UNIVERSITY OF GHANA

COLLEGE OF BASIC AND APPLIED SCIENCES

CHARACTERIZING PATHOGENIC RISK FACTORS ASSOCIATED WITH BREAST CANCER IN GHANAIAN WOMEN

MICHELLE ABENA BUCKMAN

DEPARTMENT OF BIOCHEMISTRY, CELL AND MOLECULAR BIOLOGY

JULY 2019

UNIVERSITY OF GHANA

COLLEGE OF BASIC AND APPLIED SCIENCES

CHARACTERIZING PATHOGENIC RISK FACTORS ASSOCIATED WITH BREAST CANCER IN GHANAIAN WOMEN

BY

MICHELLE ABENA BUCKMAN

(10636933)

A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES IN PARTIAL FULFILLMENT FOR THE REQUIREMENT OF THE AWARD OF DEGREE OF MASTER OF PHILOSOPHY IN MOLECULAR CELL BIOLOGY OF INFECTIOUS DISEASES

DEPARTMENT OF BIOCHEMISTRY, CELL AND MOLECULAR BIOLOGY

JULY 2019

DECLARATION

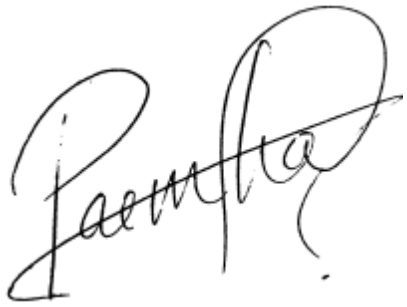
I, Michelle Abena Buckman, do hereby declare that, with the exception of cited references, this thesis is my original research work which was done under the supervision of Dr Lily Paemka in the Department of Biochemistry, Cell and Molecular Biology, University of Ghana.

1st October, 2020

…… ……

Michelle Abena Buckman Date

(Student)



1st October, 2020

…… ……

Dr Lily Paemka Date

(Supervisor)

ABSTRACT

Globally, the second leading cause of cancer death in women is breast cancer. It is responsible for 16% of all cancer cases and the most common cancer in Ghana. Breast cancer pathogenesis has not been fully elucidated but is however known to have a complex aetiology involving both genetic and environmental factors. Pathogens are one of the environmental factors that have been linked with breast cancer development. Epstein-Barr Virus (EBV), Human Papilloma Virus (HPV) and Mouse Mammary Tumour virus (MMTV) have been implicated in breast cancer. Research has showed that the local microbiome of the host could modulate breast cancer risk, yet it is unknown what microbes (pathogenic or probiotic) inhabit breast tumour tissues in Ghana.

The objectives of this study were to detect the presence of HPV, EBV, MMTV and bacteria in breast tumour tissues from Ghanaian women and determine the effect of bacteria on DNA in HeLa and MCF-7 cells.

Formalin-Fixed Paraffin-Embedded (FFPE) tissues and fresh breast cancer tissues were obtained from 204 breast cancer patients at the Department of Surgery, Korle-bu Teaching Hospital. Nucleic acid was extracted from the samples and amplified using standard Polymerase Chain Reaction (PCR) to detect the viruses. HPV-positive tumours were sequenced using the Sanger sequencing method. EBV typing of EBV positive samples was done by a nested PCR reaction using EBV-1 and EBV-2 specific primers. Bacteria from fresh breast cancer tissues were obtained and identified using basic biochemical tests and Sanger sequencing. The DNA damage potential of the isolates was tested on HeLa and MCF-7 cells using the histone-2AX phosphorylation assay.

HPV was detected in 27 (13.2%) of the 204 cases analyzed. EBV was identified in 66 (32.4%) of the samples. Out of the EBV positive cases, 2 (1%) were positive for EBV-1, 30 (14.7%) were

positive for EBV-2, 26 (12.7%) were positive for both EBV-1 and EBV-2 and 8 (3.9%) could not be genotyped using available methods. Co-infection of both HPV and EBV was detected in 11 (5.4%) of the cases. MMTV was not detected in any of the samples analyzed. Microbiome analysis showed relatively high evidence of bacteria belonging to the *Staphylococcus* and *Bacillus* species. *Staphylococcus sciuri*, *Staphylococcus epidermidis*, *Staphylococcus lugdunensis* obtained from breast tumours induced DNA double-strand breaks in MCF-7 cells.

The diversity and the probable role of the microbiome in breast cancer were also identified. Therefore, the presence of these pathogens showed probable involvement in breast cancer carcinogenesis.

DEDICATION

I dedicate this work to the Almighty God for being with me throughout this period and to my family for their love, care and support.

ACKNOWLEDGEMENTS

My sincerest gratitude goes to my supervisor, Dr Lily Paemka for her expertise, dedication, mentorship and supervision. I would also like to offer my special thanks to Dr Vincent Armah for his valuable and constructive suggestions for this research work. I'm particularly grateful to Mr Joshua Ayelazuno for his assistance with part of this work. I would also like to express my deepest appreciation to Mr Isawumi Abiola and Ms Molly Abban for their guidance and assistance in the laboratory and also with my data analysis.

Special thanks to the West African Centre for Cell Biology of Infectious Pathogens (WACCBIP) for funding my postgraduate studies and to the staff of the Department of Biochemistry, Cell and Molecular Biology for their significant role in my training. I very much appreciate the staff of the Department of Surgery, KBTH and my study participants for the provision of patients' data and clinical samples. I would also like to thank Dr Lawrence Edusei (Department of Pathology, KBTH) for his help during the sample collection.

Many thanks to Mrs. Keren Attiku, Mr. Emmanuel Tagoe, Mr. Jacob Donkor, Mr. Emmanuel Quaye and Mr. George Yankson for the diverse assistance they offered. I also wish to thank members of the Paemka lab (BCMB), Virology lab (BCMB) and Aikins lab (BCMB) for their support.

Finally, I would like to express my sincere thanks to my colleagues for their support, motivation and help during the entire course of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

EBV	Epstein-Barr virus
HPV	Human papilloma virus
MMTV	Mouse mammary tumour virus
BLV	Bovine leukaemia virus
H2AX	Histone-2AX
IDC-NOS	Invasive ductal carcinoma, not otherwise specified
IDC-NST	Invasive ductal carcinoma, no special type
ER	Estrogen receptor
PR	Progesterone receptor
HER2	Human epidermal growth factor 2
TNBC	Triple-negative breast cancer
DNA	Deoxyribonucleic acid
PCR	Polymerase chain reaction
FFPE	Formalin-fixed paraffin-embedded
RNA	Ribonucleic acid
EMB	Eosin-methylene blue
BHI	Brain heart infusion

MSA	Mannitol salt agar
TSI	Triple sugar iron
SIM	Sulphur indole motility
L1	Late protein 1
EBNA	Epstein-Barr nuclear antigen
ORF	Open reading frame
TNF	Tumour necrosis factor
LMP	Late membrane protein
IL	Interleukin
BRCA	Breast cancer gene
TP53	Tumour protein 53
MFI	Mean fluorescence intensity
MOI	Multiplicity of infection

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background

Cancer is a major health burden worldwide as a result of its high frequency of mortality and morbidity. The most frequent cancer globally among women is breast cancer with a high death rate of about 458,000 deaths per year. Breast cancer is a multi-factorial disease which acts in sequence or simultaneously to initiate and/or promote tumour growth (Ritchie *et al*., 2001). The origin of breast cancer is poorly understood but it has been associated with are risk factors such as obesity, gender, family history and consumption of alcohol (Pai *et al*., 2018).

Studies on some cancers have suggested the association of certain microbes with some cancers, indicating that these viruses and bacteria may persist in the specialized niche provided by the tumour microenvironment (Banerjee *et al*., 2018; Glenn *et al*., 2012; Naushad *et al*., 2017; Urbaniak *et al*., 2016; Xuan *et al*., 2014). Thus, 17.8% of cancers are linked to infectious agents and 12.1% of this number may be as a result of viral infections (Akhter *et al*., 2014). Whole genome sequencing of breast invasive breast cancers identified about 40 different viruses (Lawson & Glenn, 2017). However, the key viruses that have been linked to breast cancer are Bovine Leukaemia Virus (BLV), Epstein-Barr virus (EBV), Mouse Mammary Tumour Virus (MMTV) and Human Papilloma Virus (HPV).

HPVs are able to transform mammary epithelial cells (Dimri *et al*., 2005). According to Lawson and Glenn (2017), the E6 and E7 HPV oncoproteins in breast cancers have a low level of transcription, suggesting that HPV might have an indirect influence on breast cancer.

Cell-to-cell contact of the epithelial cells to the breast can lead to infection with EBV which then predisposes the epithelial cells to malignant transformation (Tsao *et al*., 2012). EBV targets cell cycle proteins and uses its viral proteins to take over the cellular pathways that regulate normal cellular processes (El-Naby *et al*., 2017).

Mouse Mammary tumour virus (MMTV) is the major cause of mammary tumours in mice. However, in the malignant epithelial cells of male and female human breast cancers, high MMTV-like sequences have been identified (Bindra *et al*., 2007). Other studies done by *in situ* PCR in human breast cancers also identified MMTV gene sequences compared to healthy controls which the *env* sequences were absent (Ford *et al*., 2004; Heng *et al*., 2009; Mazzanti *et al*., 2011). Since the link between oncogenic viruses and breast cancer tumorigenesis is arguable, it is of great interest to identify the presence and effect of these viruses in breast cancer development.

Numerous body sites are colonized by bacteria which are essential for human development. However, an alteration in the composition of the microbiota (dysbiosis) may promote disease progressions as seen in inflammatory bowel disease, diabetes, asthma and colorectal cancer (Estep *et al*., 2007; Mira-Pascual *et al*., 2015). These diseases have a microbiota that differs from that of healthy individuals. The most abundant phylum in the breast tissue is Proteobacteria (Urbaniak *et al*., 2016). Surprisingly, increased relative abundances of *Bacillus*, *Enterobacteriaceae* and *Staphylococcus* were identified in breast cancer tissues. *Staphylococcus epidermidis* and *E. coli* cultured from breast tumour samples in that study caused DNA double-strand breaks in HeLa cells (Urbaniak *et al*., 2016).

Since breast cancer studies in Ghana have not highlighted the roles of pathogens in breast cancer pathogenesis, the strains of bacteria and viruses inhabiting (microbiome) breast cancer tissue are

therefore unknown. In this study, the strains of bacteria and viruses in breast cancer tissues were identified to determine the relationship between pathogens and breast cancer risk.

1.2 Problem statement

Breast cancer aetiology involves both genetic and environmental factors (Sewani 2001; Lichtenstein, 2000). Infectious agents have been implicated as direct carcinogens or promoters of certain cancers. Therefore, the association of viral infections with different cancers has been extensively studied. Some of these oncogenic viruses have been linked with breast cancer development and/or progression (Amarante & Watanabe, 2009). The main viruses that have been implicated include EBV, high-risk HPV subtypes and MMTV (Alibek *et al*., 2013). However, the identification of these viruses in both malignant tissues is inconsistent. This is due to the prevalence of these viruses across different geographical locations and the variations in the detection methods for the viral genomes across the different studies that have been done on them.

Also, bacteria inhabit the body at numerous sites and a change in the composition of the microbiota may initiate disease progression (Xuan *et al*., 2014). Bacterial profiles have been shown to be distinctly different between healthy breast cancer tissues. Bacteria with the ability to damage DNA *in vitro* have been detected in women with breast cancer, which shows that the mammary microbiome might modulate breast cancer development (Urbaniak *et al*., 2012). However, the presence of these bacteria and viruses may vary due to different ethnic and geographical factors. Also, since there is no published study on the pathogens associated with breast cancer in Ghana, this study seeks to identify the specific viruses and bacteria in Ghanaian women with breast cancer.

1.3 Hypothesis

Pathogenic viruses and bacteria in mammary breast epithelial cells promote tumorigenesis and/or progression of breast cancer.

1.4 Aim

To identify and characterize bacteria and viruses found in breast cancer tissues of Ghanaian women.

1.5 Specific objectives

- To identify viruses (EBV, MMTV and HPV) and bacteria in breast cancer tissues of Ghanaian women.
- To investigate the DNA-damaging effects of identified bacteria in HeLa and MCF-7 cell lines.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Cancers

Globally, cancer is a key cause of mortality and morbidity (Ferlay *et al*., 2015). The disease burden in developing countries is greatly shifting from infectious to non-infectious diseases such as cancer (Bray *et al*., 2018). Lung cancer, which is the most commonly diagnosed cancer, is accountable for 11.6% of all cancer cases and is also the principal cause of cancer mortality (18.4%) worldwide. This is followed by female breast cancer (11.6%), colorectal cancer (10.2%) and prostate cancer (7.1%) (Bray *et al*., 2018).

2.2 Breast cancer

One of the three cancers that are most commonly diagnosed worldwide is female breast cancer and it's also the second leading cause of cancer death (Bray *et al*., 2018). Globally, it represents 29% of new cancers in women (Paul *et al*., 2017).

Breast cancer mortality in Africa occurs in over 50% of affected women (Torre *et al*., 2015). The risk of breast cancer mortality in African American women is higher than in women with breast cancer from other geographical regions (Satram-Hoang *et al*., 2019). It occurs frequently in African American women (Sadler *et al*., 2007). Breast cancer in Africans occurs in the premenopausal period which is contrary to the postmenopausal period in Caucasians (Adesunkanmi *et al*., 2006; Jiagge *et al*., 2016).

In Ghana, breast cancer is the commonest female cancer and accounts for about 16% of all cancers (Naku *et al*., 2016). In addition, it is the commonest cause of cancer death in Ghanaian women (Wiredu & Armah, 2006). The mean age for breast cancer in Ghana is 50.3 years which is lower

than that of the Caucasians which is 62 years (DeSantis *et al*., 2016; Edmund *et al*., 2013). In Ghana, the percentage of male breast cancer is 2.9% and this is higher than reported in other parts of the world which is 1% of all breast cancers (Korde *et al*., 2010; Quayson *et al*., 2014).

2.2.1 Types of breast cancer

Breast cancer is biologically and clinically grouped into histological grades and histological types (Ellis *et al*., 1992; Elston & Ellis, 2002; Gujar *et al*., 2018).

2.2.1.1 Histological subtypes

This type of classification is based on the morphological characteristics and the growth pattern of the tumours (Tao *et al*., 2015; Weigelt *et al*., 2010). Most cases of breast cancer occur in the ducts, some in the cells that line that lobules while a lesser percentage of cases begin in the other breast tissues (Sharma *et al*., 2010). It can be grouped into non-invasive (if cells in the breast do not break into neighbouring connective and fatty tissues and are confined in the duct) and invasive (if cells in the breast move into the neighbouring connective and fatty tissues and break through the walls of the lobules and the ducts) (Sharma *et al*., 2010).

The least occurring histological subtypes are infiltrating lobular carcinoma, followed by infiltrating ductal carcinoma, lobular carcinoma in situ and then ductal carcinoma in situ (Sharma *et al*., 2010). The most frequently occurring histological type is the invasive ductal carcinoma, of no special type (IDC-NST) or not otherwise specified (IDC-NOS) (Rakha *et al*., 2008). The less commonly occurring types are the mucinous carcinoma, tubular carcinoma and medullary carcinoma. Also, phyllodes tumour inflammatory breast cancer and Paget's disease are special types that are responsible for approximately 1% of breast cancer cases (Sharma *et al*., 2010).

2.2.1.2 Molecular subtypes

Classification can also be done according to the expression of the human epidermal growth factor receptor (HER2) oncogene, progesterone receptor (PR) and estrogen receptor (ER) (Tao *et al*., 2015). They are:

i. Luminal A breast cancer- This is HER2 negative, PR and/or ER positive and has reduced levels of Ki-67, which is a proliferation marker (Polyak, 2007; Varga *et al*., 2019)).

ii. Luminal B breast cancer- This is either HER2 positive or HER2 negative, PR and/or ER positive and has higher levels of Ki-67 (Inic *et al*., 2014; Polyak, 2007).

iii. HER2-enriched breast cancer- With this subtype, ER and PR are not expressed but HER2 is highly expressed (Allison, 2012; Nishimukai *et al*., 2015).

iv. Triple-negative breast cancer- This subtype is HER2, PR and ER negative (Allison, 2012; Voduc *et al*., 2010)

v. Normal-like breast cancer- This is identical to Luminal A but has a more severe prognosis (Feng *et al*., 2018)

The most prevalent molecular subtypes in Africa are the triple-negative and hormone receptor-negative breast cancers (Vanderpuye *et al*., 2017).

2.2.2 Aetiology of breast cancer

Breast cancer risk factors make it more likely for an individual to develop breast cancer. The main risk factor is being a woman and this is due to the continuous exposure of women to the growth-promoting abilities of estrogen and progesterone (Dumitrescu & Cotarla, 2005). Others include

genetics such as a family history of breast and cervical cancer, *TP53*, *BRCA1* and *BRCA2* gene mutations (Feng *et al*., 2018; Mann *et al*., 2006).

Hormonal risk factors include reproductive history, menopausal status (late menopause), menstrual history (early age at menarche), using oral contraceptives and exogenous hormones (hormone replacement therapy) (Bray *et al*., 2018; Sellers *et al*., 1992; Weiss *et al*., 2002).

Also, enhancing the local exposure of the breast tissue to high estrogen levels has been proposed to influence oncogenesis (Thompson *et al*., 2017; Yager & Davidson, 2006). However, breastfeeding and physical activity serve as protective factors against breast cancer development (Bray *et al*., 2018). In Ghana, breast cancer risk factors include geographical location and occupational settings (Armah & Gyeabour, 2013; Laryea *et al*., 2014). However, there is still limited knowledge about the relationship between specific aetiological factors and the role of geographic or temporal variations (Bray *et al*., 2018).

2.3 The Breast Microbiome

The quality and quantity of microbes in the human body play a very essential role in the well-being of the body (Francescone *et al*., 2014). There is a close relationship between the microbiota, which is the taxonomy and abundance of microbes in an environment and the host (Fernandez *et al*., 2018). Therefore, an imbalance in the host regulatory pathways responsible for homeostasis results in an increase in the risk for disease development which could also affect responses to cancer therapy and toxicity profiles of chemotherapeutics (Wallace *et al*., 2010).

Studies have shown that factors such as antibiotic usage, dietary cleanse and regimes, and changes in microbial composition due to travelling may affect the abundance of these microbes and not the presence and composition of the microbes found in a particular part of the body (Rajilic-Stojanovic

et al., 2012). A comparison of the human intestinal microbiota from children characterized by a rural diet and a western diet discovered a unique abundance of bacteria in each group, indicating the diverse microbial communities of both ancient and modern communities worldwide (De Filippo *et al*., 2010).

There are about six to eight openings of the human mammary duct at the surface of the nipple which serve as entry points for microbes from the skin, mouth and the environment to enter the breast (Going & Moffat, 2004). In addition, bacteria can be translocated from the mucous membrane through the gut epithelium and into the bloodstream through gap junctions located between the epithelial cells, uptake by dendritic cells or invasion of cells (Urbaniak *et al*., 2012).

2.4 Infectious Agents and Breast Cancer

Studies in the past 30 years have shown that nearly 15-20% of all cancers are linked to viral and bacterial infections (de Martel *et al*., 2012; Masrour-Roudsari & Ebrahimpour, 2017; Parkin, 2006; Pogorzelski *et al*., 2014). These infectious agents have been implicated either as promoters or direct carcinogens through the expression of oncoproteins that can lead to uncontrolled growth and immortalization and cells (Salman *et al*., 2017).

Most cancers caused by infectious agents are caused by viruses, for example, Epstein-Barr virus, Human Papilloma Virus, Human Immunodeficiency Virus, Herpes Simplex 2 Virus, Hepatitis C Virus and Hepatitis B virus. Recently, mechanisms by some selected bacteria such as *Helicobacter pylori* and parasites such as *Schistosoma haematobium* and *Opisthorchis viverrini* are becoming evident in cancer cases (Almeida *et al*., 2016; Haswell-Elkins *et al*., 1994; Roderfeld *et al*., 2018; Sriamporn *et al*., 2004; Sripa *et al*., 2012; Toller *et al*., 2011). Furthermore, studies have suggested that these viruses and bacteria may persist in the specialized niche provided by the tumour microenvironment (Banerjee *et al*., 2018; Glenn *et al*., 2012; Naushad *et al*., 2017; Urbaniak *et al*.,

2016; Xuan *et al*., 2014). Usually, the effect of a single pathogen on carcinogenesis is different from a community of microbes in carcinogenesis. An example is *H. pylori* a carcinogenic pathogen which causes fewer tumours in *H. pylori* mono-infected transgenic mice models (Parida & Sharma, 2019) The contribution of microbes to tumour growth has not been broadly studied and established in breast cancer as compared to the cellular and molecular biology of breast tumours.

These pathogenic microbes can trigger unrestrained adaptive immune responses, induce chronic inflammation, alter the balance of host cell proliferation and death thereby promoting malignancy (Chan *et al*., 2016). Nevertheless, the presence of certain microbes may be beneficial to the carcinogenesis process since they can promote antitumour immunity and immune surveillance and also increase estrogen levels (Xuan *et al*., 2014).

Oncogenic viruses have grown strategies for evading the immune system of their hosts and are therefore capable of establishing persistent long-term infections in their hosts (Kincaid & Sullivan, 2012). Host immunity, chronic inflammations and host cellular mutations are also involved in transforming the normal cells to malignant cells (Hodge *et al*., 2005). Also, the geographical distribution or location of the virus in the host is involved in the development of these cancers.

Studies on some small DNA tumour viruses such as adenoviruses, polyomaviruses and papillomaviruses have been helpful in elucidating the molecular mechanisms of transformation induced by viruses (McLaughlin-Drubin & Munger, 2008). Identification of tumour suppressor genes has also been accomplished through the study of these viruses to help in identifying mechanisms for controlling human mammalian cell growth (Shaw & Cantley, 2006).

2.5 Viruses and Breast Cancer

Viral aetiology is a recently hypothesized theory behind the pathophysiology of breast cancer (El-Naby *et al*., 2017). According to Lawson and Glenn, whole genome sequencing of invasive breast cancers led to the identification of 40 different viruses, but only 4 MMTV, BLV, EBV and HPV have probable oncogenic roles in breast cancer (Lawson & Glenn, 2017).

2.5.1 Human Papilloma Virus

2.5.1.1 Taxonomy, structure and biology

Human Papilloma Viruses (HPVs) are a group of small, non-enveloped double-stranded DNA viruses that are classified under the Papillomaviridae family of viruses (Bernard *et al*., 2010; Janicek & Averette, 2001). The DNA of HPV is circular, containing about 8 kbp genome with about 8 open reading frames (Graham, 2012). The genome encodes 6 early proteins and 2 viral structural late proteins (Doorbar & Gallimore, 1987). The early region consists of the E1, E2, E4, E5, E6 and E7 proteins (Figure 2.1). These code for transactivation of viral transcription, non-structural proteins responsible for viral replication and cellular proliferation and activation (Herbst *et al*., 2009). The L1 late protein encodes the major capsid protein and the L2 late protein encodes the minor capsid protein (Guan *et al*., 2017) (Figure 2.1). There is a noncoding region known as the upstream regulatory region which regulates cellular transcription, early gene transcription and viral amplification (Harari *et al*., 2014).

HPVs infect the squamous epithelia of organisms (De Villers, 2001). The most conserved gene within the HPV genome is the *L1* gene (De Villiers *et al*., 2004). Taxonomy is based on sequence identities in the *L1* nucleotide and also the topological position in the phylogenetic tree of Papillomaviruses (Bernard *et al*., 2010)

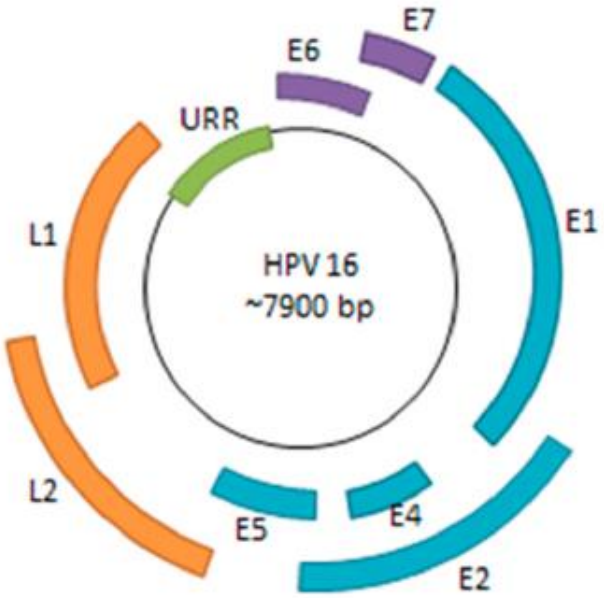


Figure 2.1: **The HPV Genome**. It is made up of six early genes (*E1, E2, E4, E5, E6* and *E7*) and two late genes (*L1* and *L2*) which encode the early and late proteins respectively (Shanmugasundaram & You, 2017).

2.5.1.2 Life cycle of HPV

The virus infects basal cells through the invasion of damaged epithelia. After entry, uncoating and transportation into the nucleus, the DNA of HPV is kept at a low-copy number in the basal cells (Moody & Laimins, 2010). Viral replication is aided by the HPV E6 and E7 proteins which inactivate p53 and retinoblastoma protein (pRb), respectively whilst the E1, E2, E4 and E5 proteins are involved in the transcription and production of non-structural proteins necessary for viral replication (Münger *et al*., 2004). The expression of the L1 and L2 HPV proteins leads to the initiation viral DNA encapsidation and finally the release of newly formed infectious virions, increasing the viral genome copy number (Doorbar & Gallimore, 1987, Pinidis *et al*., 2016; Reinson *et al*., 2015).

HPV DNA has the potential to establish latent infections within cells, which is characteristic of the virus. Papillomaviruses can persist asymptomatically or cause neoplasms when they infect the

epithelial membrane (Bernard *et al*., 2010). Most HPV infections resolve naturally in 1-2 years due to their asymptomatic nature (Serrano *et al*., 2018).

2.5.1.3 HPV subtypes

HPVs are characterized according to their tissue tropism (Tulay & Serakinci, 2016). An HPV 'type' is assigned when there is a greater than 10% dissimilarity of all known types in the nucleotide sequence of the *L1* open reading frame (ORF) (De Villiers *et al*., 2004). However, for isolates of the same HPV type, when nucleotide sequences of the *L1* gene are different by less than 10%, they are referred to as variants with each variant demonstrating various degrees of genomic diversity (Chen *et al*., 2011). GP5+/GP6+ and MY09/11 PCR primers which target extremely conserved regions within the *L1* ORF are used for HPV identification (Harari *et al*., 2014). PCR amplicons from these PCR assays can be sequenced to aid in alignment to known HPV types which facilitates classification of genotypes, identify new species, types and variant lineages (Bernard *et al*., 2010; Burk *et al*., 2013; Qu *et al*., 1997). Phylogenetic analysis of HPV suggests host cell tropism, carcinogenic risk and its associated pathology (Schiffman *et al*., 2005).

HPVs can also be further divided into five main genera which are; *Alphapapillomavirus*, *Betapapillomavirus*, *Gammapapillomavirus*, *Nupapillomavirus* and *Mupapillomavirus* according to their disease associations (Bernard *et al*., 2010; Schiffman *et al*., 2005). About 150 genotypes of HPV have been characterized (Tsakogiannis *et al*., 2015). Depending on the malignant ability of the virus and its potential to promote the proliferation of infected cells, it can be divided epidemiologically into those with low and high-risk oncogenic potentials (Chen *et al*., 2011; Muñoz *et al*., 2003; Wang *et al*., 2017). The low-risk HPVs are HPV 6, 11, 40, 42, 43, 44, 54, 61, 62, 70, 71, 72, 81, 83, 84 and 89 (Kaspersen *et al*., 2011; Muñoz *et al*., 2003). The high-risk HPVs include HPV 16, 18, 31, 35, 39, 45, 51, 52, 56, 58, 59 and 68 (Schiffman *et al*., 2005; Tulay &

Serakinci, 2016). One difference between low and high-risk HPVs is that low-risk HPVs do not stimulate cell proliferation while high-risk types stimulate cell proliferation (Kranjec *et al*., 2017). The carcinogenic HPV types are 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 66 (Cogliano *et al*., 2005).

2.5.1.4 Epidemiology of HPV

The estimated global prevalence of HPV is 11.7% (Bruni *et al*., 2010). However, there are variations in this prevalence according to studies done on a regional basis showing a relatively higher prevalence in Africa and Oceania (Serrano *et al*., 2018). HPV infection worldwide is higher in women less than 25 years and then declines in older women (Franceschi *et al*., 2006). The most frequent oncogenic type observed globally is HPV 16, then HPV 18, HPV 52, HPV 31 and HPV 58 (de Martel *et al*., 2012). HPV 16 can be found in more than 50% of all cancers (Middleton *et al*., 2003).

The frequency of HPV in Ghana was found to be 10.7% with a higher prevalence in more advanced age groups (Domfeh *et al*., 2008). Also, in Ghana, HPV DNA was identified in 98% of 50 women diagnosed with cervical cancer (Obiri-Yeboah *et al*., 2017). As stated by Awua *et al*. (2016), the predominant high-risk HPV genotypes in Ghana are HPV 18, HPV59 and HPV 45.

2.5.1.5 HPV and Breast Cancer

About 90% of HPV infections are typically asymptomatic and the immune system eliminates them within two years. Malignancy occurs during a sustained HPV infection in the presence of appropriate risk factors (Malekpour *et al*., 2018). A frequency of 2.3% different high-risk HPV subtypes was found in 855 breast cancer cases by using RNA-seq data from The Cancer Genome Atlas (TCGA) (Lawson *et al*., 2015). A study on Iranian women with breast cancer identified HPV DNA in 48.6% of breast cancer patients and showed a link between HPV infection and breast

cancer (Khodabandehlou *et al*., 2019). Using PCR for genotyping breast carcinoma samples, HPV 16 was found in 29.4% of the cases (Sigaroodi *et al*., 2012). Functional studies have also revealed that HPV -16 and -18 are able to immortalize and change the proliferative nature of human breast epithelial cells (Dimri *et al*., 2005).

The incorporation of HPV 16 into the genome of the host leads to the constitutive expression of E6 oncoproteins and the extensive propagation of infected epithelial cells (Wang *et al*., 2016). According to Yasmeen *et al*. (2007), induction of cell invasion and metastasis in MCF-7 and BT-20 non-invasive breast cancer cells can be attributed to HPV 16 E6 and E7 oncoproteins.

