins as a benign tumour, such as a polyp, and develops into a malignant tumour over time. Colorectal cancer (CRC) specifically affects the colon and rectum and is, one of the most widespread types of cancer in the world (Rawla *et al.*, 2019). Depending on the tumour's location, size, and presence of metastases; abdominal pain, rectal bleeding, changed bowel habits and involuntary weight loss are all common symptoms (Adelstein *et al.*, 2011).

Numerous elements have been linked to the occurrence of colorectal cancer, and persons with certain illnesses or diseases have an 80% probability of developing the disease. Obesity, inactivity, poor diet, smoking and excessive alcohol consumption contribute to over 80% of all instances of colon cancer (Haggar and Boushey, 2009). There are also several aspects to consider; including age, the presence of polyps, inflammatory bowel disease, lifestyle, genetic background and family history (Rawla *et al.*, 2019). Lynch syndrome (hereditary nonpolyposis colorectal cancer, HNPCC) and familial adenomatous polyposis (FAP) both have a genetic link to colon cancer susceptibility.

In most African research, colorectal cancer accounts for 3-6% of all malignancies, which looks rare and is especially common in young Africans (Adesanya and Darocha Afodu, 2000), where males have a much higher incidence and fatality rate than females

(Siegel *et al.*, 2012). According to some studies, colorectal cancer is on the decline in affluent countries but on the rise across Africa (Siegel *et al.*, 2011). In countries like Kenya, Nigeria and Zimbabwe, the prevalence is gradually rising (Katsidzira *et al.*, 2017). Since 1991, CRC incidence has been steadily rising in Zimbabwe at a rate of about 4% per year (Chokunonga *et al.*, 2013)

In Ghana's capital Accra, colorectal cancer ranks third among the most frequent cancers, and also ranked 10$^{th}$ among cancer death in the country (Biritwum *et al.*, 2000). In both men and women, it is the eighth and ninth major cancer death respectively. (Wiredu & Armah, 2006). Increase in colorectal adenocarcinoma in Ghana, according to Naeder and Acheampong (1994), is related to an increase in population life expectancy and a minor improvement in the disease detection.

Dakubo *et al.*, (2010) also discovered that, as the Accra population ages, the prevalence of colorectal cancer has increased, with adenocarcinoma being the most common histological form. Indirect evidence suggests that genetic predisposition contributes to the development of colon cancer in Africa, although features are yet to be identified, and many genetic alterations are also yet to be examined and profiled. (Raskin *et al.*, 2013; Cronje *et al.*, 2009).

Molecular markers have been used extensively to investigate the biology and prognosis of CRC. It is believed that one of these markers is DNA mismatch repair (MMR); which contributes to the onset and prognosis of CRC (Gelsomino *et al.*, 2016). Approximately 10-20% of sporadic CRCs are linked to DNA MMR gene

malfunction (Hou *et al.*, 2018). Several MMR genes, including hMSH2, hMSH6, hPMSH1, hMLH1, hMSH3 and hPMSH2 have been discovered to cause microsallite

instability (MSI) formation(Ogino and Goel, 2008). Over 90% of MSI development is attributed to the hMLH1 andhMSH2 genes. While their association to CRC disease progression and prognosis has been extensively researched, the findings are mixed (Aparicio *et al.,* 2013; Russo *et al.,* 2009; Wu *et al.,* 2013).

The MLH1 gene mutation accounts for the majority of hereditary non-polyposis colorectal cancer (HNPCC) cases reported (Peltomak**,**2001), with some mutations in the other genes being identified. In fact, the MLH1 gene has been linked to over 300 germ cell alterations. Again, the gene has also been found as having mutations by several researchers, and its existence varies by geographical location (Domingo and Schwartz, 2005). Exons16 and 18 of the MLH1 gene, on the other hand, are designated as mutant gene hotspots (Dominguez Valentin *et al.,* 2013), which corresponds to evidence from other populations (Plazzer *et al.,* 2013; Peltomaki *et al.,* 1997; Nilbert *et al.,* 2009).

DNA damage accumulates in cells over time as a result of exposure to exogenous chemicals and physical agents (i.e., benzo[a]pyrene, polychlorinated biphenyls, dioxin, cigarette smoke, asbestos, ultraviolet light, radon), as well as endogenous reactive metabolites including reactive oxygen and nitrogen species (ROS and NOS). Errors that occur during normal DNA breakdown or inappropriate DNA treatment activities,

such as DNA replication, recombination, and repair are another type of DNA damage. The nucleotide groove generates DNA incompatibilities when DNA is manufactured at variable rates, especially when DNA-specific polymers are used (Li, 2008).

Replication fidelity is rather high in replicative DNA polymerases (McCulloch and Kunkel, 2008), but lower in translesion DNA polymerases that particularly suffer DNA damage (Andersen *et al.,* 2005).

If untreated, DNA damage can cause mutations in somatic or germline cells, changing their biological phenotype and resulting in disease. Cells have a number of mechanisms for repairing DNA damage and consequently preventing mutations to avoid such adverse effects and to safeguard genomic integrity. The important process known as DNA mismatch repair is one such system (Ruddon, 2007).

Mismatch repair genes during DNA replication produce enzymes which are able to detect and correct mismatched base pairs. A lack of function of the repair gene causes microsatellite instability (MSI), which is the accumulation of significant changesin the length of tiny, repetitive DNA sequences ( Faustino and Oliveira, 2020).

There is therefore a rapid accumulation of gene mutations in proto-oncogenes and cancer suppressor genes as a result of the faulty MMR, which may disrupt normal cell growth and aid the development of cancer (Popat *et al.*, 2005; Kerr and Midgley, 2010; Sargent *et al.*, 2010). MMR deficiencies enhance the frequency of spontaneous mutations because they minimizes replication mistakes (Li, 2008).

Hereditary and sporadic malignancies have been linked to MMR deficiency in human cells (Kolodner & Marsischky, 1999; Modrich & Lahue, 1996), and the MMR system is essential for cell cycle arrest and/or programmed cell death in response to certain types of cancer or DNA damage (Li, 2008; Stojic *et al.*, 2004). MMR genes are involved in the DNA damage response system, which eliminates severely damaged cells and suppresses both short-term and long-term mutagenesis and oncogenesis, with focus on colorectal cancer.

## 1.2 Problem Statement

Indirect evidence suggests that genetic predisposition contributes to the development of colon cancer in Africa, although features are yet to be identified. Many genetic alterations are also yet to be examined and profiled (Raskin*et al.,* 2013; Cronje *et al.,* 2009).

According to Chalya *et al*. (2013), late admission is attributable in part to a lack of local data on current colorectal cancer trends, as well as a lack of public knowledge about the necessity of early admission for early diagnosis and treatment.

Secondly, malignancies can have diverse genetic origins and express different proteins from one patient to the next, making precision and individualized treatment a viable and preferred option, since the most effective treatment solutions are tailored to that particular patient or group.

CRC due to its heterogeneity can serve as a model disease to fully implement the concept of precision and personalize medicine (PPM) (Krzyszczyk *et al.*,2018). However, the tools needed to put it into action are still lacking. Numerous attempts are therefore made to gather PPM data such as mutations, changes in gene expression, changes in metabolites, and more recently in the immune nucleus, in order to define the molecular variations among tumours and also; to identify a patient or group of patients. Even that, little is still known about the genetic profile, molecular pattern and mutations that are unique to Ghanaians suffering from CRC, for which the PPM concept can be fully realized.

### 1.3 Justification

This study would generate molecular data unique to CRC patients in Ghana, and ultimately aid in the proper management of cases observed in Ghana. Data from this study could also be translated into the development of biomarkers for early diagnosis / prognosis of better CRC treatment in Ghanaians and other African descent.

Furthermore, the medical community would be able to assess the impact of precision medicine as a result of this research. In that, knowledge of molecular analysis would aid in the improvement of present treatment regimens, allowing patients to live longer.

Finally, being the first to determine the possible mutations in exon 16 of the MLH1 gene among Ghanaian CRC patients attending Korle bu teaching hospital, this study would also serve as a baseline data for other studies in this area.

### 1.4 Hypothesis

There is no mutation in exon 16 of MLH1 gene in CRC patients in Ghana

### 1.5 Main Aim of Research:

To investigate exon 16 of the MLH1 gene for possible mutations that contributed to the development of colorectal cancer in some Ghanaian patients attending the Korle Bu Teaching Hospital.

### 1.6 Specific Objectives of Research:

i To amplify exon 16 of MLH1 gene of CRC patients and healthy controls

ii. To determine exon 16 DNA sequence for cases and healthy controls.

iii. To carry out bioinformatics analysis (nucleotide variation) on the sequenced exon 16 of MLH1 gene in cases and healthy controls.

iv. To evaluate mutation frequencies in cases compared to the healthy control.

v. To evaluate colorectal cancer risk factors of the participants.

# CHAPTER TWO

## 2.0 Literature Review

### 2.1 Large Intestine's Anatomical Features

The large intestine is the last section of the digestive tract, connecting the terminal ileum and the anal canal (Scanlon and Sanders, 2014). In human adults, cecum, ascending colon, liver flexion, transverse colon, spleen flexion, descending colon, and sigmoid colon are all tube structures measuring roughly 1.5 meters in length. The rectum starts at the inflection of the peritoneum and repeats the flexion of the sacrum, finishing at the anal canal in an adult (figure 2.1). Rectal tumours are less than 16 cm from the margin of the anus and are at least partially in the superior rectal artery's blood supply, which act largely as a retention reservoir (Scanlon and Sanders, 2014).

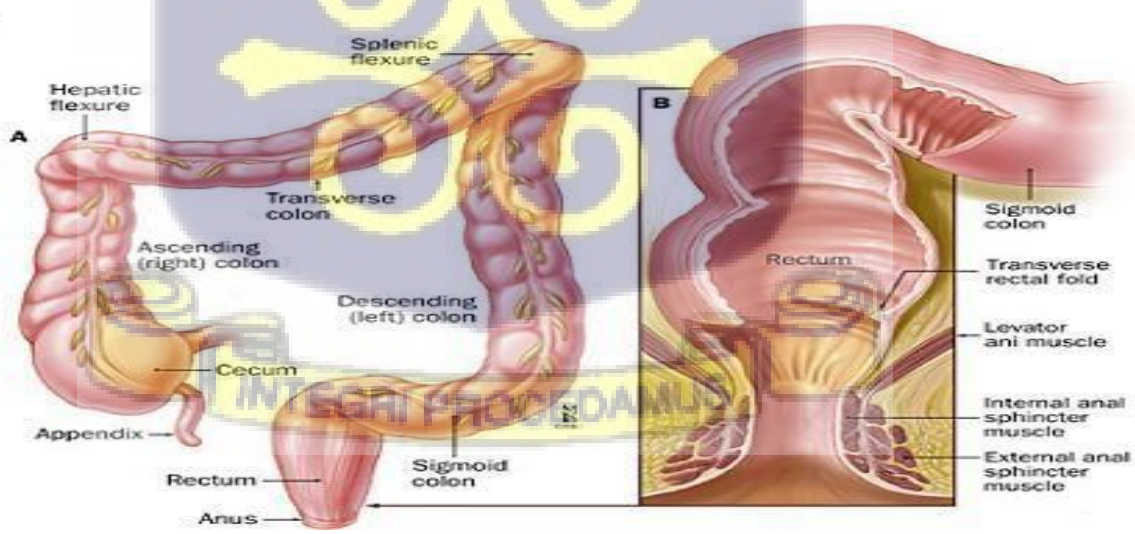



**Figure 2.1: [A]: Normal Colorectal Anatomy of the Colon, and [B]: Rectum**

file:///C:/Users/hop/Documents/EGDownloads/sporadic_nonhereditary_colorectal_cancer.pdf :23/05/2021, 03:05 pm

## 2.2 Large Intestine's Microscopic and Histological Features

Mucosa, submucosa, outer / intrinsic muscular layer (including round and smooth muscle layers) and serosa were discovered through histological and microscopic investigation of the colon (Yeatman, 2001). Due to the continuous proliferation of normal cells to replace those released from the intestinal wall into the lumen, the early genetic alterationsoccur mostly in the mucosal cells. Apart from the plentiful goblet cells, the rectal mucosaand colon are similar (Araki *et al*., 1996).

Blood veins, lymph vessels and terminal nerve fibers are all found within the submucosa (figure 2.2).This is a crucial layer that has an impact on cancer's development. This is an important factor in the emergence of cancer because when cancer invades this part of the intestinalwall, it can infiltrate the circulatory and lymphatic systems, allowing it to move throughout the body across great distances (Yeatman, 2001).

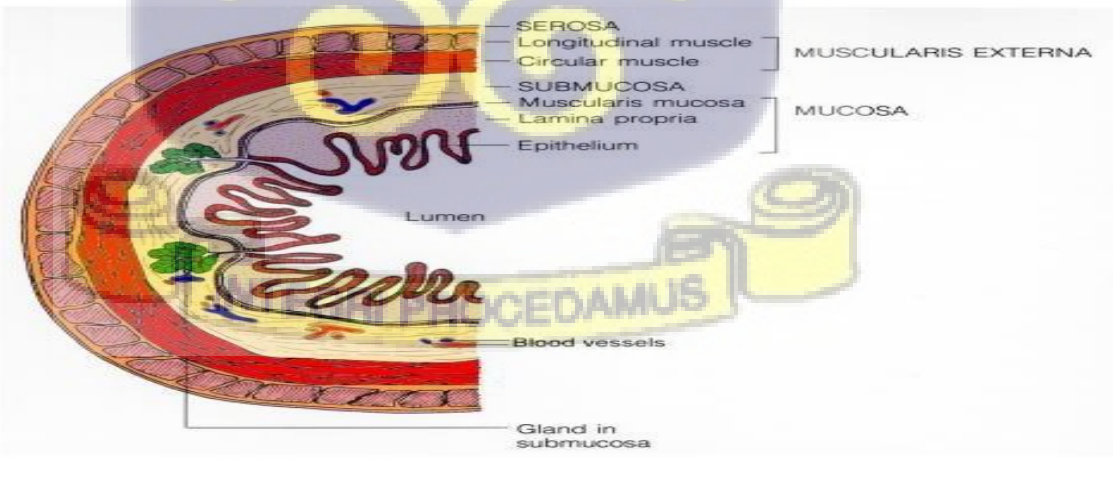



**Figure 2.2: Microscopic layers of the colon**

Image source; https://www.cancerquest.org/patients/cancer-type/colon-and-rectal-cancer.html; 26/05/2021; 10:25am
;

## 2.3 Colorectal Cancer

Both colon and rectal malignancies are classified as colorectal cancer. Colorectal cancer is a malignant growth in the colon that can remain confined for a longtime before spreading to the lymph nodes and other body parts through the gut wall. The caecum to the sigmoid colon approximately 15 cm above the anus's border is where colon cancer spreads, while the recto-sigmoid colon to the anus'sborder is where rectal cancer spreads (figure 2.3) (de Heer, 2007).

