Detection of HPV, EBV, MMTV- like DNA Sequences in Breast Cancer Patients

By

UMM-E-AIMAN

NUST ID 00000119796

MS (HEALTHCARE BIOTECHNOLOGY)

DEPARTMENT OF HEALTH CARE BIOTECHNOLOGY

ATTA-UR-RAHMAN SCHOOL OF APPLIED BIOSCIENCES,

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD, PAKISTAN.

2017

Detection of HPV, EBV, MMTV- like DNA Sequences in Breast Cancer Patients

Submitted by

Umm-e-Aiman

NUST ID 00000119796

MS (Healthcare Biotechnology)

A thesis submitted in partial fulfillment of the requirements for the degree of M.Phil Healthcare Biotechnology

Supervisor:

Dr. Hajra Sadia

DEPARTMENT OF HEALTHCARE BIOTECHNOLOGY

ATTA-UR-RAHMAN SCHOOL OF APPLIED BIOSCIENCES,

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD, PAKISTAN

FORM TH-4

National University of Sciences & Technology

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision

y: Umm-e-Aiman, Regn no: 00000119795,

itled: 'Viral Etiology of Breast Cancer' be accepted in partial fulfillment of the

equirements for the award of Masters MS (Grad&A) degree.

Examination Committee Members

1. Name: Dr. Peter John Signature: _____

2. Name: Dr. Attya Bhatti Signature: _____

3. Name: Dr. Jahangir Sarwar Signature: _____

Supervisor's name: Dr. Hajra Sadia Signature: _____

Date: _____

Head of Department 6/12/17 Date

COUNTERSINGED

Dr. Peter John
Principal
Atta ur Rahman
Applied Sciences
NUST Islamabad

Dean/Principal

ate: 07/12/17

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil Thesis written by Ms. Umm-e-Aiman (Registration No.00000119796), of Atta-ur-Rehman School of Applied Biosciences has been vetted by under signed, found complete in all respects as per NUST Statutes/Regulation, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: _____

Name of Supervisor: **Dr. Hajra Sadia**

Date: _____

Signature (HOD): _____

Date: 6/12/17

Signature (Dean/Principal): _____

Date: 07/12/17

Dr. Peter John
Principal
Atta ur Rahman S
Applied Sciences

TABLE OF CONTENTS

LIST OF ACRONYMS

ER	Estrogen receptor
PR	Progesterone receptor
$MgCl_2$	Magnesium Chloride
mL	milli liter
mM	milli Molar
WHO	World Health Organization
Taq	Thermus aquaticus
µL	Micro liter
ng	nanogram
nm	nanometer
L1	Late Protein 1
L2	Late Protein 2
Kb	Kilo Base
LMP	Latent Membrane Protein
Bp	Base Pairs
CIS	Common Integration Site
dH_2O	distilled H_2O

List of Acronyms

dNTP	deoxynucleotide Triphosphate
dsDNA	Double Stranded DNA
EBV	Epstein Barr Virus
EDTA	Ethylenediaminetetraacetic acid
FGF	Fibroblast Growth Factor
HCV	Hepatitis C Virus
HERV	Human Endogenous Retrovirus
HPV	Human Papilloma Virus
HRE	Hormone Responsive Elements
HTLV	Human T-cell Lymphoma Virus
IL-2	Interleukine 2
IL-4	Interleukine 4
LTR	Long Terminal Repeats
ENV	Envelope
MMTV	Mouse Mammary Tumor Virus
mRNA	messenger RNA
NF	Nuclease Free

List of Acronyms

PCR	Polymerase Chain Reaction
rpm	Revolutions per minute
RTC	Reverse Transcription Complex
SAg	Super Antigen
SDS	Sodium Dodecyl Sulfate
ssRNA	Single Stranded RNA
TBE	Tris Boric EDTA
tfr1	Transferrin 1
TGFβ	Tumor Growth Factor β
tRNA	Transfer RNA
UV	Ultra Violet
Int-1	Insertional site 1
IARC	International Agency for Research on Cancer
BRCA1	Breast Cancer type 1
c-myc	cellular myelocytomatosis
Gag	Group Specific Antigen
PI3K	Phosphoinositide 3-kinase

List of Acronyms

MMP	Matrix Mettalo Proteinase
ERK	Extracellular Signal Regulated Kinases
tcf	Transcription factor
gp	Glycoprotein
WNT	Wingless Type
Rspo3	R- spondin 3
%	Percentage
°C	degree Celsius

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

All the praise to the Almighty Allah guided me and filled me with courage and confidence throughout this research work. All this is nothing but a manifestation of Allah's blessings upon me, all Praise is to Him, for blessing me with a family that always encouraged me to explore my abilities.