A study showed a decreased expression of *p53, pRB, BRCA1* and *BCRA2* in breast cancer patients with HPV infection as compared to breast cancer patients without HPV infection and healthy controls (Khodabandehlou *et al*., 2019). HPV E6 and E7 oncoproteins interrupt the cell cycle by targeting p53 and RB proteins which control the cell cycle to initiate malignancy and further lead to the development of tumours (Moody & Laimins, 2010). Also, BRCA1 and BRCA2 proteins are expressed to repair DNA damage in cells (Zhang *et al*., 2009). HPV E6 and E7 proteins act as antagonists to BRCA1 which also interacts with p53-mediated transcription by acting as a co-activator and aids in RB function in the G1-checkpoint of the cell cycle (Zhang *et al*., 2005). This interaction leads to an alteration in the activity of BRCA1. Thus, a decreased expression of these genes could suggest a mechanism for HPV-mediated breast cancer carcinogenesis. The expression of reactive oxygen and nitrogen species, IL-1, IL-6, IL-17, TGF-β, NF-κB and TNF-α inflammatory cytokines showed higher levels in breast cancers positive for HPV in comparison with breast cancers negative for HPV and healthy controls, indicating that the inflammation may be related to viral infections (Khodabandehlou *et al*., 2019).

A microarray technique showed that HPV positive cervical samples are more likely to have tumours with positive progesterone receptors (Kouloura *et al*., 2018). Yet, the identification of HPV DNA in other studies have been unsuccessful (Kouloura *et al*., 2018).

2.5.2 Epstein-Barr virus

2.5.2.1 Taxonomy, structure and biology of EBV

The Epstein-Barr virus is a double-stranded γ-herpes virus that infects B cells and epithelial cells (Matsuura *et al*., 2010). *In vivo* studies have proven the association of EBV in malignancies involving the two cell types. The virus latently infects B-lymphocytes and periodically reactivates its lytic replication (Matsuura *et al*., 2010). The EBV genome is a double-stranded, linear molecule about 172 kb in size and encodes about 85 genes. The ORF of EBV is divided into latent and lytic genes, that are translated into proteins (Djavadian *et al*., 2018). The coding capacity of the viral genome is as a result of the unique long and short sequence domains that are represented as internal repeat sequences. In addition, the genome of the virus also contains a succession of 0.5 kb terminal direct repeats which help identify if the EBV-infected cells are from the same progenitor (Niedobitek *et al*., 2001).

Upon infection of a cell by EBV, the viral DNA circularizes and per the number of identical sequences that repeat in the genome of the parent, the viral DNA circularizes with an introduction of some variations during replication. However, for latent infections, impending generations will have episomes with the same number of identical repeat sequences (Thompson & Kurzrock, 2004). EBV produces six nuclear antigen proteins (EBNA-1, EBNA-2, EBNA-3A, EBNA-3B, EBNA-3C and EBNA-LP) and three latent membrane proteins (LMP1, LMP2A and LMP2B) which are expressed during different stages of latency as shown in figure 2.2 (Middeldorp *et al*., 2003).

The *LMP1*, *EBNA2* and *EBNA3* EBV oncogenes prevent cell death and facilitate cell division (van Beek *et al*., 2003). Lytic infection in EBV-infected epithelial cells could be turned on by cellular stress which includes hypoxia and differentiation. The EBV viral proteins are responsible for mimicking a number of anti-apoptotic, growth and transcription factors that take control of cellular pathways and control various homeostatic cellular functions (Thompson & Kurzrock, 2004).

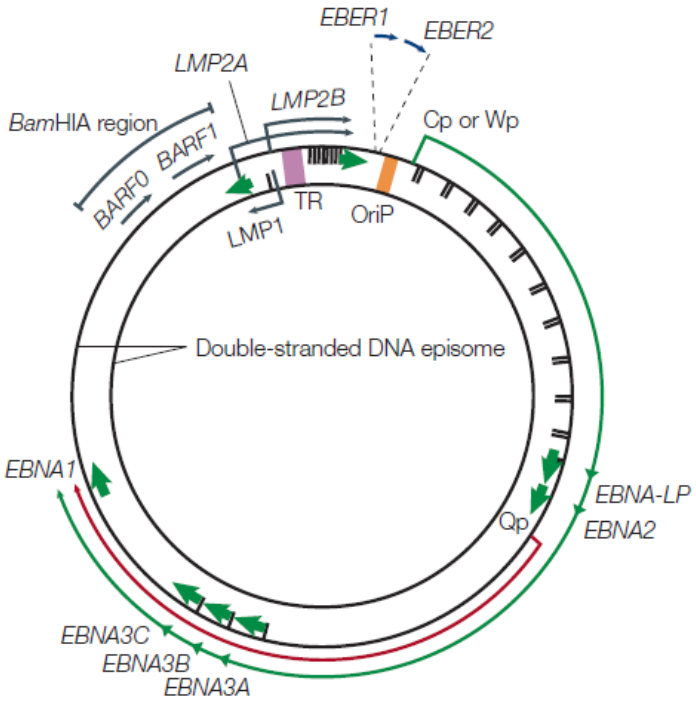


Figure 2.2: **The EBV genome.** The EBV genome indicating the location and transcription of the latent genes (Young & Rickinson, 2004).

2.5.2.2 Life cycle of EBV

EBV infects resting B cells resulting in the unregulated polyclonal expansion of latently infected lymphoblasts. Regulation of the EBV nuclear antigen 2 (EBNA-2) transcription factor leads to the expression of nine latent EBV proteins (Thorley-Lawson & Gross, 2004). The resting memory B cells of healthy carriers do not express the EBV viral proteins even though they are latently infected with the virus (Hochberg *et al*., 2004).

EBV enters the Waldeyer tonsillar ring in the oropharynx via the saliva and starts a lytic infection which brings about viral amplification (McKenzie & El-Guindy, 2015). These then infect naïve B cells in the underlying lymphoid tissues and are converted into activated lymphoblasts with the help of the growth transcription program (latency III). Three latency proteins in this stage (EBNA-3A, EBNA-3B and EBNA-3C) negatively autoregulate the growth program. The cells then migrate to the follicles to begin a reaction in the germinal centre and create the default transcription program (latency II) which sends rescue or survival signals to enable the cell to exit the germinal centre as a memory B cell. The virus is then able to persist for extended periods because memory B cells seldom die and the immune system cannot easily detect them since they do not express viral proteins. Also, the growth-promoting genes are not expressed at this stage and therefore do not pose a risk to the host (Thorley-Lawson & Gross, 2004).

However, EBNA-1 protein expression as a result of the division of latently infected memory cells leads to the expression of the which results in the replication of the DNA of the virus. This is controlled by the cell as part of the usual memory B-cell homeostasis mechanisms because the growth-promoting latent proteins are not expressed at this stage (Hochberg *et al*., 2004).

EBV causes a latent infection in circulating B lymphocytes which if not controlled subclinically, can lead to malignant tumours under favourable conditions (Crawford, 2001). EBV usually causes tumour outgrowth and fatal lymphoproliferative diseases in immunocompromised individuals (Matsuura *et al*., 2010). *In vitro* infection of human B cells with EBV leads to their immortalization, hence reflecting the oncogenic potential of EBV (Frisan *et al*., 2001).

2.5.2.3 EBV Subtypes

EBV-1 and EBV-2 are the two main subtypes of EBV which are known to cause infections in humans (Stanfield & Luftig, 2017). They are distinguished by the organization of the *ENBA-2,*

EBNA-3a, *EBNA-3b* and *EBNA-3c* genes that code for the EBV nuclear antigens (Dolan *et al*., 2006). Also, *in vitro* studies have shown that EBV-1 is able to transform B cells more efficiently than EBV-2 (Thompson & Kurzrock, 2004). The EBV-1 strains have a higher prevalence in Asia, Europe and North America while EBV-2 strains are more predominant in Africa, Papua New Guinea and Alaska (Chen *et al*., 1992; Janani *et al*., 2015; Sixbey *et al*., 1989). EBV-2 and coinfection with both subtypes are usually seen in immunocompromised individuals. This suggests that the reactivation or persistence of EBV subtypes in individuals might be controlled by immunity or infection with EBV-2 occurs when an individual is in an immunocompromised state (Chang *et al*., 2009).

2.5.2.4 Epidemiology of EBV

About 90% of the world's populations are believed to be carriers of the virus (Perkins *et al*., 2006). Primary infection with EBV occurs more in children and is generally asymptomatic. However, in adults, an infection could lead to infectious mononucleosis (Santpere *et al*., 2014). EBV is usually associated with malignancies such as nasopharyngeal carcinoma (NPC), post-transplant lymphomas, Burkitt lymphoma and Hodgkin's disease (Pai *et al*., 2018).

In Ghana, EBV DNA was identified in 25% of NPC biopsies using PCR (Asante *et al*., 2017). Also, a study in Ghanaian HIV-AIDS patients showed that the seroprevalence of EBV in these individuals was 87.2% (Adjei *et al*., 2008).

2.5.2.5 EBV and breast cancer

EBV can immortalize human B lymphocytes in culture and this explains its ability to cause human diseases, particularly autoimmune diseases and cancer (Atkinson & Samter, 2001; Rickinson, 2006; Rickinson, 2007).

The first association between infection with EBV and breast cancer was identified in 1995 where EBV DNA was found in 21% of primary breast carcinomas using PCR and of the PCR-positive samples, RNA was detected in 31.5% of them using in situ hybridization (Labrecque *et al*., 1995). Also, a study in primary invasive breast cancers (PIBC) identified EBV in 45% and 28% of Egyptian and Iraqi women, respectively, indicating that EBV might promote the development of PIBCs or increase tumour aggressiveness in these patients (Zekri *et al*., 2012).

In addition to these, different studies have also explored EBV presence in different types of breast cancer among different groups of people in different geographical regions using different methods of detection such as PCR (Bensaber *et al*., 2017; El-Naby *et al*., 2017; Marrao *et al*., 2014; Mazouni *et al*., 2011; Murray *et al*., 2003; Perrigoue *et al*., 2005; Thorne *et al*., 2005).

A study indicated that infection with EBV predisposes mammary cells to malignant transformation, but the virus is no longer necessary once malignant transformation has occurred which shows a probable involvement of EBV in breast cancer aetiology (Hu *et al*., 2016).

2.5.3 Mouse Mammary Tumour Virus

2.5.3.1 Taxonomy, structure and biology

MMTV/HMTV is a β-retrovirus that induces mammary tumours in mice. It is an RNA molecule with a genome size of approximately 9 kb which is flanked by 5' and 3' long terminal repeats (LTR) and encodes 7 genes (Ross, 2010). The genome of the virus is made up of two matching strands of RNA which encode the capsid/nucleocapsid Gag proteins used for replication of the genome and the env envelope proteins that bind to the surface receptors of the cell needed for viral entry into cells (Akhter *et al*., 2014).

2.5.3.2 Life cycle of MMTV

The virus infects lymphocytes, where it is expressed and can cause mammary tumours (Czarneski *et al*., 2003; Ross, 2000).

MMTV infection of host cells is done by binding its env glycoprotein to a cellular receptor known as transferrin receptor 1 (TfR1) (Ross *et al*., 2002). The virus is endocytosed to a compartment with a low pH where the viral membrane fuses with the host and leads to the entry of the capsids to the cytoplasm. The virus replicates by reverse-transcription of the RNA genome and forms a double-stranded DNA which migrates to the nucleus. The provirus then fuses with the cellular genome where the viral genes are transcribed and sent to the cytoplasm. The viral proteins are translated and the RNA genome is packaged. This then buds out of the infected cell to continue the infection process as shown in figure 2.3 (Dudley *et al*., 2016).

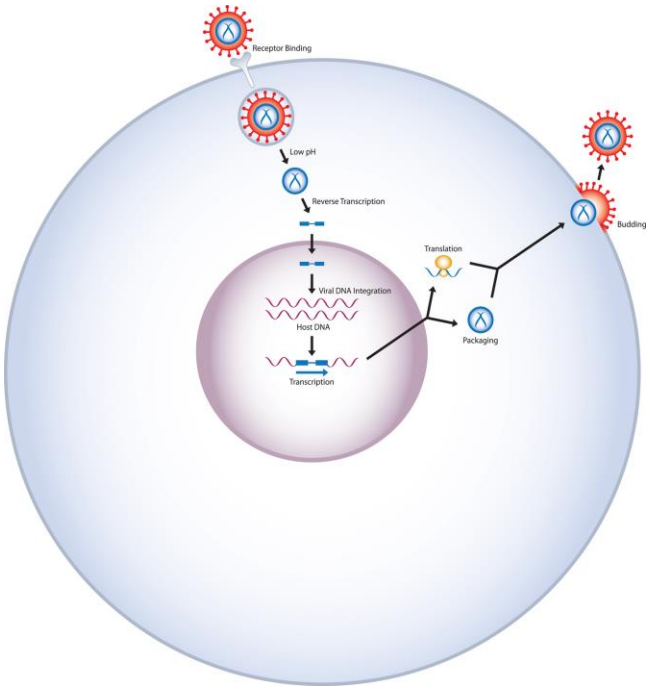


Figure 2.3: **MMTV infection.** The virus binds to transferrin receptor 1 (TfR1) and is internalized. The viral genome is reverse transcribed and integrated into the genome (Dudley *et al*., 2016)

2.5.3.3 Epidemiology of MMTV

Studies have suggested that the free-roaming wild house mouse known as *Mus Domesticus* is the probable reservoir for the human tropic strain of MMTV (Faedo *et al*., 2007; Hsu *et al*., 2010). MMTV is acquired through the milk of infected females which is then passed on to suckling pups (Dudley *et al*., 2016). It is responsible for adenocarcinomas of adult mammary epithelial tissues (Lawson & Glenn, 2017).

There is limited information on the epidemiology of MMTV because the house mice of the genus *Mus* are usually found in Asian, European and North African countries (Stewart *et al*., 2000). A study to identify the prevalence of MMTV in free-roaming, wild house mice in Australia showed positive PCR results for 97% of the mice (Faedo *et al*., 2007).

2.5.3.4 MMTV and Breast Cancer

Env is an envelope protein of MMTV which is absent in normal tissues and present in breast cancer tissues (Pogo *et al*., 2010). Its expression leads to morphological changes with expression consistent with transformation with normal mouse and human mammary epithelial cells *in vitro* (Katz *et al*., 2005). The initial investigation of MMTV in human breast cancers identified MMTV *env* gene-like sequences in 38-40% of human breast cancers (Wang *et al*., 1995). Further reports have shown that in human breast cancers, MMTV-like sequences with at least 95% identity to MMTV are highly expressed (Etkind *et al*., 2000; Melana *et al*., 2007; Nartey *et al*., 2017; Wang *et al*., 2001). Research on primary cell cultures has also identified these MMTV-like sequences in viral particles obtained from the cell cultures (Indik *et al*., 2007). This is due to the similarities between the morphologies of MMTV-positive mouse mammary tumours and MMTV-like positive breast cancer (Lawson & Heng, 2010).

A human to human transmission of the virus was suggested after identifying MMTV-like viral sequences in the saliva of 57% of adults with breast cancer, 11% healthy adults and 27% healthy children (Mazzanti *et al*., 2011).

Australian researchers identified MMTV-like sequences in breast tissues from Australian women preceding the development of breast cancer positive for MMTV (Nartey *et al*., 2017). Also, *in situ* PCR of FFPE tissues detected RNA transcripts from the *env* gene in cancerous epithelial cells in 80% of gynecomastia, 78% of DNA-positive tumours and none in normal tissues (Ford *et al*., 2004).

2.6 Bacterial Communities in the Breast

The human breast tissue is not sterile but is made up of a very distinct and diverse community of bacteria which is unique from all other body parts (Donnet-Hughes *et al*., 2010; Urbaniak *et al*., 2014). The breast is made up of glandular and lobular nodules, therefore, the resident bacteria in the ducts and breast milk might differ from those in the breast tissues due to the different nutrients that may be present in those different sites. This is because the nutritional requirements for each species of bacteria differ along with its preferred niche (Pereira & Berry, 2017).

Urbaniak *et al*. (2012) hypothesized that the existence of microbes in breast tissues may be due to the existence of bacteria in human milk. Also, bacteria with potential pathogenic properties are able to process fats that are present in abundance in the human breast (Xuan *et al*., 2014).

A concurrent collection of breast tissue, skin tissue and skin swab samples obtained under aseptic conditions during surgery showed that surgically sterile breast tissues have a distinct microbiome which is different from the overlying skin (Hieken *et al*., 2016). One study of the breast

microbiome by Urbaniak *et al*. identified the predominant phyla in the breast as Proteobacteria and Firmicutes (Urbaniak *et al*., 2014).

2.7 Bacteria and breast cancer

Bacteria that can be found in and around tumour sites are described as part of the tumour microenvironment. The microbiome has been suggested to colonize tumours and can affect all the stages of cancer development (Bashiardes *et al*., 2017). In the escape phase of cancer development, T regulatory (Treg) and myeloid suppressor cells are recruited to the tumour site to contribute to local immune suppression (Bashiardes *et al*., 2017). At this stage, the microbiome may regulate Treg induction and also the activity of tumour-infiltrating myeloid cells (Bashiardes *et al*., 2017). Other ways by which bacteria are able to colonize tumours are through the leakage of bacteria into the tissue due to angiogenesis and the recognition of cancer-specific glycosylation moieties (Abed *et al*., 2017; Baban *et al*., 2010).

2.7.1 Bacteria in Breast Tumours

A comprehensive analysis of the breast cancer microbiota using next-generation sequencing identified the predominant phylum of bacteria as Proteobacteria (Thompson *et al*., 2017; Urbaniak *et al*., 2016).

Another study also found that bacteria belonging to the family *Alcaligenaceae* are present at a higher abundance in breast cancer tissues compared to non-cancer tissues (Wang *et al*., 2017). This study additionally detected augmented levels of gram-positive bacteria in the urinary microbiome of cancer patients. (Wang *et al*., 2017).

A study to identify some enriched microbial biomarkers in Chinese malignant breast cancer tumours using Illumina HiSeq system to sequence the V1-V2 region of the *16SrRNA* gene found

some ethno-specific bacteria belonging to the genus *Propionodoinas* and bacteria belonging to the family *Micrococcaceae*, *Caulobacteriaceae*, *Rhodobacteriaceae*, *Norcardioidaceae* and *Methylobacteriaceae* (Meng *et al*., 2018).

Analysis of the gene expression and breast cancer microbiota identified *Listeria spp*. to be more enriched in gene correlations associated with epithelial to mesenchymal transitions (Thompson *et al*., 2017). *In vitro* studies demonstrated that epithelial to mesenchymal transitions (EMTs) contribute to breast tumour metastasis (Wu *et al*., 2016). Therefore, modulation of this species of bacteria might result in a decreased expression of some genes involved in the EMT pathway.

Also, investigating the microbial signatures of different breast cancer types using a pan-pathogen array which detects low copy number and fragmented genomes of FFPE samples indicated the presence of diverse genera of bacteria for each breast cancer subtype (Banerjee *et al*., 2018). However, in all the subtypes analyzed (triple-negative, triple-positive, PR positive and ER positive), dominant signatures were observed first for Proteobacteria and then Firmicutes as reported (Hieken *et al*., 2016; Thompson *et al*., 2017; Urbaniak *et al*., 2014) specifically, *Brevundimonas*, *Mobiluncus* and *Actinomyces* (Banerjee *et al*., 2018). *Brevundimonas* acts as an opportunistic pathogen in immunocompromised individuals to cause nosocomial infections (Ryan & Pembroke, 2018). This bacterium has been reported to cause bacteremia in cancer patients (Han & Andrade, 2005; Lee *et al*., 2011). *Mobiluncus* species are associated with bacterial vaginosis and are potentially pathogenic since it has been isolated from breast abscesses (Edmiston *et al*., 1990; Hill *et al*., 1998; Sturm, 1989; Weinbren *et al*., 1986).

A further comparison of the differences in the microbiome of different histological grades of malignant tissues identified an increase in the genus *Agrococcus* during malignancy (Meng *et al*., 2018). Another study also discovered that women with malignant breast cancer have different

microbiome from those with benign disease (Hieken *et al*., 2016). In malignancy, bacteria belonging to the genus *Fusobacterium*, *Atopium*, *Gluconacterobacter*, *Hydrogenophaga* and *LactoBacillus* were significantly enriched (Hieken *et al*., 2016).

One study to compare the microbiome of breast cancer patients at different body sites relative to non-cancer controls showed that individuals with cancer have a decreased abundance of the genus *Methylobacterium* in which was also observed in tumours with a greater invasive potential (Wang *et al*., 2017). Contrary to this, an earlier study by Xuan *et al*. (2014) found a relatively high enrichment of the bacterium *Methylobacterium radiotolerans* in breast cancer tissues while the bacteria *Sphingomonas yanoikuyae* was highly enriched in the paired normal tissues. *Sphingomonas yanoikuyae* is able to degrade bis(4-hydroxyphenyl) methane (bisphenol F), an aromatic hydrocarbon and use it as its only carbon source (Inoue *et al*., 2008). Exposure of polyaromatic hydrocarbons (PAH) can be through contact with organic materials burning of wood, tobacco smoke and eating smoked and grilled foods which can be subsequently absorbed through the dermis (Korsh *et al*., 2015). *In vitro* studies and animal studies have proven that breast cancer can be induced by exposure to PAH (El-Bayoumy *et al*., 1995; Morris & Seifter, 1992).

A number of bacteria such as *Acinetobacter radioresistens*, *Actinomyces sp HPA0247*, *Citrobacter koseri*, *E. coli*, *Enterococcus gallinarum*, *Erwinia amylovora*, *Salmonella enterica*, *Shewallena putrefaciens* and *Fusobacterium nucleatum* have been identified in postmenopausal breast cancer women (Zhu *et al*., 2018). Comparing the faecal microbiota of newly diagnosed premenopausal women with breast cancer and controls revealed that the gut microbiota may modulate the risk of breast cancer through methylo-independent pathways (Goedert *et al*., 2015). However, microbes can also play a beneficial role in tumorigenesis by modulating estrogen levels or by promoting antitumour immunity and immune surveillance (Xuan *et al*., 2014). A case-control study which

compared the faecal microbiota of postmenopausal women awaiting treatment for biopsy-proven breast cancer with similar women without breast cancer as controls found a less diverse and compositionally different faecal microbiota in the breast cancer patients (Goedert *et al*., 2015). Altered gut microbiome may modulate breast cancer development and/or progression due to a relatively low abundance of *Roseburia inulinivorans* which is a butyrate-producing bacterium (Zhu *et al*., 2018). In intestinal epithelial cells, the activation of nuclear factor κB (NF-κB) is prevented by butyrate, which is an anti-inflammatory agent (Inan *et al*., 2000).

This is, however, a protective mechanism for the host which ensures direct and continuous contact with the immune system. A comparison of the nipple aspirate fluid from ductal carcinoma breast cancer survivors and healthy controls by sequencing the V4 hypervariable region of the *16SrRNA* gene revealed the presence of the genus *Alistepes* in all the breast cancer samples and complete absence of this bacteria in the healthy controls (Chan *et al*., 2016). A review on the role of colonic microbiota in colorectal carcinogenesis found *Alistepes* to be one of the highly augmented bacteria in tumorigenesis (Baxter *et al*., 2014).

2.7.2 Bacteria in Non-Cancerous Adjacent Tumors

Different stages of breast cancer progression have similar bacterial profiles using normal adjacent tissues of breast cancer patients (Urbaniak *et al*., 2016). Analysis of non-cancerous adjacent tissues identified *Actinobacteria* as the most abundant phylum (Thompson *et al*., 2017). However, a previous investigation of bacterial taxa in normal adjacent cancer tissues discovered a higher abundance of *Bacillus, Staphylococcus, Enterobacteriaceae, Comamondaceae* and *Bacteroidetes*, none of which belong to the Actinobacteria phylum (Urbaniak *et al*., 2016).

About 85-90% of all *Enterobacteriaceae* found in clinical specimen are known to be *E. coli*, *Klebsiella pneunomeae* or *Proteus mirables* (Farmer *et al*., 1985). *E. coli* isolates have been shown

to induce DNA double-strand breaks in cancer cells (Cuevas-Ramos *et al*., 2010; Urbaniak *et al*., 2016). One of the most common phyla that can be found in individuals with early stage oesophagal squamous cell carcinoma (ESCC) is Bacteroidetes, during a dysbiosis in the oesophagal microbiota (Nasrollahzadeh *et al*., 2015; Patel *et al*., 2016).

Similar microbiota has been found in cancerous and surrounding normal tissues, indicating that microbial dysbiosis is a precursor for carcinogenesis as it establishes a microenvironment which makes the body susceptible to cancer (Costantini *et al*., 2018).

2.8 Proposed Mechanisms for Breast Microbes Pathogenicity

Understanding the pathogenic nature of these microbes has led to the proposal of several mechanisms by which these microbes might lead to carcinogenesis.

2.8.1 Regulation of Chronic Inflammation and Immunity

Some bacteria produce high molecular weight exopolysaccharides called capsules which they use to evade immune clearance, thereby serving as a source of protection from the host inflammatory response which involves complement activation and phagocyte mediated killing (Wilson *et al*., 2002). Gram-positive bacteria are known to induce more Interleukin-12 (IL-12) which is known to activate cytotoxicity and Interferon-γ (IFN-γ) secretion by Natural Killer (NK) cells and T cells (Hessle *et al*., 2000). IFN-γ is a probable therapeutic tool in breast cancer because, in the early stages of tumour development, tumour cells are able to escape recognition by IFN-γ (Garcia-Tunon *et al*., 2007).

The transcription factor nuclear factor NF-κB mediates the host inflammatory response to pathogenic bacteria and other stress signals which coordinates several aspects of immune function required for resistance (O'Hara & Shanahan, 2006; Tato & Hunter, 2002). Innate immune

responses to intestinal bacteria which require the activation of tumour necrosis factor TNF-α cause breast cancer in Rag2 deficient Min mice lacking lymphocytes (Rao *et al*., 2006). The expression of TNF is associated with a wide range of tumours including breast cancer. It influences the survival, growth, proliferation, differentiation and movement of the tumour and stromal cells (Balkwill, 2005). A number of animal cancer studies have shown that TNF produced in the tumour microenvironment may promote the spread and development of cancer (Knight *et al*., 2000; Komori *et al*., 1993; Selinsky & Howell, 2000).

2.8.2 Metabolic Function

Studies involving metabolic pathways in breast cancer primarily involve progesterone and estrogen metabolism, metabolism of cysteine and methionine, C5-branched dibasic acid metabolism, fatty acid biosynthesis and metabolism and glycerol transferases (Cavuoto & Fenech, 2012; Fernandez *et al*., 2018). Bacterial strains such as *Escherichia* and *Clostridium* have been linked with an increase in the risk of breast cancer by their abilities to increase circulating levels of estrogen by deconjugating sulphonated estrogens through the activities of β-glucuronidases (Kwa *et al*., 2016).

2.8.3 DNA Damage and Genomic Stability

Most pathogenic bacteria produce genotoxins and induce inflammatory responses in host cells (Thompson *et al*., 2017). Therefore, pathogenic strains of *E. coli* can produce a genotoxin known as colibactin which encodes a *pks* genomic pathogenicity island. The colibactin toxin is a known inducer of DNA double-strand breaks, chromosomal aberrations and cell-cycle arrest in the G2/M phase (Nougayrede *et al*., 2006). A double-strand break is the most detrimental damage to the DNA that if left unrepaired can lead to genomic instability and carcinogenesis (Liu *et al*., 2017). Also, in the double-strand break repair pathway, genomic instability and higher mutation rates can

favour the survival of cancer cells and also make them susceptible to stress due to genotoxins (Srivastava & Raghavan, 2015). Also, *E. coli* strains which express colibactin promote carcinogenesis by altering the sumoylation of p53 which induces senescence in the host cells and produces tumorigenic growth factors (Dalmasso *et al*., 2014).

Studies have shown that *Klebsiella pneumoniae, Enterobacter aerogenes* and *Citrobacter koseri* which are members of the Enterobacteriaceae family also harbour the *pks* pathogenicity island (Fais *et al*., 2018). This was confirmed by an assessment of the DNA damage ability of cultured isolates of normal adjacent breast cancer tissues which showed that in HeLa cells, some *E. coli* and *Staphylococcus epidermidis* isolates are able to induce DNA double-strand breaks (Urbaniak *et al*., 2016).

However, all these studies produce extensive, but no definitive evidence for the involvement of pathogens in breast cancer disease development and progression; suggesting an association between viral DNA, the breast microbiome and the breast cancer subtypes (Corbex *et al*., 2014; Naushad *et al*., 2017; Urbaniak *et al*., 2016).

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Materials

Blood agar base (Oxoid, USA), eosin-methylene blue agar (Oxoid, USA), brain heart infusion agar (Oxoid, USA), MacConkey agar (Oxoid, USA), mannitol salt agar (Oxoid, USA), LB broth (Lennox), Simmons citrate agar (Oxoid, USA), Sulfide indole motility agar (Oxoid, USA), Triple sugar iron agar (Oxoid, USA), PCR primers (IDT, USA), Choice Taq DNA polymerase (Denville Scientific, USA), Choice Taq master mix (Denville Scientific, USA), Ethidium bromide (Sigma Aldrich), 50X TAE buffer (Invitrogen, USA), 1 kb molecular weight marker (New England Biolabs), 100 bp molecular weight marker (New England Biolabs), UltraPure Agarose (Thermofisher, USA), Phosphate buffered saline (Gibco, USA), Dulbecco's modified Eagle's medium (Gibco, USA), Fetal bovine serum (Sigma-Aldrich, UK), Trypan blue (Gibco, USA), HCS DNA damage kit (Invitrogen, USA), MCF-7 cells (ATCC) and HeLa cells (ATCC® CCL-2™).

3.1.1 Ethical Consent

Ethical approvals were obtained from the Institutional Review Boards of the Noguchi Memorial Institute for Medical Research (NMIMR) and the Korle-bu Teaching Hospital (KBTH). Eligible patients were recruited into the study and breast cancer tissues were collected into labelled sterile tubes for subsequent testing and storage. A copy of the KBTH ethical clearance is attached in the appendix as Appendix 3.