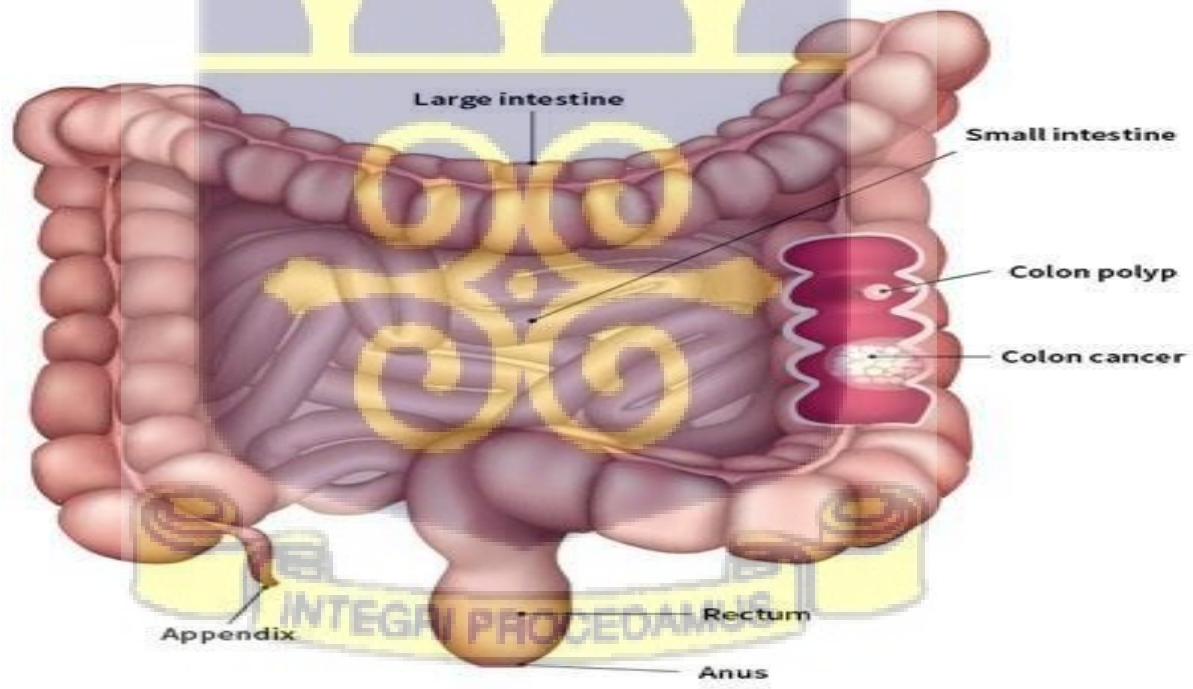



**Figure 2.3: Alimentary canal with colon cancer**

Source: https://www.inovanewsroom.org/ifh/2019/06/get-screened-for-colorectal-cancer/ 26/05/2021. .

## 2.4 Colorectal Cancer Aetiology

Colorectal cancer has a complicated aetiology that involves interactions between environmental variables and inherited vulnerability (Rattray *et al.*, 2017). Spontaneous CRC and inherited CRC are the two forms of colorectal cancer. About two-thirds (70-75%) of new CRC cases are spontaneous / sporadic (i.e., develop among persons who have no history of the illness in their family) (Yamagishi *et al.*, 2016). Due to aetiological factors such as age, lifestyle, and environmental effects, required genetic modifications arise de novo in these rare cases.

## 2.5 Colorectal Cancer that is not Sporadic (Hereditary)

Five to ten percent of all cases of colorectal cancer, or almost one-third of all cases, are hereditary (Cherry, 2011). Since hereditary CRC runs in families, first-degree relatives of patients with newly discovered adenomas or invasive CRCs are at higher risk. Some of these family clusters are related to germline mutations in certain genes, while others are related to similar behaviour or environmental exposure across family members (Cherry, 2011).

Hereditary Nonpolyposis Colorectal Cancer Syndrome (HNPCC) (Lynch's syndrome) and familial adenomatous polyposis syndrome (FAP), are the two most frequent hereditary disorders (Segelman, 2012).These syndromes frequently result in cancer that emerges at a younger age. Gardner's syndrome, Turkot's syndrome, Old Field syndrome, Peitz-Yegers disease, Juvenile polyposis and Cowden's disease are some of the less prevalent syndromes and illnesses (Scott 2003).

## 2.6. Inherited Genetic Syndrome In Relation To CRC

### 2.6.1 HNPCC Syndrome (Hereditary Non-Polyposis Colorectal Cancer)

HNPCC is an autosomal dominant cancer susceptibility disorder with high penetrance (80–85%) (Lynch and Smyrk, 1996), and mutations in five mismatch repair genes: hMSH2, hMLH1, hPMS1, hPMS2, and GTBP (MSH6) have been identified (Leach *et al.*, 1993; Nicolaides *et al.*, 1994; Akiyama *et al.*, 1997; Miyaki *et al.*, 1997; Papadopoulos et al., 1994). Despite the fact that a prospective study in Finland discovered that only 10 (2%) out of 509 colorectal cancer patients had either hMLH1 or hMSH2 germline mutations (Aaltonen *et al.*, 1998), previously HNPCC was assumed to be responsible for 5-10% of all colorectal malignancies (Thorson *et al.*, 1999). But this was thought to be overestimated since the Finnish foundation mutation 1 was present in five out of the ten individuals (Jass, 2006).

It is now widely acknowledged with the progress of molecular genetics that, the distinctions between Lynch I and II syndromes are arbitrary and do not represent two illnesses caused by mutations at different genetic loci. Right colon cancer is more common in patients with Lynch I syndrome (70 percent of colorectal cancer occurs in the spleen flexion area). They are also more likely to get synchronous and metachronic cancers (Lynch and Smyrk, 1996; Lynch *et al.*, 1997). Lynch II syndrome patients have colorectal cancer with a similar histology, but they also have colon cancer.

The majority (70%) of HNPCC patients had an hMSH2 or HMLH germline mutation (Liu *et al.*, 1996), increasing their lifetime risk of colorectal cancer by roughly 80% (Dunlop *et al.*, 1997; Vasen *et al.*, 1996). According to findings of a significant

cooperative analysis, 83% of hMSH2 mutations in HNPCC patients were either nonsense or frameshift mutations, in contrast to 49% of lineage hMLH1 mutants (Peltomaki and Vasen, 1997). Erroneous alteration mutations account for 31% of hMLH1 germline mutations. Between the two mismatch repair genes the mutations were uniformly distributed.

### 2.6.2 FAP- Familial Adenomatous Polyposis Syndrome

Due to mutations in the adenomatous polyposis coli (APC) gene on chromosome 5p22.2, FAP is an autosomal dominant syndrome characterized by hundreds to thousands of adenomatous polyps throughout the colon. Attenuated FAP (AFAP) is a kind of FAP with less than 100 colon polyps and a later onset of polyposis and malignancy than FAP. The APC gene produces a big protein (2,843 amino acids) that plays a variety of roles, including signaling, cell-to-cell adhesion, and microtubule assembly and stabilization. Congenital hypertrophy of the retinal pigment epithelium (CHRPE) (60% of families), upper gastrointestinal tract cancers (particularly peripapillary carcinoma), epidermal cysts, osteomas, and desmoids tumours (Goss and Groden, 2000; Fearnhead *et al.*, 2001) are just a few of the many variable traits connected to FAP. The APC gene has around 1,500 FAP mutations, the bulk of which are nonsense (28%) or frameshift (46%) or insertion (10%) mutations, resulting in a shortened and damaged protein. About 10-15% of mutations are caused by large deletions or duplications of the APC gene. Furthermore, new APC mutations are anticipated to occur at a rate of roughly 20%.

## 2.7 Human Mismatch Repair Genes Discovery

After using linkage analysis to explore large familial related populations on chromosome 2p16 and chromosome 3p21, the HNPCC susceptibility loci were found (Peltomaki *et al.*, 1993; Lindbloom *et al.*, 1993).

Microsatellite instability has been researched extensively in bacteria and yeast, and this was where human mutS gene homologue (hMSH2) on the 2p chromosome was identified (Fishel *et al.*, 1993; Leach *et al.*, 1993) and mutL (hMLH1 is a human MutL homologue) on chromosome 3p (Fishel *et al.*, 1993). Despite the discovery of two other mutL homologues (hPMS1 on chromosome 2q and hPMS2 on chromosome 7q) and a large number of mutations associated to a small number of HNPCC families (Eshleman and Markowitz, 1996), mutations in hMSH2 andhMLH1 account for the great majority of HNPCC cases observed (Eshleman and Markowitz, 1996).

Two more mutS homologues (hMSH3 and GTBP/hMSH6) have also been cloned. Mutations in GTBP have been discovered in the HNPCC strain recently (Akiyama *et al.*, 1997: Miyaki *et al.*, 1997), as well as previously reported somatic mutations in GTBP in colorectal cancer (Risinger *et al.*, 1996; Palombo *et al.*, 1995). Eukaryotic mismatch repair is controlled by these genes and the proteins they encode.

## 2.8 The Mut L Homolog 1 (MLH1) Mismatch Repair Gene

The protein MLH1 is involved in mismatch repair after DNA replication: one for MutS homologues (MSH2, MSH3, MSH6) and the other for PMS2, MLH3 or PMS1. MLH1 has an ATPase domain and two contact domains (Domingo and Schwartz, 2005). Despite the fact that it can link to PMS1 or MLH3, it has no known enzymatic function and forms a heterodimer with PMS2 known as MutLa. This heterodimeric complex detects DNA damage by adhering to the MutSa (MSH2 and MSH6) or MutSb (MSH2 and MSH3) heteroduplexes. The recruitment of proteins involved in excision and repair synthesis is dependent on the heterodimer of MLH1.

MLH1 is a homolog of the bacterial MutL gene, particularly in the N-terminal domain, and MLH1 homologs exist in eukaryotes (e.g., *Mus musculus, Drosophila melanogaster, Caenorhabditis elegans or Saccharomyces cerevisae*).

Exon 16 of the MLH1 gene is the most often mutated exon, with a base pair size of 165. Human MLH1 is located on chromosome 3p22 with nineteen (19) exons and 57.36 kb in length. According to the GenomeReference Consortium Human Build 38 (GRCh38 / hg38Dec 2013), the localization policy pair starts at 36,993,350 and finishes at 37,050,846 bps from pter. The length of the transcribed mRNA is 2524 bps. It has 756 amino acids and a molecular weight of 84.6 kDa.

In the gene that causes HNPCC, more than 300 germline MLH1 mutations have been found (Domingo and Schwartz, 2005). This gene has been discovered by several studies, and its presence varies geographically. Nucleotide substitutions (nonsense, missense, or splicing errors) or insertions/deletions (big or minuscule) that may be discovered by

automated sequencing are among the mutations that are not present in a single hotspot or gene zone (Momma *et al.*, 2019). The bulk of these mutations lead to a shorter protein.

Additionally, in some groups founder mutations account for a large proportion of HNPCC tumours (Domingo and Schwartz, 2005).Two Finnish variants, for example, eliminate exons 16 and 6 respectively. Exons and introns have both been found to have non-pathogenic germline genetic alterations.

## 2.9 Exon 16 of Mut L Homolog 1 (MLH1) Gene

Colorectal cancer (CRC), is caused by the inactivation or mutation of the mismatch repair genes in human cells (Kolodner and Marsischky, 1999; Modrich and Lahue, 1996). Although mutations in other MMR genes have also been discovered, the MLH1 gene has been associated with the bulk of HNPCC cases (Peltomak, 2001).

Despite the fact that there are over 300 MLH1 germline mutations known to induce HNPCC in the MLH1 gene, none of them are found in any one hot region or gene zone. Published in many articles including Momma *et al.* 2019, exons 16 and 18 of the MLH1 gene are listed as genetic mutation hotspots by Dominguez Valentin *et al.*, (2013). According to Dominguez Valentin *et al.*, 2013, exons 16 and 18 of the MLH1 gene are hotspots for mutations, which matches evidence from other populations (Plazzer *et al.*, 2013; Peltomaki *et al.*, 1997; Nilbert *et al.*, 2009). In other populations, exons 16 and 18 of the MLH1 gene have been discovered as a genetic hotspot, accounting for 26% of all MLH1 mutations documented here (Li *et al.*, 2011).

In addition, a study published in Sweden in 2016, Lagerstedt-Robinson *et al.,* discovered the Finnish founder mutation MLH1 c.1732-? 1896+?del, which had the whole deletion of MLH1 exon 16, in 6% of the Lynch syndrome families that were selected for the studies.

A heterozygous genotype for a mutation was revealed in a participant, in a study done among the Mexican Lynch group in 2016. Two distinct sequence variants in exon 16 of MLH1 were discovered in patient LS-41. The c.1852 1854delAAG (p.K618del) allele had an in-frame codon deletion, while the c.1852 1853delinsGC allele had a two-nucleotide deletion/insertion (p.K618A) allele (Moreno-Ortiz *et al.*, 2016).

Finally, Momma and colleagues in Japan analyzed the chromosomal rearrangement of the MLH1 and MSH2 genes in a Lynch syndrome family that had been tracked for over 45 years in 2016. In this family, they discovered widespread deletions (c. (306+ 1 307–1) (*193?) del of MLH1 exons 4-19, which included exon 16. (Momma *et al.*, 2019).

## 2.10 Other Mismatch Repair Genes

### 2.10.1 Gene for Mut S Homolog 2 (MSH2)

The MSH2 gene has sixteen (16) exons and measures 80.10 kb in length. It is found on the chromosome 2p16. In HNPCC, MSH2 and MLH1 jointly account for more than 64% of germline mutations (Liu *et al.,* 1996; Wijnen *et al.,* 1997).

### 2.10.2 Gene for Mut L Homolog 6 (MLH6)

The MLH6 gene is found on chromosome 2p15 and has ten exons (Akiyama *et al.,* 1997). Germline mutations in these genes induce atypical lynch syndrome, which has recently

been recognized as aprevalent cause. The clinical presentation of affected families was different from that of those with classic Lynch syndrome caused by MLH1 and MSH2 mutations, according toearly MLH6 mutation investigations. Endometrial cancer is the most common clinical manifestation in women who have the MLH6 mutation, but it appears to have a lower penetration. In MLH6 mutations which mostly occur in mononucleotide sequences, microsatellite instability is similarly infrequent. (Akiyama *et al.*, 1997; Plaschke *et al.*, 2000; Kolodner *et al.*, 1999; Wu *et al.*, 1999; Wijnen *et al.*, 1999)

### 2.10.3 The PMSI and PMS2 Genes

The PMSI and PMS2 genes are related. The PMS1 and PMS2 genes are 16 kb and 15 exons long each, and found on chromosomes 2q31q33 and 7p22, respectively. Both constituted 5% of all Lynch syndrome patients (Nakagawa *et al.*, 2004). Despite the importance of the PMS2 gene in the repair mechanism, Lynch syndrome or Turkot syndrome mutations are rare (Jass, 2006). The mechanism through which PMS2 and PMS1 may play a role in cancer risk is unknown (Nakagawa *et al.*, 2004). Despite the fact that the complex has been discovered, its function is yet to be determined in DNA repair or in the development of colorectal cancer.

## 2.11 Types of Mutations Spotted in Colorectal Cancer

Single base changes, such as adenine changing to thymine; eg. in sickle cell disease which caused sickle cell anaemia where (Glu → Val) are called point mutations (figure 2.4). A point mutation is the most common type of mutation and there are two types of

point mutations.