I want to express my gratitude to my supervisor, **Dr. Hajra Sadia**, who has been a great supervisor and great mentor at ASAB. I want to thank **Dr. Peter John,** Principal, ASAB and **Dr. Attya Bhatti** HOD Healthcare Biotechnology, for maintaining a research-pro environment at ASAB. I would like to acknowledge NUST for provision of funds, **Dr. Jahangir Sarwer**, my GEC member and Head of Surgery Department in Holy Family Hospital Rawalpindi, for providing us samples of breast cancer patients. My lab associates, **Behroz Naeem**, **Sara Hassan** and **Faiza Saleem** for their continuous support, **Gohar Ayub** and **Aqsa Ikram** for their help. I am especially thankful to my seniors **Orooj surriya**, **Wasifa Naushad, Maliha Malik**, **Rubina Daad, Mr.Sajid Mansoor** for being a mentor and for their constant guidance during my research, **Mrs. Fozia Parveen and Huma** for always being concerned and helpful. I am also grateful to my friends especially **Fatima-tuz-Zahra** who went out of her way to help me out, **Yumna Farukh, Ifrah Rafaqat, Javeria mirza, Sidra Hussain, Farwa Amin** and **Arun Asif** for their continuous moral support. And finally my utmost respect and love for my parents my brother **Mazhar Fareed**. I don't have words to thank them for their support.

UMM-E-AIMAN

ABSTRACT

Breast cancer is known to affect every ninth women in Pakistan. Besides, genetic predisposition being a leading cause of this carcinoma, the oncogenic potential of Mouse mammary tumor virus (MMTV), Human Papilloma Virus (HPV), and Epstein Barr Virus (EBV) is also a reported cause behind the cancer. MMTV has a significant role as it is speculated to be a causative agent to breast cancer. On the contrary, HPV and EBV are not known to be directly causing the cancer. However, these viruses, upon integration, result in the down regulation of tumor suppressor genes whereas in the up regulation of proto oncogenes, thus causing breast carcinoma.

The direct role of viruses in the disease is still a debatable argument. The study was therefore designed by keeping this in mind and therefore both tissue and blood samples of the patients were analyzed for the presence of the aforementioned viruses, to ensure credibility of the results.

Mean Mono-infection of these viruses was observed in 39 % cases and co-infection of MMTV, EBV and HPV was 32 %, suggesting that co-infection as well as mono infection of EBV, MMTV LTR and HPV is significantly associated with studied samples of breast carcinoma. The coinfection of EBV is also observed to play a significant role in cancer metastasis through lymph node, however, no significant association of these viruses was found with tumor grade, breast cancer types and age of patients. A limitation of the study was its lesser sample size; however, further studies in this area encompassing a larger sample size will help us analyze the prevalence of these viruses in a better way.

CHAPTER 1

INTRODUCTION

The history of breast cancer dates back to 1600BC, where it was first diagnosed and studied by the Egyptian health professionals. Female all across the globe are prone to this malignant tumor which metastasize at different sites, leading to deaths. This disease infects the epithelium of the mammary glands which is a derivation of the ectodermal mammary fat pad of the embryonic stage. Various symptoms of breast cancer include change in shape, tumorous growth in breast, and leakage of fluid from nipple. A study in 2004 gave an estimate of about 40,000 deaths from a total of 216000 patients of breast cancer, throughout.

According to World health Organization (WHO) breast cancer is responsible for 324,000 deaths per annum so 25% of the global cancer burden is due to breast cancer (Ferlay *et al.*, 2015). Incident rate of breast cancer is quite high in Pakistan as compared to other countries like Iran and India one in every ninth women is suffering from breast cancer. It is also reported that in Pakistani female cancer patients prevalence of breast cancer is 45.9 (Badar *et al.*, 2011).