3.2 Methods

3.2.1 Study Design

This study was a cross-sectional study involving adult female breast cancer patients from the Department of Surgery at the Korle-bu Teaching Hospital (KBTH). It involved collection of breast cancer tissues from breast cancer patients undergoing surgery at the KBTH. Participants were selected on the breast cancer clinic days and informed consent was sought from them for sample collection on their clinic-scheduled surgery days.

3.2.2 Study Sites

The study site was the Department of Surgery at the Korle-bu Teaching Hospital in Accra, Ghana.

3.2.3 Subjects/ Study Population

3.2.3.1 Inclusion criteria

1. Female breast cancer patients at the Department of Surgery, Korle-bu Teaching Hospital. Samples were collected from patients; 18 years and above.

3.2.3.2 Exclusion criteria

1. Breast cancer patients undergoing radiotherapy or chemotherapy.

3.2.4 Sample size determination

Breast cancer represents 16% of all cancers in Ghana (Clegg Lamptey *et al*., 2009). The sample size was calculated based on this prevalence with the formula $n = Z^2 \times p(1-p)/ e^2$, where p is the prevalence, Z is the statistic corresponding to the level of confidence and equals 1.96 with a level of confidence of 95%, n is the sample size and e is precision and equals 5%. Therefore, the calculated sample size was $n=1.96^2 \times 0.16\ (1-0.16)/ 0.05^2 = 207$.

3.2.5 Participant Selection

A maximum of 4 patients who met the inclusion criteria were selected for each clinic day. This was to provide enough time to assess the parameters and increase variability in patient characteristics. Folders of all patients who met the inclusion criteria were collected and labelled serially and they were included in the study for that day.

3.2.6 Sample collection

FFPE and fresh breast cancer tissues were obtained from Korle-bu Teaching Hospital. The samples were collected from invasive lobular carcinoma, invasive ductal carcinoma and invasive carcinoma breast cancer patients from grades I-III of breast cancer development. Participants were recruited into the study on the breast clinic days of the Korle-bu Teaching Hospital. Fresh tissues from the participants were collected on the scheduled surgery days for each of them. Sample collection was done over the period of a year, from 2018 to 2019.

Informed consent was sought from each of the recruited participants. A structured questionnaire was administered to each of the participants under supervision to provide information on demographic facts and clinical data.

Pre-prepared FFPE samples which are routinely processed by the Korle-bu pathologists were collected from the Pathology department, KBTH. The fresh tumour samples were sectioned by a pathologist, then transferred into sterile vials containing PBS and transported on ice to the lab at the Department of Biochemistry, Cell and Molecular Biology, University of Ghana for analysis.

3.2.7 Genomic DNA Extraction

The FFPE tissues were sectioned and processed for DNA extraction using the QIAamp DNA FFPE tissue kit (Qiagen). Briefly, for each sample, the block was sectioned into a microcentrifuge tube, deparaffinized by adding xylene and centrifuged. The supernatant was removed and ethanol was added to the pellet and mixed carefully by vortexing. The ethanol was removed and it was allowed to evaporate at 37^{o}C after spinning down the sample. Buffer ATL and proteinase K were added to the pellet and incubated for 1 hour at 56^{o}C. The solution was incubated again for 1 hour at 90^{o}C and centrifuged. Buffer AL and ethanol were added to the pellet, the mixture was transferred into a spin column and centrifuged. Buffer AW1 was added to the column, centrifuged and the flow-through was discarded. The same was done for buffer AW2. Then to elute the DNA bound to the column, buffer ATE was added to the column, incubated and centrifuged. The flow-through containing the DNA was collected in a separate microcentrifuge tube.

The fresh tumours were flash-frozen for DNA extraction using the Qiagen DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions. Briefly, for each sample, the tissue was homogenized in PBS with a mortar and pestle. The homogenate was transferred to a microcentrifuge tube, centrifuged and the PBS was removed. Buffer ATL and Proteinase K were added to the pellet and incubated overnight at 56^{o}C. After, buffer AL and ethanol were added to the lysed tissues and mixed thoroughly. The mixture was transferred into a DNeasy mini spin column. Buffers AW1 and AW2 were added to the column individually and the flow-through was discarded after each addition. The DNA was eluted by adding buffer AE to the column and incubating at room temperature for 5 minutes.

The concentration of each extracted DNA was measured using the nanodrop spectrophotometer and also using the Qubit assay. A 1% agarose gel was run to confirm the presence of DNA in the samples.

3.2.8 *β-Globin* gene Amplification

The integrity of each DNA was confirmed by standard PCR targeting the *β-Globin* housekeeping gene (table 3.4). The overall reaction volume was 25 µl which was made up of 2 µl of DNA (the negative control was nuclease-free water), 0.5 µM of each oligonucleotide primer and 12.5 µl of 2X Choice Taq Mastermix. The reaction was carried out in a Techne Prime thermal cycler using the conditions shown in table 3.1. The PCR amplicons were visualized on a 2% ethidium bromide stained-agarose gel. Twenty microliters of the amplicons were mixed with 2 µl of the 6X loading dye and run on the gel together with a 100 bp molecular weight marker at 100V for 90 minutes. The images were captured using a GE Healthcare Amersham Imager600 gel dock. Samples which showed a positive amplification for the *β-Globin* gene were used for further analyses.

Table 3.1: **Cycling conditions for *β-Globin, L1, ENBA-1* and *env* amplification**

Step	Temperature (°C)	Time	Cycles
Initial denaturation	94	2 minutes	1
Denaturation	94	45 seconds	
Annealing	50	30 seconds	31
Extension	72	2 minutes	
Final extension	72	10 minutes	1

3.2.9 HPV Screening

A nested PCR was carried out targeting the *L1* gene using GP6+ and GP5+ primers (table 3.4) (Portugal *et al*., 2019). The overall reaction volume was 25 µl which was made up of 2 µl of DNA (the negative control was nuclease-free water), 0.5 µM of each oligonucleotide primer, 100 µM dNTPs and 2 µl of 10X Choice Taq Buffer containing 1.5 mM $MgCl_2$. The positive control was DNA extracted from HeLa cells. The reaction was carried out in a Techne Prime thermal cycler

using the conditions shown in table 3.1. The second PCR had the same reaction components, reaction volume and thermal cycling conditions as the first PCR. The PCR amplicons were visualized on a 1.5% ethidium bromide stained-agarose gel. Ten microliters of the amplicons were mixed with 2 µl of the 6X loading dye and run on the gel together with a 100 bp molecular weight marker at 100V for 90 minutes. The images were captured using a GE Healthcare Amersham Imager600 gel dock. A 1.5% agarose gel was run after the reaction to visualize the results.

3.9.9.1 HPV Sequencing

The samples which showed a positive amplification for HPV screening were sequenced using the Sanger sequencing method. Sequences were trimmed using the Finch TV software and consensus sequences were generated using the CLC Main Workbench 8 software (Eftekhaar *et al*., 2017; Mishra *et al*., 2010). The consensus sequences were blasted using the NCBI BLASTn software and sequences were identified based on the highest maximum scores. (McGinnis & Madden, 2004).

3.2.10 EBV Screening

Standard PCR was carried out using *EBNA-1* primers (table 3.4) (Lay *et al*., 2010). The overall reaction volume was 25 µl which was made up of 2 µl of DNA (the negative control was nuclease-free water), 0.5 µM of each oligonucleotide primer, 100 µM dNTPs and 2 µl of 10X Choice Taq Buffer containing 1.5 mM $MgCl_2$. The positive control was a plasmid containing the *EBNA-1* gene. The reaction was carried out in a Techne Prime thermal cycler using the conditions shown in table 3.1. The PCR amplicons were visualized on a 1.5% ethidium bromide stained-agarose gel. Twenty microliters of the amplicons were mixed with 2 µl of the 6X loading dye and run on the gel together with a 100 bp molecular weight marker at 100V for 90 minutes.

3.2.10.1 EBV Typing

Nested PCR was used to genotype EBV as described by (Hassan *et al*., 2006). The first reaction involved amplifying a common region of *EBNA2* using *EBNA-2F* sense and *EBNA-2I* antisense primers (table 3.4). The overall reaction volume was 25 µl which was made up of 2 µl of DNA (the negative control was nuclease-free water), 0.5 µM of each oligonucleotide primer, 100 µM dNTPs and 2 µl of 10X Choice Taq Buffer containing 1.5 mM $MgCl_2$. The positive control was a plasmid containing the *EBNA-1* gene. The reaction was carried out in a Techne Prime thermal cycler using the conditions shown in table 3.2. The primers were replaced with *EBNA-2C* nested sense primer, *EBNA-2G* nested antisense type-1 primer and *EBNA-2B* nested antisense type-2 primer (table 3.4). The total volume and the other reaction components were the same as for the first PCR except for the templates for the second PCR which were amplicons from the first PCR. The reaction was carried out in a Techne Prime thermal cycler using the conditions shown in table 3.3. The amplicons were visualized on a 1.5% agarose gel and images were captured using the GE Healthcare Amersham Imager 600 gel dock.

Table 3.2: **Cycling conditions for first EBV typing PCR**

Step	Temperature (ºC)	Time	Cycles
Initial denaturation	94	2 minutes	1
Denaturation	94	60 seconds	
Annealing	52	90 seconds	35
Extension	72	4 minutes	
Final extension	72	10 minutes	1

Table 3.3: **Cycling conditions for second EBV typing PCR**

Step	Temperature (ºC)	Time	Cycles
Initial denaturation	94	2 minutes	1
Denaturation	94	30 seconds	
Annealing	52	60 seconds	35
Extension	72	2 minutes	
Final extension	72	2 minutes	1

3.2.11 MMTV Screening

Amplification was done using primers targeting the *env* gene (table 3.4) (Wang *et al*., 1995). The overall reaction volume was 25 µl which was made up of 2 µl of DNA (the negative control was nuclease-free water), 0.5 µM of each oligonucleotide primer, 100 µM dNTPs and 2 µl of 10X Choice Taq Buffer containing 1.5 mM $MgCl_2$. The positive control for MMTV screening was DNA extracted from MCF-7 cells. The reaction was carried out in a Techne Prime thermal cycler using the conditions shown in table 3.1. A 1.5% gel was run visualize the PCR products after which the gel was captured using the GE Healthcare Amersham Imager 600 gel dock.

Table 3.4: **Primers used for the detection of viruses, bacteria and *ß-Globin* amplification**

Gene	Primer Code	Sequence	Reference
β-Globin	β-Globin F	ACACAACTGTGTTCACTAGC	(Naushad *et al*., 2017)
	β-Globin R	CAACTTCATCCACGTTCACC	
EBNA-1	QP1	GCCGGTGTGTTCGTATATGG	(Lay *et al*., 2010)
	QP2	CAAAACCTCAGCAAATATATGAG	
EBNA-2	EBNA-2F	TGGAAACCCGTCACTCTC	(Hassan *et al*., 2006)
	EBNA-2I	TAATGGCATAGGTGGAATG	
	EBNA-2C	AGGGATGCCTGGACACAAGA	
	ENBA-2G	GCCTCGGTTGTGACAGAG	
	EBNA-2B	TTGAAGAGTATGTCCTAAGG	
HPV *L1*	GP5+	TTTGTTACTGTGGTAGATACTAC	(Portugal *et al*., 2019)
	GP6+	GAAAAATAAACTGTAAATCATATTC	
MMTV-*Env*	Env-F	CCTCACTGCCAGATC	(Wang *et al*., 1995)
	Env-R	CTATCTGTGGCATACCT	
16SrRNA	pA	AGAGTTTGATCCTGGCTCAG	(Urbaniak *et al*., 2016)
	pH	AAGGAGGTGATCCAGCCGCA	

3.2.12 DNA Damage assay of bacterial strains

3.2.12.1 Media Preparation

3.2.12.1.1 Nutrient agar

Nutrient agar was prepared according to the manufacturer's instructions by weighing 28g of the dehydrated media into 1L of distilled water. The solution was dissolved briefly by boiling. This was then sterilized by autoclaving at 121^{o}C for 15 minutes. The mixture was left to cool to 50^{o}C in a water bath and then poured into sterile Petri dishes in 20-25 ml volumes to obtain a depth of 4 mm. The plates were left to dry at room temperature. A sterility check was done for each of the plates by leaving them overnight at room temperature. Then, the plates were examined for any contamination in the form of growth on the plates. Sterile plates were packaged, sealed and stored at 4^{o}C if not to be used immediately.

3.2.12.1.2 MacConkey Agar

MacConkey agar was prepared by following the manufacturer's instructions. Briefly, 52g of the dehydrated media was suspended in 1L of distilled water. The solution was boiled briefly to dissolve completely and autoclaved at 121^{o}C for 15 minutes. The mixture was left to cool at 50^{o}C in a water bath and then dispensed into sterile Petri dishes in 20-25 ml volumes. They were left to dry at room temperature. Next, the plates were incubated overnight at room temperature to check for sterility. After they were examined for contamination and the sterile plates were packaged, sealed and stored at 4^{o}C if not to be used immediately.

3.2.12.1.3 Blood Agar

Blood agar was prepared according to the manufacturer's instructions by suspending 39g of the Columbia blood agar base in 1L of distilled water and boiled briefly to dissolve. The suspension was sterilized by autoclaving at 121^{o}C for 15 minutes. This was allowed to cool to 50^{o}C in a water

bath and then 5% sterile defibrinated sheep blood was added. This was mixed thoroughly by inverting several times. Then 20-25 ml of the solution was immediately dispensed into sterile Petri dishes to obtain a depth of 4 mm. This was then allowed to solidify at room temperature to check for sterility. Subsequently, the plates were sealed, packaged and stored at 4^{o}C if not used immediately.

3.2.12.1.4 Eosin-methylene Blue Agar (EMB)

The media was prepared according to the manufacturer's instructions by suspending 37.5g of the dehydrated media into 1L of distilled water and dissolved briefly by boiling. The solution was sterilized by autoclaving at 121^{o}C for 15 minutes. This was allowed to cool to 60^{o}C. The medium was shaken vigorously to oxidize the methylene blue and to suspend the essential part of the medium, which is the precipitate. Then, 20-25 ml of the solution was immediately dispensed into sterile Petri dishes. This was then allowed to solidify at room temperature and then checked for sterility. Next, the sterile plated were sealed, packaged and stored at 4^{o}C if not used immediately.

3.2.12.1.5 Mannitol Salt Agar (MSA)

Mannitol salt agar was prepared according to the manufacturer's instructions. Briefly, 111g of the dehydrated media was suspended in 1L of distilled water and dissolved briefly by boiling. This was then sterilized by autoclaving at 121^{o}C for 15 minutes and allowed to cool to 50^{o}C in a water bath. Next, 20-25 ml of the media was poured into sterile Petri dishes and allowed to solidify at room temperature and then checked for sterility after which the sterile plates were sealed, packaged and stored at 4^{o}C of not used immediately.

3.2.12.1.6 Brain Heart Infusion Agar (BHI)

The media was prepared according to the manufacturer's instructions by briefly suspending 47g of the dehydrated media in 1L of distilled water and boiled to dissolve completely. The solution

was sterilized by autoclaving at 121^{o}C for 15 minutes and cooled to 50^{o}C in a water bath. Then 20-25 ml cooled media was dispensed into sterile Petri dishes, allowed to cool at room temperature and checked for sterility. Sterile plates were sealed, packaged and stored at 4^{o}C if not used immediately.

3.2.12.1.7 LB Broth

LB broth was prepared according to the manufacturer's instructions by dissolving 1 capsule (20g) of dehydrated media in 1L of distilled water. This was mixed by gentle agitation and autoclaved at 121^{o}C for 15 minutes. The media was allowed to cool to room temperature and the pH was adjusted to 7.2 using 10 mM of Tris HCl. The broth was then stored at 4^{o}C.

3.2.12.2 Identification of Bacterial Strains

Fresh, sectioned tumours were enriched in 6 ml LB broth aerobically for 24 hours at 37^{o}C in an incubator. The enriched samples were cultured on Nutrient agar, MacConkey agar, Blood agar, Eosin-methylene blue agar (EMB), Mannitol Salt Agar (MSA) and Brain Heart Infusion (BHI) agar for 24 hours at 37^{o}C. Successive streaking of individual colonies was done to obtain pure colonies on their respective agar plates. A frozen stock of the pure colonies was prepared by inoculating each pure colony in a solution of 50% glycerol and 50% broth. The glycerol stocks were stored at -20^{o}C.

3.2.12.2.1 Gram Staining of Bacterial Isolates

Preliminary identification of pure colonies was done by using the Gram's staining technique to classify the isolates into Gram-positive and Gram-negative bacteria. Smears of pure colonies were prepared on clean microscope slides and heat fixed. The smears were flooded with the primary stain (crystal violet) for 60 seconds and washed off under running water. Next, they were covered with Lugol's iodine for 60 seconds to fix the crystal violet into the bacterial cell wall. The iodine

was washed off and a decolourization step was performed by covering the smears with a solution of acetone-alcohol which was washed off immediately. Safranin stain was added for 2 minutes and washed off. The slides were air-dried and observed using 1000X magnification under oil immersion (Cheesbrough, 1984). Colonies were considered Gram-positive if they stained purple and Gram-negative if they stained red.

3.2.12.2.2 Biochemical Analysis of Bacterial Isolates

Basic biochemical tests were also performed to further identify the isolates as described by Miranda *et al*., 2008.

3.2.12.2.2.1 Triple sugar Iron Agar test (TSI)

This test was used to differentiate *Enterobacteriaceae* according to their ability to ferment lactose, sucrose and dextrose, and to produce hydrogen sulphide. The TSI agar was prepared by weighing 65g of the dehydrated powder into 1L distilled water. The suspension was boiled to dissolve completely. This was autoclaved at 121^{o}C for 20 minutes and dispensed into test tubes. After, the medium was allowed to set in a slope. A well-isolated colony was picked with a straight wire loop and emulsified at the side of the tube. The butt was stubbed and the slope of the tube was streaked with the wire loop. An alkaline slant with no change in butt (K/NC) was identified as a glucose, lactose and sucrose non-fermenter; an alkaline slant with an alkaline butt (K/K) was identified as a glucose, lactose and sucrose non-fermenter; an alkaline slant with an acidic butt (K/A) was identified as a glucose fermenter, and an acidic slant with an acidic butt (A/A) was interpreted as a glucose, lactose and/or sucrose fermenter. Also, a break or production of bubbles in the slope showed gas production and darkening of the agar showed the production of hydrogen sulphide (Miranda *et al*., 2008).

3.2.12.2.2.2 Sulfide Indole Motility Test (SIM)

This medium is used to differentiate *Enterobacteriaceae* by their motility, indole reactions and hydrogen sulfide producing abilities. The SIM medium was prepared by weighing 30g of the dehydrated powder into 1L of distilled water. This was dissolved by heating at 100^{o}C and allowed to cool to a temperature of between 50-55^{o}C. Phenol red solution was added, mixed thoroughly and dispensed into 4ml screw-capped tubes. The tubes were sterilized by autoclaving at 121^{o}C for 15 minutes. They were allowed to cool and each tube was stabbed several times with inoculating loops containing the cultured isolates. The tubes were capped and incubated overnight at 37^{o}C. Then, 2-3 drops of Kovac's reagent was added and incubated for 24 hours at 37^{o}C. The formation of a red ring on the surface of the tube showed a positive indole test, a black colouration showed hydrogen sulfide production and a motile organism was classified by the ability of the organism to move away from the stab (Miranda *et al*., 2008).

3.2.12.2.2.3 Simmons Citrate Test

This test was used to differentiate members of the *Enterobacteriaceae* family according to their ability to utilize citrate as the sole source of carbon. The Simmons Citrate agar was prepared by weighing 28g of the dehydrated medium into 1L of distilled water. This was sterilized by autoclaving at 121^{o}C for 15 minutes and allowed to cool. The medium was dispensed into test tubes and allowed to set on a slant. A well-isolated colony was picked and inoculated on the slant of the agar and incubated at 37^{o}C for 24 hours. An organism was classified as citrate positive if there was visible growth on the surface of the slant and the medium changed colour to an intense Prussian blue. An organism was classified as citrate negative if the medium remained deep forest green with no trace or visible growth (Miranda *et al*., 2008).

3.2.12.2.2.4 Catalase Test

This test was used to differentiate catalase producing bacteria from non-catalase producing bacteria. A drop of sterile, distilled water was placed on a clean microscope slide. A well-isolated colony was picked and emulsified onto the slide to form a suspension. A drop of hydrogen peroxide was added. Active bubbles formation was recorded as a positive test and organisms which had no release of bubbles were classified as negative for the catalase test (Miranda *et al*., 2008).

3.2.12.2.2.5 Oxidase Test

The oxidase test was done to identify organisms which produce oxidase. A piece of filter paper was placed in a petri dish and soaked with a few drops of freshly prepared oxidase reagent. A well-isolated colony was smeared on the filter paper. The formation of a blue purple colour within 10 seconds indicated an oxidase-positive organism and a colourless test showed an oxidase negative organism (Miranda *et al*., 2008).

3.2.12.2.3 DNA Extraction from Bacterial Isolates

DNA was extracted from the pure colonies using the Qiagen DNeasy DNA extraction kit (Qiagen) based on the results from the Gram-staining. Briefly, for each isolate was grown in LB broth for 48 hours and the cells were harvested by centrifugation. The pellets were re-suspended in buffer ATL and proteinase K were added. Incubation was done overnight at 56^{o}C for Gram-positive bacteria and incubation was done for three hours at 56^{o}C for Gram-negative isolates. After, buffer AL and ethanol were added to the lysed cells and mixed thoroughly. The mixture was transferred into a DNeasy mini spin column. Buffers AW1 and AW2 were added to the column individually and the flow-through was discarded after each addition. The DNA was eluted by adding buffer AE to the column and incubating at room temperature (25^{o}C) for 5 minutes.

3.2.12.2.3.1 Amplification of the *16SrRNA* gene by Polymerase Chain Reaction (PCR)

The DNA extracted from the colonies were amplified using eubacterial primers pA/pH which amplify the complete *16SrRNA* gene (table 3.4). The PCR was performed in a 25 µl reaction containing 2 µl of DNA (nuclease-free water was used as the negative control), 2.5 µl of 10X PCR buffer with 1.5 mM $MgCl_2$ (Choice Taq DNA Polymerase), 100 µM dNTP mix (Invitrogen), 50 µM of each primer and 2.5U of Choice Taq DNA Polymerase. Amplification was done was shown in table 3.5. Two microliters of the amplicons were mixed with 1 µl of the 6X loading dye and run on the gel together with a 1 kb molecular weight marker at 100V for 90 minutes. The images were captured using a GE Healthcare Amersham Imager600 gel dock. A 1.5% agarose gel was run after the reaction to visualize the results.

Table 3.5: **Cycling conditions for *16SrRNA* PCR**

Step	Temperature (ºC)	Time	Cycles
Initial denaturation	94	2 minutes	1
Denaturation	94	45 seconds	
Annealing	62.5	30 seconds	31
Extension	72	2minutes	
Final extension	72	10 minutes	1

3.2.12.2.3.2 Analysis of *16SrRNA* sequences

The *16SrRNA* sequences were identified using the NCBI BLAST microbes software (McGinnis & Madden, 2004). Taxonomy was assigned based on the highest maximum scores.

3.2.12.3 HeLa and MCF-7 Cell Culture

HeLa and MCF-7 cells were cultured in antibiotics-free DMEM supplemented with 10% fetal bovine serum (FBS).

3.2.12.4 DNA Damage Assay

The DNA damage of bacterial isolates was assessed using the HCS DNA Damage Kit (Invitrogen). The bacterial isolates were prepared for infection by growing 200 µl of the isolates from the glycerol stock in 1 mL LB broth for 48 hours at 37^{o}C.

After 24 hours, 24-well plates containing sterile coverslips were seeded with 0.5 ml of 1×10^{5} cells/ml, resulting in 5×10^{4} cells/well and incubated for 24 hours.

Next, 500 µl of the LB culture was inoculated into 9.5 ml pre-warmed DMEM and grown at 37^{o}C with shaking to reach OD_{600nm}=0.4 to 0.5. One unit of OD corresponds to 5×10^{8} bacteria/ml.

The cells were washed three times with 500 µl PBS and 500 µl antibiotic-free DMEM only was added to each well. They were then infected at a multiplicity of infection (MOI) of 100 or 50 for each bacterial isolate. The 24-well plates were incubated at 37^{o}C with 5% CO_2 for 4 hours. The cells were also infected with *E. coli* at an MOI of 100 or treated with 50 µl of 10% etoposide to serve as positive controls. After 4 hours, 100 µl of the Image-iT dead green cell viability stain was added, and the plate was again incubated for 30 minutes under the same conditions. The medium was removed and 200 µl of the fixative solution was added to each well and incubated for 15 minutes at room temperature. The fixative solution was removed, and the cells were washed with PBS. The cells were incubated with 100 µl of the permeabilization solution for 15 minutes and rinsed once with PBS. Then, 200 µl of the blocking buffer was added to each well and incubated for 60 minutes at room temperature. The blocking solution was removed, 100 µl of the primary antibody solution was added to the wells and incubated for 60 minutes. The primary antibody was removed, and the cells were washed three times with PBS and 100 µl of the secondary antibody solution counterstained with Hoechst was added. This was incubated for 60 minutes at room temperature, away from light. The secondary antibody solution was removed, and the cells were

washed three times with PBS. The coverslips were removed and placed on sterile microscope slides containing 10 µl of mounting medium. The coverslips on the microscope slides were sealed with nail polish.

3.2.12.5 Imaging and Analysis

The DNA damage was observed using the Olympus Fluorescent microscope with a magnification of 1000X. Three fields of view for each MOI was captured. Images were analyzed using the Image J software to calculate the mean fluorescence intensity of each phosphorylated-H2AX- stained cell (Urbaniak *et al*., 2016). The Graph Pad Prism software was used to check for significant differences between treatment groups (Elliott & Hynan, 2011).

CHAPTER FOUR

4.0 RESULTS

4.1 *β-Globin* amplification

A total of 400 FFPE and 17 fresh breast cancer tissues were collected and DNA extracted, and using the nanodrop spectrophotometer, Qubit assay and gel electrophoresis, the quality and concentration of the DNA were determined. Out of the 417 breast cancer tissue samples, 229 were determined to be of good concentration and quality. The *β-globin*, a housekeeping gene was then amplified in these 229 samples to determine the integrity the DNA. The *β-globin* gene was amplified in 204 out of 219 DNA extracts, and these 204 samples were used for further analysis (table 4.1).

Table 4.1: **Number of *β-Globin* amplified samples**

Total number of samples analyzed	***β-Globin* amplified**	***β-Globin* not amplified**
229	204	25

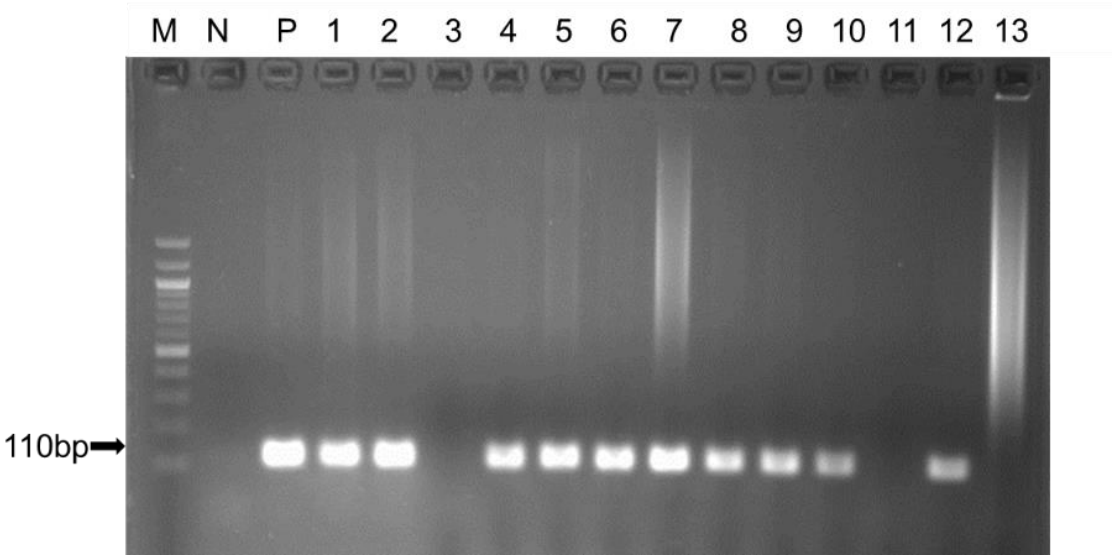


Figure 4.1: **Gel image of *β-Globin* amplification.** M represents the 100 bp molecular weight marker, N represents the negative control, P represents the positive control, 1-13 represent the sample IDs. The expected band size of the *β-Globin* gene was 110 bp.

4.2 Clinical and Pathological Data of Participants

All 204 samples which showed a positive amplification for the *β-Globin* gene were used for further analyses. Out of the 204 samples, 78 (38.2%) were from the left breast, 84 (41.2%) were from the right breast, 2 (1.0%) was from both breasts and 40 (19.6%) had no available data for the breast affected as shown in table 4.2. For the tumour grades, 13 (6.4%) were grade I, 108 (52.9%) were grade II, 50 (24.5%) were grade III and the grades of 33 (16.2%) were not reported. Data on diagnosis was available on 134 (65.7%) for invasive ductal carcinoma, 3 (1.5%) for invasive lobular carcinoma 52 (25.5%) for invasive carcinoma, while the diagnosis of 15 (7.4%) was not available.

Table 4.2: **Frequency of clinicopathological parameters of breast cancer patients**

Pathological description	Number of samples (%)
Breast Affected	
Left	78 (38.2)
Right	84 (41.2)
Both	2 (1.0)
Not reported	40 (19.6)
Grade	
Grade I	13 (6.4)
Grade II	108 (52.9)
Grade III	50 (24.5)
Not reported	33 (16.2)
Diagnosis	
Invasive ductal carcinoma	134 (65.7)
Invasive lobular carcinoma	3 (1.5)
Invasive carcinoma	52 (25.5)
Not reported	15 (7.4)

4.3 HPV Detection and Genotyping

All 204 samples that had the *β-Globin* gene amplified were used for HPV screening. Primers targeting the HPV *L1* gene with an expected band size of 150 bp was used for identification (figure 4.2a). Out of the 204 samples screened, HPV was detected in 27 samples (figure 4.2b). HPV genotyping was done by sequencing which revealed that all the samples that showed positive amplification for HPV were HPV 18. However, 15 were HPV 18 isolate 3, 5 were HPV 18 isolate BB, 1 was HPV 18 strain CNTZ36, 1 was HPV 18 isolate MJR16 and 1 was HPV 18 isolate F.