When one purine is replaced by another purine, or when one pyrimidine is replaced by another pyrimidine, a transition occurs (figure 2.4).

When purine replaces pyrimidine or pyrimidine replaces purine, transversion occurs.

http://www2.csudh.edu/nsturm/CHEMXL153/DNAMutationRepair.

## 2.12 Effects of Point Mutations in DNA Sequence

Point mutations in DNA sequences that code for proteins might be silent, irrelevant (unimportant) or relevant.

**2.12.1 Silent**: A synonymous codon is likely to be formed when a rule change occurs at the third codon location. As such, the gene sequence of the amino acid doesn't change, and the mutation goes undetected (figure 2.4).

**2.12.2 Missense:** When a base is substituted in a codon, a different amino acid is identified and as a result a different polypeptide sequence is produced. The missense mutation can be conservative or non-conservative depending on the type of amino acid change. Conservative mutations are those in which the structure and properties of a substituted amino acid are almost identical to those of the original amino acid, and are unlikely to impact the structure or function of subsequent proteins. The mutation is non-conservative if it results in an amino acid with a dramatically altered structure and properties, and the resulting structure / function is likely to be bad (i.e., sickle cell mutation).

**2.12.3 Nonsense:** Translation is sped up when a nucleotide change results in a stop codon and the protein becomes non-functional (figure 2.4).

## 2.13 Deletions

A frame shift deletion occurs when one or more base pairs in the DNA are removed. When one or two bases are dropped, the translation frame shifts, leading to a muddled message and a non-operational product. When three or more sockets are taken out, the reading frame is still intact. One or more amino acids are lost when one or more codons are deleted from a protein. It may or may not be dangerous. http://www2.csudh.edu/nsturm/CHEMXL153/DNAMutationRepair ; 06/05/2023

## 2.14 Insertions

Depending on whether multiples of three base pairs are introduced, adding more base pairs can cause the frame to change. When one or more base pairs are added to a DNA sequence, an insertion mutation happens. The size of an insertion mutation can range from a single base pair inserted into a DNA sequence to the insertion of an entire chromosome. There is a chance that the protein will be flawed as a result (figure 2.4). http://www2.csudh.edu/nsturm/CHEMXL153/DNAMutationRepair.

## 2.15 Frameshift Mutation

An addition or deletion of many nucleotides in a DNA sequence that is not divisible by three results in a frameshift mutation, also known as a framing error or reading frame shift. The insertion or deletion might alter the reading frame (the grouping of the codons) because of the triplet nature of gene expression via codons, which leads to a completely different translation from the original. The degree of protein alteration increases with the position of the deletion or insertion in the sequence (Losick *et al.*, 2008). Generally speaking, a frameshift mutation causes the codons read after the mutation to code for

alternative amino acids. The first stop codon (UAA, UGA, or UAG) that is met in the sequence will likewise be altered by the frameshift mutation. It's possible that the newly produced polypeptide will be excessively long or short, and it won't likely operate properly (figure 2.4), http://www2.csudh.edu/nsturm/CHEMXL153/DNAMutationRepair.

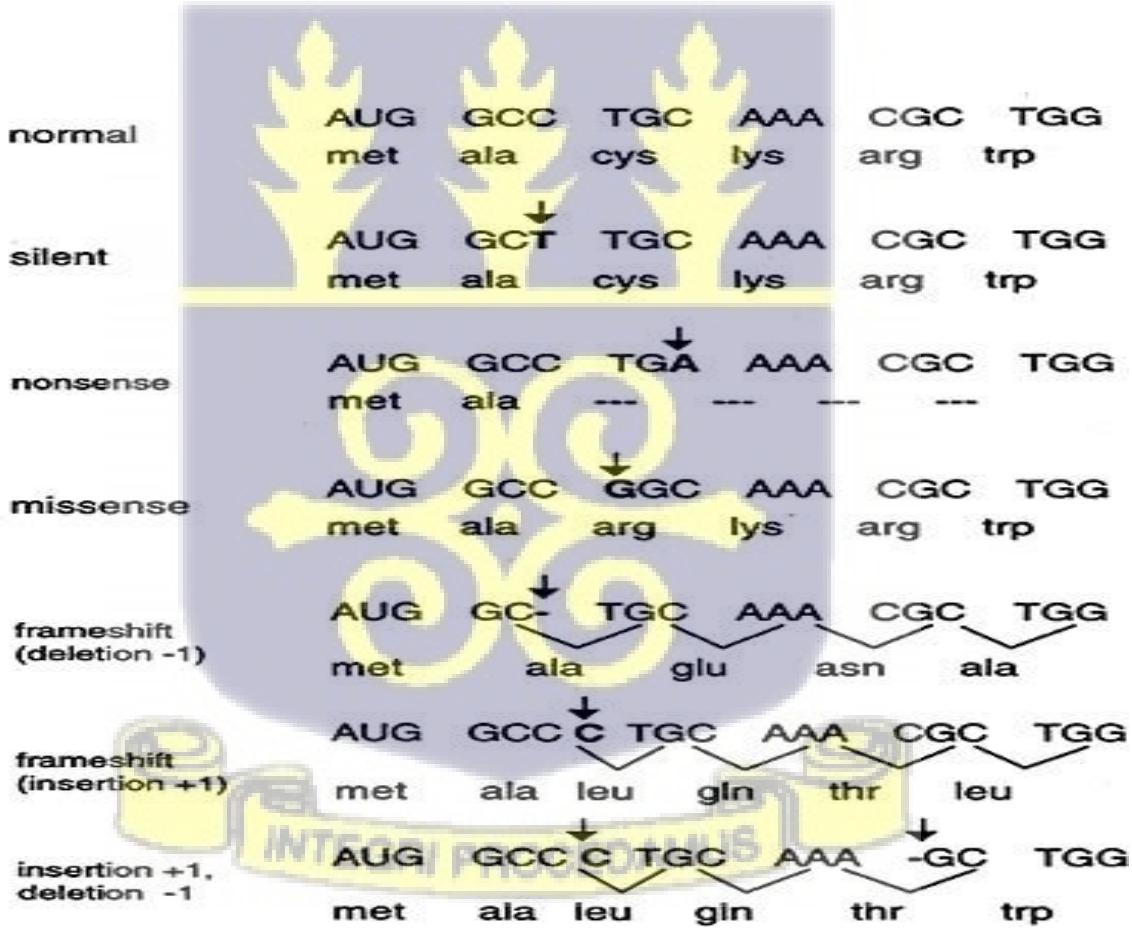



**Figure 2.4: Demonstration of the Different Possible Mutations in CRC**

http://www2.csudh.edu/nsturm/CHEMXL153/DNAMutationRepair; 24/06/2021

## 2.16 Functions and Mechanisms of Mismatch Repair Genes

### 2.16.1 Functions

The DNA mismatch repair machinery main job is to fix base deletions and insertion mismatches/ mistakes caused by DNA polymerase during DNA synthesis. Mismatch repair genes are involved in the early stages of apoptosis in reaction to many types of DNA damage, as well as mending errors in DNA synthesis, ensuring the accuracy of genetic recombination, and participating in the early stages of apoptosis in response to various types of DNA damage (https://doi.org/10.3389/fmolb.2020.00122 ; 05/10/2023)

### 2.16.2 Mechanisms

MSH2 / GTBP (MutSα) and MSH2 / MSH3 (MutSβ), two heterodimeric complexes of MutS-related proteins detect aberrant bases in eukaryotic DNA first (Fig. 2.5) (Greene and Jinks Robertson, 1997; Kolodner, 1996; Marsischky *et al.,* 1997). MutSβ is engaged in the repair of larger mismatched insertion /deletion pairs despite the fact that MutSα is assumed to be the cause of mismatched base pairs (Kolodner, 1996; Sia *et al.,* 1997).

MutSα and MutSβ are unlikely to be able to duplicate the identical base's aberrant insertion/deletion pairs. Mismatched base pair repair as well as mismatched insertion / deletion pairs with up to 12 mismatched bases have been demonstrated using MSH2 / GTBP complexes (Genschel *et al.,* 1998).MSH2 / GTBP excels at recognizing and combining an invalid G: T combination as wellas an invalid insertion / deletion +1 pair, whereas MSH2 / MSH3 excels at detecting and combining an improper insertion / deletion +1 and more (Palombo *et al.,* 1995; Habraken *et al.,* 1996; Drummond *et al.,* 1995; Alani, 1996).

Two other mutS homologues (hMSH3 and GTBP / hMSH6) have been cloned (Risinger *et al.*, 1996; Palombo *et al.*, 1995), and GTBP mutations have recently been described in related HNPCCs (Akiyama *et al.*, 1997: Miyaki *et al.*, 1997), with the previously described somatic mutation in GTBP in the colon cancer cell line (Papadopoulos *et al.*, 1995). In eukaryotes, these genes and the proteins they encode are responsible for rectifying the mismatch.

The heterodimeric complex of MutL-associated proteins (MLH1 / PMS1 (PMS2 in humans)) interacts with MutS-associated proteins that are already mismatched when a mismatch is discovered (Prolla *et al.*, 1994; Li and Modrich *et al.*, 1995).

The MSH2 / MSH3 mismatch complex is bound by MLH1 / PMS1, which transforms it into a higher molecular weight structure (Habraken *et al.*, 1997). In the MSH2 / MSH3 pathway, MLH1 forms a complex with MLH3 (PMS2 in humans) which plays a periodic role in the repair of mismatched insertion/deletion pairs (Flores Rozas and Kolodner, 1998). The MLH1/PMS1 complex improves MutS-associated proteins' ability to detect mismatches (Drotschmann *et al.*, 1998; Habraken *et al.*, 1997.

DNA polymerase, replication protein A (RPA), proliferating cell nuclear antigen (PCNA), replication factor C (RFC), and exonuclease, to mention a few, are mismatch repair proteins. Exonuclease-mediated degradation of DNA from a "break" located 1-2 kb from the mismatch occurs when the heterodimeric complexes of MutS and MutL are

linked to the mismatch DNA (Sarkar, 1999). Until the misplaced stem is removed, the deterioration will continue. DNA polymerase δ fills the resulting lengthy excision path by inserting the proper nucleotide into the sequence.

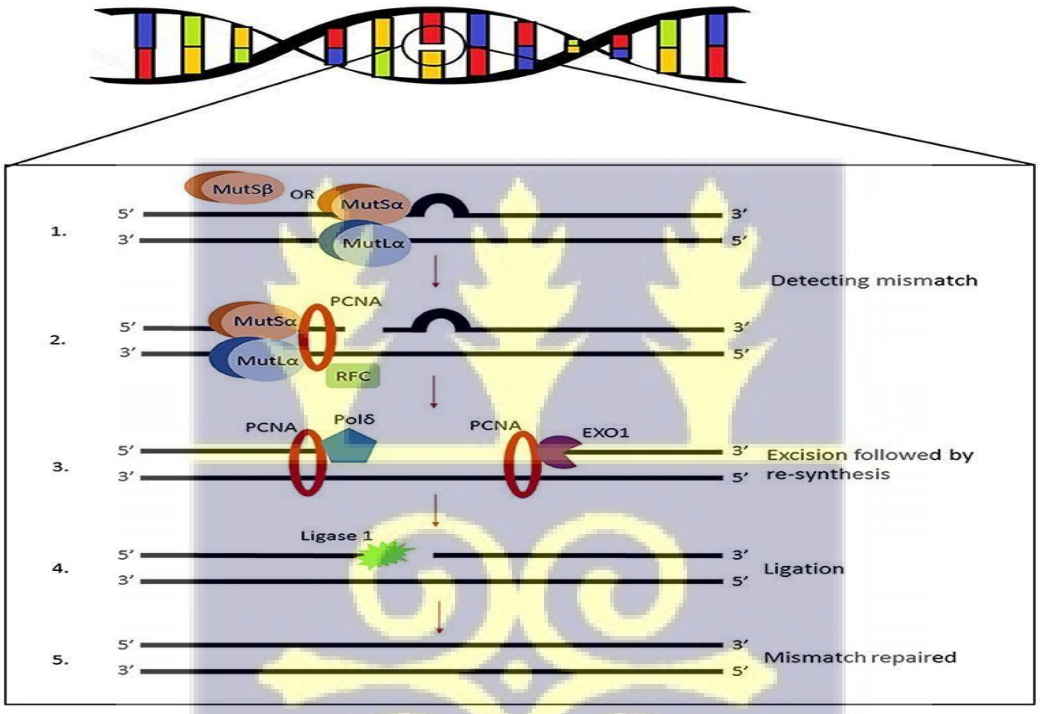



**Figure 2.5: Eukaryotic MMR System**; Source: doi: 10.3389/fmolb.2020.00122 24/09/2022

MutS (MSH2/MSH6 heterodimer, which is dominant in human cells) or MutS (MSH2/MSH6 heterodimer): the heterodimer initiates DNA repair by recognizing and binding mismatches. MutL (MLH1/PMS2 heterodimer), as well as proliferating cell nuclear antigen (PCNA) and replication factor C, are all part of the complex (RFC). The team activates the PMS2 endonuclease, which generates single strand breaks around the mismatch and opens the exonuclease 1 (EXO1) entry sites, allowing the DNA damage to be dissociated.

## 2.17 Microsatellite Instability

Microsatellites are recurrent genomic loci (1-5 bp, repeated 15-30 times), whereas microsatellite instability (also known as positive replication error, RER +) is characterized by frequent modifications of any length. These loci slide during DNA replication as duplicate entities are added or removed, resulting in a tiny loop in the template or nascent DNA strand. They are usually repaired, but in the absence of a suitable mismatch repair function, these "loops" may become permanent, resulting in the generation of alleles of varying length in the subsequent replication cycle.

Hundreds of thousands of microsatellites can be identified across the human genome, all of which are susceptible to insertion and deletion changes. Because mutations in gene introns are widespread even in normal cells, they frequently result in polymorphic changes (Aaltonen *et al.*, 1993; Thibodeau and colleagues, 1993).

# CHAPTER THREE

## 3.0 Methodology

### 3.1 Study Design

A case-control design was used in this investigation, which employed quantitative method to describe the possible mutations that exist among Ghanaian colorectal cancer patients managed at the Korle Bu Teaching Hospital. Patients with colorectal cancer treated at the Korle bu facility were included in the study; and some Ghanaians who had undergone colonoscopy and proven not to have the colorectal cancer, also served as controls. Questionnaires were used to capture patients' information on demographics, nutrition, life-style, date of diagnosis and family history. Blood samples were collected from the patients and DNA extracted from buffy coat. The DNAs extracted were amplified, sequenced and analysed bioinformatically for any available nucleotide variants. Before recruitment, study participants consent were obtained.

### 3.2 Study Population and Sampling Technique

The case population was recruited by enrolling consecutive Ghanaian patients who reported with CRC from September 2014- October 2019, to the Department of Surgery at the KBTH. The controls subjects were also obtained by enrolling patients with no CRC as per colonoscopy at the endoscopy unit of the KBTH during the same period, i.e. September 2014- October 2019.

### 3.3 Study Site

Both case and control individuals were recruited at the Department of Surgery, Korle-Bu Teaching Hospital. Korle-Bu Teaching Hospital is Ghana's most prestigious tertiary hospital. It is country's major public university hospital in the south and connected to the University of Ghana Medical School. Korle Bu Teaching Hospital (KBTH) has grown from 200 beds to 2,000 beds since it first opened its doors in October 9, 1923 https://kbth.gov.gh/departments-centres/department-of-surgery/.

The Department of Surgery also saw the introduction of units, including the colorectal unit, which specializes in the care of a large number of patients with colon and rectal cancer. The department of surgery also houses the doctors who attend to all cancer patients requiring chemotherapy and radiotherapy, including patients with colon and rectal cancers. KBTH is now Africa's third largest hospital and Ghana's most important national referral centre https://kbth.gov.gh/departments-centres/department-of-surgery.

### 3.4 Sample Size

The minimum sample size for each study group (case and control) required in each arm of the study to achieve a power of 99% in two- sided, one to one matching design was 71. This was determined using the formula for quantitative case-control study:

Sample size (N) =$[(r + 1)r * [SD^2 (Z\beta + Z\alpha/2)^2 / d^2]$ (Charan & Biswas, 2013)

where N= sample size; SD = Standard deviation = 2.76 (Puchner *et al.*, 2011); r = Ratio of control to cases, 1 for equal number of case and control ; $Z\beta$ = Standard normal variate for power = 90% = 1.28 ; $Z\alpha/2$ = Standard normal variate for level of significance =95%

= 1.96; d = Expected mean difference between cases and controls = 1.5 (Puchner *et al.*, 2011)

$N = \underline{(1+1)/1}\ \underline{[(2.76)^2(1.28+1.96)^2]\ /\ (1.5)\ =}\ 71.08 \approx 71$

### 3.5 Criteria for Cases

i. Ghanaians with CRC who received medical care at the KBTH and consented.

ii. Ghanaians who have the CRC disease after having done colonoscopy and consented.

### 3.6 Criteria for Control

i. Ghanaians who had undergone colonoscopy to determine their CRC status but were negative.

### 3.7 Data Collection and Sample Processing:

Patients provided demographic information such as age, sex, nationality, educational level, residence, and favourite food via a standardized questionnaire (appendix II), with a unique codeallocated to each participant. About 5ml of whole blood was collected from both case and control participants into an EDTA tube using syringe and allowed to settle, then centrifuged at 3500 rpm for 5 minutes and afterwards separated into various well-

labelled Eppendorf tubes. These were stored in -20$^{o}$C freezer in the Medical Biochemistry Department while awaiting other laboratory procedures such as DNA extraction and quantification, PCR and DNA sequencing to be carried out as recruitment went on.

### 3.8 Laboratory Procedures and Analyses

#### 3.8.1 Extraction of DNA

The DNA QiaAmp (QIAGEN GmbH, D-40724 Hilden, Germany) kit was used to extract DNA from white blood cells (buffy coats), following the blood and body fluid technique in the manufacturer's instructions that came with the kit.

Twenty microliters (20µl) of Qiagen Protease were pipetted into a 1.5 millilitre (mL) centrifugation tubes containing 200 microliters (µl) of buffy coats after separation of whole blood into individual components as plasma, buffy coat and blood cells was done. After that, 200 µl of buffer (Al) was added and vortex pulsed for 15 seconds (S). The solutionwas incubated in a water bath at 56 degrees Celsius (°C) for ten minutes and a quick centrifugation done to remove drips from the inside of the cover.

200 µl of 100% ethanol was added to the tube contents and mixed for 15 seconds with pulse votexing. The mixture was then carefully transferred to a QiaAmp Mini-column (in a 2 ml manifold tube). The tube content was centrifuged for 1 minute at 8000 rpm. The QiaAmp tube was then cleaned and the filtrate was discarded.

500µl of buffer (AW1) was added to the QiaAmp tube without soaking the rim. After sealing the lid, the tube contents was centrifuged at 8000 rpm for 1 minute. After cleaning

the QiaAmp tube and discarding the filtrate, the contents were centrifuged at 8000 rpm for 1 minute after closing the lid. The previous tube containing the filtrate was discarded and the mini QiaAmp column was transferred to a clean collecting tube.

Without wetting the rim, another 500µl of buffer (AW2) was added and centrifuged for 3 minutes at 14,000 rpm. The centrifugation column was then transferred to a fresh 2 mL collection tube, and the previous collection tube discarded with the filtrate. The collection tube was centrifuged at 14000rpm for 1 minute with the spin column.

The QiaAmp small column was placed in a clean 1.5 mL micro centrifugation tube and 200µL of elusion buffer (AE) was added after throwing away the filter tube. To allow DNA to be eluted from the QiaAmp Mini column into the micro centrifugation tube, the tube content was incubated at room temperature for 1 minute and then centrifuged at 8000 rpm for 1 minute. The QiaAmp Mini column was discarded and the DNA extracted.

### 3.8.2 DNA Quantification

NanoDrop one (NanoDrop One UV-Vis Spectrophotometer from Thermofisher Scientific, USA) was used to check the concentration and purity of each extracted DNA samples (71cases, 68 controls). To quantify each DNA sample, the NanoDrop machine was first blanked with small distilled water and a 0.18 µL of the DNA sample was pipetted and dropped carefully onto the measurement pedestal in a raised position. The pedestal was then carefully lowered to initiate the quantification process, with concentration and purification values been displayed on the screen afterwards. Records of values was kept on a pen drive and also in note book for future reference.

### 3.8.3 Polymerase Chain Reaction (PCR)

Conventional PCR was performed on all the extracted DNA samples using the primer pair: forward primer -5'-TCTTGCTTCTTCCTAGGAGCC-3'; reverse primer 5'-AATGGCTGTCACACCTCATCA-3' to amplify target gene exon 16 of MLH1. PCR for each sample was run in triplicate. For the third amplification, a fresh DNA samples were used. The reaction was carried out on a Proflex PCR System (Thermofisher Scientific, Singapore) using the PCR cycling and thermal conditions in table (1). After that, the amplicons were stored at -40°C for further analysis.

To perform PCR, the final reaction volume of 10 µL constituted of 2µL of gDNA extract, 5 µL of water, 2.5 µL of Qiagen master mix and 0.25 µL (10 µM) each of the forward and reverse MLH1 exon 16 primers. The amplicons were held at 15°C and later stored at -4°C for further analysis (table 1).

**Table 3.1: PCR Cycling and Thermal Conditions used in Amplification of Study Samples. MLH1-Exon16 Assay – PCR for 10 µl Rxn Vol Calculations for 1 Plate (96 Wells).**

| **Reagents** | **Concentration** | **1Rxn Vol (µl)** | **10x Reaction** | **108 Rxns** | **Final Conc.** |
|---|---|---|---|---|---|
| MLH1-Ex16-F1 | 10µM | 0.25 | 2.5 | 27 | 0.25µM |
| MLH1-Ex16-R1 | 10µM | 0.25 | 2.5 | 27 | 0.25µM |
| Qiagen Master Mix (x2) | x2 | 2.5 | 25 | 270 | 0.5x |
| Milliq Water | | 5 | 50 | 540 | |
| gDNA template/water | 20ng/µl | 2 | | | 2ng |
| Final Vol | | 10 | 80 | 864 | |
| | | | Each tube =8 µl | | |
| **MLH1-Exon 16 Assay PCR Conditions** | | | | | |
| **PCR Step** | **Description** | **Temp (C)** | | **Time** | |
| 1 | Initial Denaturation | 95 | 95 | 5 mins | 4.5 mins |
| 2 | Denaturation | 95 | 96 | 5 s | 30 s |
| 3 | Annealing | 63 | 48 50 52 54 56 58 60 63 | 6 s | 1 min 14 s |
| 4 | Extension | 68 | 69 | 3 min 15 s | 2 min 10 s |
| 5 | Repeat | Go to Step 2, x34-39 | Go to Step 2, x34-39 | | |
| 6 | Final Extension | 72 | 72 | 1 min | 1 min |
| 7 | Hold/End | 15 | 15 | Hold | Hold |
| Pippette 5µl to each well containing PCR product | | | | | |
| | | | | | |
| **MLH1-Exon16 assay – Gel Electrophoresis** | | | | | |
| Prepare 3% agarose gel | | | | | |
| Add 2 µl loading dye to digested PCR product (to make 12µl final volume now)<br>Load 12 µl for digested product to 3 % agarose gel<br>Run for 2 hours at 100 V | | | | | |

### 3.8.4 Agarose Gel Electrophoresis

Using 100ml of IX Tris-acetate EDTA (TAE) buffer, a 3% (3g) agarose gel stained with gel red DNA staining dye (4.5ul ethidium bromide) was prepared. In a 2ul of 6X Blue / orange loading dye, 10ul of PCR amplified products were added. To aid in the confirmation of the expected band size of 195 base pairs, 3ul of a one hundred base pair (100 bp) DNA ladder (GeneRuler, Thermofisher Scientific, USA) was utilized. For 1 hours, the gel was run at a voltage of 100 volts (V) and a current of 500 milliamps (mA). A Gel Imager (Luminescent Image Analyser: GE healthcare biosciences, Japan) was utilized to see whether the amplified products were present or absent on the gel.

### 3.8.5 Post Quantification of PCR Products

The PCR products (amplicons) were afterwards quantified to know the concentration and purity of the individual amplicons and finally to know whether to dilute the samples' content or not,per the sequencing pipeline that will be used.

### 3.8.6 DNA Sequencing

The PCR amplicons were submitted to a company (Apical Scientific Laboratories) based in Malaysia for sequencing. Sanger Sequencing was used to determine order of nucleotide in the case samples and the control samples. The primer pair used were: forward primer -5'-TCTTGCTTCTTCCTAGGAGCC-3'; reverse primer 5'-AATGGCTGTCACACCTCATCA-3'. The sequencing done was bidirectional.

### 3.9 Statistical Analysis

The data from the study was entered into MS Excel Sheet and then exportedto SPSS version 20.0 for analysis. To summarize data on demographic and clinical characteristics ofresearch participants, descriptive statistics (percentages, means, standard deviations, etc.) were utilized. The significance of the predictor variables was assessed by determining the *p* values and confidence intervals; *p* values below 0.05 were considered significant.

### 3.10 Bioinformatics Analysis

The sequenced products were analyzed using bioinformatics tool for detection of any possible sequence variance mutation. BLAST analysis was done for the sequences. The analysis was performed on Geneious Prime Trial Version 2022 (Geneious by Dotmatics, New Zealand). The analysis was done by simply trimming of low quality regions and alignment of cases and controls sequences to the MLH1 exon 16 reference sequence, so as to detect any possible single nucleotide polymorphisms (SNPS) present. The reference sequence was downloaded from Gene bank.

### 3.11 Ethics and Confidentiality

Ethical approval was obtained from the College of Health Sciences Ethics and Protocol Review Committee, University of Ghana. Furthermore, all potential study participants consented before they were enrolled in the study. Patient information obtained through

the questionnaire was anonymized and entered into a Microsoft Excel sheet, securely saved and password protected.

# CHAPTER FOUR

## 4.0 Results

### 4.1 Demographic Characteristics of Study Population

A total of 71 CRC (cases) and 68 non-CRC (controls) blood samples were collected from the participants from 2014- 2019.

In all, a total of 139 participants were recruited; 71/139 (51.08%) males and 68/139 (48.90%) females, giving a male: female participant ratio of 1:1. Most of the participants were within the age ranged 61-70 years, with a median age range of 56.5 years for cases and 47 years for controls (Table 4.1).

Also, majority of the participants recruited in the study were Christians by religion 126/139 (90.65%). Most were of Akan ethnic group 77/139 (55.40%), a large percentage had formal education 105/137 (75.5%), and 80/137 (57.6%) were formally employed. A majority of the participants lived in Accra city 99/139 (71.20%), and most of the participants ate mostly home prepared meals 91/137 (65.47%).

### 4.2 Evaluation of CRC Risk Factors of the Participants Recruited

From the assessment, majority 138/139 (99.3%) and 116/139 (83.5%) of the study participants recruited did not smoke and did not drink alcohol, respectively. 6/71 (8.5%) of the case population had a family history of cancer, the remaining either 'did not know or 'did not have' a family history of cancer (figure 4.2). From those with a family history of malignancy, CRC (3/6) topped the list, followed by breast and stomach cancer (2/6); and lastly pancreatic cancer (1/6).

**Table 4. 1: Distribution of Sex and Age of the CRC Participants Recruited**

| Variables | Cases(71) n (%) | Controls(68) n (%) | P value |
|---|---|---|---|
| **Sex** | | | |
| Female | 34(47.9) | 34(50.0) | 0.805 |
| Male | 37(52.1) | 34(50.0) | |
| **Age range (years)** | | | |
| 21-30 | 2(2.82) | 5(7.4) | |
| 31-40 | 5(7.04) | 11(16.18) | |
| 41-50 | 8(11.27) | 11(16.18) | |
| 51-60 | 15(21.13) | 12(17.65) | |
| 61-70 | 27(38.03) | 23(33.82) | |
| 71-80 | 8(11.27) | 6(8.82) | |
| 81-90 | 6(8.45) | 0(0.0) | |
| **Median age (range)** | 56.8 years | 47 years | |

**Table 4.2: Distribution of CRC Risk Factors of the Participants Recruited**

| | Cases | Controls | $X^2$ |
|---|---|---|---|
| | n (%) | n (%) | |
| **Smoking status** | | | |
| No | 70(98.6) | 68(100.0) | 0.965 |
| Yes | 1(1.4) | 0(0.0) | |
| **Alcohol status** | | | |
| Yes | 2(2.8) | 0(0.0) | 2.503 |
| No | 60(84.5) | 56(82.4) | |
| Occasionally | 9(12.7) | 12(17.6) | |
| **Family history** | | | |
| Don't know | 30(42.3) | 0(0.0) | 43.566** |
| None | 35(49.3) | 67(98.5) | |
| Yes | 6(8.5) | 1(1.5) | |

*Chi square ($X^2$) **statistically significant at $p<0.001$*

### 4.3 DNA Quantification Result

The average DNA concentrations calculated were 15.21 ± (13.81) ng/µl and 23.21± (10.14) ng/µl for case and control samples, respectively. The DNA concentrations determined in the samples ranged between 4-30 ng/µl for the cases and 16-50 ng/µl for the controls. Majority (75%) of the DNA extracts met the ($A_{260}/A_{280}$) DNA purification criteria.

### 4.4 Polymerase Chain Reaction Result

Out of the 71 case samples, 28(39.4%) were detected by the ethidium bromide to be successfully amplified to the expected band size of 195 bps. All the 68 (100%) control samples were amplified to the band size of 195 bps after the gel electrophoresis run. (Figures 4.1- 4.3 show sample gel images).