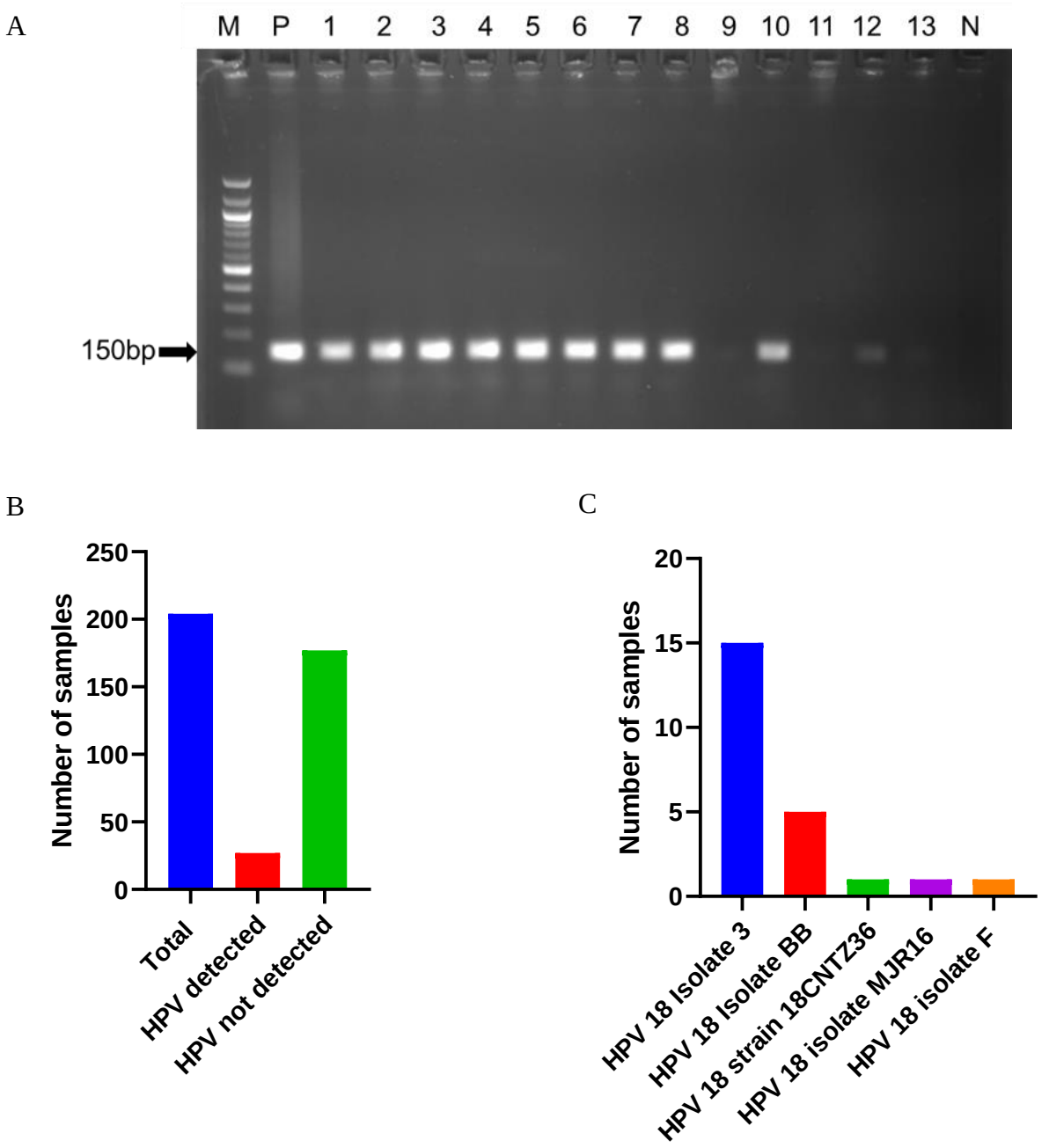


Figure 4.2: **Amplification of the HPV *L1* gene.** (A) A representative gel showing amplification of the HPV *L1* gene. M represents the 100 bp DNA ladder, P represents the positive control, N represents the negative control and 1-13 represent the sample IDs. (B) A graphical representation of the amplification of the *L1* gene. (C) Distribution of HPV subtypes in HPV positive samples.

4.4 EBV Detection and Genotyping

All 204 samples were screened for EBV using PCR and visualized by gel electrophoresis (figure 4.3a). EBV was detected in 66 samples as shown in figure 4.3b. Nested PCR was done to genotype EBV into EBV-1 and EBV-2 (figure 4.4a). From the EBV typing, 2 samples were positive for EBV-1, 30 for EBV-2, 26 for both EBV-1 and EBV-2 and 8 samples could not be grouped into either type 1 or type 2 (figure 4.4b). Comparing the tumour grades and infection with EBV subtypes (table 4.3), the highest EBV infection in the grade I tumours was EBV-2, which was 3 (23.1%), followed by EBV-1 and 2 co-infection which was 2 (15.4%) then EBV-1 which was 1 (7.7%). For grade II, 17 (15.7%) were EBV-2, 12 (11.1%) were co-infected with both subtypes and 1 (0.9%) was EBV-1. For the grade III tumours, 6 (12%) were EBV-2 which was the same for EBV-1 and 2 co-infection. In the invasive ductal carcinoma cases, 16 (11.9%) were EBV-2 which was same for EBV-1 and EBV-2 co-infection and 2 (1.5%) were EBV-1. Only 1 (33.3%) invasive lobular carcinoma tumour was EBV-2. Eleven (21.2%) invasive carcinoma tumours were EBV-2 and 8 (15.4) were co-infected with both subtypes.

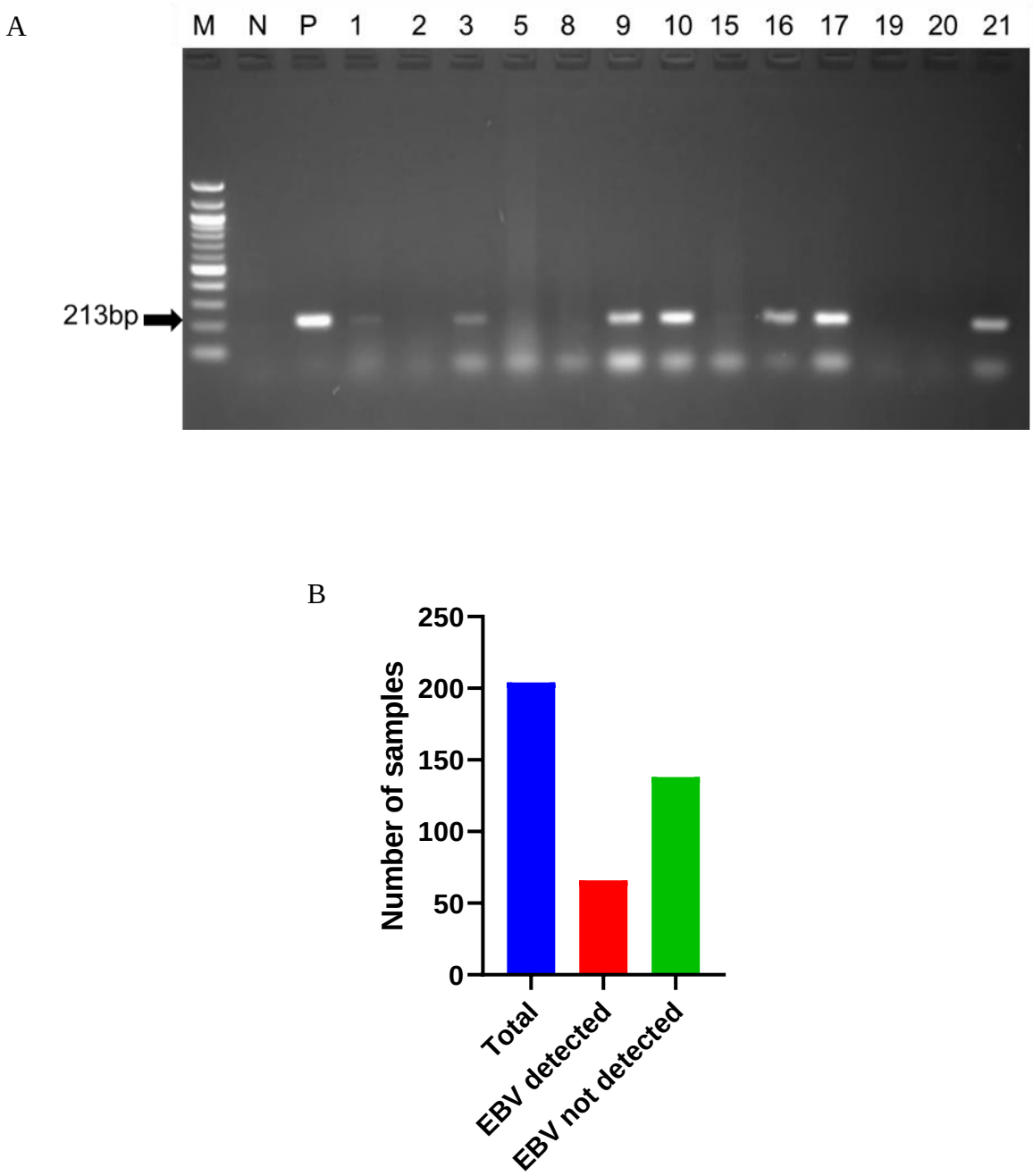
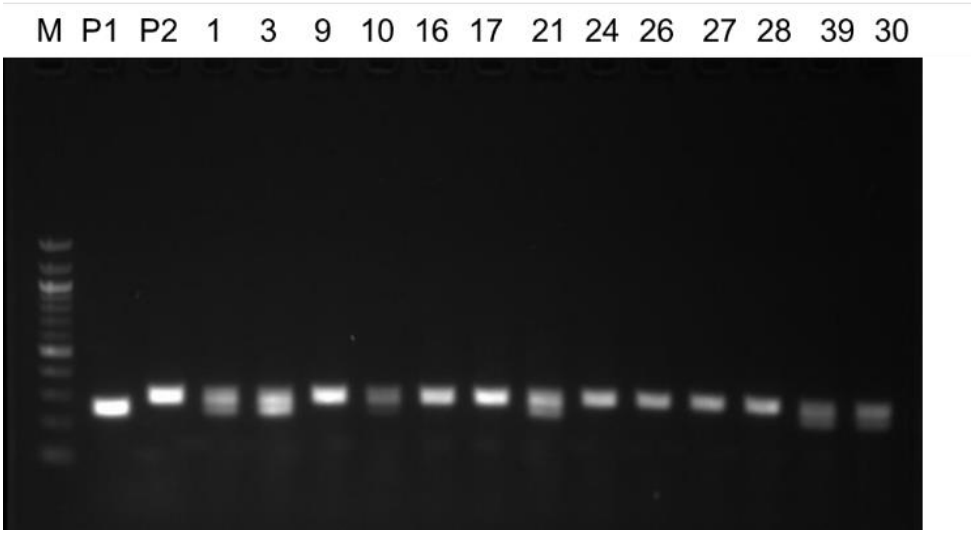


Figure 4.3: **Amplification of the *EBNA-1* gene.** (A) A representative gel showing EBV screening of samples with an expected band size of 213 bp. M=100 bp DNA ladder, N= negative control, P= positive control, 1-21= Sample IDs. (B) A graphical representation of EBV detection in the samples.

A



B

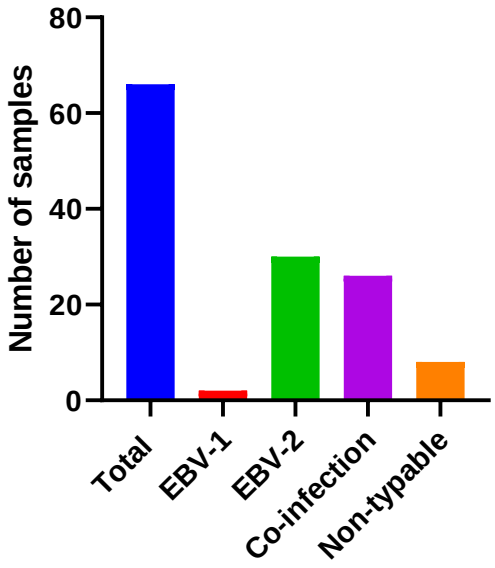


Figure 4.4: **EBV genotyping of samples that showed positive amplification for *EBNA-1*.** (A) A representative gel showing the products of the nested PCR EBV genotyping. The expected band sizes of the nested PCR products were 250 bp (EBV-1) and 300 bp (EBV-2). M represents the 100 bp DNA ladder, P1 represents the positive control for type 1, P2 represents the positive control for type 2, 1-30 represent sample IDs. (B) Graphical representation of EBV genotypes.

Table 4.3: **Pathological outcomes and EBV subtype infections**

Pathological description	No. of cases	EBV subtype infection		
		EBV-1	EBV-2	EBV-1 & EBV-2 co-infection
Grade				
Grade I	13	1 (7.7)	3 (23.1)	2 (15.4)
Grade II	108	1 (0.9)	17 (15.7)	12 (11.1)
Grade III	50	0 (0.0)	6 (12.0)	6 (12.0)
Diagnosis				
Invasive ductal carcinoma	134	2 (1.5)	16 (11.9)	16 (11.9)
Invasive lobular carcinoma	3	0 (0.0)	1 (33.3)	0 (0.0)
Invasive carcinoma	52	0 (0.0)	11 (21.2)	8 (15.4)

4.5 MMTV Detection

The samples were screened for MMTV using PCR and the results were visualized by gel electrophoresis (figure 4.5). Of all the 204 samples screened, MMTV was not detected in any of them.

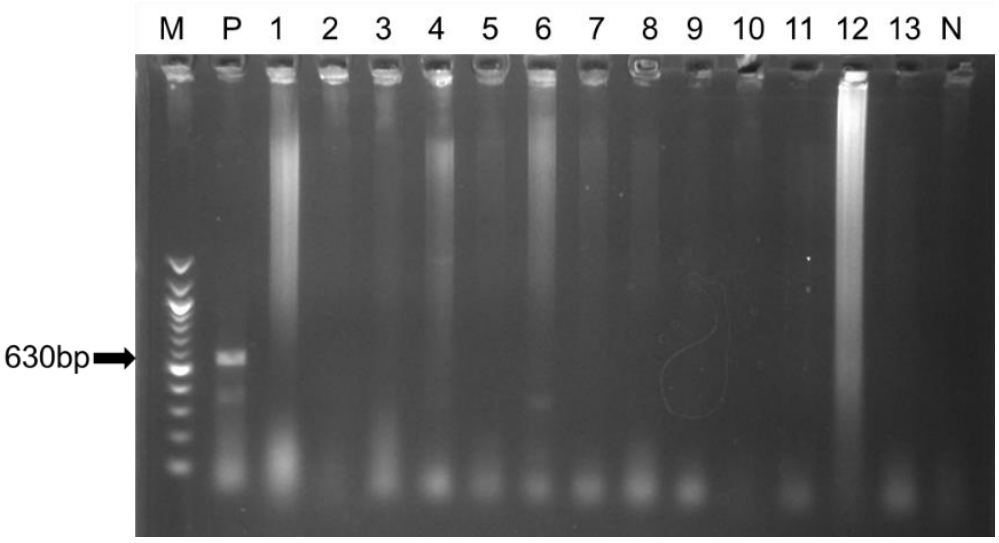


Figure 4.5: **A representative gel of MMTV screening for samples.** M=Molecular weight marker, P=positive control, N=negative control, 1-13=Sample IDs. The expected band size was 630 bp.

4.6 Prevalence of Viral Infections in Breast Cancer Tumours

Out of 204 samples analyzed, EBV only was detected in 55 (27%) tumours, HPV alone was detected in 16 (7.8%) samples while co-infection of EBV and HPV was detected in 11 (5.4%) samples (figure 4.6). After analysis, 133 (65.2) of tumours were negative for viral sequences and 71 (34.8%) were positive for viral sequences (table 4.4). HPV was detected in 27 (13.2%) of the samples. EBV was detected in 66 (32.4%) of the samples, EBV-1 was detected in 2 (1%) of the total number of samples analyzed, EBV-2 was detected in 30 (14.7%) of the samples, EBV-1 and EBV-2 co-infection was seen in 26 (12.7%) of the samples and 8 (3.9%) of the samples were EBV positive but could not be genotyped by the PCR method used.

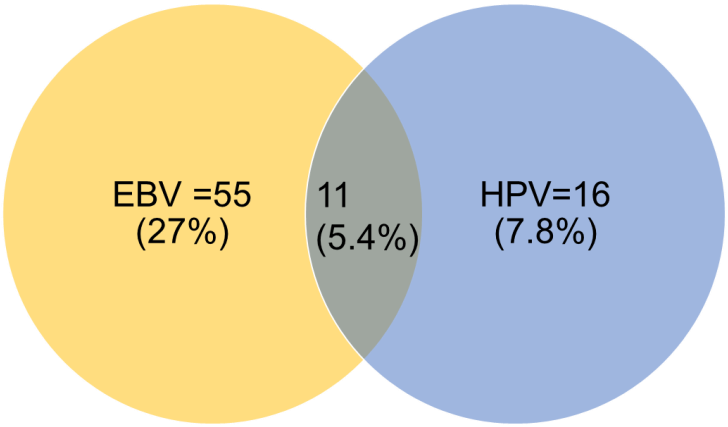


Figure 4.6: Graphical distribution of EBV only, HPV only and EBV+HPV only in breast cancer tissues.

Table 4.4: **Prevalence of HPV, EBV, EBV Subtypes and Co-infections in Breast Cancer Patients**

Total number of samples	Number of samples negative for viral sequences	Number of samples positive for viral sequences	HPV	EBV	EBV-1	EBV-2	EBV-1 and EBV-2 co-infection	Non-typable EBV	HPV and EBV co-infection
N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
204 (100)	133 (65.2)	71 (34.8)	27 (13.2)	66 (32.4)	2 (1)	30 (14.7)	26 (12.7)	8 (3.9)	11 (5.4)

4.7 Bacterial Colonies on Different Media After Incubation for 24 hours.

Growth of bacteria was observed on different media after plating enriched tissues on BHI agar, EMB agar, MacConkey agar, Mannitol agar, nutrient agar and blood agar (figure 4.7). Morphologically different colonies were selected and purified by successive streaking to obtain pure colonies (figure 4.8).

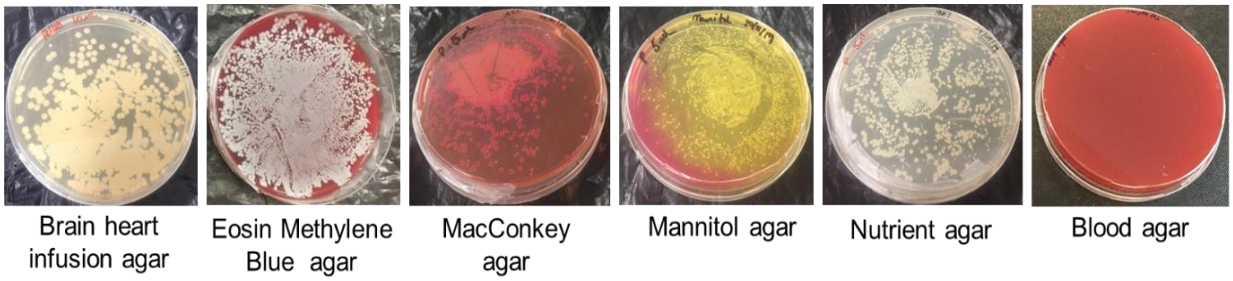


Figure 4.7: **Initial cultures of tissues after incubation for 24 hours.** Fresh tumour samples were cultured on blood agar, brain heart infusion agar, MacConkey agar, eosin-methylene blue agar, nutrient agar and mannitol salt agar. The plates were incubated for 24 hours and observed for growth.

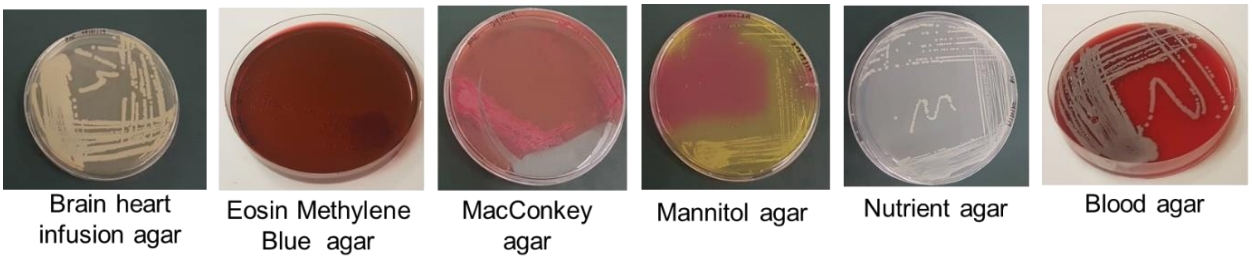
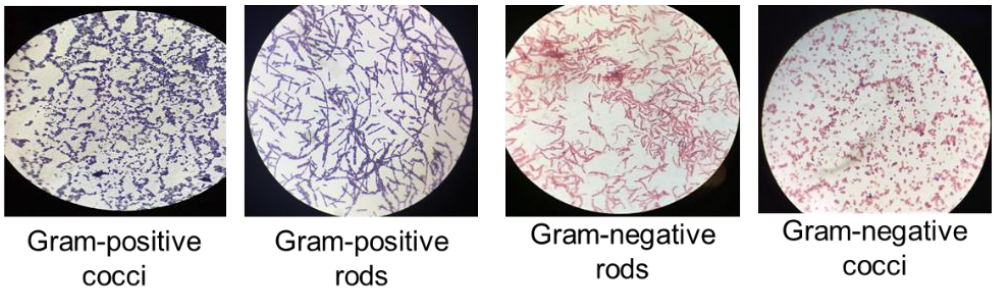


Figure 4.8: **Purified colonies from initial cultures after successive streaking for 24 hours.** Selected colonies from the initial 24-hour cultures were subcultured on their respective plates to obtain pure cultures.

4.8 Identification of Cultured Isolates

4.8.1 Gram-staining of Isolates

Preliminary identification of the isolates was done by Gram-staining (figure 4.9). Out of 190 isolates cultured from 10 fresh breast cancer samples, 55 were Gram-positive cocci, 56 were Gram-positive rods, 23 were Gram-negative rods and 56 were Gram-negative cocci (figure 4.10).



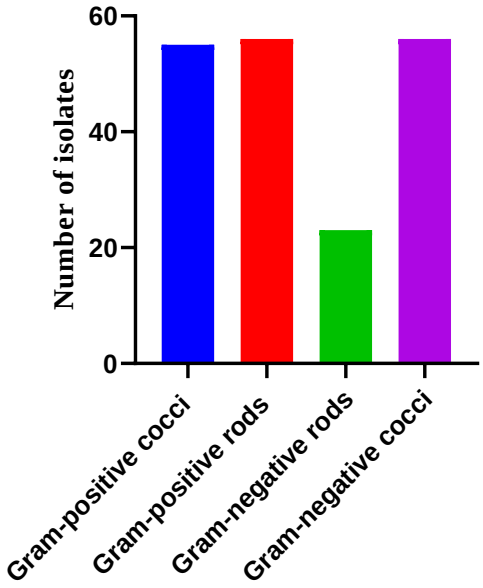


Figure 4.10: **Graphical representation of Gram-staining.** Cultured isolates were identified by Gram-staining of pure colonies.

4.8.2 Biochemical Identification of Isolates

Biochemical tests were then performed to further identify the bacteria (figure 4.11). Five biochemical tests were performed on 141 isolates from the tumour samples and 6 isolates from the PBS controls (table 4.5). For the catalase test, 84 isolates were positive while 32 isolates were negative. Also, 6 isolates were oxidase-positive and 6 were negative for the oxidase test. The isolates which showed positive results for the Simmons Citrate test were 6 while 5 were negative. For the SIM test, 4 isolates were positive while 7 isolates were negative. The TSI test results showed 18 glucose, sucrose and lactose non-fermenters (K/K) and 1 glucose fermenter only (K/A).

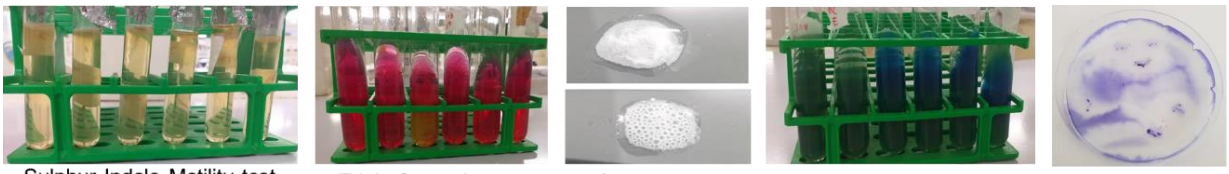


Figure 4.11: **Images from biochemical tests to identify cultured bacteria.** Biochemical analyses of selected isolates for bacterial identification.

Table 4.5: **Biochemical analysis of identified bacteria**

Sample number	Number of isolates	Gram staining								
		Catalase		Oxidase		Simmons Citrate agar		SIM test		TSI
		+	-	+	-	+	-	+	-	
001	28	27	-	-	1	-	-	-	-	-
002	13	8	5	3	2	4	1	2	3	12 K/K 1 K/A
002 control	6	-	6	3	3	2	4	2	4	6 K/K
003	20	20	-							
004	24	17	7							
005	12	10	2							
006	14	2	12							
005 and 006 control	4	-	4							

+: Positive, -: Negative, N/A- Not analysed.

4.8.3 *16SrRNA* Identification of Bacteria

The *16SrRNA* gene of bacteria was amplified using PCR and sequencing to identify the isolates obtained from the cultures (figure 4.12). The bacteria identified by Sanger sequencing of the *16SrRNA* gene were *Staphylococcus sciuri* (7), *Megasphaera cerevisiae* (2) , *Staphylococcus capitis* (4), *Herbaspirillum chlorophenolicum* (1), *Staphylococcus warneri* (1), *Microbacterium chocolatum* (2), *Acinetobacter pittii* (7), *Staphylococcus epidermidis* (7), *Bacillus pumilus* (1), *Staphylococcus lugdenensis* (1) and *Bacillus thuringiensis* (6) as seen in figure 4.13.

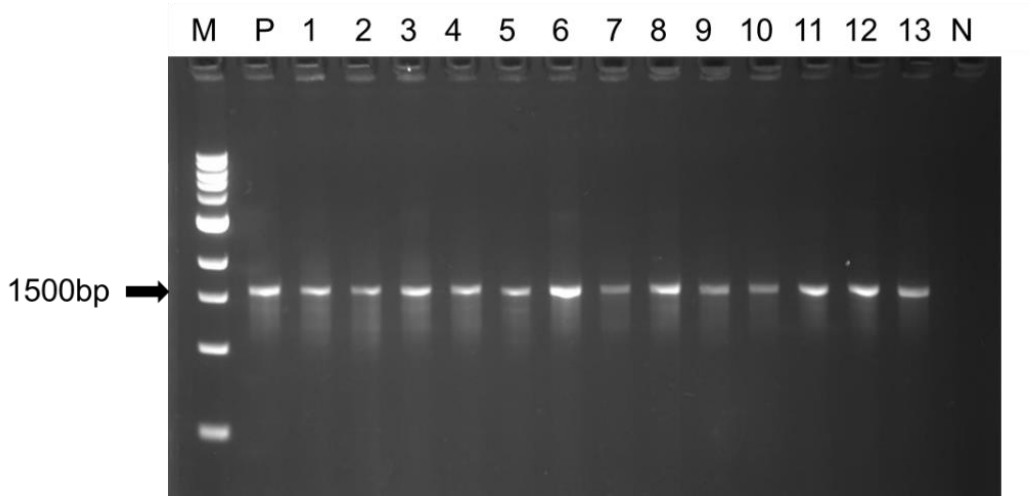


Figure 4.12: **A representative gel showing amplification of the *16SrRNA* gene**. M: 1 kb molecular weight marker, P: positive control, N: negative control, 1-13: sample IDs. The expected band size of the *16SrRNA* gene is 1500 bp.

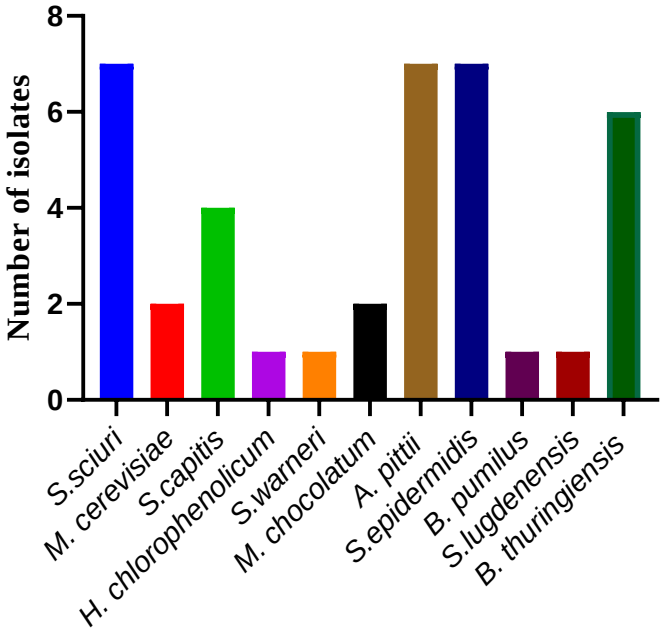


Figure 4.13: **Distribution of bacteria after sequencing**. A graph of the number of bacteria obtained after sequencing cultured isolates. *S. scuiri* represents *Staphylococcus scuiri*, *M. cerevisiae* represents *Megasphaera cerevisiae*, *S. capitis* represents *Staphylococcus capitis*, *H. chlorophenolicum* represents *Herbaspirillum chlorophenolicum*, *S. warneri* represents *Staphylococcus warneri*, *M. chocolatum* represents *Microbacterium chocolatum*, *A. pittii* represents *Acinetobacter pittii*, *S. epidermidis* represents *Staphylococcus epidermidis*, *B. pumilus* represents *Bacillus pumilus*, *S. lugdensis* represents *Staphylococcus lugdenensis* and *B. thuringiensis* represents *Bacillus thuringiensis*.