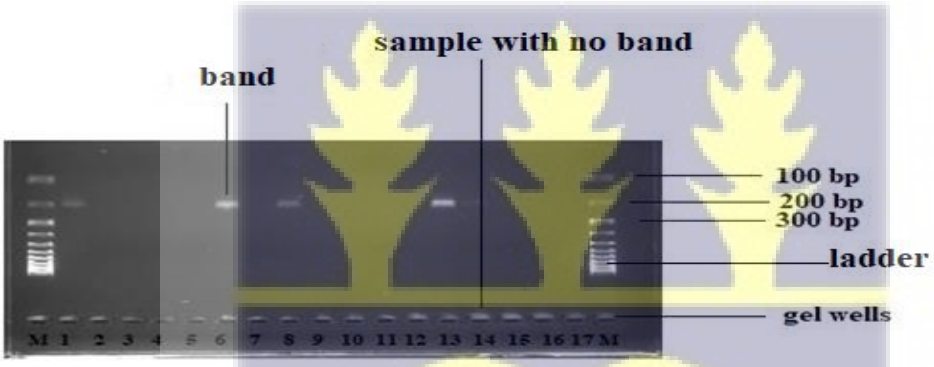



**Figure 4.1: Case samples (1-17) gel image showing PCR amplification result for exon 16 of the MLH1 gene.** The Restriction Fragment Length Polymorphism (RFLP) 100 bp molecular weight (M) ladder was used for referencing.

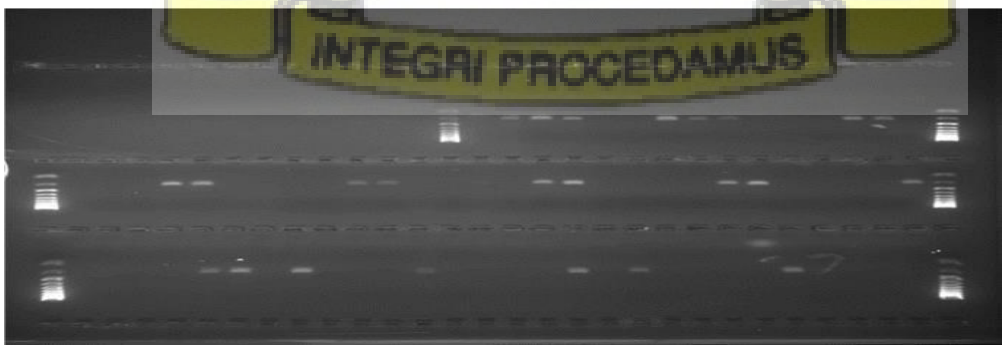



**Figure 4.2: Case samples (18-71) gel image showing PCR amplification result for exon 16 of the MLH1 gene**. Bands and no bands for some of the case samples. The RFLP 100bp molecular weight ladder was used.

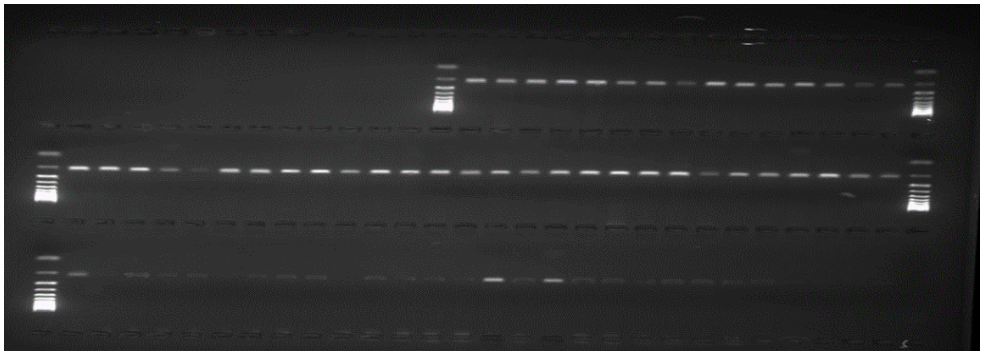


**Figure 4.3: Control samples (1-68) gel image showing PCR amplification result for exon 16 of the MLH1 gene**. Gel image showed bands for all control samples, though first five were positive but looked faint. Faint bands were more visible on the gel imager. The RFLP 100bp molecular weight ladder was used.

### 4.5 Nucleotide Sequencing and Bioinformatics Analysis Results

Ninety-six samples (71 cases, 25 controls) were submitted for sequencing.

For the cases: 14(19.7%) out of the 71 case samples were without nucleotide sequence (figure 4.10). Thirty 30(42.3%) of the case samples had clear good nucleotide sequences. Two cases (C041 and C054) of the clear good sequences analyzed using bioinformatics had a single guanine (g) nucleotide insertion at codon 596 of the coding sequence of the exon 16 MLH1protein (figure 4.8) , and as a result, changed the next amino acid after that position, from Tryptophan (W) to Leucine (L) (figure 4.7). This introduced an earlier stop codon that truncated the protein at codon 608 (figure 4.7). The mutant protein was therefore 148 amino acids shorter than the normal MLH1 protein sequence which was 756 amino acid long.

Twenty-five 25(35.2%) case samples had sequences that were longer than the expected length with poor chromatograms, with ~90% of the sequence having a high probability

for wrong call (0.05%) and this left those sequences with less than 50 bases .

For the control samples; all 25(100%) samples submitted for sequencing had a fairly good nucleotide sequence, but with no variants in good quality control (QC) regions when the sequences were compared with the MLH1 exon 16 reference sequence (figure 4.9).

**Table 4.3: Frequencies of Mutations for Participants.**

| Type of mutation | Frequencies | |
|---|---|---|
| | Case n (%) | Control n (%) |
| Whole exon deletion | 14(19.7) | None seen |
| Single nucleotide (guanine) insertion frameshift | 2(2.8) | None seen |
| | | |

## 4.6 Graphical Representation of the Guanine Insertion Frameshift Mutation found in Exon 16 of the MLH1 Protein

**Sequenced Exon**

**TCTTGCT TCTTCCTAGG AGCC**AGCACC GCTCTTTGAC CTTGCCATGC TTGCCTTAGA TAGTCCAGAG AGTGGCTGGA CAGAGGAAGATGGTCCCAAA GAAGGACTTG CTGAATACAT TGTTGAGTTT CTGAAGAAGA AGGCTGAGAT GCTTGCAGAC TATTTCTCTT TGGAAAT**TGA TGAGGTGTGA CAGCCATT**

**(Figure 4.4): Part of the sequenced exon 16 with flanking sequences.**

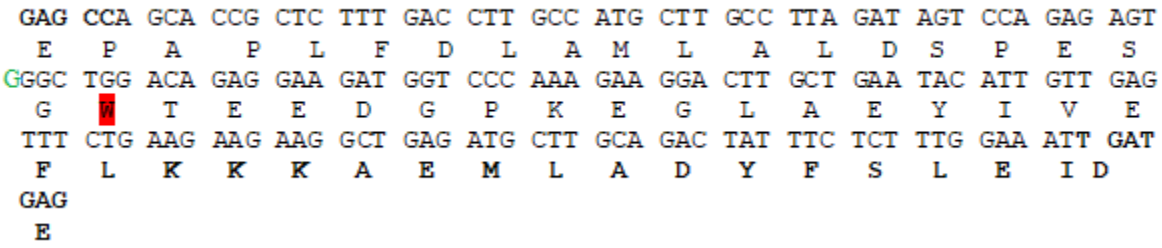
**NB**: The first four nucleotides at the beginning of the exon is highlighted in yellow and the last 4 nucleotides of the exon is shaded in green. The primer regionsare underlined and in bold.

**MLH1 Exon 16 Reference Sequence (Accession Np_000240 Version Np_000240.1)**

gag cca gca ccg ctc ttt gac ctt gcc atg ctt gcc tta gat agt cca gag agt ggc tgg
E P A P L F D L A M L A L D S P E S G W
aca gag gaa gat ggt ccc aaa gaa gga ctt gct gaa tac att gtt gag ttt ctg aag aag
T E E D G P K E G L A E Y I V E F L K K
aag gct gag atg ctt gca gac tat ttc tct ttg gaa att gat gag
K A E M L A D Y F S L E I D E

**( Figure 4.5) MLH1 Exon 16 Reference Sequence**

NB: The exact nucleotide number of exon 16 is 165 and with 55 amino acids.




**(Figure 4.6): Insertion of the mutated variant (guanine nucleotide) into exon 16**

Written in green is the guanine nucleotide at codon 596 of exon 16 of MLH1 protein.

Highlighted in red is the original amino acid before the point of mutation.

**gag cca gca ccg ctc ttt gac ctt gcc atg ctt gcc tta gat agt cca gag agt ggg ctg**
**E P A P L F D L A M L A L D S P E S G L**
**gac aga gga aga tgg tcc caa aga agg act tgc tga ata cat tgt tga gtt tct gaa gaa**
**D R G R W S Q R R T C - I H C - V S E E**
**gaa ggc tga gat gct tgc aga cta ttt ctc ttt gga aat tga tga**
**E G - D A C R L F L F G N - -**

**(Figure 4.7) Nucleotide insertion resulted in a frameshift mutation and truncated protein.**

As there was a shift in the reading frame this resulted in the different amino acid productions. Example; tgg (W) → ctg (L) highlighted in yellow.

The single (G) nucleotide insertion led to a shift in the ribosome reading frame where (i.e. Tryptophan changed to Leucine) highlighted in yellow. The coding sequence therefore shifted from Codon 596 onwards, producing not only different amino acids but more importantly leading to an earlier stop codon that truncated the protein at codon 608. The mutant protein is therefore 148 amino acids shorter (figure 4.7) than the normal MLH1 protein which is 756 amino acids (aa) long.

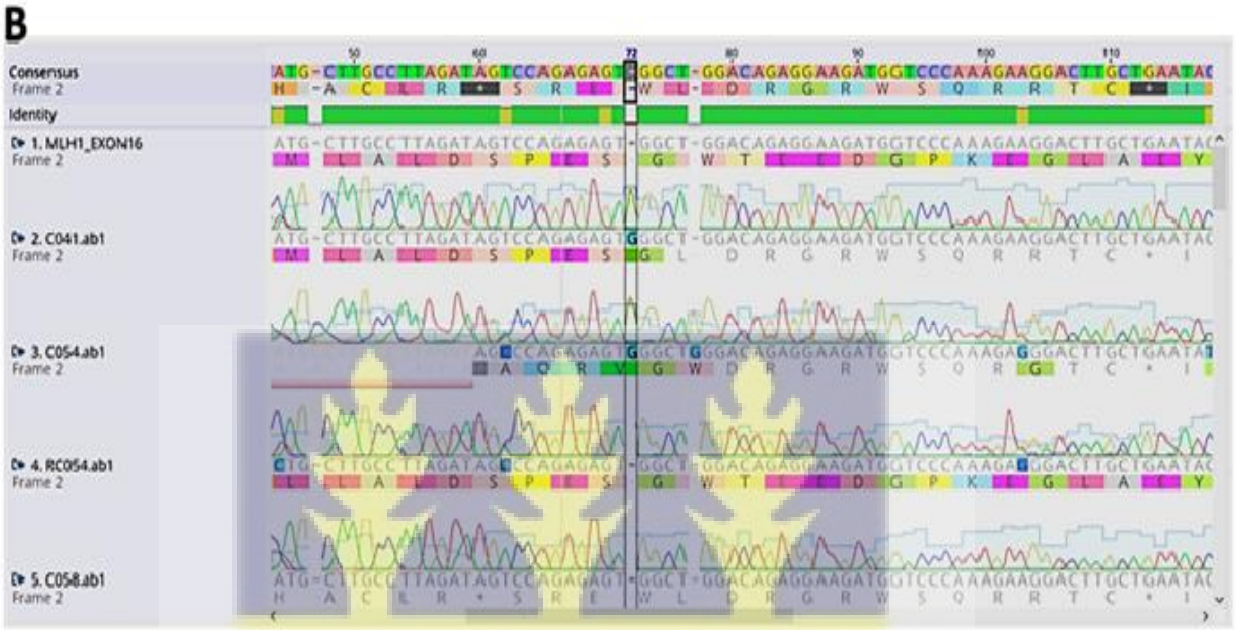



**Figure 4.8: Alignment of the two case sample sequences with mutation against the MLH1 Exon 16 reference sequence.**

Image showing guanine (G) insertion at codon 596 of exon 16 of the MLH1 protein in two samples (C041 and C054).