4.9 Detection of DNA Damage and Cytotoxicity in HeLa cells

4.9.1 Infection at an MOI of 100

Cultured HeLa cells were infected with bacteria isolated from fresh breast cancer tissues at an MOI of 100 for 4 hours, the DNA damage assay was done and images captured using the fluorescent microscope. The pH2AX antibody was used as a marker for DNA damage, nuclear morphology was observed with the Hoechst dye and cytotoxicity was observed with a cell viability stain (figure 4.14). Cells treated with etoposide; a known inducer of double-strand breaks had a significantly high Mean Fluorescent Intensity (MFI) when compared with untreated cells (figure 4.15). There were no statistically significant differences in cells treated with the other bacteria. Etoposide and *E. coli* were included as positive controls.

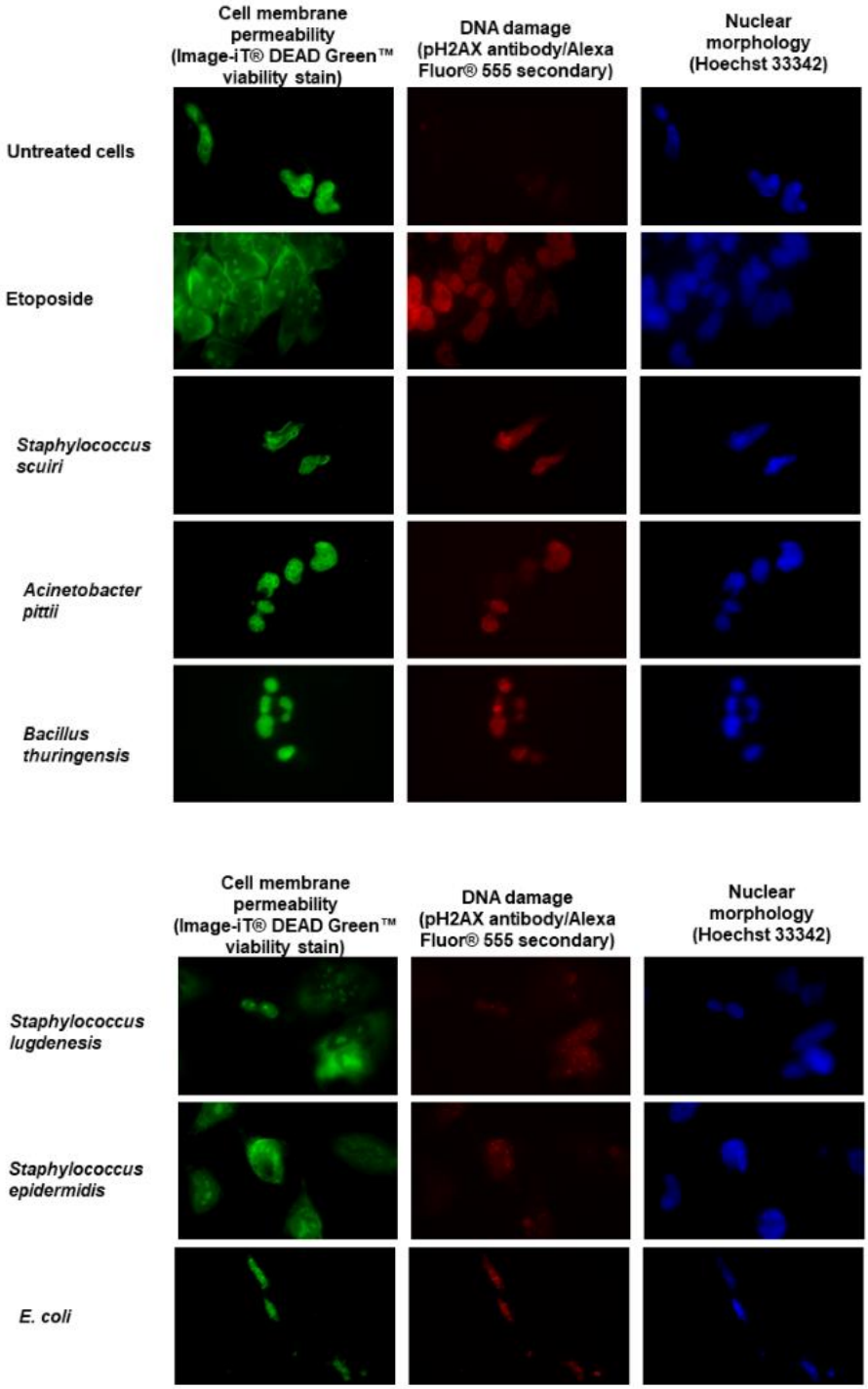


Figure 4.14**: Immunofluorescence images of the DNA damage assay in HeLa cells at an MOI of 100**. Bacteria were isolated from breast cancer tissues and the DNA damage abilities of the bacteria were tested using the pH2AX DNA damage assay. Cells were treated with the bacteria and stained for pH2AX, cytotoxicity and Hoechst. Images were captured and viewed at a magnification of 1000X. Scale bar = 0.5 µM. Etoposide was included as a positive control since is a known inducer of DNA double-strand breaks.

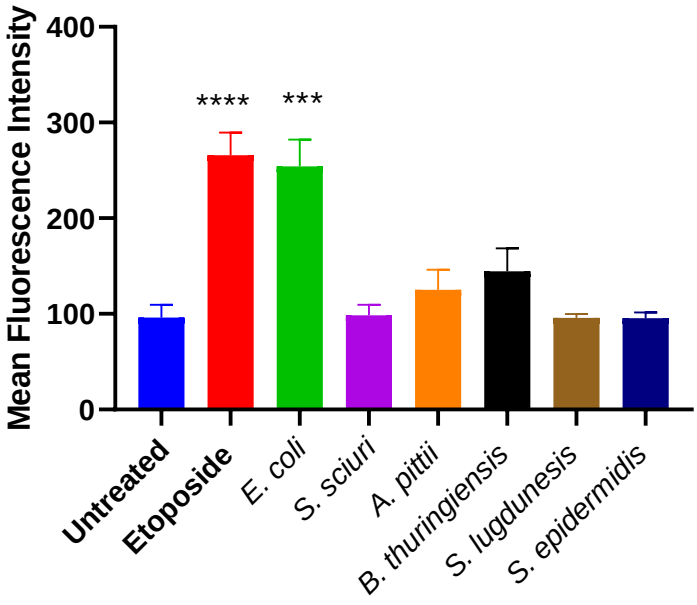


Figure 4.15**: Mean fluorescent intensities of images from the DNA damage assay of HeLa cells**. Image J was used to calculate the mean fluorescent intensities of pH2AX from the images. The statistical significance of the mean fluorescence intensity of the isolates was calculated using the Dunn's multiple comparisons test. The data represented for each bacterium on the graph shows the mean and the standard deviations of three cells from three fields of view for each of the isolates tested, showing nine cells for three fields of view for each bacterium. *** $p<0.001$. **** $p<0.0001$.

4.9.2 Infection at an MOI of 50

HeLa cells were cultured and treated with different bacteria at an MOI of 50. Images from the infection assay were captured using a fluorescent microscope (figure 4.16). The mean fluorescent intensities were calculated using the Image J software. Apart from etoposide treated cells, no other treatment had a statistically significant intensity when compared with the untreated cells (figure 4.17).

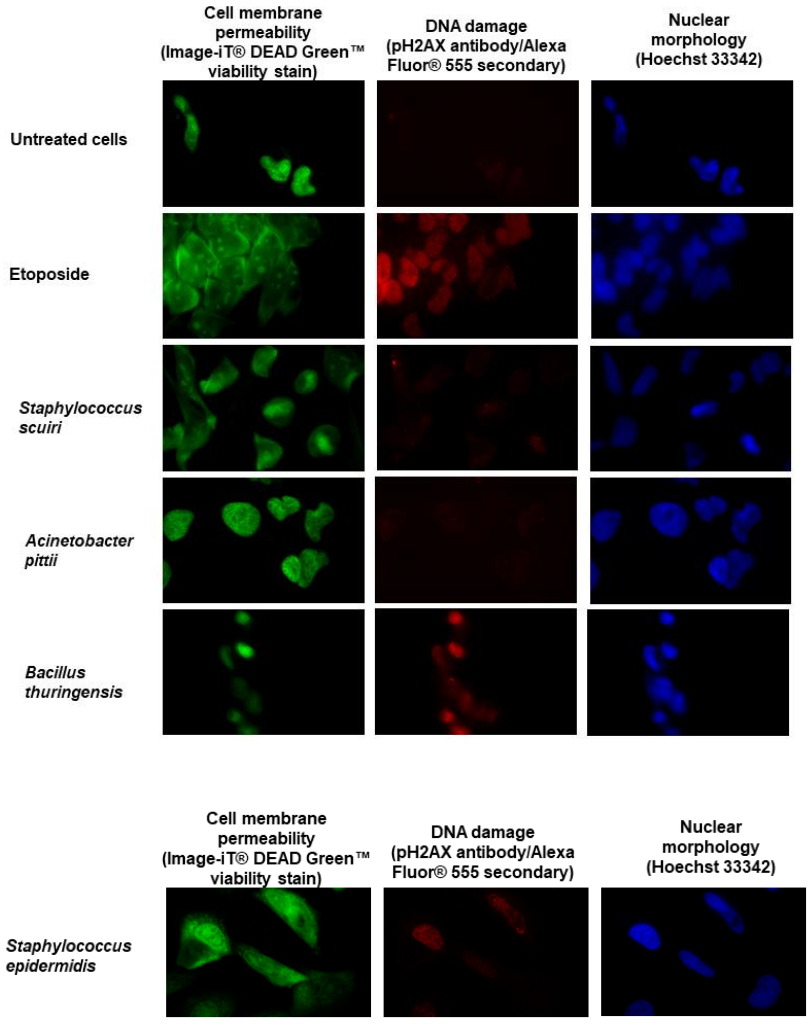


Figure 4.16: **DNA damage assay of HeLa cells at an MOI of 50**. Bacteria were isolated from breast cancer tissues and the DNA damage abilities of the bacteria were tested using the pH2AX DNA damage assay. Images for HeLa cells treated with bacteria were captured using a fluorescent microscope at a 1000X magnification. Scale bar = 0.5 µM. Etoposide was included as a positive control since is a known inducer of DNA double-strand breaks.

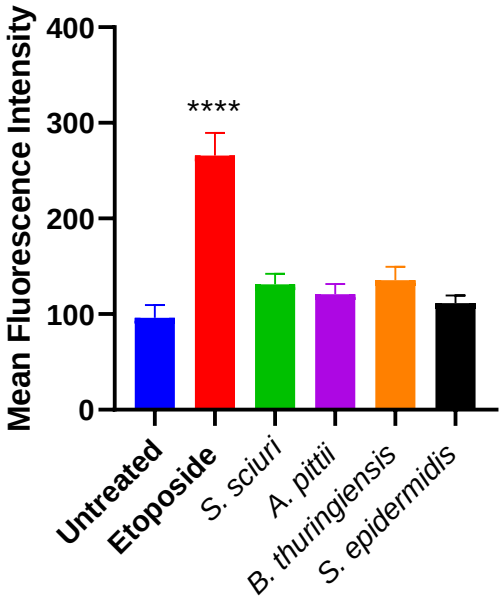


Figure 4.17**: Image J analysis of the mean fluorescent intensities of pH2AX**. The significance of the mean fluorescence intensity of each isolate was calculated using the Dunn's multiple comparisons test. The data represented for each bacterium on the graph shows the mean and the standard deviations of three cells from three fields of view for each of the isolates tested, showing nine cells for three fields of view for each bacterium. **** $p<0.0001$.

4.10 Detection of DNA Damage and Cytotoxicity in MCF-7 cells

4.10.1 Infection at an MOI of 100

Cultured MCF-7 cells were incubated with various bacteria isolated from breast tumours. Cellular levels of pH2AX were measured and used as a marker for double-strand breaks. Cell membrane permeability was determined using a cell viability stain, pH2AX levels were also measured and nuclear morphology was also detected using the Hoechst dye (figure 18). Out of 5 isolates tested, cells exposed to *Staphylococcus sciuri* had a statistically significantly higher pH2AX MFI when compared with the untreated cells (figure 19). Also, etoposide had a statistically significantly high MFI as expected since it was the positive control.

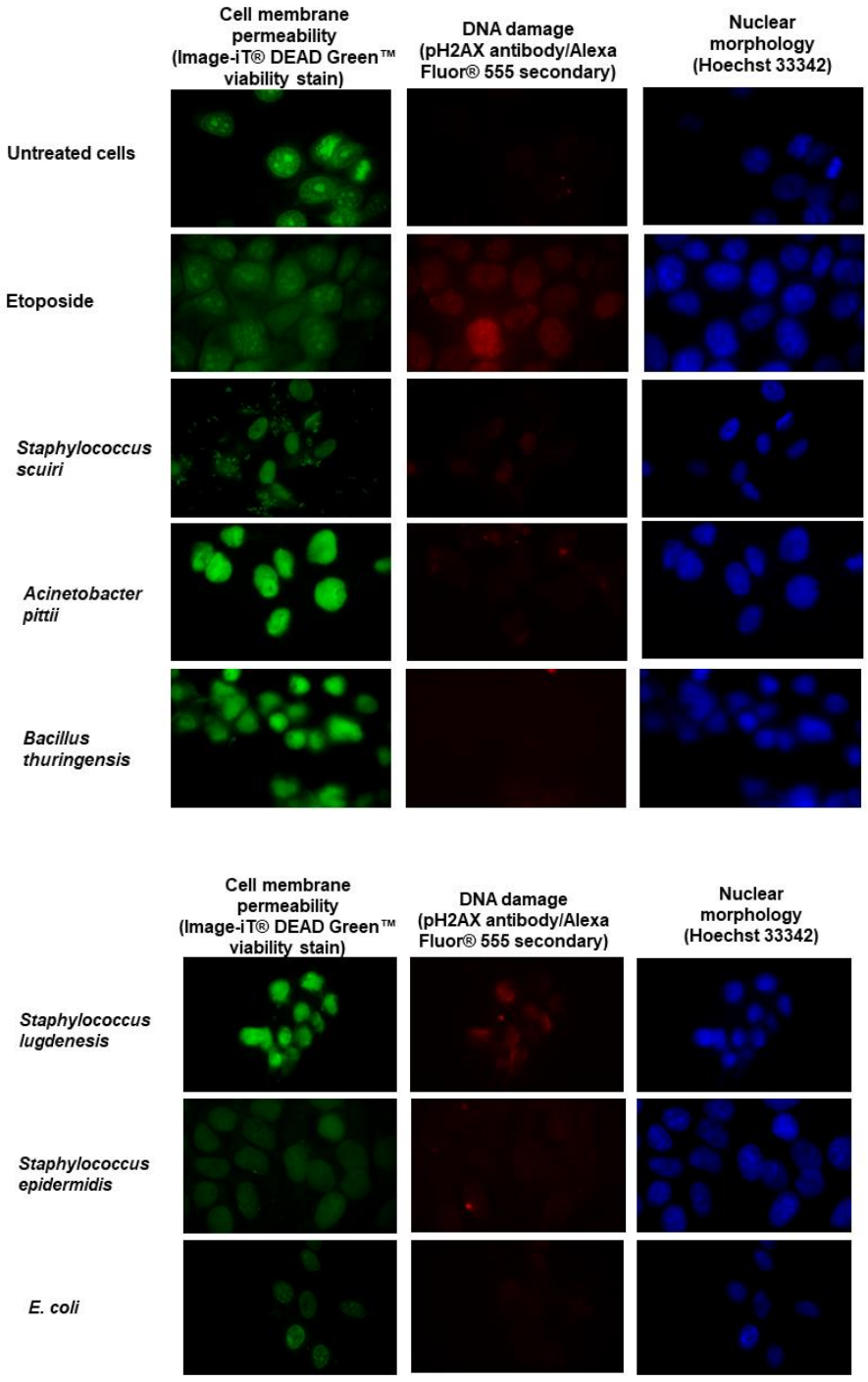


Figure 4.18: **Immunofluorescence images captured from the infection assay of MCF-7 cells at an MOI of 100**. Bacteria were isolated from breast cancer tissues and the DNA damage abilities of the bacteria were tested using the pH2AX DNA damage assay. Images were visualized using the fluorescent microscope at a magnification of 1000X. Scale bar= 0.5 µM. Etoposide was included as a positive control since is a known inducer of DNA double-strand breaks.

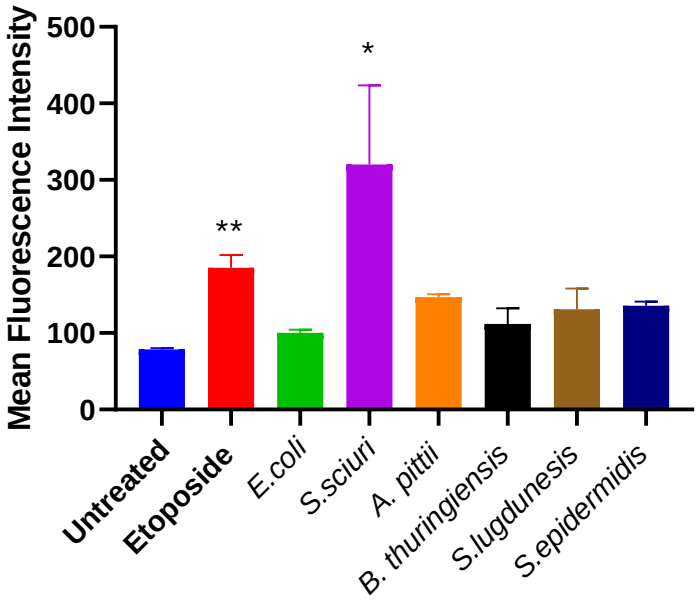


Figure 4.19**: Bacterial strains induce DNA damage in MCF-7 cells at an MOI of 100**. Fluorescent images captured were analyzed using the Image J software and the mean fluorescent intensities of pH2AX were calculated for each bacterium. The data represented for each bacterium on the graph shows the mean and the standard deviations of three cells from three fields of view for each of the isolates tested, showing nine cells for three fields of view for each bacterium. Statistical significance was tested using the Dunn's multiple comparisons test. * $p < 0.05$, ** $p < 0.01$.

4.10.2 Infection at an MOI of 50

MCF-7 cells were cultured and incubated with different bacteria to determine the abilities of the bacteria to cause DNA damage *in vitro*. This was visualized using a fluorescent microscope after 4 hours (figure 4.20). Cells treated with *Staphylococcus sciuri*, *Staphylococcus lugdunensis* and *Staphylococcus epidermidis* were the only isolates which had statistically significantly higher pH2AX intensities when compared with the untreated cells and the others. Etoposide, the positive control also had a significant MFI for pH2AX (figure 4.21).

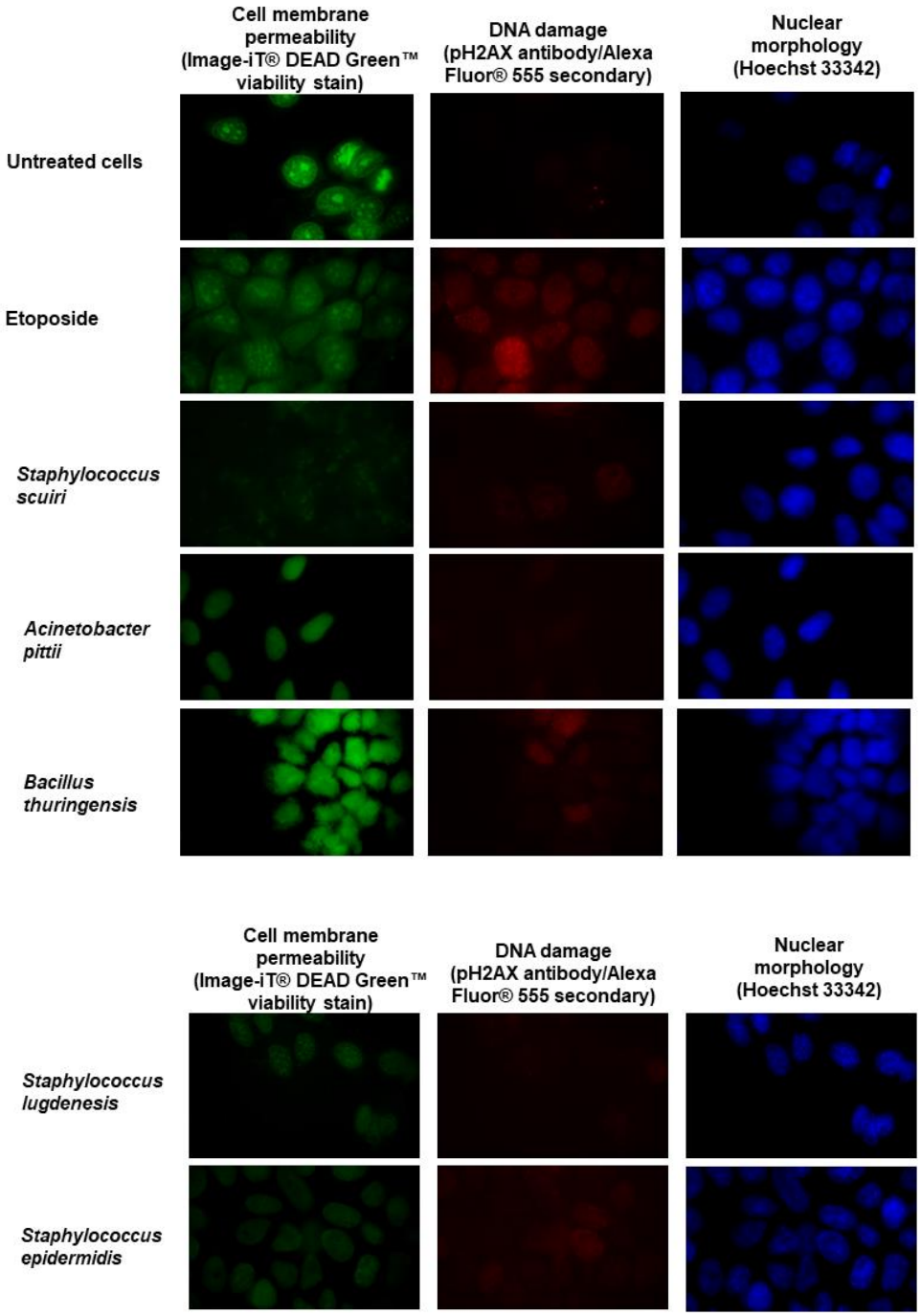


Figure 4.20**: Immunofluorescence images of MCF-7 cells at an MOI of 100**. Bacteria were isolated from breast cancer tissues and the DNA damage abilities of the bacteria were tested using the pH2AX DNA damage assay. Images were visualized using the fluorescent microscope at a magnification of 1000X. Scale bar= 0.5 µM. Etoposide was included as a positive control since is a known inducer of DNA double-strand breaks.

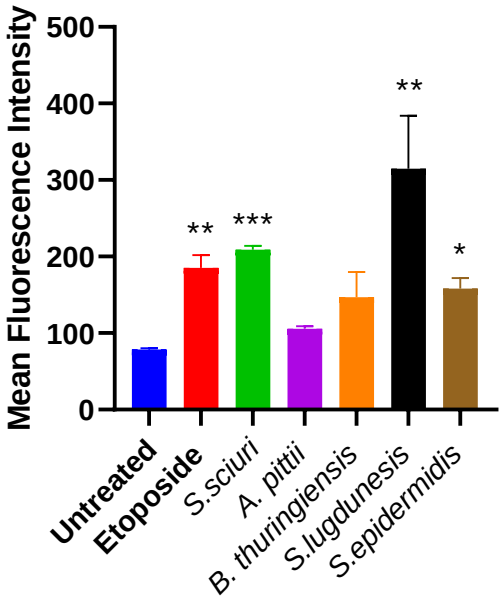


Figure 4.21: **Bacterial strains induce DNA damage at an MOI of 50**. Fluorescent images were analyzed using the Image J software and the mean fluorescent intensities of cells stained with the pH2AX antibody were measured for each bacterium. The data represented for each bacterium on the graph shows the mean and the standard deviations of three cells from three fields of view for each of the isolates tested, showing nine cells for three fields of view for each bacterium. Statistical significance was calculated using the Dunn's multiple comparisons test. * $p < 0.05$, ** $p < 0.01$, *** $p<0.001$.

CHAPTER FIVE

5.0 DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.1 Discussion

The increasing incidence of breast cancer has led to the identification of probable novel factors associated with the malignancy which might serve as risk factors for carcinogenesis. Also, the multifactorial nature of breast cancer emphasizes the need to identify more novel aetiological factors associated with the development of breast cancer. Therefore, this has led to the investigation of biological carcinogens such as bacteria and viruses in breast cancer since they contribute to approximately 18-20% of all cancers (Bouvard *et al*., 2009).

In this study, the *β-Globin* housekeeping gene was amplified in 187 out of 212 FFPE tissues and 17 fresh tissues. This was done to check the integrity of the DNA extracted. A positive amplification shows the DNA is of good quality and can be used for further experiments (Naushad *et al*., 2017). Hence, a total of 204 samples which were positive for the *β-Globin* gene were used for further analyses.

Out of the 204 samples, 66.3% were invasive ductal carcinomas (IDC) with most of them having high grades. This is consistent with other breast cancer studies in Ghanaian women which have shown the IDC histological subtype as the most prevalent type in their study populations (Edmund *et al*., 2013; Naku *et al*., 2016; Quayson *et al*., 2014). Also, another study showed that high-grade tumours (grades II and III) accounted for 85.2% of their study population, an observation also consistent with the Ghanaian population (Adjei, 2012).

Identification of HPV and genotyping

In this study, HPV was detected in 27 out of 204 samples representing 13.2% of the cases. The observations from this study is consistent with findings from other studies in different geographical

regions which reported 18.1%, 5.5% and 8.7% of breast cancer tissues being positive for HPV in women from Pakistan, Iran and Chile, respectively (Aguayo *et al*., 2011; Ghaffari *et al*., 2018; Naushad *et al*., 2017). Some other research groups have also identified a higher prevalence of between 47% and 72% in breast cancer cases in women from the UK and Australia, respectively (Lawson & Glenn, 2017; Salman *et al*., 2017). However, others have also not found any evidence of HPV aetiology in breast cancer patients (Ahangar-Oskouee *et al*., 2014; de Cremoux *et al*., 2008; Lindel *et al*., 2007). These studies use different methods of detection such as PCR, sequencing, dot blot analysis, Western blot and immunohistochemistry due to the low viral load of HPV in breast cancer (Glenn *et al*., 2012; Salman *et al*., 2017). Also, the geographical distribution of the pathogens also accounts for the inconsistent detection of HPV in breast cancer (Obiri-Yeboah *et al*., 2017; Serrano *et al*., 2018)

HPV genotyping was also done to detect the high-risk carcinogenic HPV subtypes present in the cancer cases. Sequencing of all the HPV positive cases identified them as HPV 18. A study in cervical cancer cases in Ghanaian women also identified HPV 18 as the most prevalent subtype (Awua *et al*., 2016). Similarly, a cohort study using RNA-seq data from The Cancer Genome Atlas (TCGA), USA identified HPV 18 was the most identified subtype detected in 55% of breast cancer cases, followed by HPV 16 in 13% of breast cancer cases (Heng *et al*., 2009; Lawson *et al*., 2015). In Australian women, HPV 18 sequences were identified in breast cancer tissues 1-11 years prior to the development of cancer in benign breast tissues (Lawson & Glenn, 2017). However, in another study in the UK, HPV 39 was the most prevalent in 20% of the cases, followed by HPV 18 and 45 in 12% of the cases (Salman *et al*., 2017). A number of reports in women from Chile, Japan and Thailand have also identified HPV 16 as the most prevalent subtype (Aguayo *et al*., 2011; Khan *et al*., 2008; Ngamkham *et al*., 2017).

Breast cancer can originate from the breast milk epithelial cells, making it 'glandular' (Heng *et al*., 2009). HPV18 has been shown to have an affinity or tropism to glandular cells as compared to squamous epithelial cells (Clifford & Franceschi, 2008). Therefore, this could be the reason for the high prevalence of this subtype in some breast cancer studies.

Identification of EBV and genotyping

EBV is classified as a class1 carcinogen by the International Agency for Research on cancer and is implicated in a number of cancers such as nasopharyngeal carcinoma and gastric cancer (Pai *et al*., 2018; Yahia *et al*., 2014). In this study, EBV was identified in 66 (32.4%) of the cases. Consistent with this study include detection of EBV in 6.5% to 26.6% of breast cancer cases according to studies done in women from Chile, Algeria, Egypt, Australia, Pakistan and Iraq. (Aguayo *et al*., 2011; Bensaber *et al*., 2017; El-Naby *et al*., 2017; Glenn *et al*., 2012; Naushad *et al*., 2017; Zekri *et al*., 2012).

According to Perkins *et al*. EBV was identified using minor groove binding (MGB)-TaqMan real-time PCR in both peripheral blood and tumour samples of American participants. A high number of breast cancer biopsies were EBV positive (46%) but they had an extremely low viral load which was similar to the observation from the blood of breast cancer patients (Perkins *et al*., 2006). However, another study in France also measured EBV genome levels in EBV positive breast cancer specimens and found a high viral load in breast biopsies (Arbach *et al*., 2006).

Nevertheless, some studies in breast cancer patients in American women have reported the absence of EBV in their study populations (Baltzell *et al*., 2012; Deshpande *et al*., 2002; Glaser *et al*., 1998; Lespagnard *et al*., 1995; Perrigoue *et al*., 2005).

With these inconsistent findings and heterogenous EBV distribution in cells, studies have tried to find the actual role of EBV in breast cancer. Some concluded that;

i. EBV acts as a carcinogen that is able to maintain telomerase activity *in vivo*, making it a protagonist cocarcinogen in some breast cancers (Xue *et al*., 2003).

ii. EBV infection can lead to epigenetic silencing of tumour suppressors (Yahia *et al*., 2014).

iii. The virus can modulate inflammatory cytokines such as IL-6 and TNF-α secretion, both of which play a role in breast cancer development (Marrao *et al*., 2014; Zekri *et al*., 2012).

iv. Infection with EBV predisposes mammary epithelial cells to malignant transformation into breast cancer cells but is no longer needed once transformation has occurred (Hu *et al*., 2016).

v. Paclitaxel resistance to the MDA-MB-231 metastatic breast cancer cell line is developed during infection with EBV, consequently leading to an overexpression of multidrug-resistant gene (MDR1), which might have therapeutic effects on the treatment of cancer (Arbach *et al*., 2006).

The different results obtained from numerous studies might be due to the geographical variation in the incidence of EBV, the variation in methods used for the detection and the different EBV proteins or nucleic acids targeted (Glaser *et al*., 1998). Some of the methods used for EBV detection are PCR, immunohistochemistry, *in situ* hybridization, laser capture microdissection and Southern blot hybridization (Glaser *et al*., 2004). However, the drawback with the use of PCR is that it cannot differentiate EBV in tumour cells from EBV in surrounding lymphocytes thereby making it difficult to localize the viral genome, even though it is a highly sensitive and specific method for EBV detection (Baltzell *et al*., 2012; Pai *et al*., 2018). Thus, "the gold standard" for detecting EBV latent infection in tumour cells is *in situ* hybridization for EBER-1 (Pai *et al*., 2018; Zekri *et al*., 2012).

EBV typing identified the prevalent subtypes in Ghanaian breast cancer patients. EBV-1 was identified in 2 (1%) of the cases, EBV-2 in 30 (14.7%), EBV-1 and EBV-2 co-infection in 26

(12.7%) of the cases. In general, limited studies have been done on EBV typing, most of which were associated with cancers such as nasopharyngeal carcinoma and Burkitt lymphoma (Hassan *et al*., 2006; Middleton *et al*., 2003; Peh *et al*., 2003). These studies show a high prevalence of EBV-2 and a coinfection with both subtypes in the Equitorial African population which is consistent with the results obtained from this study (Sixbey *et al*., 1989). EBV-1 is however associated with individuals in Asia, Europe and North America (Chen *et al*., 1992; Janani *et al*., 2015). Also, from this study, some EBV positive cases could not be typed which could have been identified if they were sequenced to confirm their EBV positivity.

Co-infection of HPV and EBV

In this study, EBV and HPV co-infection was observed in 5.4% of the study population. Similarly, a study in Pakistani women detected a co-infection of EBV and HPV in 9.2% of breast cancer patients (Naushad *et al*., 2017).

The effects of viral co-infection on oncogenesis are still unclear since studies have shown that in normal breast tissues, EBV alone may not be oncogenic but may become oncogenic in the presence of other viruses (Glenn *et al*., 2012). Also, infection by one virus can occur when there is a chronic inflammation as a result of the other virus (Rickinson, 2014). In addition to that, the co-infection reports have shown that virus-infected cells release inflammatory cytokines that provide a favourable microenvironment for another viral infection (Naushad *et al*., 2017).

Identification of MMTV

MMTV is an oncogenic virus that causes breast cancer in mice. In this study, MMTV was not detected in any of the samples analyzed. Apart from this study, other studies in Japanese and Australian women have also reported the absence of MMTV-like sequences (Fukuoka *et al*., 2008; Park *et al*., 2011; Pogo *et al*., 2011).

Contrary to this, other studies on breast cancer among Saudi Arabian, Australian, American and Iranian women have identified MMTV-*env* like sequences in breast cancer tissues using methods like Western blots, fluorescence-activated cell sorting, PCR and sequencing (Al Dossary *et al*., 2018; Ford *et al*., 2004; Melana *et al*., 2007; Shariatpanahi *et al*., 2017). The first study to discover MMTV-like sequences which was done in American women identified the sequences in 38.5% of breast cancer samples, 6.9% of fibroadenomas. 1.8% of tissues from breast reduction mammoplasties and had no evidence of this virus in normal controls (Wang *et al*., 1995). Another study also suggested that MMTV has the ability to transform epithelial cells and may modulate epithelial to mesenchymal transition of mammary cells (Naushad *et al*., 2017).

Studies have shown that in areas where the *Mus domesticus* species is the resident mouse, there is a higher incidence of breast cancer suggesting a zoonotic transmission of the virus (Stewart *et al*., 2000). Therefore, the absence of this virus in this study could be due to the absence of this species of mice in Ghana (Stewart *et al*., 2000).

Identification of cultured bacteria

Microscopy and biochemical methods were used as the basis for selecting bacterial isolates to sequence. Therefore, 50 out of 190 isolates were identified by sequencing. The bacterial species identified in this study are mainly commensals which can become pathogenic and cause nosocomial infections. They are *Staphylococcus sciuri, Megasphaera cerevisiae, Staphylococcus capitis, Herbaspirillum chlorophenolicum, Staphylococcus warneri, Microbacterium chocolatum, Acinetobacter pittii, Staphylococcus epidermidis, Staphylococcus lugdunensis, Bacillus pumilus* and *Bacillus thuringiensis*.

One of the most abundant bacteria, *Staphylococcus sciuri* is a coagulase-negative commensal bacterium belonging to the phylum Firmicutes (Svec *et al*., 2016). It is an animal-associated

species which has a range of habitats including domestic and wild animals, the environment and humans (Schleifer *et al*., 1983). It causes infections such as endocarditis, urinary tract infections, periodontitis, pelvic inflammatory disease, septic shock and surgical wound infections (Nemeghaire *et al*., 2014; Severin *et al*., 2010; Stepanovic *et al*., 2005). *S. sciuri* has been classified as a pathogen because it has been shown to harbour virulence and antimicrobial resistance genes for other Staphylococci (Nemeghaire *et al*., 2014).

Megasphaera cerevisiae is a gram-negative bacterium belonging to the phylum Firmicutes and was first isolated in 3 to 7% of beer spoilage cases in Europe between 1980 and 2002 (Kutumbaka *et al*., 2015). *M. cerevisiae* can catabolize arginine due to the arc APC operon encoded by its draft genome. Putrescine, a polyamine is a product of arginine catabolism which is involved in the proliferation of ER negative and highly invasive breast tumour cells (Cervelli *et al*., 2014).

Staphylococcus capitis is a subgroup of coagulase-negative *Staphylococci* which is known to harbour several antibiotic resistance genes from non-clinical environments (Nwibo *et al*., 2019). It is usually found on the human skin, mucosal membranes and gut (Li *et al*., 2014). *S. capitis* possesses some virulence factors such as fibronectin-binding protein A which has been shown to act synergistically with collagen by ensuring a constantly altered signalling cascade which leads to the invasiveness of breast cancer (Wang *et al*., 2017).

Herbaspirillum chlorophenolicum is a bacterial species belonging to the phylum Proteobacteria and can thrive on fluoranthene as its main source of carbon and energy (Xu *et al*., 2016). Fluoroanthane is a polycyclic aromatic hydrocarbon which has been shown as an occupational risk factor for sporadic breast cancer development (Brody *et al*., 2007; Korsh *et al*., 2015). *H. chlorophenolicum* has the ability to degrade fluoranthene, thereby playing a beneficial role in breast cancer (Xu *et al*., 2016).

Staphylococcus warneri is able to produce *warneri*cn RK, a haemolytic bactericidal cationic peptide which is cytotoxic to cancer cells but not healthy mononuclear cells (Loiseau *et al*., 2016).

Microbacterium chocolatum belongs to the phylum Actinobacteria and has the ability to produce bioemulsifiers which induce apoptosis in MCF-7 breast cancer cells through a ROS/JNK-mediated mitochondrial response pathway (Cao *et al*., 2010; Cappello *et al*., 2016).

Another bacterial isolate known as *Acinetobacter pittii* belongs to the phylum Proteobacteria and is known to cause most of the *Acinetobacter* infections (Fu *et al*., 2014). It harbours several antimicrobial resistance genes and causes nosocomial infections (Pagano *et al*., 2015).

Staphylococcus lugdunensis is an opportunistic pathogen which infects the skin and soft tissues (Manica & Cohen, 2017). It can cause infections such as peritonitis, soft tissue infections, vascular graft infections, breast and cerebral abscesses, endocarditis and periprosthetic joint infections (Hellbacher *et al*., 2006; Herchline & Ayers, 1991; Lourtet-Hascoët *et al*., 2016)

Staphylococcus epidermidis is an opportunistic pathogen that causes nosocomial infections via its capacity to form biofilms (Zhang *et al*., 2003). *S. epidermidis* also has genes coding for some virulence factors such as fibronectin-binding proteins which might modulate breast cancer development and/or progression (Williams *et al*., 2002).

Bacillus thuringiensis (*Bt*) is a Gram-positive insecticidal bacterium belonging to the phylum firmicutes which is usually isolated from the environment (Ohba *et al*., 2009; Raymond *et al*., 2010). *Bt* is able to produce insecticidal crystal proteins which have cytotoxic effects on breast cancer cells (Brasseur *et al*., 2015). Thereby serving as potential therapeutic and diagnostic tools in cancer therapy.

Bacterial strains induce DNA damage in MCF-7 cells

The DNA damage ability of breast tissue isolates was first assessed by Urbaniak *et al*. They isolated *E. coli* strains belonging to the B2 phylotype which harbour the *pks* pathogenicity island known to encode the machinery for the production of the genotoxin colibactin (Urbaniak *et al*., 2016). These *E. coli* strains are able to induce DNA double-strand breaks and cause chromosomal instabilities according to studies done in colon cancer (Buc *et al*., 2013; Cuevas-Ramos *et al*., 2010; Nougayrede *et al*., 2006).

Therefore, in this study, implication of the presence of these bacteria was identified by determining the DNA damage abilities of breast tissue isolates from Ghanaian women. Four (*S. sciuri, Acinetobacter pittii, Bacillus thuringiensis* and S. *epidermidis*) of the five main isolates were used for the assay together with one of the least abundant isolates (*S. lugdenensis*). *S. sciuri, S. lugdenensis* and *S. epidermidis* induced DNA double-strand breaks in MCF-7 breast cancer cells. The only other study on bacterial isolates from breast cancer as DNA damage agents reported that *E. coli* and *S. epidermidis* isolated from breast cancer patients were able to induce DNA double-strand breaks in HeLa cells using the γ-H2AX phosphorylation assay (Urbaniak *et al*., 2016). *H. pylori* infection has also been shown to cause DNA damage in gastric cancer (Handa *et al*., 2011).

DNA double-strand breaks are the most dangerous type of DNA damage to the cell because they can lead to the loss of very large chromosomal regions (Chang *et al*., 2017). The two major pathways of repairing DSBs are non-homologous end-joining and homologous recombination (Shibata & Jeggo, 2014). Homologous recombination is the pathway that is closely linked to the risk of developing human cancers especially through the involvement of *BRCA1* and *BRCA2* genes (Scully *et al*., 2019). Together with the appropriate risk factors, errors in the DNA damage repair mechanisms can lead to mutations and to genomic instability in human hereditary diseases such

as developmental disorders, premature ageing and cancer (Bernstein *et al*., 2013; Minchom *et al*., 2018). In addition, the risk of cancer is increased in individuals with a germline mutation in a DNA repair or DNA damage response gene (Scully *et al*., 2019). In such individuals, the protein encoded by the gene will not be expressed or will be expressed in a mutated form (Bernstein *et al*., 2013). This eventually leads to the accumulation of the damage, then mutations that might lead to cancer.

Therefore, women who have these impaired DNA repair or DNA checkpoints are more susceptible to DNA damage induced by bacteria and may have an elevated risk of developing breast cancer compared with those without these mutations.

5.2 Conclusion

HPV and EBV were identified in breast cancer tissues of Ghanaian women. A number of bacterial species were also identified, some of which induced double-strand breaks in MCF-7 cells.

However, there is a constant debate on the implications of these infectious agents as aetiological risk factors for breast carcinogenesis and therefore the association of these pathogens and breast cancer still remains controversial.

Nonetheless, the involvement of these agents in the pathogenesis of even a small proportion of breast cancer could have important implications in understanding the aetiology of the disease which might also help in the fight against a considerable number of cancer cases.

5.3 Recommendations

i. Studies should be done to detect viral proteins whose presence will provide detailed information on the association of viruses with breast cancer.

ii. More breast cancer bacterial isolates should be tested to determine their effects on DNA.

iii. A normal cell line should be included in the DNA damage assay to serve as a control for the assay.

iv. The bacterial isolates should be screened for the *pks* pathogenicity island responsible for DNA damage in cancer cells to help strengthen the evidence.

v. The sample size should be increased to help make definitive associations between the pathogens and the risk of developing breast cancer.

vi. Normal breast tissues should be included to serve as controls for both bacterial and viral detection.

vii. Bacterial profiles in breast cancer patients should be identified using *16SrRNA* amplicon next generation sequencing to identify bacterial profiles in breast cancer patients.

REFERENCES

Abed, J., Maalouf, N., Parhi, L., Chaushu, S., Mandelboim, O., & Bachrach, G. (2017). Tumor targeting by Fusobacterium nucleatum: a pilot study and future perspectives. *Frontiers in Cellular and Infection Microbiology, 7*, 295.

Adesunkanmi, A., Lawal, O., Adelusola, K., & Durosimi, M. (2006). The severity, outcome and challenges of breast cancer in Nigeria. *The Breast, 15*(3), 399-409.

Adjei, A. A., Armah, H. B., Gbagbo, F., Boamah, I., Adu-Gyamfi, C., & Asare, I. (2008). Seroprevalence of HHV-8, CMV, and EBV among the general population in Ghana, West Africa. *BioMed Central Infectious Diseases, 8*(1), 111.

Adjei, E. (2012). Breast cancer in Kumasi, Ghana. *Ghana Medical Journal, 46*(1).

Aguayo, F., Khan, N., Koriyama, C., Gonzalez, C., Ampuero, S., Padilla, O., & Akiba, S. (2011). Human papillomavirus and Epstein-Barr virus infections in breast cancer from chile. *Infectious Agents and Cancer, 6*(1), 7.

Ahangar-Oskouee, M., Shahmahmoodi, S., Jalilvand, S., Mahmoodi, M., Ziaee, A. A., Esmaeili, H.-A., & Mollaei-Kandelous, Y. (2014). No detection of 'high-risk'human papillomaviruses in a group of Iranian women with breast cancer. *Asian Pacific Journal of Cancer Prevention, 15*(15), 4061-4065.

Akhter, J., Aziz, M. A. A., Al Ajlan, A., Tulbah, A., & Akhtar, M. (2014). Breast cancer: is there a viral connection? *Advances in Anatomic Pathology, 21*(5), 373-381.

Al Dossary, R., Alkharsah, K. R., & Kussaibi, H. (2018). Prevalence of Mouse Mammary Tumor Virus (MMTV)-like sequences in human breast cancer tissues and adjacent normal breast tissues in Saudi Arabia. *BioMed Central Cancer, 18*(1), 170.

Alibek, K., Kakpenova, A., Mussabekova, A., Sypabekova, M., & Karatayeva, N. (2013). Role of viruses in the development of breast cancer. *Infectious Agents and Cancer, 8*(1), 32.

Allison, K. H. (2012). Molecular pathology of breast cancer: what a pathologist needs to know. *American Journal of Clinical Pathology, 138*(6), 770-780.

Almeida, G. F. G., Mattos, L. A. R., Oliveira Filho, J. B., Carvalho, P., de Brito, B. R., Alcantara, E. R., & Freitas, M. A. L. (2016). DNA repair defect and RAS mutation in Schistosoma mansoni-associated colorectal cancer patients: Carcinogenesis steps or mere coincidence? *Journal of Clinical Oncology, 34*(15), 23279-23279

Amarante, M. K., & Watanabe, M. A. E. (2009). The possible involvement of virus in breast cancer. *Journal of Cancer Research and Clinical Oncology, 135*(3), 329-337.

Arbach, H., Viglasky, V., Lefeu, F., Guinebretiere, J. M., Ramirez, V., Bride, N., & Joab, I. (2006). Epstein-Barr virus (EBV) genome and expression in breast cancer tissue: effect of EBV infection of breast cancer cells on resistance to paclitaxel (Taxol). *Journal of Virology, 80*(2), 845-853.

Armah, F. A., & Gyeabour, E. K. (2013). Health risks to children and adults residing in riverine environments where surficial sediments contain metals generated by active gold mining in Ghana. *Toxicological Research, 29*(1), 69.

Asante, D.-B., Asmah, R. H., Adjei, A. A., Kyei, F., Simpong, D. L., Brown, C. A., & Gyasi, R. K. (2017). Detection of human papillomavirus genotypes and Epstein-Barr virus in nasopharyngeal carcinomas at the korle-bu teaching hospital, Ghana. *The Scientific World Journal, 2017*(9), 2356-6140 .

Atkinson, J. P., & Samter, M. (2001). *Samter's Immunologic Diseases* (Vol. 1, pp. 970-985). Philadelphia, Lippincott Williams & Wilkins.

Awua, A. K., Sackey, S. T., Osei, Y. D., Asmah, R. H., & Wiredu, E. K. (2016). Prevalence of human papillomavirus genotypes among women with cervical cancer in Ghana. *Infectious Agents and Cancer, 11*(1), 4.

Baban, C. K., Cronin, M., O'Hanlon, D., O'Sullivan, G. C., & Tangney, M. (2010). Bacteria as vectors for gene therapy of cancer. *Bioengineered Bugs, 1*(6), 385-394.

Balkwill, F., Charles, K. A., & Mantovani, A. (2005). Smoldering and polarized inflammation in the initiation and promotion of malignant disease. *Cancer Cell, 7*(3), 211-217.

Baltzell, K., Buehring, G. C., Krishnamurthy, S., Kuerer, H., Shen, H. M., & Sison, J. D. (2012). Epstein–Barr virus is seldom found in mammary epithelium of breast cancer tissue using in situ molecular methods. *Breast Cancer Research and Treatment, 132*(1), 267-274.

Banerjee, S., Tian, T., Wei, Z., Shih, N., Feldman, M. D., Peck, K. N., & Robertson, E. S. (2018). Distinct Microbial Signatures Associated With Different Breast Cancer Types. *Frontiers in Microbiology, 9*, 951.

Bashiardes, S., Tuganbaev, T., Federici, S., & Elinav, E. (2017). The microbiome in anti-cancer therapy. *Seminars in Immunology, 32*, 74-81.

Baxter, N. T., Zackular, J. P., Chen, G. Y., & Schloss, P. D. (2014). Structure of the gut microbiome following colonization with human feces determines colonic tumor burden. *Microbiome, 2*(1), 1-11.

Bensaber, H. S., Bicout, D. J., Medjamia, M., Bensnouci, A. M., Comez, A., Chebloune, Y., & Kébir, F. Z. E. (2017). Molecular Detection of Epstein Barr Virus in Women with Breast Cancer in the West Algeria. *Journal of Cancer Therapy, 08*(03), 234-249.

Bernard, H.-U., Burk, R. D., Chen, Z., van Doorslaer, K., Zur Hausen, H., & de Villiers, E.-M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology, 401*(1), 70-79.

Bernstein, C., Prasad, A. R., Nfonsam, V., & Bernstein, H. (2013). DNA damage, DNA repair and cancer. *New Research Directions in DNA Repair*, 413-465.

Bindra, A., Muradrasoli, S., Kisekka, R., Nordgren, H., Warnberg, F., & Blomberg, J. (2007). Search for DNA of exogenous mouse mammary tumor virus-related virus in human breast cancer samples. *Journal of General Virology, 88*(6), 1806-1809.

Bouvard, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F., & Galichet, L. (2009). A review of human carcinogens—Part B: biological agents. *The Lancet Oncology, 10*(4), 321-322.

Brasseur, K., Auger, P., Asselin, E., Parent, S., Côté, J. C., & Sirois, M. (2015). Parasporin-2 from a new *Bacillus thuringiensis 4R2* strain induces caspases activation and apoptosis in human cancer cells. *PLoS One*, *10*(8), e0135106.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 68*(6), 394-424.

Brody, J. G., Moysich, K. B., Humblet, O., Attfield, K. R., Beehler, G. P., & Rudel, R. A. (2007). Environmental pollutants and breast cancer: epidemiologic studies. *Cancer, 109*(12 Suppl), 2667-2711.

Bruni, L., Diaz, M., Castellsagué, M., Ferrer, E., Bosch, F. X., & de Sanjosé, S. (2010). Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *Journal of Infectious Diseases, 202*(12), 1789-1799.

Buc, E., Dubois, D., Sauvanet, P., Raisch, J., Delmas, J., Darfeuille-Michaud, A., & Bonnet, R. (2013). High prevalence of mucosa-associated *E. coli* producing cyclomodulin and genotoxin in colon cancer. *PLoS One, 8*(2), e56964.

Burk, R. D., Harari, A., & Chen, Z. (2013). Human papillomavirus genome variants. *Virology, 445*(1-2), 232-243.

Cao, X. H., Wang, A. H., Wang, C. L., Mao, D. Z., Lu, M. F., Cui, Y. Q., & Jiao, R. Z. (2010). Surfactin induces apoptosis in human breast cancer MCF-7 cells through a ROS/JNK-mediated mitochondrial/caspase pathway. *Chemo-Biological Interactions, 183*(3), 357-362.

Cappello, S., Volta, A., Santisi, S., Morici, C., Mancini, G., Quatrini, P., & Torregrossa, M. (2016). Oil-degrading bacteria from a membrane bioreactor (BF-MBR) system for treatment of saline oily waste: Isolation, identification and characterization of the biotechnological potential. *International Biodeterioration & Biodegradation, 110*, 235-244.

Cavuoto, P., & Fenech, M. F. (2012). A review of methionine dependency and the role of methionine restriction in cancer growth control and life-span extension. *Cancer Treatment Reviews, 38*(6), 726-736.

Cervelli, M., Pietropaoli, S., Signore, F., Amendola, R., & Mariottini, P. (2014). Polyamines metabolism and breast cancer: state of the art and perspectives. *Breast Cancer Research and Treatment, 148*(2), 233-248.

Chan, A. A., Bashir, M., Rivas, M. N., Duvall, K., Sieling, P. A., Pieber, T. R., & Lee, D. J. (2016). Characterization of the microbiome of nipple aspirate fluid of breast cancer survivors. *Scientific Reports, 6*, 28061.

Chang, C. M., Kelly, J. Y., Mbulaiteye, S. M., Hildesheim, A., & Bhatia, K. (2009). The extent of genetic diversity of Epstein-Barr virus and its geographic and disease patterns: a need for reappraisal. *Virus Research, 143*(2), 209-221.

Chang, H. H. Y., Pannunzio, N. R., Adachi, N., & Lieber, M. R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology, 18*(8), 495-506.

Chen, X., Pepper, S. d., & Arrand, J. R. (1992). Prevalence of the A and B types of Epstein-Barr virus DNA in nasopharyngeal carcinoma biopsies from southern China. *Journal of General Virology, 73*(2), 463-466.

Chen, Z., Schiffman, M., Herrero, R., DeSalle, R., Anastos, K., Segondy, M., & Burk, R. D. (2011). Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLoS One, 6*(5), e20183.

Clegg-Lamptey, J. N. A., Dakubo, J. C. B., & Attobra, Y. N. (2009). Psychosocial aspects of breast cancer treatement in Accra, Ghana. *East African Medical Journal, 86*(7).

Cheesbrough, M. (1984). *Medical Laboratory Manual for Tropical Countries,* In *Microbiology* (Vol 2) . Norfolk, Tropical Health Technology.

Clifford, G., & Franceschi, S. (2008). Members of the human papillomavirus type 18 family (alpha-7 species) share a common association with adenocarcinoma of the cervix. *International Journal of Cancer, 122*(7), 1684-1685.

Cogliano, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., & El Ghissassi, F. (2005). Carcinogenicity of human papillomaviruses. *The Lancet Oncology, 6*(4), 204.

Corbex, M., Bouzbid, S., Traverse-Glehen, A., Aouras, H., McKay-Chopin, S., Carreira, C., & Gheit, T. (2014). Prevalence of papillomaviruses, polyomaviruses, and herpesviruses in

triple-negative and inflammatory breast tumors from algeria compared with other types of breast cancer tumors. *PLoS One, 9*(12), e114559.

Costantini, L., Magno, S., Albanese, D., Donati, C., Molinari, R., Filippone, A., & Merendino, N. (2018). Characterization of human breast tissue microbiota from core needle biopsies through the analysis of multi hypervariable 16S-rRNA gene regions. *Scientific Reports, 8*(1), 16893.

Crawford, D. H. (2001). Biology and disease associations of Epstein–Barr virus. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 356*(1408), 461-473.

Cuevas-Ramos, G., Petit, C. R., Marcq, I., Boury, M., Oswald, E., & Nougayrede, J. P. (2010). *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America, 107*(25), 11537-11542.

Czarneski, J., Rassa, J. C., & Ross, S. R. (2003). Mouse mammary tumor virus and the immune system. *Immunologic Research, 27*(2-3), 469-479.

Dalmasso, G., Cougnoux, A., Delmas, J., Darfeuille-Michaud, A., & Bonnet, R. (2014). The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes, 5*(5), 675-680.

de Cremoux, P., Thioux, M., Lebigot, I., Sigal-Zafrani, B., Salmon, R., & Sastre-Garau, X. (2008). No evidence of human papillomavirus DNA sequences in invasive breast carcinoma. *Breast Cancer Research and Treatment, 109*(1), 55-58.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., & Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America, 107*(33), 14691-14696.

de Martel, C., Ferlay, J., Franceschi, S., Vignat, J., Bray, F., Forman, D., & Plummer, M. (2012). Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *The Lancet Oncology, 13*(6), 607-615.

De Villiers, E.-M., Fauquet, C., Broker, T. R., Bernard, H.-U., & Zur Hausen, H. (2004). Classification of papillomaviruses. *Virology, 324*(1), 17-27.

DeSantis, C. E., Fedewa, S. A., Goding Sauer, A., Kramer, J. L., Smith, R. A., & Jemal, A. (2016). Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA: A Cancer Journal for Clinicians, 66*(1), 31-42.

Deshpande, C., Badve, S., Kidwai, N., & Longnecker, R. (2002). Lack of expression of the Epstein-Barr Virus (EBV) gene products, EBERs, EBNA1, LMP1, and LMP2A, in breast cancer cells. *Laboratory Investigation, 82*(9), 1193.

Dimri, G., Band, H., & Band, V. (2005). Mammary epithelial cell transformation: insights from cell culture and mouse models. *Breast Cancer Research, 7*(4), 171.

Djavadian, R., Hayes, M., & Johannsen, E. (2018). CAGE-seq analysis of Epstein-Barr virus lytic gene transcription: 3 kinetic classes from 2 mechanisms. *PLoS Pathogens, 14*(6), e1007114.

Dolan, A., Addison, C., Gatherer, D., Davison, A. J., & McGeoch, D. J. (2006). The genome of Epstein–Barr virus type 2 strain AG876. *Virology, 350*(1), 164-170.

Domfeh, A., Wiredu, E., Adjei, A., Ayeh-Kumi, P., Adiku, T., Tettey, Y., & Armah, H. (2008). Cervical human papillomavirus infection in Accra, Ghana. *Ghana Medical Journal, 42*(2).

Donnet-Hughes, A., Perez, P. F., Doré, J., Leclerc, M., Levenez, F., Benyacoub, J., & Schiffrin, E. J. (2010). Potential role of the intestinal microbiota of the mother in neonatal immune education. *Proceedings of the Nutrition Society, 69*(3), 407-415.

Doorbar, J., & Gallimore, P. H. (1987). Identification of proteins encoded by the L1 and L2 open reading frames of human papillomavirus 1a. *Journal of Virology, 61*(9), 2793-2799.

Dudley, J. P., Golovkina, T. V., & Ross, S. R. (2016). Lessons learned from mouse mammary tumor virus in animal models. *Institute for Laboratory Animal Research Journal, 57*(1), 12-23.

Dumitrescu, R., & Cotarla, I. (2005). Understanding breast cancer risk-where do we stand in 2005? *Journal of Cellular and Molecular Medicine, 9*(1), 208-221.

Edmiston Jr, C. E., Walker, A. P., Krepel, C. J., & Gohr, C. (1990). The nonpuerperal breast infection: aerobic and anaerobic microbial recovery from acute and chronic disease. *Journal of Infectious Diseases, 162*(3), 695-699.

Edmund, D. M., Naaeder, S. B., Tettey, Y., & Gyasi, R. K. (2013). Breast cancer in Ghanaian women: what has changed? *American Journal of Clinical Pathology, 140*(1), 97-102.

Eftekhaar, N. S., Niya, M. H. K., Izadi, F., Teaghinezhad-S, S., & Keyvani, H. (2017). Human papillomavirus (HPV) genotype distribution in patients with recurrent respiratory papillomatosis (RRP) in Iran. *Asian Pacific Journal of Cancer Prevention, 18*(7), 1973.

El-Bayoumy, K., Chae, Y.-H., Upadhyaya, P., Rivenson, A., Kurtzke, C., Reddy, B., & Hecht, S. S. (1995). Comparative tumorigenicity of benzo [a] pyrene, 1-nitropyrene and 2-amino-1-methyl-6-phenylimidazo [4, 5-b] pyridine administered by gavage to female CD rats. *Carcinogenesis, 16*(2), 431-434.

El-Naby, N. E. H., Hassan Mohamed, H., Mohamed Goda, A., & El Sayed Mohamed, A. (2017). Epstein-Barr virus infection and breast invasive ductal carcinoma in Egyptian women: A single center experience. *Journal of the Egyptian National Cancer Institute, 29*(2), 77-82.

Elliott, A. C., & Hynan, L. S. (2011). A SAS® macro implementation of a multiple comparison post hoc test for a Kruskal–Wallis analysis. *Computer Methods and Programs in Biomedicine, 102*(1), 75-80.

Ellis, I., Galea, M., Broughton, N., Locker, A., Blamey, R., & Elston, C. (1992). Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up. *Histopathology, 20*(6), 479-489.

Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology, 19*(5), 403-410.

Estep, A. L., Palmer, C., McCormick, F., & Rauen, K. A. (2007). Mutation analysis of BRAF, MEK1 and MEK2 in 15 ovarian cancer cell lines: implications for therapy. *PLoS One, 2*(12), e1279.

Etkind, P., Du, J., Khan, A., Pillitteri, J., & Wiernik, P. H. (2000). Mouse mammary tumor virus-like *ENV* gene sequences in human breast tumors and in a lymphoma of a breast cancer patient. *Clinical Cancer Research, 6*(4), 1273-1278.

Faedo, M., Hinds, L. A., Singleton, G. R., & Rawlinson, W. D. (2007). Prevalence of mouse mammary tumor virus (MMTV) in wild house mice (Mus musculis) in southeastern Australia. *Journal of Wildlife Diseases, 43*(4), 668-674.