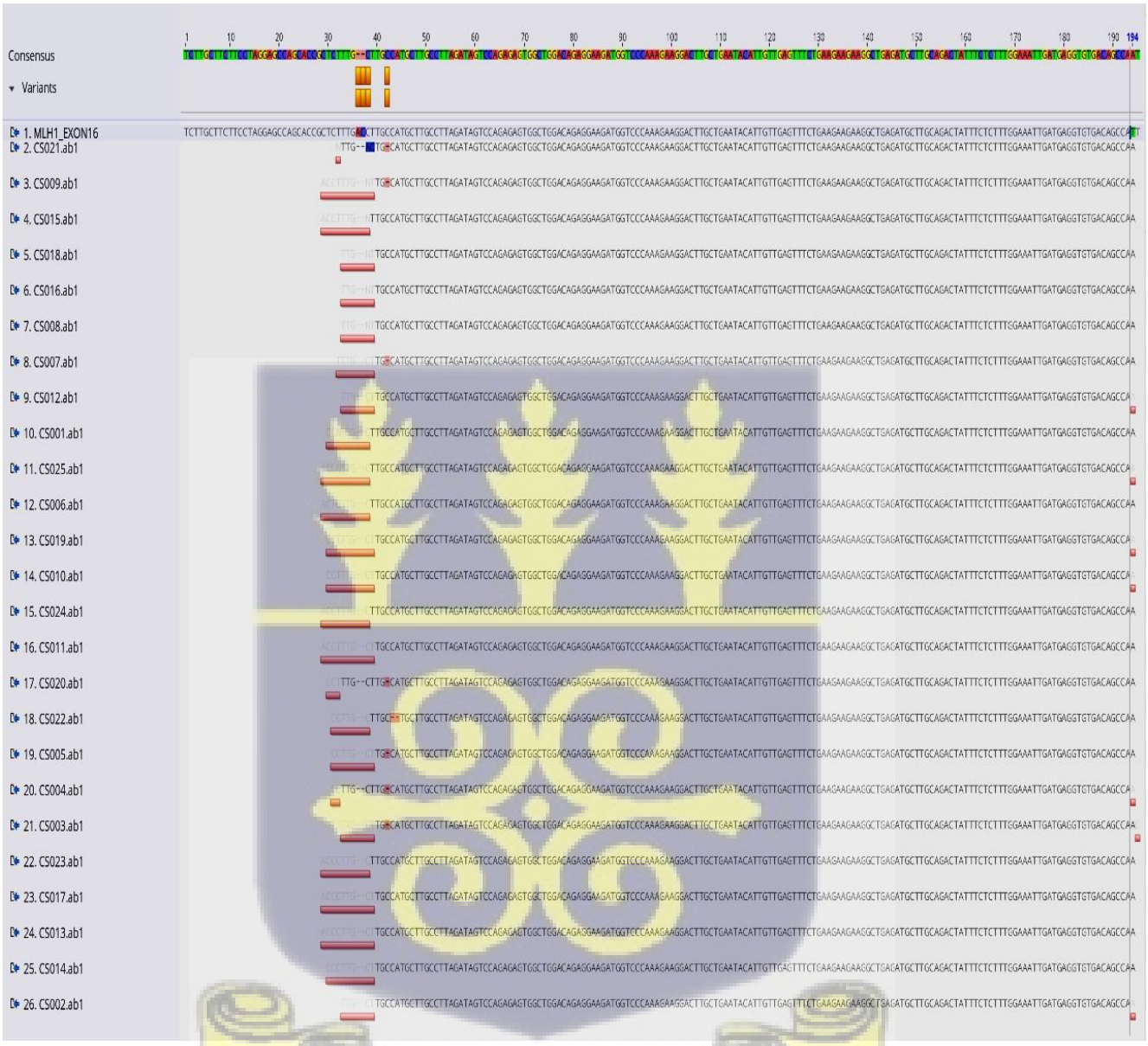
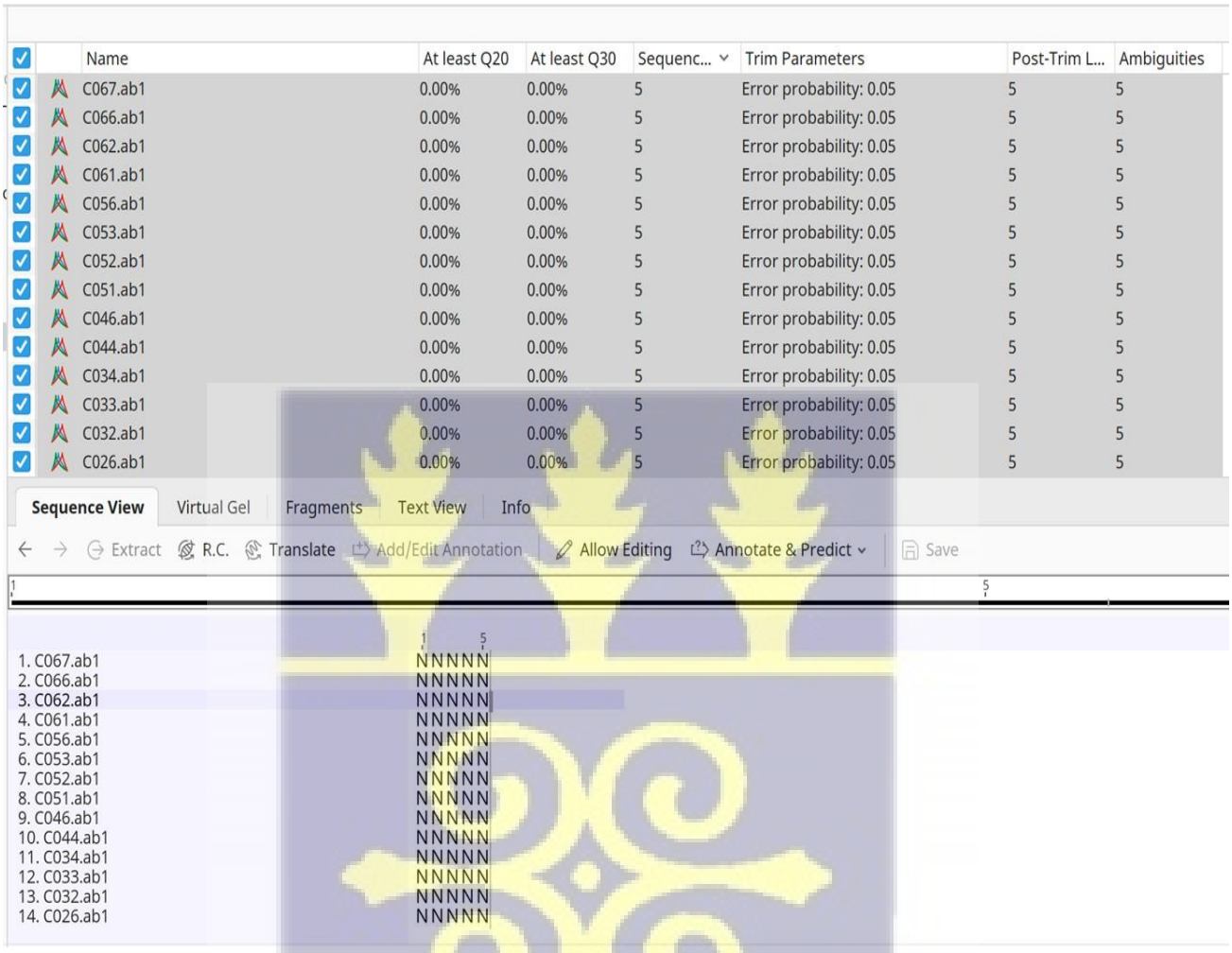



**Figure 4.9: Alignment of control sequences with the MLH1 Exon16 reference sequence (No variants detected in good QC regions).**

Control sequence compared to the reference sequence shows no mutations in the control group. Positions coloured signify potential mutations, but they could not be verified on the traces because of the error probability.



| Name | At least Q20 | At least Q30 | Sequenc... | Trim Parameters | Post-Trim L... | Ambiguities |
|---|---|---|---|---|---|---|
| C067.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C066.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C062.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C061.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C056.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C053.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C052.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C051.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C046.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C044.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C034.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C033.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C032.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |
| C026.ab1 | 0.00% | 0.00% | 5 | Error probability: 0.05 | 5 | 5 |



**Figure. 4.10 - The 14 Case Samples with No Nucleotide Sequence.**

# CHAPTER FIVE

## 5.0 DISCUSSION AND CONCLUSION

### 5.1 Discussion

Regional differences in the clinical pathological patterns of colorectal cancer have reportedly been identified. In the west, colorectal cancer is typically detected in older people after the age of 50 years (Center *et al.*, 2009a). However in the African population, CRCtends to present early with an aggressive, advanced disease and a dismal prognosis (Sule and Mandong, 1999).

The median ages for cases and controls in this study were 56.5 years and 47 years respectively; this is comparable to the results of other African studies, which reported mean ages of 53 years in Nigeria (Akinola and Arigbabu, 1994), 58.8 years in Tunisia (Missaoui *et al.*, 2010), 58 years in Ghana (Dakubo *et al.*, 2014), 50 years in Central Sudan (Ntagirabiri *et al.*, 2016). It should be noted that the median ages recorded in this study (56.5 years and 47 years) were lower than the ages described in most developed countries (Koo et al., 2008; Altekruse *et al.*, 2010). From 2003 to 2007, the median age at diagnosis was 70 years in the United States (Koo *et al.*, 2008). The likelihood of receiving a colon cancer diagnosis rises after age 40 years, then gradually after that age, and then suddenly after age 50 years (Ries *et al.*, 2008).

From this study, colorectal cancer prevalence among the young (<40 years) was 21.1% (15/71) and that for the aged (>50 years) was 78.9% (56/71); which is similar but lower when compared to the finding in Egypt by Gado *et al.* (2014) , where the prevalence of CRC in patients aged less than 40 years was 25%. This finding is also similar to Ries *et*

*al.*, 2008 findings, where more than 90% of colorectal cancer cases occur in people aged 50 or older.

Although this study shows a reduced incidence of CRC in young people, it has been shown that CRC in this age group poses diagnostic and therapeutic challenges, and the prognosis is often less favourable. (Sule and Mandong, 1999). Family screening for genetic mutations (HNPCC) is necessary as a result of the rising frequency of CRC in young people in low-risk communities since genetic variables mayplay a significant role in the development of this illness. Hence, the molecular componentof this work.

Similar to previous studies in Ghana (Dakubo *et al.*, 2010a; Dakubo *et al.*, 2014), and other African countries (Boytchev *et al.*, 1999), the gender (male: female) prevalence ratio of both case and control groups in this study was the same 1:1(Table1), instead of the 1.4:1 ratio proposed by Ferlay *et al.*, 2010, in favour of the men. Prevalence rates are noticeably greater in men than in women over the world (Siegel *et al.*, 2012). Men experience the global CRC prevalence at a rate of 23.4 cases per 100,000 people in 2020, which is 44% greater than women's incidence at 16.2 cases per 100,000 people, which is 30% higher (Xi and Xu, 2021). This has been linked to the increased use of risk behaviors for colorectal cancer such as smoking, drinking excessive amounts of alcohol, and male obesity (Center *et al.*, 2009a; Missaoui *et al.*, 2010).

The role of female hormones has also been connected to the declining occurrence in both humans and female animals. Through its affinity to estrogen receptors, estrogen controls

the growth of colonic mucosal cells and prevents the division of colonic tumor cells, according to in vitro studies (Issa *et al.*, 1994; Campbell-Thompson *et al.*, 2001). The present male to female ratio of 1:1 may be as a result of sampling, or it could also be that at age 50 when most of the diagnosis are made the female has lost the advantage of the estrogen effect because of menopause.

A review of the family's history of cancer in the recruited individuals serves to validate the findings of Haggar and Boushey (2009) and Rawla *et al.* (2017). A number of factors have been linked to the development of colorectal cancer, and those who indulge in ridk life styles like smoking, excessive alcohol use, poor diet and have risk factors like family history, inflammatory bowel disease, inactivity, genetic make-up, and specific illnesses or diseases have an 80% chance of getting the disease. Therefore it was not surprising to see a statisticallysignificant association ($p<0.001$) between the family history of the studied participants and the CRC disease. 6/71 of the case studied participants had a family history of cancers;with colorectal cancer leading at a prevalence of 50.0% (3/6), as compared to breast and stomach cancer 33.3% (2/6) and pancreatic cancer 16.7% (1/6).

Following through with the Chi squared analysis (table 4.2), it could be said that there wasno statistically significant association between smoking, drinking alcohol and CRC; making this research finding to be inconsistent with Rawla *et al.*, 2019 findings at this point.

Thorough observation within the case group suggested that, 2.8% (2/71) of the case participants consume some form of alcohol while there was no report of any within the

control group. Again, while there was no report of smoking within the control group, there was 1 (1.4%) participant within the case group who smoked. Therefore in comparing the studied cases to controls, it could be deduced easily that, those participants who smoked and drank risked developing the CRC disease and surely did; since they were all found within the case domain.

Analysis of the mutation spectrum of the exon 16 of the MLH1 gene among the Ghanaian CRC population revealed the following:

That 14/71 (19.7%) of the case population probably had no exon 16 sequence after sequencing. Their MLH1 exon 16 sequence may have been deleted and this has confirmed the gel results obtained (figure 4.1 and 4.2) as there were no bands for these same case samples. The deletion could have occurred as a result of mutations at the splice site of intron 15. This could be possible if the conserved AG at the 3' site of intron 15 is mutated,and this resulted in the site not being recognized by the spliceosomes, thereby leading to inclusion of exon 16 as part of intron 15 resulting to its deletion.

The percentage of whole exon 16 deletion mutation observed is 19.7% from this study, which is similar but rather on the higher side to a study conducted in Sweden in 2016, by Lagerstedt-Robinson *et al.;* where MLH1 exon 16 deletion was identified to be 6% in the Sweden lynch syndrome family. The geographical location and race could account for the variability in the differences in the two percentages.

Again, similar to a study done in Japan in 2016, where Momma and colleagues analyzed the chromosomal rearrangement of the MLH1 and MSH2 genes in a Lynch syndrome family that had been tracked for over 45 years. In this family, they discovered widespread deletions (c. (306+ 1 307–1) (*193?) del of the MLH1 exons 4-19.

This study is again in line with Dominguez Valentin et al., 2013 findings. Accordingto his findings, exons 16 and 18 of the MLH1 gene are genetic hotspots for mutations, with frameshift mutations and nonsense mutations being the most common types of mutations and accounted for 26% of all the MLH1 mutations (Li *et al.*, 2011).

Establishment of a linked between these 14 case participants, and their documented risk factors (alcoholism, smoking, and family history of cancer) showed that, these patients rather lived a supposedly healthy lifestyle in terms of the risk factors mentioned; and none had a family history of cancer too. From literature it is noted that about two-thirds (70-75%) of new CRC cases are spontaneous / sporadic (i.e., these cancers develop among persons who have no history of the illness in their family) (Yamagishi *et al.*, 2016), but rather arise 'de novo' in these rare cases due to aetiological factors such as age, lifestyle and environmental effects.

It could however be speculated that, not all these 14 cases could be classified as sporadic due to their ages .Approximately, 78.6% (11/14) of the cases in this category were within the age ranged (52-82) years; this might have predisposed them to the condition they had since 'age' is one of the risk factors of CRC according to Yamagishi *et al.*, 2016. It could also be that, they acquired these genetic modifications de novo; although this speculation could not be supported with additional data from the study.

In this regard; this finding could therefore be seen to support Dakubo *et al.*, findings in

2010, when they observed that, as the Accra population ages, the prevalence of colorectal cancer has increased, with adenocarcinoma being the most common histological form. The remaining three 3/14 (21.4%) case participants who aged between 28-49 years can be said to have inherited the condition genetically- this was indicative of their young ages. This was difficult to prove since collected data was not enough to make such a conclusive statement.

Hereditary colorectal cancer accounts for 5% to 10% of all CRC cases, or nearly one-third of all CRC cases (Cherry, 2011). First-degree relatives of patients with newly diagnosed adenomas or invasive CRCs are at higher risk since hereditary CRC runs in families and is also linked to germline mutations in specific genes (Cherry, 2011).
These 3 case individuals were oblivious of their family's cancer history with their 'don't know" response; even though they would have had a family history of HNPCC or FAP syndrome; this is difficult to prove with data though. These syndromes frequently result in cancers that emerge at a younger age (Segelman, 2012).

To explain the concept of carcinogenesis in these individuals, mismatch repair genes during DNA replication produce protein enzymes which are able to detect and correct mismatched base pairs. A lack of function of the repair genes (example deletion of MLH1 exon 16) causes MSI which leads to accumulation of significant changes in the length of tiny repetitive DNA sequences. As a result, the malfunctioning MMR would cause a rapid accumulation of gene mutations in proto- oncogenes and cancer suppressor genes which may disrupt normal cell growth and eventually, aid development of the CRC (Popat *et al.*, 2005; Kerr and Midgley, 2010; Sargent *et al.*, 2010).

Single guanine nucleotide insertion frameshift mutation was also identified in two ofthe case individuals; participants (C041) and (C054).

A single (G) nucleotide insertion at codon 596 of the coding sequence of the exon 16 MLH1 protein was called out during the bioinformatics analysis; which obviously affected the ribosome reading frame as there was a nucleotide reading frameshift, resulting in a premature stop codon at codon 608 of the MLH1 and as such produced a truncated non- functional protein, which was 148 amino acids shorter than the normal MLH1 protein sequence which is 756 amino acid long.

Interesting as it may seem, this finding is not consistent with the findings of any of these researchers (Dominguez Valentin *et al.*, 2013, Lagerstedt-Robinson *et al.*, 2016; Moreno-Ortiz *et al.*, 2016; and Momma *et al.*, 2019); because they had not reported guanine (G) insertion at the same position/codon site in their publications. This finding could therefore be said to be unique and novel, due to the guanine insertion frameshift mutation at codon 596 of the coding sequence of exon 16 MLH1 protein. Although there was an insertion frameshift, it may seem difficult to classify this mutation as novel among the Ghanaian CRC population, since the research data at disposal is limited in terms of sample size.

There were no SNPs called within the QC regions of the control samples. And this was expected because these were supposedly CRC free healthy Ghanaians who stood at no risk of developing the disease and therefore it not expected to see any form of mutation in their MLH1 exon 16.

## 5.2 LIMITATIONS

1. This study failed to sequence the whole MLH1 genome and as such reduced drastically the possibility of finding novel mutations for the Ghanaian CRC population.
2. Again because the study was centered on finding mutations in only exon 16 of the MLH1 gene, it failed to give off knowledge and information about the mutation spectrums of the other exons within the MLH1 gene.
3. Difficulty in referring information since there was few studies (baseline data) done in this particular area in the country, Ghana.
4. Although nearly all oncology patients in Ghana's southern and central regions are sent to KBTH for management, the study only included patients who were assessed and treated at a single facility, which may not accurately represent the whole community in this region.

## 5.3 CONCLUSION

This study is the first compiled Molecular data on the Ghanaian CRC population. The study reported two types of mutation; whole exon 16 deletion and a single nucleotide (G) insertion frameshift mutation at codon 596 of the coding sequence of exon 16 of MLH1 protein.The whole exon (16) deletion in this study accounted for 19.7 % (14/71) of the total mutations found within the MLH1 gene while insertion frameshift mutation accounted for 2.8% (2/71) of the total mutation discovered.

## RECOMMENDATIONS

1. Further studies to amplify and sequence the whole MLH1 gene using a primer with a larger flanking region: as this would help to call out more mutations that were prevailing within the Ghanaian CRC population. Once more, the knowledge (data) gleaned from this mutation spectrum would be helpful for identifying prognostic markers for efficient treatment interventions in the Ghanaian CRC population, as well as for profiling particular genetic alterations.

2. The government should strengthen the capabilities of Ghanaian healthcare practitioners by providing the essential diagnostic equipment for the quick detection and referral of colorectal cancer cases. National screening programs should be put in place for the illness prevention and early detection.

## REFERENCES

Adelstein, B. A., Macaskill, P., Chan, S. F., Katelaris, P. H., & Irwig, L. (2011). Most bowel cancer symptoms do not indicate colorectal cancer and polyps: a systematic review. *BMC gastroenterology*, *11*, 1-10.

Adesanya, A. A., & da Rocha-Afodu, J. T. (2000). Colorectal cancer in Lagos: a critical review of 100 cases. *The Nigerian postgraduate medical journal*, *7*(3), 129-136.

Akiyama, Y., Sato, H., Yamada, T., Nagasaki, H., Tsuchiya, A., Abe, R., & Yuasa, Y. (1997). Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. *Cancer research*, *57*(18), 3920-3923.

AO, P., Omodele, F. O., Oludara, M. A., NA, I., AI, I., & SBL, M. (2013). Prevalence and pattern of cancers among adults attending a tertiary health institution in Lagos, Nigeria. *Journal of Dental and Medical Sciences*, *6*(3), 68-73.

Araki, T., & Milbrandt, J. (1996). Ninjurin, a novel adhesion molecule, is induced by nerve injury and promotes axonal growth. *Neuron*, *17*(2), 353-361.

Biritwum, R. B., Welbeck, J., & Barnish, G. (2000). Incidence and management of malaria in two communities of different socio—economic level, in Accra, Ghana. *Annals of Tropical Medicine & Parasitology*, *94*(8), 771-778.

Boytchev, H., Marcovic, S., & Oettle, G. J. (1999). The characteristics of large bowel cancer in the low-risk black population of the Witwatersrand. *Journal of the Royal College of Surgeons of Edinburgh*, *44*(6), 366-370.

Center, M. M., Jemal, A., Smith, R. A., & Ward, E. (2009). Worldwide variations in colorectal cancer. *CA: a cancer journal for clinicians*, *59*(6), 366-378.

Calys-Tagoe, B. N., Yarney, J., Kenu, E., Amanhyia, N. A. K. O., Enchill, E., & Obeng,

I. (2014). Profile of cancer patients' seen at Korle Bu teaching hospital in Ghana (A cancer registry review). *BMC research notes*, *7*(1), 1-6.

Chalya, P. L., Mchembe, M. D., Mabula, J. B., Rambau, P. F., Jaka, H., Koy, M., ... & Masalu, N. (2013). Clinicopathological patterns and challenges of management of colorectal cancer in a resource-limited setting: a Tanzanian experience. *World journal of surgical oncology*, *11*(1), 1-9.

Charan, J., & Biswas, T. (2013). How to calculate sample size for different study designs in medical research? *Indian journal of psychological medicine*, *35*(2), 121.

Cherry, L. M. (2011, April). The genetic etiology of familial and nonfamilial colorectal cancer. In *Baylor University Medical Center Proceedings* (Vol. 24, No. 2, pp. 139-141). Taylor & Francis.

Chokunonga, E., Borok, M. Z., Chirenje, Z. M., Nyakabau, A. M., & Parkin, D. M. (2013). Trends in the incidence of cancer in the black population of Harare, Zimbabwe 1991–2010. *International journal of cancer*, *133*(3), 721-729.

Cronje, L., Paterson, A., & Becker, P. (2009). Colorectal cancer in South Africa: a heritable cause suspected in many young black patients. *South African Medical Journal*, *99*(2), 103-106.

Dakubo, J. C. B., Naaeder, S. B., Tettey, Y., & Gyasi, R. K. (2010). Colorectal carcinoma: an update of current trends in Accra. *West African journal of medicine*, *29*(3).

Domingo, E; Schwartz, S Jr, MLH1 (human mutL homolog 1), Atlas Genet Cytogenet oncol haematol. 2005 ;( 2):120-122. *http://atlasgeneticsoncology.org › Genes › GC_MLH1*

Dominguez-Valentin, M., Nilbert, M., Wernhoff, P., López-Köstner, F., Vaccaro, C.,

Sarroca, C., & Rossi, B. M. (2013). Mutation spectrum in South American Lynch syndrome families. *Hereditary cancer in clinical practice*, *11*(1), 1-9.

Drotschmann, K., Fritz, H. J., Aronshtam, A., & Marinus, M. G. (1998). The Escherichia coli MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucleic acids research*, *26*(4), 948-953.

Drummond, J. T., Li, G. M., Longley, M. J., & Modrich, P. (1995). Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science*, *268*(5219), 1909-1912.

Dunlop, M. G., Farrington, S. M., Carothers, A. D., Wyllie, A. H., Sharp, L., Burn, J., & Vogelstein, B. (1997). Cancer risk associated with germline DNA mismatch repair gene mutations. *Human molecular genetics*, *6*(1), 105-110.

Eshleman, J. R., & Markowitz, S. D. (1996). Mismatch repair defects in human carcinogenesis. *Human molecular genetics*, *5*(Supplement_1), 1489-1494.

Fang, Y. J., Wu, X. J., Zhao, Q., Li, L. R., Lu, Z. H., Ding, P. R., ... & Wan, D. S. (2013). Hospital-based colorectal cancer survival trend of different tumor locations from 1960s to 2000s. *PloS one*, *8*(9), e73528.

Fearnhead, N. S., Britton, M. P., & Bodmer, W. F. (2001). The abc of apc. *Human molecular genetics*, *10*(7), 721-733.

Fishel, R., Lescoe, M. K., Rao, M. R. S., Copeland, N. G., Jenkins, N. A., Garber, J., ... & Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, *75*(5), 1027-1038.

Flores-Rozas, H., & Kolodner, R. D. (1998). The Saccharomyces cerevisiae MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. *Proceedings of the National Academy of Sciences*, *95*(21), 12404-12409.

Faustino-Rocha, A. I., & Oliveira, P. A. (2020). Chemical carcinogens.

Genschel, J., Littman, S. J., Drummond, J. T., & Modrich, P. (1998). Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα. *Journal of Biological Chemistry*, *273*(31), 19895-19901.

Genschel, J., Littman, S. J., Drummond, J. T., & Modrich, P. (1998). Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα. *Journal of Biological Chemistry*, *273*(31), 19895-19901.

Gooch, M. J. (2016). Influence of Mutation Frequency on Mutation Profile in Colon Cancer.

Goss, K. H., & Groden, J. (2000). Biology of the adenomatous polyposis coli tumor suppressor. *Journal of clinical oncology*, *18*(9), 1967-1979.

Greene, C. N., & Jinks-Robertson, S. (1997). Frameshift intermediates in homopolymer runs are removed efficiently by yeast mismatch repair proteins. *Molecular and Cellular Biology*, *17*(5), 2844-2850.

Gelsomino, F., Barbolini, M., Spallanzani, A., Pugliese, G., & Cascinu, S. (2016). The evolving role of microsatellite instability in colorectal cancer: a review. *Cancer treatment reviews*, *51*, 19-26.

Gado, A., Ebeid, B., Abdelmohsen, A., & Axon, A. (2014). Colorectal cancer in Egypt is commoner in young people: Is this cause for alarm?. *Alexandria Journal of Medicine*, *50*(3), 197-201.

Habraken, Y., Sung, P., Prakash, L., & Prakash, S. (1996). Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Current Biology*, *6*(9), 1185-1187.

Habraken, Y., Sung, P., Prakash, L., & Prakash, S. (1997). Enhancement of MSH2–MSH3-

mediated mismatch recognition by the yeast MLH1–PMS1 complex. *Current Biology*, *7*(10), 790-793.

Haggar, F. A., & Boushey, R. P. (2009). Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clinics in colon and rectal surgery*, *22*(04), 191-197.

Hamilton, S. R., Liu, B., Parsons, R. E., Papadopoulos, N., Jen, J., Powell, S. M., ... & Kinzler, K. W. (1995). The molecular basis of Turcot's syndrome. *New England Journal of Medicine*, *332*(13), 839-847.

Hou, J. T., Zhao, L. N., Zhang, D. J., Lv, D. Y., He, W. L., Chen, B., ... & Chen, X. L. (2018). Prognostic Value of Mismatch Repair Genes for Patients with Colorectal Cancer: Meta-Analysis. *Technology in cancer research & treatment*, *17*, 1533033818808507.

Hegde, P., Qi, R., Gaspard, R., Abernathy, K., Dharap, S., Earle-Hughes, J., ... & Quackenbush, J. (2001). Identification of tumor markers in models of human colorectal cancer using a 19,200-element complementary DNA microarray. *Cancer Research*, *61*(21), 7792-7797

Jass, J. R. (2006). Hereditary non-polyposis colorectal cancer: the rise and fall of a confusing term. *World journal of gastroenterology: WJG*, *12*(31), 4943.

Kamangar, F., Dores, G. M., & Anderson, W. F. (2006). Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *Journal of clinical oncology*, *24*(14), 2137-2150.

Katsidzira, L., Gangaidzo, I., Thomson, S., Rusakaniko, S., Matenga, J., & Ramesar, R. (2017). The shifting epidemiology of colorectal cancer in sub-Saharan Africa. *The*

*Lancet Gastroenterology & Hepatology*, *2*(5), 377-383.

Kerr, D. J., & Midgley, R. (2010). Defective mismatch repair in colon cancer: a prognostic or predictive biomarker?

Kolodner, R. (1996). Biochemistry and genetics of eukaryotic mismatch repair. *Genes & development*, *10*(12), 1433-1442.

Kolodner, R. D., Tytell, J. D., Schmeits, J. L., Kane, M. F., Gupta, R. D., Weger, J., ... & Li, F. P. (1999). Germ-line msh6 mutations in colorectal cancer families. *Cancer Research*, *59*(20), 5068-5074.

Kolodner, R. D., & Marsischky, G. T. (1999). Eukaryotic DNA mismatch repair. *Current opinion in genetics & development*, *9*(1), 89-96.

Krzyszczyk, P., Acevedo, A., Davidoff, E. J., Timmins, L. M., Marrero-Berrios, I., Patel, M., & Yarmush, M. L. (2018). The growing role of precision and personalized medicine for cancer treatment. *Technology*, *6*(03n04), 79-100.

Kudryavtseva, A. V., Lipatova, A. V., Zaretsky, A. R., Moskalev, A. A., Fedorova, M. S., Rasskazova, A. S., ... & Krasnov, G. S. (2016). Important molecular genetic markers of colorectal cancer. *Oncotarget*, *7*(33), 53959.

Kunkel, T. A., & Erie, D. A. (2005). DNA mismatch repair. *Annu. Rev. Biochem.*, *74*, 681-710.

Lagerstedt-Robinson, K., Rohlin, A., Aravidis, C., Melin, B., Nordling, M., Stenmark-Askmalm, M., & Nilbert, M. (2016). Mismatch repair gene mutation spectrum in the Swedish Lynch syndrome population. *Oncology reports*, *36*(5), 2823-2835.

Lakatos, P. L., & Lakatos, L. (2008). Risk for colorectal cancer in ulcerative colitis: changes, causes and management strategies. *World journal of gastroenterology: WJG*, *14*(25), 3937.

Leach, F. S., Nicolaides, N. C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., & Vogelstein, B. (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell*, *75*(6), 1215-1225.

Li, G. M. (2008). Mechanisms and functions of DNA mismatch repair. *Cell research*, *18*(1), 85-98.

Lindblom, A., Tannergård, P., Werelius, B., & Nordenskjöld, M. (1993). Genetic mapping of a second locus predisposing to hereditary non–polyposis colon cancer. *Nature genetics*, *5*(3), 279-282.

Liu, B., Parsons, R., Papadopoulos, N., Nicolaides, N. C., Lynch, H. T., Watson, P., & Kinzler, K. W. (1996). Analysis of mismatch repair genes in hereditary non–polyposis colorectal cancer patients. *Nature medicine*, *2*(2), 169-174.

Lynch, H. T., & De la Chapelle, A. (1999). Genetic susceptibility to non-polyposis colorectal cancer. *Journal of medical genetics*, *36*(11), 801-818.

Lynch, H. T., & Lynch, J. F. (1998). Genetics of colonic cancer. *Digestion*, *59*(5), 481-492.

Lynch, H. T., Smyrk, T., & Lynch, J. (1997). An update of HNPCC (Lynch syndrome). *Cancer genetics and cytogenetics*, *93*(1), 84-99.

Marcus, D. M., Marcus, R. P., Prabhu, R. S., Owonikoko, T. K., Lawson, D. H., Switchenko, J., & Beitler, J. J. (2012). Rising incidence of mucosal melanoma of the head and neck in the United States. *Journal of skin cancer*, *2012*.

Marsischky, G. T., Filosi, N., Kane, M. F., & Kolodner, R. (1996). Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes & development*, *10*(4), 407-420.

McCulloch, S. D., & Kunkel, T. A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell research*, *18*(1), 148-161.

Miyaki, M., Konishi, M., Tanaka, K., Kikuchi-Yanoshita, R., Muraoka, M., Yasuno, M., & Mori, T. (1997). Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. *Nature genetics*, *17*(3), 271-272.

Modrich, P., & Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annual review of biochemistry*, *65*(1), 101-133.

Momma, T., Gonda, K., Akama, Y., Endo, E., Ujiie, D., Fujita, S., & Nomizu, T. (2019). MLH1 germline mutation associated with Lynch syndrome in a family followed for more than 45 years. *BMC medical genetics*, *20*(1), 1-6.

Moreno-Ortiz, J. M., Ayala-Madrigal, M. D. L. L., Corona-Rivera, J. R., Centeno-Flores, M., Maciel-Gutiérrez, V., Franco-Topete, R. A., & Gutiérrez-Angulo, M. (2016). Novel mutations in MLH1 and MSH2 genes in Mexican patients with Lynch syndrome. *Gastroenterology research and practice*, *2016*.

Macfarlane, S., & Macfarlane, G. T. (2003). Regulation of short-chain fatty acid production. *Proceedings of the Nutrition Society*, *62*(1), 67-72.