Farmer, J., Davis, B. R., Hickman-Brenner, F. W., McWhorter, A., Huntley-Carter, G. P., Asbury, M. A., & Fanning, G. R. (1985). Biochemical identification of new species and biogroups

of *Enterobacteriaceae* isolated from clinical specimens. *Journal of Clinical Microbiology*, *21*(1), 46-76.

Fais, T., Delmas, J., Barnich, N., Bonnet, R., & Dalmasso, G. (2018). Colibactin: More Than a New Bacterial Toxin. *Toxins*, *10*(4).

Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., & Luo, W. (2018). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes & Diseases*, *5*(2), 77-106.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., & Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, *136*(5), E359-E386.

Fernandez, M. F., Reina-Perez, I., Astorga, J. M., Rodriguez-Carrillo, A., Plaza-Diaz, J., & Fontana, L. (2018). Breast Cancer and Its Relationship with the Microbiota. *International Journal of Environmental Research and Public Health*, *15*(8).

Ford, C. E., Faedo, M., & Rawlinson, W. D. (2004). Mouse mammary tumor virus-like RNA transcripts and DNA are found in affected cells of human breast cancer. *Clinical Cancer Research*, *10*(21), 7284-7289.

Franceschi, S., Herrero, R., Clifford, G. M., Snijders, P. J., Arslan, A., Anh, P. T. H., & Lazcano-Ponce, E. (2006). Variations in the age-specific curves of human papillomavirus prevalence in women worldwide. *International Journal of Cancer*, *119*(11), 2677-2684.

Francescone, R., Hou, V., & Grivennikov, S. I. (2014). Microbiome, inflammation, and cancer. *Cancer Journal*, *20*(3), 181-189.

Frisan, T., Levitsky, V., & Masucci, M. (2001). Generation of lymphoblastoid cell lines (LCLs). In *Methods in Molecular Biology* (Vol 174, pp. 125-127). Totowa, Humana Press.

Fu, Y., Jiang, J., Zhou, H., Jiang, Y., Fu, Y., Yu, Y., & Zhou, J. (2014). Characterization of a novel plasmid type and various genetic contexts of *bla OXA-58* in *Acinetobacter spp*. from multiple cities in China. *PLoS One*, *9*(1), e84680.

Fukuoka, H., Moriuchi, M., Yano, H., Nagayasu, T., & Moriuchi, H. (2008). No association of mouse mammary tumor virus-related retrovirus with Japanese cases of breast cancer. *Journal of Medical Virology*, *80*(8), 1447-1451.

Garcia-Tunon, I., Ricote, M., Ruiz, A. A., Fraile, B., Paniagua, R., & Royuela, M. (2007). Influence of IFN-gamma and its receptors in human breast cancer. *BioMed Central Cancer*, *7*, 158.

Ghaffari, H., Nafissi, N., Hashemi-Bahremani, M., Alebouyeh, M. R., Tavakoli, A., Javanmard, D., & Monavari, S. H. (2018). Molecular prevalence of human papillomavirus infection among Iranian women with breast cancer. *Breast Disease*, *37*(4), 207-213.

Glaser, S. L., Ambinder, R. F., DiGiuseppe, J. A., Horn-Ross, P. L., & Hsu, J. L. (1998). Absence of Epstein-Barr virus EBER-1 transcripts in an epidemiologically diverse group of breast cancers. *International Journal of Cancer*, *75*(4), 555-558.

Glaser, S. L., Hsu, J. L., & Gulley, M. L. (2004). Epstein-Barr virus and breast cancer: state of the evidence for viral carcinogenesis. *Cancer Epidemiology and Prevention Biomarkers*, *13*(5), 688-697.

Glenn, W. K., Heng, B., Delprado, W., Iacopetta, B., Whitaker, N. J., & Lawson, J. S. (2012). Epstein-Barr virus, human papillomavirus and mouse mammary tumour virus as multiple viruses in breast cancer. *PLoS One*, *7*(11), e48788.

Goedert, J. J., Jones, G., Hua, X., Xu, X., Yu, G., Flores, R., & Ravel, J. (2015). Investigation of the association between the fecal microbiota and breast cancer in postmenopausal women:

a population-based case-control pilot study. *Journal of the National Cancer Institute, 107*(8).

Going, J. J., & Moffat, D. F. (2004). Escaping from Flatland: clinical and biological aspects of human mammary duct anatomy in three dimensions. *The Journal of Pathology, 203*(1), 538-544.

Graham, S. V. (2010). Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies. *Future Microbiology, 5*(10), 1493-1506.

Guan, J., Bywaters, S. M., Brendle, S. A., Ashley, R. E., Makhov, A. M., Conway, J. F., & Hafenstein, S. (2017). Cryoelectron microscopy maps of human papillomavirus 16 reveal L2 densities and heparin binding site. *Structure, 25*(2), 253-263.

Gujar, S., Pol, J. G., & Kroemer, G. (2018). Heating it up: Oncolytic viruses make tumors 'hot' and suitable for checkpoint blockade immunotherapies. *Oncoimmunology, 7*(8), e1442169.

Han, X. Y., & Andrade, R. A. (2005). Brevundimonas diminuta infections and its resistance to fluoroquinolones. *Journal of Antimicrobial Chemotherapy, 55*(6), 853-859.

Handa, O., Naito, Y., & Yoshikawa, T. (2011). Redox biology and gastric carcinogenesis: the role of *Helicobacter pylori*. *Redox Report, 16*(1), 1-7.

Hassan, R., White, L. R., Stefanoff, C. G., de Oliveira, D. E., Felisbino, F. E., Klumb, C. E., & Zalcberg, I. R. (2006). Epstein-Barr virus (EBV) detection and typing by PCR: a contribution to diagnostic screening of EBV-positive Burkitt's lymphoma. *Diagnostic Pathology, 1*(1), 17.

Haswell-Elkins, M. R., Mairiang, E., Mairiang, P., Chaiyakum, J., Chamadol, N., Loapaiboon, V., & Elkins, D. B. (1994). Cross-sectional study of *Opisthorchis viverrini* infection and cholangiocarcinoma in communities within a high-risk area in northeast Thailand. *International Journal of Cancer, 59*(4), 505-509.

Hellbacher, C., Törnqvist, E., & Söderquist, B. (2006). *Staphylococcus lugdunensis*: clinical spectrum, antibiotic susceptibility, and phenotypic and genotypic patterns of 39 isolates. *Clinical Microbiology and Infection, 12*(1), 43-49.

Heng, B., Glenn, W., Ye, Y., Tran, B., Delprado, W., Lutze-Mann, L., & Lawson, J. (2009). Human papilloma virus is associated with breast cancer. *British Journal of Cancer, 101*(8), 1345.

Herbst, L. H., Lenz, J., Van Doorslaer, K., Chen, Z., Stacy, B. A., Wellehan Jr, J. F., & Burk, R. D. (2009). Genomic characterization of two novel reptilian papillomaviruses, Chelonia mydas papillomavirus 1 and Caretta caretta papillomavirus 1. *Virology, 383*(1), 131-135.

Herchline, T., & Ayers, L. W. (1991). Occurrence of *Staphylococcus lugdunensis* in consecutive clinical cultures and relationship of isolation to infection. *Journal of Clinical Microbiology, 29*(3), 419-421.

Hessle, C., Andersson, B., & Wold, A. E. (2000). Gram-positive bacteria are potent inducers of monocytic interleukin-12 (IL-12) while gram-negative bacteria preferentially stimulate IL-10 production. *Infection and Immunity, 68*(6), 3581-3586.

Hieken, T. J., Chen, J., Hoskin, T. L., Walther-Antonio, M., Johnson, S., Ramaker, S., & Degnim, A. C. (2016). The Microbiome of Aseptically Collected Human Breast Tissue in Benign and Malignant Disease. *Scientific Reports, 6*, 30751.

Hill, D., Seaton, R., Cameron, F., McLellan, A., Brown, R., & France, A. (1998). Severe sepsis caused by *Mobiluncus curtisii* subsp. curtisii in a previously healthy female: case report and review. *Journal of Infection, 37*(2), 194-196.

Hochberg, D., Middeldorp, J. M., Catalina, M., Sullivan, J. L., Luzuriaga, K., & Thorley-Lawson, D. A. (2004). Demonstration of the Burkitt's lymphoma Epstein-Barr virus phenotype in dividing latently infected memory cells in vivo. *Proceedings of the National Academy of Sciences, 101*(1), 239-244.

Hodge, D. R., Hurt, E. M., & Farrar, W. L. (2005). The role of IL-6 and STAT3 in inflammation and cancer. *European Journal of Cancer, 41*(16), 2502-2512.

Hsu, W.-L., Lin, H.-Y., Chiou, S.-S., Chang, C.-C., Wang, S.-P., Lin, K.-H., & Chang, S.-C. (2010). Mouse mammary tumor virus-like nucleotide sequences in canine and feline mammary tumors. *Journal of Clinical Microbiology, 48*(12), 4354-4362.

Hu, H., Luo, M. L., Desmedt, C., Nabavi, S., Yadegarynia, S., Hong, A., & Wulf, G. M. (2016). Epstein-Barr Virus Infection of Mammary Epithelial Cells Promotes Malignant Transformation. *EBioMedicine, 9*, 148-160.

Indik, S., Gunzburg, W. H., Kulich, P., Salmons, B., & Rouault, F. (2007). Rapid spread of mouse mammary tumor virus in cultured human breast cells. *Retrovirology, 4*, 73.

Inan, M. S., Rasoulpour, R. J., Yin, L., Hubbard, A. K., Rosenberg, D. W., & Giardina, C. (2000). The luminal short-chain fatty acid butyrate modulates NF-κB activity in a human colonic epithelial cell line. *Gastroenterology, 118*(4), 724-734.

Inic, Z., Zegarac, M., Inic, M., Markovic, I., Kozomara, Z., Djurisic, I., & Jancic, S. (2014). Difference between luminal A and luminal B subtypes according to Ki-67, tumor size, and progesterone receptor negativity providing prognostic information. *Clinical Medicine Insights: Oncology, 8*, CMO. S18006.

Inoue, D., Hara, S., Kashihara, M., Murai, Y., Danzl, E., Sei, K., Tsunoi, S., Fujita, M., & Ike, M. (2008). Degradation of Bis(4-Hydroxyphenyl)methane (bisphenol F) by *Sphingobium yanoikuyae* strain FM-2 isolated from river water. *Applied and Environmental Microbiology, 74*(2), 352–358. doi: 10.1128/AEM.01708-07.

Janani, M. K., Malathi, J., Biswas, J., Sridharan, S., & Madhavan, H. N. (2015). Genotypic detection of Epstein Barr virus from clinically suspected viral retinitis patients in a tertiary eye care centre, India. *Ocular Immunology and Inflammation, 23*(5), 384-391.

Janicek, M. F., & Averette, H. E. (2001). Cervical cancer: prevention, diagnosis, and therapeutics. *CA: A Cancer Journal for Clinicians, 51*(2), 92-114.

Jiagge, E., Oppong, J. K., Bensenhaver, J., Aitpillah, F., Gyan, K., Kyei, I., & Toy, K. (2016). Breast cancer and African ancestry: Lessons learned at the 10-year anniversary of the Ghana-Michigan research partnership and International Breast Registry. *Journal of Global Oncology, 2*(5), 302-310.

Kaspersen, M. D., Larsen, P. B., Ingerslev, H. J., Fedder, J., Petersen, G. B., Bonde, J., & Höllsberg, P. (2011). Identification of multiple HPV types on spermatozoa from human sperm donors. *PLoS One, 6*(3), e18095.

Katz, E., Lareef, M. H., Rassa, J. C., Grande, S. M., King, L. B., Russo, J., Monroe, & J. G. (2005). MMTV Env encodes an ITAM responsible for transformation of mammary epithelial cells in three-dimensional culture. *Journal of Experimental Medicine, 201*(3), 431-439.

Khan, N., Castillo, A., Koriyama, C., Kijima, Y., Umekita, Y., Ohi, Y., & Tsuji, T. (2008). Human papillomavirus detected in female breast carcinomas in Japan. *British Journal of Cancer, 99*(3), 408.

Khodabandehlou, N., Mostafaei, S., Etemadi, A., Ghasemi, A., Payandeh, M., Hadifar, S., & Moghoofei, M. (2019). Human papilloma virus and breast cancer: the role of inflammation and viral expressed proteins. *BioMed Central Cancer, 19*(1), 61.

Kincaid, R. P., & Sullivan, C. S. (2012). Virus-encoded microRNAs: an overview and a look to the future. *PLoS Pathogens, 8*(12), e1003018.

Knight, B., Yeoh, G. C., Husk, K. L., Ly, T., Abraham, L. J., Yu, C., & Fausto, N. (2000). Impaired preneoplastic changes and liver tumor formation in tumor necrosis factor receptor type 1 knockout mice. *Journal of Experimental Medicine, 192*(12), 1809-1818.

Komori, A., Yatsunami, J., Suganuma, M., Okabe, S., Abe, S., Sakai, A., & Fujiki, H. (1993). Tumor necrosis factor acts as a tumor promoter in BALB/3T3 cell transformation. *Cancer Research, 53*(9), 1982-1985.

Korde, L. A., Zujewski, J. A., Kamin, L., Giordano, S., Domchek, S., Anderson, W. F., & Bergh, J. (2010). Multidisciplinary meeting on male breast cancer: summary and research recommendations. *Journal of Clinical Oncology, 28*(12), 2114.

Korsh, J., Shen, A., Aliano, K., & Davenport, T. (2015). Polycyclic Aromatic Hydrocarbons and Breast Cancer: A Review of the Literature. *Breast Care, 10*(5), 316-318.

Kouloura, A., Nicolaidou, E., Misitzis, I., Panotopoulou, E., Smyrniotis, V., Corso, G., & Arkadopoulos, N. (2018). HPV infection and breast cancer. Results of a microarray approach. *The Breast, 40*, 165-169.

Kranjec, C., Holleywood, C., Libert, D., Griffin, H., Mahmood, R., Isaacson, E., & Doorbar, J. (2017). Modulation of basal cell fate during productive and transforming HPV-16 infection is mediated by progressive E6-driven depletion of Notch. *The Journal of Pathology, 242*(4), 448-462.

Kutumbaka, K. K., Pasmowitz, J., Mategko, J., Reyes, D., Friedrich, A., Han, S., & Samadpour, M. (2015). Draft Genome Sequence of the Beer Spoilage Bacterium *Megasphaera cerevisiae* Strain PAT 1T. *Genome Announcements, 3*(5).

Kwa, M., Plottel, C. S., Blaser, M. J., & Adams, S. (2016). The Intestinal Microbiome and Estrogen Receptor-Positive Female Breast Cancer. *Journal of the National Cancer Institute, 108*(8).

Labrecque, L. G., Barnes, D. M., Fentiman, I. S., & Griffin, B. E. (1995). Epstein-Barr virus in epithelial cell tumors: a breast cancer study. *Cancer Research, 55*(1), 39-45.

Laryea, D. O., Awuah, B., Amoako, Y. A., Osei-Bonsu, E., Dogbe, J., Larsen-Reindorf, R., & Konney, T. O. (2014). Cancer incidence in Ghana, 2012: evidence from a population-based cancer registry. *BioMed Central Cancer, 14*(1), 362.

Lawson, J. S., & Glenn, W. K. (2017). Multiple oncogenic viruses are present in human breast tissues before development of virus associated breast cancer. *Infectious Agents Cancer, 12*, 55.

Lawson, J. S., Glenn, W. K., Salyakina, D., Delprado, W., Clay, R., Antonsson, A., & Whitaker, N. J. (2015). Human Papilloma Viruses and Breast Cancer. *Frontiers in Oncology, 5*, 277.

Lawson, J. S., & Heng, B. (2010). Viruses and breast cancer. *Cancers (Basel), 2*(2), 752-772.

Lay, M.-L. J., Lucas, R. M., Ratnamohan, M., Taylor, J., Ponsonby, A.-L., & Dwyer, D. E. (2010). Measurement of Epstein-Barr virus DNA load using a novel quantification standard containing two EBV DNA targets and SYBR Green I dye. *Virology Journal, 7*(1), 252.

Lee, M., Huang, Y., Liao, C.-H., Chuang, T., Lin, C., Lee, S., & Hsueh, P. (2011). Bacteremia caused by *Brevundimonas* species at a tertiary care hospital in Taiwan, 2000–2010. *European Journal of Clinical Microbiology & Infectious Diseases, 30*(10), 1185-1191.

Lespagnard, L., Cochaux, P., Larsimont, D., Degeyter, M., Velu, T., & Heimann, R. (1995). Absence of Epstein-Barr virus in medullary carcinoma of the breast as demonstrated by immunophenotyping, in situ hybridization and polymerase chain reaction. *American Journal of Clinical Pathology, 103*(4), 449-452.

Li, X., Lei, M., Song, Y., Gong, K., Li, L., Liang, H., & Jiang, X. (2014). Whole genome sequence and comparative genomic analysis of multidrug-resistant *Staphylococcus capitis subsp. urealyticus* strain *LNZR-1*. *Gut Pathogens, 6*(1), 45.

Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M., & Hemminki, K. (2000). Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England Journal of Medicine, 343*(2), 78-85.

Lindel, K., Forster, A., Altermatt, H. J., Greiner, R., & Gruber, G. (2007). Breast cancer and human papillomavirus (HPV) infection: no evidence of a viral etiology in a group of Swiss women. *The Breast, 16*(2), 172-177.

Liu, X., Li, F., Huang, Q., Zhang, Z., Zhou, L., Deng, Y., & Li, C. Y. (2017). Self-inflicted DNA double-strand breaks sustain tumorigenicity and stemness of cancer cells. *Cell Research, 27*(6), 764-783.

Loiseau, C., Augenstreich, J., Marchand, A., Harte, E., Garcia, M., Verdon, J., & Berjeaud, J. M. (2016). Specific Anti-Leukemic Activity of the Peptide *Warneri*cin RK and Analogues and Visualization of Their Effect on Cancer Cells by Chemical Raman Imaging. *PLoS One, 11*(9), e0162007.

Lourtet-Hascoët, J., Bicart-See, A., Félicé, M., Giordano, G., & Bonnet, E. (2016). *Staphylococcus lugdunensis*, a serious pathogen in periprosthetic joint infections: comparison to *Staphylococcus aureus* and *Staphylococcus epidermidis*. *International Journal of Infectious Diseases, 51*, 56-61.

Malekpour Afshar, R., Balar, N., Mollaei, H. R., Arabzadeh, S. A., & Iranpour, M. (2018). Low Prevalence of Human Papilloma Virus in Patients with Breast Cancer, Kerman; Iran. *Asian Pacific Journal of Cancer Prevention, 19*(11), 3039-3044.

Manica, L. A. H., & Cohen, P. R. (2017). *Staphylococcus lugdunensis* infections of the skin and soft tissue: a case series and review. *Dermatology and Therapy, 7*(4), 555-562.

Mann, G. J., Thorne, H., Balleine, R. L., Butow, P. N., Clarke, C. L., Edkins, E., & Gattas, M. (2006). Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Research, 8*(1), R12.

Marrao, G., Habib, M., Paiva, A., Bicout, D., Fallecker, C., Franco, S., & Drouet, E. (2014). Epstein-Barr virus infection and clinical outcome in breast cancer patients correlate with immune cell TNF-alpha/IFN-gamma response. *BioMed Central Cancer, 14*, 665.

Masrour-Roudsari, J., & Ebrahimpour, S. (2017). Causal role of infectious agents in cancer: An overview. *Caspian Journal of Internal Medicine, 8*(3), 153.

Matsuura, H., Kirschner, A. N., Longnecker, R., & Jardetzky, T. S. (2010). Crystal structure of the Epstein-Barr virus (EBV) glycoprotein H/glycoprotein L (gH/gL) complex. *Proceedings of the National Academy of Sciences of the United States of America, 107*(52), 22641-22646.

Mazouni, C., Fina, F., Romain, S., Ouafik, L., Bonnier, P., Brandone, J. M., & Martin, P. M. (2011). Epstein-Barr virus as a marker of biological aggressiveness in breast cancer. *British Journal of Cancer, 104*(2), 332-337.

Mazzanti, C. M., Al Hamad, M., Fanelli, G., Scatena, C., Zammarchi, F., Zavaglia, K., & Bevilacqua, G. (2011). A mouse mammary tumor virus env-like exogenous sequence is strictly related to progression of human sporadic breast carcinoma. *The American Journal of Pathology, 179*(4), 2083-2090.

McGinnis, S., & Madden, T. L. (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research, 32*, 20-25.

McLaughlin-Drubin, M. E., & Munger, K. (2008). Viruses associated with human cancer. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1782*(3), 127-150.

McKenzie, J., & El-Guindy, A. (2015). Epstein-Barr virus lytic cycle reactivation. In *Epstein Barr Virus* (Vol. 2, pp. 237-261). Cham, Springer.

Melana, S. M., Nepomnaschy, I., Sakalian, M., Abbott, A., Hasa, J., Holland, J. F., & Pogo, B. G. (2007). Characterization of viral particles isolated from primary cultures of human breast cancer cells. *Cancer Research, 67*(18), 8960-8965.

Meng, S., Chen, B., Yang, J., Wang, J., Zhu, D., Meng, Q., & Zhang, L. (2018). Study of Microbiomes in Aseptically Collected Samples of Human Breast Tissue Using Needle Biopsy and the Potential Role of *in situ* Tissue Microbiomes for Promoting Malignancy. *Frontiers in Oncology, 8*, 318.

Middeldorp, J. M., Brink, A. A., Van den Brule, A. J., & Meijer, C. J. (2003). Pathogenic roles for Epstein–Barr virus (EBV) gene products in EBV-associated proliferative disorders. *Critical Reviews in Oncology/Hematology, 45*(1), 1-36.

Middleton, K., Peh, W., Southern, S., Griffin, H., Sotlar, K., Nakahara, T., & Hibma, M. (2003). Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *Journal of Virology, 77*(19), 10186-10201.

Minchom, A., Aversa, C., & Lopez, J. (2018). Dancing with the DNA damage response: next-generation anti-cancer therapeutic strategies. *Therapeutic Advances in Medical Oncology, 10*.

Mira-Pascual, L., Cabrera-Rubio, R., Ocon, S., Costales, P., Parra, A., Suarez, A., & Collado, M. (2015). Microbial mucosal colonic shifts associated with the development of colorectal cancer reveal the presence of different bacterial and archaeal biomarkers. *Journal of Gastroenterology, 50*(2), 167-179.

Miranda, C. A., Martins, O. B., & Clementino, M. M. (2008). Species-level identification of *Bacillus* strains isolates from marine sediments by conventional biochemical, *16SrRNA* gene sequencing and inter-tRNA gene sequence lengths analysis. *Antonie van Leeuwenhoek, 93*(3), 297-304.

Mishra, K. N., Aaggarwal, A., Abdelhadi, E., & Srivastava, D. (2010). An efficient horizontal and vertical method for online dna sequence compression. *International Journal of Computer Applications, 3*(1), 39-46.

Moody, C. A., & Laimins, L. A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer, 10*(8), 550.

Morris, J., & Seifter, E. (1992). The role of aromatic hydrocarbons in the genesis of breast cancer. *Medical Hypotheses, 38*(3), 177-184.

Münger, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M. & Huh, K. (2004). Mechanisms of human papillomavirus-induced oncogenesis. *Journal of Virology, 78*(21), 11451-11460.

Muñoz, N., Bosch, F. X., De Sanjosé, S., Herrero, R., Castellsagué, X., Shah, K. V., & Meijer, C. J. (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New England Journal of Medicine, 348*(6), 518-527.

Murray, P. G., Lissauer, D., Junying, J., Davies, G., Moore, S., Bell, A., & Reynolds, G. M. (2003). Reactivity with a monoclonal antibody to Epstein-Barr virus (EBV) nuclear antigen 1

defines a subset of aggressive breast cancers in the absence of the EBV genome. *Cancer Research, 63*(9), 2338-2343.

Naku, G. J. F., Anyanful, A., Eliason, S., Mohammed Adamu, S., & Debrah, S. (2016). Pattern of Breast Cancer Distribution in Ghana: A Survey to Enhance Early Detection, Diagnosis, and Treatment. *International Journal of Breast Cancer, 2016*, 3645308.

Nartey, T., Mazzanti, C. M., Melana, S., Glenn, W. K., Bevilacqua, G., Holland, J. F., & Pogo, B. G. (2017). Mouse mammary tumor-like virus (MMTV) is present in human breast tissue before development of virally associated breast cancer. *Infectious Agents and Cancer, 12*, 1.

Nasrollahzadeh, D., Malekzadeh, R., Ploner, A., Shakeri, R., Sotoudeh, M., Fahimi, S., & Ye, W. (2015). Variations of gastric corpus microbiota are associated with early esophageal squamous cell carcinoma and squamous dysplasia. *Scientific Reports, 5*, 8820.

Naushad, W., Surriya, O., & Sadia, H. (2017). Prevalence of EBV, HPV and MMTV in Pakistani breast cancer patients: A possible etiological role of viruses in breast cancer. *Infection, Genetics and Evolution, 54*, 230-237.

Nemeghaire, S., Argudin, M. A., Fessler, A. T., Hauschild, T., Schwarz, S., & Butaye, P. (2014). The ecological importance of the *Staphylococcus sciuri* species group as a reservoir for resistance and virulence genes. *Veterinary Microbiology, 171*(3-4), 342-356.

Ngamkham, J., Karalak, A., Chaiwerawattana, A., Sornprom, A., Thanasutthichai, S., Sukarayodhin, S., & Laochan, N. (2017). Prevalence of Human Papillomavirus Infection in Breast Cancer Cells from Thai Women. *Asian Pacific Journal of Cancer Prevention, 18*(7), 1839-1845.

Niedobitek, G., Meru, N., & Delecluse, H. J. (2001). Epstein-Barr virus infection and human malignancies. *International Journal of Experimental Pathology, 82*(3), 149-170.

Nishimukai, A., Yagi, T., Yanai, A., Miyagawa, Y., Enomoto, Y., Murase, K., & Hatada, T. (2015). High Ki-67 expression and low progesterone receptor expression could independently lead to a worse prognosis for postmenopausal patients with estrogen receptor-positive and HER2-negative breast cancer. *Clinical Breast Cancer, 15*(3), 204-211.

Nougayrede, J. P., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G., & Oswald, E. (2006). *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science, 313*(5788), 848-851.

Nwibo, D. D., Panthee, S., Hamamoto, H., & Sekimizu, K. (2019). Molecular characterization of multi-drug resistant coagulase negative cocci in non-hospital environment. *Drug Discoveries and Therapeutics 13*(3):145-149.

O'Hara, A. M., & Shanahan, F. (2006). The gut flora as a forgotten organ. *EMBO Reports, 7*(7), 688-693.

Ohba, M., Mizuki, E., & Uemori, A. (2009). Parasporin, a new anticancer protein group from *Bacillus thuringiensis*. *Anticancer Research*, *29*(1), 427-433.

Obiri-Yeboah, D., Akakpo, P. K., Mutocheluh, M., Adjei-Danso, E., Allornuvor, G., Amoako-Sakyi, D., & Mayaud, P. (2017). Epidemiology of cervical human papillomavirus (HPV) infection and squamous intraepithelial lesions (SIL) among a cohort of HIV-infected and uninfected Ghanaian women. *BioMed Central Cancer, 17*(1), 688.

Pagano, M., Poirel, L., Martins, A. F., Rozales, F. P., Zavascki, A. P., Barth, A. L., & Nordmann, P. (2015). Emergence of NDM-1-producing *Acinetobacter pittii* in Brazil. *International Journal of Antimicrobial Agents, 45*(4), 444-445.

Pai, T., Gupta, S., Gurav, M., Nag, S., Shet, T., Patil, A., & Desai, S. (2018). Evidence for the association of Epstein-Barr Virus in breast cancer in Indian patients using *in-situ* hybridization technique. *Breast Journal, 24*(1), 16-22.

Parida, S., & Sharma, D. (2019). The power of small changes: comprehensive analyses of microbial dysbiosis in breast cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 1871*(2), 392-405.

Park, D. J., Southey, M. C., Giles, G. G., & Hopper, J. L. (2011). No evidence of MMTV-like *env* sequences in specimens from the Australian Breast Cancer Family Study. *Breast Cancer Research and Treatment, 125*(1), 229-235.

Parkin, D. M. (2006). The global health burden of infection-associated cancers in the year 2002. *International Journal of Cancer, 118*(12), 3030-3044.

Patel, T., Bhattacharya, P., & Das, S. (2016). Gut microbiota: an Indicator to Gastrointestinal Tract Diseases. *Journal of Gastrointestinal Cancer, 47*(3), 232-238.

Paul, B., Royston, K. J., Li, Y., Stoll, M. L., Skibola, C. F., Wilson, L. S., & Tollefsbol, T. O. (2017). Impact of genistein on the gut microbiome of humanized mice and its role in breast tumor inhibition. *PLoS One, 12*(12), e0189756.

Peh, S.-C., Kim, L.-H., Mun, K.-S., Tan, E.-L., Sam, C.-K., & Poppema, S. (2003). Epstein-Barr virus (EBV) subtypes and variants in malignant tissue from Malaysian patients. *Journal of Clinical and Experimental Hematopathology, 43*(2), 61-69.

Pereira, F. C., & Berry, D. (2017). Microbial nutrient niches in the gut. *Environmental Microbiology, 19*(4), 1366-1378.

Perkins, R. S., Sahm, K., Marando, C., Dickson-Witmer, D., Pahnke, G. R., Mitchell, M., & Krueger, L. J. (2006). Analysis of Epstein-Barr virus reservoirs in paired blood and breast cancer primary biopsy specimens by real time PCR. *Breast Cancer Research, 8*(6), 1-8.

Perrigoue, J. G., Den Boon, J. A., Friedl, A., Newton, M. A., Ahlquist, P., & Sugden, B. (2005). Lack of association between EBV and breast carcinoma. *Cancer Epidemiology and Prevention Biomarkers, 14*(4), 809-814.

Pinidis, P., Tsikouras, P., Iatrakis, G., Zervoudis, S., Koukouli, Z., Bothou, A., & Vladareanu, S. (2016). Human papilloma virus' life cycle and carcinogenesis. *Maedica, 11*(1), 48.

Pogo, B. G.-T., Melana, S. M., Moran, H., & Holland, J. F. (2011). Presence of MMTV-like *env* gene sequences in human breast cancer. *Breast Cancer Research and Treatment, 125*(1), 295-297.

Pogo, B. G. T., Holland, J. F., & Levine, P. H. (2010). Human mammary tumor virus in inflammatory breast cancer. *Cancer, 116*(S11), 2741-2744.

Pogorzelski, M., Ting, S., Gauler, T. C., Breitenbuecher, F., Vossebein, I., Hoffarth, S., ... & Jawad, J. A. (2014). Impact of human papilloma virus infection on the response of head and neck cancers to anti-epidermal growth factor receptor antibody therapy. *Cell death & disease, 5*(2), e1091-e1091.

Polyak, K. (2007). Breast cancer: origins and evolution. *The Journal of Clinical Investigation, 117*(11), 3155-3163.

Portugal, M. E. G., Raboni, S. M., Nogueira, M. B., Vidal, L. R. R., Dingueleski, A. H., Kluk, E., Torres-Pereira, C. C. (2019). High frequency of multiple HPV types detection in Fanconi anemia patients oral swabs. *Transplant Infectious Disease, 21*(2), e13030.

Qu, W., Jiang, G., Cruz, Y., Chang, C. J., Ho, G., Klein, R. S., & Burk, R. D. (1997). PCR detection of human papillomavirus: comparison between MY09/MY11 and GP5+/GP6+ primer systems. *Journal of Clinical Microbiology, 35*(6), 1304-1310.

Quayson, S., Wiredu, E., Adjei, D., & Anim, J. (2014). Breast cancer in Accra, Ghana. *Journal of Medical and Biomedical Sciences, 3*(3), 21-26.

Rajilić-Stojanović, M., Heilig, H. G., Tims, S., Zoetendal, E. G., & de Vos, W. M. (2013). Long-term monitoring of the human intestinal microbiota composition. *Environmental Microbiology, 15*(4), 1146-1159.

Rakha, E. A., El-Sayed, M. E., Lee, A. H., Elston, C. W., Grainge, M. J., Hodi, Z., & Ellis, I. O. (2008). Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *Journal of Clinical Oncology, 26*(19), 3153-3158.

Ramírez-Fort, M.K., Khan, F., Rady, P,L., & Tyring, S,K. (2014). Human Papillomavirus: Bench to Bedside. In *Current Problems in Dermatology* (Vol 45, pp. 1-18). Basel, Karger Publishers.

Rao, V. P., Poutahidis, T., Ge, Z., Nambiar, P. R., Boussahmain, C., Wang, Y. Y., & Erdman, S. E. (2006). Innate immune inflammatory response against enteric bacteria Helicobacter hepaticus induces mammary adenocarcinoma in mice. *Cancer Research, 66*(15), 7395-7400.

Raymond, B., Johnston, P. R., Nielsen-LeRoux, C., Lereclus, D., & Crickmore, N. (2010). *Bacillus thuringiensis*: an impotent pathogen? *Trends in Microbiology, 18*(5), 189-194.

Reinson, T., Henno, L., Toots, M., Ustav Jr, M., & Ustav, M. (2015). The cell cycle timing of human papillomavirus DNA replication. *PLoS One, 10*(7), e0131675.

Rickinson, A. (2006). Epstein-Barr virus and its replication. *Virology, 2*, 2603-2654.

Rickinson, A. (2007). Epstein-barr virus. *Fields Virology*, 2655-2700.

Rickinson, A. B. (2014). Co-infections, inflammation and oncogenesis: future directions for EBV research. *Seminars in Cancer Biology, 26*, 99-115.

Ritchie, M. D., Hahn, L. W., Roodi, N., Bailey, L. R., Dupont, W. D., Parl, F. F., & Moore, J. H. (2001). Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *The American Journal of Human Genetics, 69*(1), 138-147.

Roderfeld, M., Padem, S., Lichtenberger, J., Quack, T., Weiskirchen, R., Longerich, T., & Tschuschner, A. (2018). *Schistosoma mansoni* egg secreted antigens activate HCC-associated transcription factors c-Jun and STAT 3 in hamster and human hepatocytes. *Hepatology 72(2):626-641*.

Ross, S. R. (2000). Using genetics to probe host-virus interactions; the mouse mammary tumor virus model. *Microbes and Infection, 2*(10), 1215-1223.

Ross, S. R. (2010). Mouse mammary tumor virus molecular biology and oncogenesis. *Viruses, 2*(9), 2000-2012.

Ross, S. R., Schofield, J. J., Farr, C. J., & Bucan, M. (2002). Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *Proceedings of the National Academy of Sciences, 99*(19), 12386-12390.

Ryan, M. P., & Pembroke, J. T. (2018). *Brevundimonas spp*: Emerging global opportunistic pathogens. *Virulence, 9*(1), 480-493.

Sadler, G. R., Ko, C. M., Cohn, J. A., White, M., Weldon, R.-N., & Wu, P. (2007). Breast cancer knowledge, attitudes, and screening behaviors among African American women: the Black cosmetologists promoting health program. *BioMed Central Public Health, 7*(1), 57.

Salman, N. A., Davies, G., Majidy, F., Shakir, F., Akinrinade, H., Perumal, D., & Ashrafi, G. H. (2017). Association of High Risk Human Papillomavirus and Breast cancer: A UK based Study. *Scientific Reports, 7*, 43591.

Santpere, G., Darre, F., Blanco, S., Alcami, A., Villoslada, P., Mar Alba, M., & Navarro, A. (2014). Genome-wide analysis of wild-type Epstein–Barr virus genomes derived from healthy individuals of the 1000 Genomes Project. *Genome Biology and Evolution*, *6*(4), 846-860.

Satram-Hoang, S., Stein, A., Cortazar, P., Momin, F., & Reyes, C. (2019). Increased Mortality Risk among Early Stage Hormone Receptor Positive Breast Cancer Patients Who Did Not Receive Adjuvant or Neoadjuvant Therapy. *Journal of Cancer Therapy*, *10*(1), 1-20.

Schiffman, M., Herrero, R., DeSalle, R., Hildesheim, A., Wacholder, S., Rodriguez, A. C., & Guillen, D. (2005). The carcinogenicity of human papillomavirus types reflects viral evolution. *Virology*, *337*(1), 76-84.

Schleifer, K. H., Geyer, U., Kilpper-Bälz, R., & Devriese, L. A. (1983). Elevation of *Staphylococcus sciuri subsp. lentus* (Kloos *et al*.) to Species Status: *Staphylococcus lentus*. *Systematic and Applied Microbiology*, *4*(3), 382-387.

Scully, R., Panday, A., Elango, R., & Willis, N. A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology*, *20*(11), 698-714.

Selinsky, C. L., & Howell, M. D. (2000). Soluble tumor necrosis factor receptor type I enhances tumor development and persistence in vivo. *Cellular Immunology*, *200*(2), 81-87.

Sellers, T. A., Kushi, L. H., Potter, J. D., Kaye, S. A., Nelson, C. L., McGovern, P. G., & Folsom, A. R. (1992). Effect of family history, body-fat distribution, and reproductive factors on the risk of postmenopausal breast cancer. *New England Journal of Medicine*, *326*(20), 1323-1329.

Serrano, B., Brotons, M., Bosch, F. X., & Bruni, L. (2018). Epidemiology and burden of HPV-related disease. *Best Practice & Research Clinical Obstetrics & Gynaecology*, *47*, 14-26.

Severin, J. A., Lestari, E. S., Kuntaman, K., Pastink, M., Snijders, S. V., Lemmens-den Toom, N., & Group, A. S. (2010). Nasal carriage of methicillin-resistant and methicillin-sensitive strains of *Staphylococcus sciuri* in the Indonesian population. *Antimicrobial Agents and Chemotherapy*, *54*(12), 5413-5417.

Sewani, A. (2001). Breast Cancer: Causes, Risk Factors and Prevention. *The Anthropologist*, *3*(2), 131-134.

Shanmugasundaram, S., & You, J. (2017). Targeting persistent human papillomavirus infection. *Viruses*, *9*(8), 229.

Shariatpanahi, S., Farahani, N., Salehi, A. R., & Salehi, R. (2017). High prevalence of mouse mammary tumor virus-like gene sequences in breast cancer samples of Iranian women. *Nucleosides, Nucleotides & Nucleic Acids*, *36*(10), 621-630.

Sharma, G. N., Dave, R., Sanadya, J., Sharma, P., & Sharma, K. (2010). Various types and management of breast cancer: an overview. *Journal of Advanced Pharmaceutical Technology & Research*, *1*(2), 109.

Shaw, R. J., & Cantley, L. C. (2006). Ras, PI (3) K and mTOR signalling controls tumour cell growth. *Nature*, *441*(7092), 424.

Shibata, A., & Jeggo, P. A. (2014). DNA double-strand break repair in a cellular context. *Clinical Oncology*, *26*(5), 243-249.

Sigaroodi, A., Nadji, S. A., Naghshvar, F., Nategh, R., Emami, H., & Velayati, A. A. (2012). Human papillomavirus is associated with breast cancer in the north part of Iran. *The Scientific World Journal*, *2012*, 837191.

Sixbey, J., Chesney, P. J., Shirley, P., Buntin, D., & Resnick, L. (1989). Detection of a second widespread strain of Epstein-Barr virus. *The Lancet*, *334*(8666), 761-765.

Sriamporn, S., Pisani, P., Pipitgool, V., Suwanrungruang, K., Kamsa-ard, S., & Parkin, D. (2004). Prevalence of *Opisthorchis viverrini* infection and incidence of cholangiocarcinoma in Khon Kaen, Northeast Thailand. *Tropical Medicine & International Health, 9*(5), 588-594.

Sripa, B., Brindley, P. J., Mulvenna, J., Laha, T., Smout, M. J., Mairiang, E., & Loukas, A. (2012). The tumorigenic liver fluke *Opisthorchis viverrini*–multiple pathways to cancer. *Trends in Parasitology, 28*(10), 395-407.

Srivastava, M., & Raghavan, S. C. (2015). DNA double-strand break repair inhibitors as cancer therapeutics. *Chemical Biology, 22*(1), 17-29.

Stanfield, B. A, & Micah, A. L. (2017) "Recent advances in understanding Epstein-Barr virus." *F1000Research, 6,* 386.

Stepanovic, S., Dakic, I., Morrison, D., Hauschild, T., Jezek, P., Petras, P., & Devriese, L. A. (2005). Identification and characterization of clinical isolates of members of the *Staphylococcus sciuri* group. *Journal of Clinical Microbiology, 43*(2), 956-958.

Stewart, T., Sage, R., Stewart, A., & Cameron, D. (2000). Breast cancer incidence highest in the range of one species of house mouse, *Mus domesticus*. *British Journal of Cancer, 82*(2), 446.

Sturm, A. (1989). *Mobiluncus* species and other anaerobic bacteria in non-puerperal breast abscesses. *European Journal of Clinical Microbiology and Infectious Diseases, 8*(9), 789-792.

Svec, P., Petras, P., Pantucek, R., Doskar, J., & Sedlacek, I. (2016). High intraspecies heterogeneity within *Staphylococcus sciuri* and rejection of its classification into *S. sciuri subsp. sciuri, S. sciuri subsp. carnaticus* and *S. sciuri subsp. rodentium*. *International Journal of Systematic and Evolutionary Microbiology, 66*(12), 5181-5186.

Tao, Z., Shi, A., Lu, C., Song, T., Zhang, Z., & Zhao, J. (2015). Breast cancer: epidemiology and etiology. *Cell Biochemistry and Biophysics, 72*(2), 333-338.

Tato, C. M., & Hunter, C. A. (2002). Host-Pathogen Interactions: Subversion and Utilization of the NF- B Pathway during Infection. *Infection and Immunity, 70*(7), 3311-3317.

Thompson, K. J., Ingle, J. N., Tang, X., Chia, N., Jeraldo, P. R., Walther-Antonio, M. R., & Kalari, K. R. (2017). A comprehensive analysis of breast cancer microbiota and host gene expression. *PLoS One, 12*(11), e0188873.

Thompson, M. P., & Kurzrock, R. (2004). Epstein-Barr virus and cancer. *Clinical Cancer Research, 10*(3), 803-821.

Thorley-Lawson, D. A., & Gross, A. (2004). Persistence of the Epstein–Barr virus and the origins of associated lymphomas. *New England Journal of Medicine, 350*(13), 1328-1337.

Thorne, L. B., Ryan, J. L., Elmore, S. H., Glaser, S. L., & Gulley, M. L. (2005). Real-time PCR measures Epstein-Barr Virus DNA in archival breast adenocarcinomas. *Diagnostic Molecular Pathology, 14*(1), 29-33.

Toller, I. M., Neelsen, K. J., Steger, M., Hartung, M. L., Hottiger, M. O., Stucki, M., & Lopes, M. (2011). Carcinogenic bacterial pathogen Helicobacter pylori triggers DNA double-strand breaks and a DNA damage response in its host cells. *Proceedings of the National Academy of Sciences, 108*(36), 14944-14949.

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. *CA: Cancer Journal of Clinicians, 65*(2), 87-108.

Tsakogiannis, D., Diamantidou, V., Toska, E., Kyriakopoulou, Z., Dimitriou, T., Ruether, I., & Markoulatos, P. (2015). Multiplex PCR assay for the rapid identification of human

papillomavirus genotypes 16, 18, 45, 35, 66, 33, 51, 58, and 31 in clinical samples. *Archives of Virology, 160*(1), 207-214.

Tsao, S.W., Tsang, C.M., Pang, P.S., Zhang. G., Chen. H., & Lo, K.W. (2012). The biology of EBV infection in human epithelial cells. *Seminars in Cancer Biology*, *22*(2), 137-143.

Tulay, P., & Serakinci, N. (2016). The role of human papillomaviruses in cancer progression. *Journal of Cancer Metastasis and Treatment , 2*, 201-213.

Urbaniak, C., Burton, J. P., & Reid, G. (2012). Breast, milk and microbes: a complex relationship that does not end with lactation. *Women's Health, 8*(4), 385-398.

Urbaniak, C., Cummins, J., Brackstone, M., Macklaim, J. M., Gloor, G. B., Baban, C. K., & Reid, G. (2014). Microbiota of human breast tissue. *Applied and Environmental Microbiology, 80*(10), 3007-3014.

Urbaniak, C., Gloor, G. B., Brackstone, M., Scott, L., Tangney, M., & Reid, G. (2016). The Microbiota of Breast Tissue and Its Association with Breast Cancer. *Applied and Environmental Microbiology, 82*(16), 5039-5048.

van Beek, J., Brink, A. A., Vervoort, M. B., van Zijp, M. J., Meijer, C. J., van den Brule, A. J., & Middeldorp, J. M. (2003). *In vivo* transcription of the Epstein–Barr virus (EBV) BamHI-A region without associated in vivo BARF0 protein expression in multiple EBV-associated disorders. *Journal of General Virology, 84*(10), 2647-2659.

Vanderpuye, V., Grover, S., Hammad, N., PoojaPrabhakar, Simonds, H., Olopade, F., & Stefan, D. C. (2017). An update on the management of breast cancer in Africa. *Infectious Agents and Cancer, 12*, 13.

Varga, Z., Li, Q., Jochum, W., Perriard, U., Rau, T., Tille, J.-C., & Ruhstaller, T. (2019). Ki-67 assessment in early breast cancer: SAKK28/12 validation study on the IBCSG VIII and IBCSG IX cohort. *Scientific reports, 9*(1), 1-11.

Voduc, K. D., Cheang, M. C., Tyldesley, S., Gelmon, K., Nielsen, T. O., & Kennecke, H. (2010). Breast cancer subtypes and the risk of local and regional relapse. *Journal of Clinical Oncology, 28*(10), 1684-1691.

Wallace, B. D., Wang, H., Lane, K. T., Scott, J. E., Orans, J., Koo, J. S., & Mani, S. (2010). Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science, 330*(6005), 831-835.

Wang, D., Fu, L., Shah, W., Zhang, J., Yan, Y., Ge, X., & Li, X. (2016). Presence of high risk HPV DNA but indolent transcription of E6/E7 oncogenes in invasive ductal carcinoma of breast. *Pathology-Research and Practice, 212*(12), 1151-1156.

Wang, K., Wu, F., Seo, B. R., Fischbach, C., Chen, W., Hsu, L., & Gourdon, D. (2017). Breast cancer cells alter the dynamics of stromal fibronectin-collagen interactions. *Matrix Biology, 60*, 86-95.

Wang, Y., Holland, J. F., Bleiweiss, I. J., Melana, S., Liu, X., Pelisson, I., & Pogo, B. G. (1995). Detection of mammary tumor virus *env* gene-like sequences in human breast cancer. *Cancer Research, 55*(22), 5173-5179.

Wang, Y., Pelisson, I., Melana, S., Go, V., Holland, J., & Pogo, B.-T. (2001). MMTV-like *env* gene sequences in human breast cancer. *Archives of Virology, 146*(1), 171-180.

Wang, Y. X., Li, Y. Z., Zhang, Z. Y., Wang, J. Q., Cui, J., & Qian, X. L. (2017). HPV16 E6 Promotes Breast Cancer Proliferation via Upregulation of COX-2 Expression. *BioMed Research International, 2017,* 2948467.

Weigelt, B., Geyer, F. C., & Reis-Filho, J. S. (2010). Histological types of breast cancer: how special are they? *Molecular oncology, 4*(3), 192-208.

Weinbren, M., Perinpanayagam, R., Malnick, H., & Ormerod, F. (1986). Mobiluncus spp: pathogenic role in non-puerperal breast abscess. *Journal of Clinical Pathology, 39*(3), 342.

Weiss, L. K., Burkman, R. T., Cushing-Haugen, K. L., Voigt, L. F., Simon, M. S., Daling, J. R., & Marchbanks, P. A. (2002). Hormone replacement therapy regimens and breast cancer risk. *Obstetrics & Gynecology, 100*(6), 1148-1158.

Williams, R. J., Henderson, B., Sharp, L. J., & Nair, S. P. (2002). Identification of a fibronectin-binding protein from *Staphylococcus epidermidis*. *Infection and Immunity, 70*(12), 6805-6810.

Wilson, J., Schurr, M., LeBlanc, C., Ramamurthy, R., Buchanan, K., & Nickerson, C. (2002). Mechanisms of bacterial pathogenicity. *Postgraduate Medical Journal, 78*(918), 216-224.

Wiredu, E. K., & Armah, H. B. (2006). Cancer mortality patterns in Ghana: a 10-year review of autopsies and hospital mortality. *BioMed Central Public Health, 6*(1), 159.

Wu, Y., Sarkissyan, M., & Vadgama, J. (2016). Epithelial-mesenchymal transition and breast cancer. *Journal of Clinical Medicine, 5*(2), 13.

Xu, H., Li, X., Sun, Y., Shi, X., & Wu, J. (2016). Biodegradation of Pyrene by Free and Immobilized Cells of *Herbaspirillum chlorophenolicum* Strain *FA1*. *Water, Air, & Soil Pollution, 227*(4).

Xuan, C., Shamonki, J. M., Chung, A., Dinome, M. L., Chung, M., Sieling, P. A., & Lee, D. J. (2014). Microbial dysbiosis is associated with human breast cancer. *PLoS One, 9*(1), e83744.

Xue, S., Lampert, I., Haldane, J., Bridger, J., & Griffin, B. (2003). Epstein–Barr virus gene expression in human breast cancer: protagonist or passenger? *British Journal of Cancer, 89*(1), 113.

Yager, J. D., & Davidson, N. E. (2006). Estrogen carcinogenesis in breast cancer. *New England Journal of Medicine, 354*(3), 270-282.

Yahia, Z. A., Adam, A. A., Elgizouli, M., Hussein, A., Masri, M. A., Kamal, M., & Hamad, K. (2014). Epstein Barr virus: a prime candidate of breast cancer aetiology in Sudanese patients. *Infectious Agents and Cancer, 9*(1), 9.

Yasmeen, A., Bismar, T. A., Kandouz, M., Foulkes, W. D., Desprez, P. Y., & Al Moustafa, A. E. (2007). E6/E7 of HPV type 16 promotes cell invasion and metastasis of human breast cancer cells. *Cell Cycle*, *6*(16), 2038-2042.

Young, L. S., & Rickinson, A. B. (2004). Epstein-Barr virus: 40 years on. *Nature Reviews Cancer, 4*(10), 757-768.

Zekri, A. R., Bahnassy, A. A., Mohamed, W. S., El-Kassem, F. A., El-Khalidi, S. J., Hafez, M. M., & Hassan, Z. K. (2012). Epstein-Barr virus and breast cancer: epidemiological and molecular study on Egyptian and Iraqi women. *Journal of the Egyptian National Cancer Institute, 24*(3), 123-131.

Zhang, F., Ma, J., Wu, J., Ye, L., Cai, H., Xia, B., & Yu, X. (2009). PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Current Biology, 19*(6), 524-529.

Zhang, Y.-Q., Ren, S.-X., Li, H.-L., Wang, Y.-X., Fu, G., Yang, J., & Wen, Y.-M. (2003). Genome-based analysis of virulence genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). *Molecular Microbiology, 49*(6), 1577-1593.

Zhang, Y., Fan, S., Meng, Q., Ma, Y., Katiyar, P., Schlegel, R., & Rosen, E. M. (2005). BRCA1 interaction with human papillomavirus oncoproteins. *Journal of Biological Chemistry, 280*(39), 33165-33177.

Zhu, J., Liao, M., Yao, Z., Liang, W., Li, Q., Liu, J., & Mo, Z. (2018). Breast cancer in postmenopausal women is associated with an altered gut metagenome. *Microbiome*, *6*(1), 136.

APPENDICES

APPENDIX 1

Sample number	Isolate number	Sequence	Description	Query cover	% identity	Accession
005	9	Forward	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	99%	98.81%	NZ_CP022046.2
		Reverse	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	99%	98.80%	NZ_CP022046.2
	14	Forward	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	99%	98.62%	NZ_CP022046.2
		Reverse	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	99%	99.08%	NZ_CP022046.2
	16	Forward	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	96%	97.40%	NZ_CP022046.2
		Reverse	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	98.03%	NZ_CP022046.2
	18	Forward	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	99.17%	NZ_CP022046.2
		Reverse	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	99.36%	NZ_CP022046.2
	20	Forward	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	98.71%	NZ_CP022046.2
		Reverse	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	99.17%	NZ_CP022046.2
	23	Forward	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	98.71%	NZ_CP022046.2
		Reverse	*Staphylococcus sciuri* strain FDAARGOS_285	98%	98.89%	NZ_CP022046.2

			chromosome, complete genome			
	27	Forward	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	98.43%	NZ_CP022046.2
		Reverse	*Staphylococcus sciuri* strain FDAARGOS_285 chromosome, complete genome	98%	99.07%	NZ_CP022046.2
007	30	Forward	*Megasphaera cerevisiae* DSM 20462 scaffold_113, whole genome shotgun sequence	98%	98.88%	NZ_LEKT01000114.1
		Reverse	*Megasphaera cerevisiae* DSM 20462 scaffold_113, whole genome shotgun sequence	98%	98.62%	NZ_LEKT01000114.1
	32	Forward	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	94%	96.67%	NZ_CP007601.1
		Reverse	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	79%	86.10%	NZ_CP007601.1
	34	Forward	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	98%	96.66%	NZ_CP007601.1
		Reverse	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	65%	77.39%	NZ_CP007601.
	37	Forward	*Herbaspirillum chlorophenolicum* strain CPW301 contig90, whole genome shotgun sequence	98%	97.93%	NZ_LFLT01000090.1
		Reverse	*Herbaspirillum chlorophenolicum* strain CPW301 contig90, whole genome shotgun sequence	99%	98.37%	NZ_LFLT01000090.1
	38	Forward	*Staphylococcus warneri* SG1, complete genome	99%	97.99%	NC_020164.1
		Reverse	*Staphylococcus warneri* SG1, complete genome	98%	97.98%	NC_020164.1
		Forward	*Microbacterium chocolatum* strain SIT 101, complete genome	98%	96.10%	NZ_CP015810.1
		Reverse	*Microbacterium chocolatum* strain SIT 101, complete genome	98%	96.52%	NZ_CP015810.1
	40	Forward	*Megasphaera cerevisiae* DSM 20462 scaffold_113,	98%	98.16%	NZ_LEKT01000114.1

			whole genome shotgun sequence			
		Reverse	*Megasphaera cerevisiae* DSM 20462 scaffold_113, whole genome shotgun sequence	99%	98.06%	NZ_LEKT0100
	41	Forward	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	97%	98.43%	NC_016603.1
		Reverse	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	97%	98.97%	NC_016603.1
008	42	Forward	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	99%	98.84%	NC_016603.1
		Reverse	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	97%	98.42%	
	44	Forward	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	98%	97.93%	NC_016603.1
		Reverse	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	97%	98.25%	NC_016603.1
	52	Forward	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	96%	98.31%	NC_016603.1
		Reverse	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	98%	96.87%	NC_016603.1
	56	Forward	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	93%	94.27%	NC_016603.1
		Reverse	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	94%	92.25%	NC_016603.1
	59	Forward	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	98%	98.22%	NC_016603.1
		Reverse	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	98%	97.44%	NC_016603.1
	60	Forward	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	97%	97.36%	NC_016603.1
		Reverse	*Acinetobacter pittii* PHEA-2 chromosome, complete genome	97%	96.73%	NC_016603.1
006	62	Forward	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	96.96%	NC_004461.1
		Reverse	*Staphylococcus epidermidis* ATCC 12228	98%	98.25%	NC_004461.1

			chromosome, complete genome			
	64	Forward	*Bacillus pumilus* strain SH-B9, complete genome	98%	98.53%	NZ_CP011007.1
		Reverse	*Bacillus pumilus* strain SH-B9, complete genome	97%	98.02%	NZ_CP011007.1
	65	Forward	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	98%	98.34%	NC_005957.1
		Reverse	Trichococcus palustris isolate Trichococcus palustris, whole genome shotgun sequence	4%	96.15%	NZ_FJNE01000017.1
	69	Forward	*Staphylococcus lugdunensis* HKU09-01, complete genome	94%	96.94%	NC_013893.1
		Reverse	*Staphylococcus lugdunensis* HKU09-01, complete genome	97%	99.30%	NC_013893.1
	81	Forward	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	97%	99.08%	NC_005957.1
		Reverse	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	92%	94.67%	NC_005957.
011	86	Forward	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	97%	98.89%	NC_005957.1
		Reverse	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	97%	96.16%	NC_005957.1
	90	Forward	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	98%	99.19%	NC_005957.1
		Reverse	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	98%	96.79%	NC_005957.1
012	98	Forward	*Bacillus thuringiensis* YBT-1518, complete genome	97%	98.34%	NC_022873.1
		Reverse	*Bacillus thuringiensis* YBT-1518, complete genome	95%	96.84%	NC_022873.1
	99	Reverse	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	97.19%	NC_004461.1

	102	Forward	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	98.24%	NC_004461.
		Reverse	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	98.57%	NC_004461.1
	105	Forward	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	98.80%	NC_004461.1
		Reverse	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	97%	98.18%	NC_004461.1
	107	Forward	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	98.24%	NC_004461.1
		Reverse	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	96%	97.42%	NC_004461.1
014	126	Forward	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	98%	98.69%	NZ_CP007601.1
		Reverse	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	98%	97.74%	NZ_CP007601.1
	136	Forward	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	98.89%	NC_004461.1
		Reverse	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	96.20%	NC_004461.1
010 and 015	146	Forward	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	96%	98.80%	NZ_CP007601.1
		Reverse	*Staphylococcus capitis* subsp. *capitis* strain AYP1020, complete genome	98%	93.39%	NZ_CP007601.1
	150	Forward	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	98%	97.71%	NC_005957.1
		Reverse	*Bacillus thuringiensis* YBT-1518, complete genome	97%	98.10%	NC_022873.

	152	Forward	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	96%	97.16%	NC_005957.1
		Reverse	[*Bacillus thuringiensis*] serovar konkukian str. 97-27 chromosome, complete genome	97%	96.90%	NC_005957.1
007 and 008 control	F	Forward	*Herbaspirillum seropedicae* strain Z67, complete genome	97%	91.63%	NZ_CP011930.1
		Reverse	*Herbaspirillum chlorophenolicum* strain CPW301 contig90, whole genome shotgun sequence	98%	98.84%	NZ_LFLT01000090.1
013 control	N	Reverse	Brevi*Bacillus* brevis NBRC 100599 DNA, complete genome	99%	97.86%	NC_012491.1
014 control	P	Forward	*Microbacterium chocolatum* strain SIT 101, complete genome	97%	96.36%	NZ_CP015810.1
		Reverse	*Microbacterium chocolatum* strain SIT 101, complete genome	98%	97.70%	NZ_CP015810.1
	Q	Forward	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	96%	94.53%	NC_004461.1
		Reverse	*Staphylococcus epidermidis* ATCC 12228 chromosome, complete genome	98%	97.83%	NC_004461.1

APPENDIX 2

Table 2: **Description of HPV subtypes obtained after Sanger sequencing**

Sample number	Description	Query cover	% identity	Accession
73(3)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
12(3)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
5(3)	Human papillomavirus type 18 isolate BB major capsid protein L1 gene, partial cds	100%	99.24%	EF140825.1
2(3)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
76(3)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
79(3)	Human papillomavirus type 18 isolate 2 L1 (L1) gene, partial cds	99%	100.00%	AF548837.1
36(4)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
37(4)	Human papillomavirus type 18 isolate BB major capsid protein L1 gene, partial cds	100%	100.00%	EF140825.1
30(3)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
32(3)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
35(3)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
10(4)	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
1F	Human papillomavirus type 18 strain 18CNTZ36 E6 (E6), E7 (E7), and L1 (L1) genes, complete cds	93%	98.39%	KY457840.1
17F	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	99.32%	AF548838.1
10F	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	99.32%	AF548838.1
7F	Human papillomavirus type 18 isolate MJR16 major capsid protein L1 gene, partial cds	100%	94.62%	KT932006.2
8F	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	99.32%	AF548838.1
9F	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
10F	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
11F	Human papillomavirus type 18 isolate BB major capsid protein L1 gene, partial cds	90%	97.50%	EF140825.1
13F	Human papillomavirus type 18 isolate 3 L1 (L1) gene, partial cds	100%	100.00%	AF548838.1
15F	Human papillomavirus type 18 isolate F nonfunctional L1 protein gene, partial sequence	99%	99.00%	DQ315392.1
16F	Human papillomavirus type 18 isolate BB major capsid protein L1 gene, partial cds	100%	97.78%	EF140825.1

14F	Human papillomavirus isolate B2-GP major capsid protein L1 (L1) gene, partial cds	100%	100.00%	HM748607.1