Mesker, W. E., Junggeburt, J., Szuhai, K., de Heer, P., Morreau, H., Tanke, H. J., & Tollenaar, R. A. (2007). The carcinoma–stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Analytical Cellular Pathology*, *29*(5), 387-398.

Missaoui, N., Jaidaine, L., Ben Abdelkader, A., Beizig, N., Anjorin, A., Yaacoubi, M. T., & Hmissa, S. (2010). Clinicopathological patterns of colorectal cancer in Tunisia. *Asian Pacific J Cancer Prev*, *11*(6), 1719-1722.

Miller, B. A., Chu, K. C., Hankey, B. F., & Ries, L. A. (2008). Cancer incidence and mortality patterns among specific Asian and Pacific Islander populations in the US. *Cancer Causes & Control*, *19*, 227-256.

Naaeder, S. B., & Archampong, E. Q. (1994). Cancer of the colon and rectum in Ghana: a 5-year prospective study. *Journal of British Surgery*, *81*(3), 456-459.

Nakagawa, H., Lockman, J. C., Frankel, W. L., Hampel, H., Steenblock, K., Burgart, L. J., ... & De La Chapelle, A. (2004). Mismatch repair gene PMS2: disease-causing germline mutations are frequent in patients whose tumors stain negative for PMS2

Ntagirabiri, R., Karayuba, R., Ndayisaba, G., Niyonkuru, S., & Amani, M. (2016). Colorectal cancer: Epidemiological, clinical and histopathological aspects in Burundi. *Open Journal of Gastroenterology*, *6*(3), 83-87.

Nicolaides, N. C., Papadopoulos, N., Liu, B., Weit, Y. F., Carter, K. C., Ruben, S. M., & Kinzler, K. W. (1994). Mutations of two P/WS homologues in hereditary nonpolyposis colon cancer. *Nature*, *371*(6492), 75-80.

Nicolaides, N. C., Papadopoulos, N., Liu, B., Weit, Y. F., Carter, K. C., Ruben, S. M., & Kinzler, K. W. (1994). Mutations of two P/WS homologues in hereditary nonpolyposis colon cancer. *Nature*, *371*(6492), 75-80.

Nilbert, M., Therkildsen, C., Nissen, A., Åkerman, M., & Bernstein, I. (2009). Sarcomas associated with hereditary nonpolyposis colorectal cancer: broad anatomical and morphological spectrum. *Familial cancer*, *8*(3), 209.

Ogino, S., & Goel, A. (2008). Molecular classification and correlates in colorectal cancer. *The Journal of Molecular Diagnostics*, *10*(1), 13-27.

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., ... & Masuho, Y. (2004). Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nature*

*genetics*, *36*(1), 40-45.

Palombo, F., Iaccarino, I., Nakajima, E., Ikejima, M., Shimada, T., & Jiricny, J. (1996). HMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Current Biology*, *6*(9), 1181-1184.

Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'Arrigo, A., & Jiricny, J. (1995). GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science*, *268*(5219), 1912-1914.

Papadopoulos, N., Nicolaides, N. C., Liu, B., Parsons, R., Lengauer, C., Palombo, F., & Kinzler, K. W. (1995). Mutations of GTBP in genetically unstable cells. *Science*, *268*(5219), 1915-1917.

Papadopoulos, N., Nicolaides, N. C., Wei, Y. F., Ruben, S. M., Carter, K. C., Rosen, C. A., ... & Vogelstein, B. (1994). Mutation of a mutL homolog in hereditary colon cancer. *Science*, *263*(5153), 1625-1629.

Peltomaki, P., Aaltonen, L. A., Sistonen, P., Pylkkanen, L., Mecklin, J. P., Jarvinen, H., & Leach, F. S. (1993). Genetic mapping of a locus predisposing to human colorectal cancer. *Science*, *260*(5109), 810-812.

Plaschke, J., Kruppa, C., Tischler, R., Bocker, T., Pistorius, S., Dralle, H., & Schackert, H. K. (2000). Sequence analysis of the mismatch repair gene hMSH6 in the germline of patients with familial and sporadic colorectal cancer. *International journal of cancer*, *85*(5), 606-613.

Plaschke, J., Kruppa, C., Tischler, R., Bocker, T., Pistorius, S., Dralle, H., & Schackert, H. K. (2000). Sequence analysis of the mismatch repair gene hMSH6 in the germline of patients with familial and sporadic colorectal cancer. *International journal of cancer*, *85*(5), 606-613.

Popat, S., Hubner, R., & Houlston, R. S. (2005). Systematic review of microsatellite instability and colorectal cancer prognosis. *Journal of clinical oncology*, *23*(3), 609-618.

Prolla, T. A., Pang, Q., Alani, E., Kolodner, R. D., & Liskay, R. M. (1994). MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. *Science*, *265*(5175), 1091-1093.

Puchner, K., Iavazzo, C., Gourgiotis, D., Boutsikou, M., Baka, S., Hassiakos, D., & Creatsas, G. (2011). Mid-trimester amniotic fluid interleukins (IL-1β, IL-10 and IL-18) as possible predictors of preterm delivery. *In vivo*, *25*(1), 141-148.

Raskin, L., Dakubo, J. C., Palaski, N., Greenson, J. K., & Gruber, S. B. (2013). Distinct molecular features of colorectal cancer in Ghana. *Cancer epidemiology*, *37*(5), 556-561.

Rattray, N. J., Charkoftaki, G., Rattray, Z., Hansen, J. E., Vasiliou, V., & Johnson, C. H. (2017). Environmental influences in the etiology of colorectal cancer: the premise of metabolomics. *Current pharmacology reports*, *3*(3), 114-125.

Rawla, Prashanth, Tagore Sunkara, and Adam Barsouk. "Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors." *Przeglad gastroenterologiczny* 14.2 (2019): 89.

Ren C, Liu Y, Wang Y, Tang Y, Wei Y, Liu C. Identification of novel Lynch syndrome mutations in Chinese patients with endometriod endometrial cancer. Cancer *Biol Med. 2020; 17: 458-467*. doi: 10.20892/j.issn. 2095-3941.2019.0295

Risinger, J. I., Umar, A., Boyd, J., Berchuck, A., Kunkel, T. A., & Barrett, J. C. (1996). Mutation of MSH3 in endometrial cancer and evidence for its functional role in heteroduplex repair. *Nature genetics*, *14*(1), 102-105.

Ruddon, R. W. (2007). *Cancer biology*. Oxford University Press.

Russo, A., Sala, P., Alberici, P., Gazzoli, I., Radice, P., Montefusco, C., & Bertario, L. (2009). Prognostic relevance of MLH1 and MSH2 mutations in hereditary non-polyposis colorectal cancer patients. *Tumori Journal*, *95*(6), 731-738.

Sancar, A. (1999). Excision repair invades the territory of mismatch repair. *Nature genetics*, *21*(3), 247-249.

Sargent, D. J., Marsoni, S., Monges, G., Thibodeau, S. N., Labianca, R., Hamilton, S. R., & Gallinger, S. (2010). Defective mismatch repair as a predictive marker for lack

Scanlon, V. C., & Sanders, T. (2014). The digestive system, Essential of Anatomy and Physiology.

Sule, A. Z., & Mandong, B. M. (1999). Malignant colorectal tumours in patients 30 years and below: a review of 35 cases. *The Central African journal of medicine*, *45*(8), 209-212.

Siegel, R., DeSantis, C., Virgo, K., Stein, K., Mariotto, A., Smith, T., ... & Ward, E. (2012). Cancer treatment and survivorship statistics, 2012. *CA: a cancer journal for clinicians*, *62*(4), 220-241.

Scott, Rodney J. "Familial adenomatous polyposis (FAP) and other polyposis syndromes." *Hereditary Cancer in Clinical Practice* 1.1 (2003): 1-12.

Segelman, J., Granath, F., Holm, T., Machado, M., Mahteme, H., & Martling, A. (2012). Incidence, prevalence and risk factors for peritoneal carcinomatosis from colorectal cancer. *Journal of British Surgery*, *99*(5), 699-705.

Sia, E. A., Kokoska, R. J., Dominska, M., Greenwell, P., & Petes, T. D. (1997). Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. *Molecular and cellular biology*, *17*(5), 2851-2858.

Soliman, A. S., Bondy, M. L., El-Badawy, S. A., Mokhtar, N., Eissa, S., Bayoumy, S., ... & Hamilton, S. R. (2001). Contrasting molecular pathology of colorectal

carcinoma in Egyptian and Western patients. *British journal of cancer*, *85*(7), 1037-1046.

Spallanzani, A., Gelsomino, F., Caputo, F., Santini, C., Andrikou, K., Orsi, G., & Cascinu, S. (2018). Immunotherapy in the treatment of colorectal cancer: a new kid on the block. *Journal of Cancer Metastasis and Treatment*, *4*.

Swale, V. J., Quinn, A. G., Bataille, V. A., Wheeler, J. M., Beck, N. E., Dove-Edwin, I., ... & Bodmer, W. F. (1999). Microsatellite instability in benign skin lesions in hereditary non-polyposis colorectal cancer syndrome. *Journal of investigative dermatology*, *113*(6), 901-905.

Syngal, S., Fox, E. A., Li, C., Dovidio, M., Eng, C., Kolodner, R. D., & Garber, J. E. (1999). Interpretation of genetic test results for hereditary nonpolyposis colorectal cancer: implications for clinical predisposition testing. *Jama*, *282*(3), 247-253.

Thibodeau, S. N., Bren, G., & Schaid, D. (1993). Microsatellite instability in cancer of the proximal colon. *Science*, *260*(5109), 816-819.

Thorson, A. G., Knezetic, J. A., & Lynch, H. T. (1999). A century of progress in hereditary nonpolyposis colorectal cancer (Lynch syndrome). *Diseases of the colon & rectum*, *42*(1), 1-9.

Vasen, H. F., Wijnen, J. T., Menko, F. H., Kleibeuker, J. H., Taal, B. G., Griffioen, G., ... & Khan, P. M. (1996). Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology*, *110*(4), 1020-1027.

Vasen, H. F., Stormorken, A., Menko, F. H., Nagengast, F. M., Kleibeuker, J. H.,Griffioen, G., & Wijnen, J. T. (2001). MSH2 mutation carriers are at higher risk of cancer than MLH1 mutation carriers: a study of hereditary nonpolyposis colorectal

cancer families.

Wang, J. Y., & Edelmann, W. (2006). Mismatch repair proteins as sensors of alkylation DNA damage. *Cancer cell*, *9*(6), 417-418.

Watson, P., & Lynch, H. T. (1993). Extracolonic cancer in hereditary nonpolyposis colorectal cancer. *Cancer*, *71*(3), 677-685.

Wheeler, J. M. D., Bodmer, W. F., & Mortensen, N. M. (2000). DNA mismatch repair genes and colorectal cancer. *Gut*, *47*(1), 148-153.

Wijnen, J., Khan, P. M., Vasen, H., van der Klift, H., Mulder, A., van Leeuwen-Cornelisse, I., ... & Fodde, R. (1997). Hereditary nonpolyposis colorectal cancer families not complying with the Amsterdam criteria show extremely low frequency of mismatch-repair-gene mutations. *The American Journal of Human Genetics*, *61*(2), 329-335.

Wiredu, E. K., & Armah, H. B. (2006). Cancer mortality patterns in Ghana: a 10-year review of autopsies and hospital mortality. *BMC public health*, *6*(1), 1-7.

Wu, Y., Berends, M. J., Mensink, R. G., Kempinga, C., Sijmons, R. H., van der Zee, A. G., ... & Hofstra, R. M. (1999). Association of hereditary nonpolyposis colorectal cancer–related tumors displaying low microsatellite instability with MSH6 germline mutations. *The American Journal of Human Genetics*, *65*(5), 1291-1298.

Xi, Y., & Xu, P. (2021). Global colorectal cancer burden in 2020 and projections to 2040. *Translational oncology*, *14*(10), 101174.

Yamagishi, H., Kuroda, H., Imai, Y., & Hiraishi, H. (2016). Molecular pathogenesis of sporadic colorectal cancers. *Chinese journal of cancer*, *35*(1), 1-8.

# APPENDICES

## APPENDIX I: CRC Questionnaire Form

**Title: Mut L Homologue 1 (MLH1) Exon 16 Mutations in Ghanaian CRC Patients Receiving Treatment at the Korle Bu Teaching Hospital**

1. Date of interview.....
2. Patient's ID…….. Sample ID number----------
3. Sex a. Male b. Female
4. Age.....
5. Educational Level.....
6. Religion a. Christian b. Muslim C. Traditionalist d. Atheist
7. Tribal group/Ethnicity----------
8. Occupation.....
9. Residence.....
10. Where is mostly your daily source of meal? a. Home b. Food Vender c. Restaurant
11. What is your favorite's food? ----------
12. What is your smoking status? a. Never b. Ex-smoker c. current smoker
13. Do you drink alcohol? a. No b. Yes
14. Do you know of anyone in your family with colorectal cancer? a. No b. Yes

    If yes describe the relationship. ----------
15. How long have you been living with the colorectal cancer disease? ----------

## APPENDIX II: Ethical Clearance –CHS

UNIVERSITY OF GHANA
COLLEGE OF HEALTH SCIENCES
ETHICAL AND PROTOCOL REVIEW COMMITTEE

Ref. No.: [illegible] EPRC/DEC/2021 December 1, 2021

Miss Agnes Sitsofe Dorbu
Department of Medical Biochemistry
University of Ghana Medical School
Korle Bu.

**ETHICAL CLEARANCE**

*Protocol Identification Number: CHS-Et/M. 3 – P 4.2/2021-2022*

**FWA: 00018577** **IORG: 0005170** **IRB: 00006220**

The College of Health Sciences Ethical and Protocol Review Committee (EPRC) on December 1 2021 reviewed and approved your research protocol.

Title of Protocol: **"Mut L Homolog 1 (MLH1) Exon 16 Mutations In Ghanaian Colorectal Cancer Patients Receiving Treatment at the Korle Bu Teaching Hospital"**

Principal Investigator: **Miss Agnes Sitsofe Dorbu**

This approval requires that you submit six-monthly review report(s) of the study to the Committee and a final full review report to the EPRC at the completion of the study. The Committee may observe, or cause to be observed, procedures and records of the study before, during and after implementation.

Please note that any significant modification(s) to this project/study must be submitted to the Committee for review and approval before its implementation.

You are required to report all serious adverse events related to this study to the EPRC within seven (7) days verbally and fourteen (14) days in writing.

As part of the review process, it is the Committee's duty to review the ethical aspects of any manuscript that may be produced from this study. You will therefore be required to furnish the Committee with any manuscript for publication.

**This ethical clearance is valid until December 1, 2022.**

Please always quote the protocol identification number in all future correspondence in relation to this protocol.

Signed: .....
**Professor Andrew Anthony Adjei**
Chair, Ethical and Protocol Review Committee

cc: Provost, CHS
Dean, UGMS
Head, Medical Biochemistry

• P. O. Box KB 52, Korle Bu, Accra, Ghana. Telephone: +233 (0) 302 665103/4 • Fax: +233 (0) 302 660762
• Email: admin.chs@ug.edu.gh / provost.chs@ug.edu.gh • Website: www.chs.ug.edu.gh