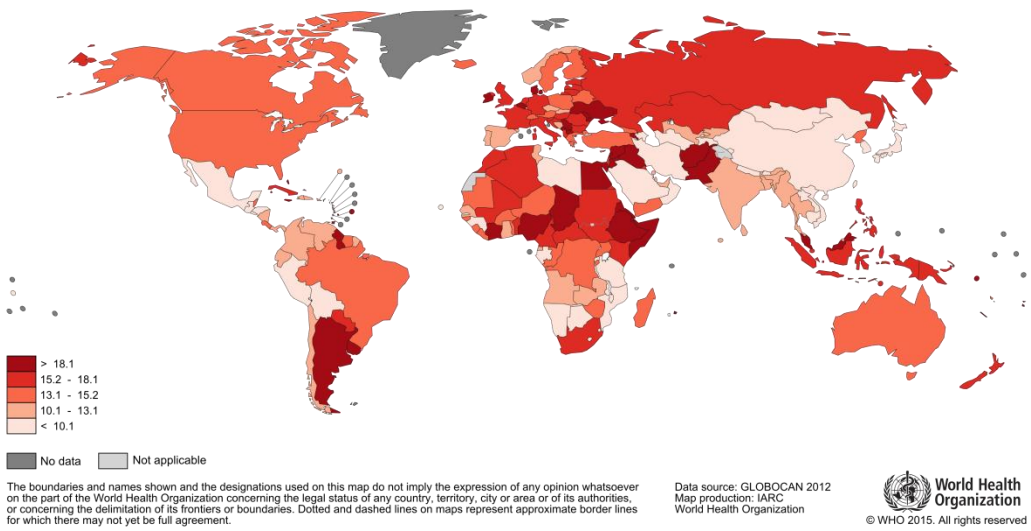

Figure 1.1: Worldwide Mortality rate of Breast Cancer. (GLOBOCAN, 2012).

Lobules and ducts are the major sites that are mostly infected by breast carcinoma. On basis of metastasis breast cancer is classified in to two major classes, i) invasive (that metastasize from its point of origin) and ii) non-invasive (don't metastasize from its point of origin) carcinoma. Various studies showed that ductal carcinoma is more prevalent then lobular carcinoma (Malhotra *et al*., 2010). Invasive Ductal carcinoma is one of the prevalent types of breast cancer followed by insitu ductal carcinoma in Pakistan and globally (Badar *et al.*, 2015).

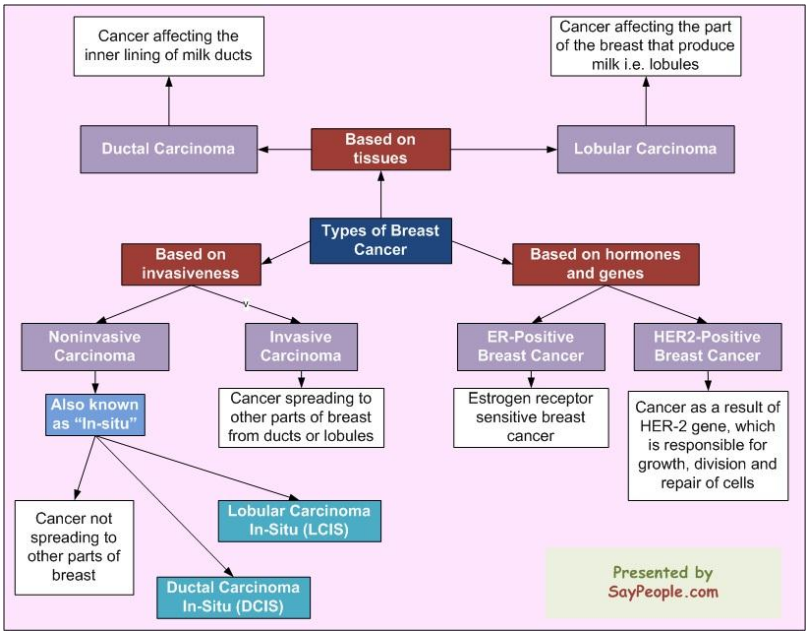

Figure 1.2: Types of breast cancer (Rabu, 2012).

Various external and internal factors lead to breast cancer. 5-10 % cases are familial and remaining 90-95% cases are sporadic. Factors that lead to breast cancer are genetic predisposition e.g. alteration in oncogenic Breast Cancer type 1 and 2 (BRCA1 and 2) and tumor suppressor genes (p53), lifestyle, family history, dense breast tissue, ethnicity, hormonal misbalance (Alibek *et al.*, 2013; Salmons *et al.*, 2013) alteration in the genes that play important role in DNA repair, estrogen production, cell cycle regulation and metabolism.

Incidence rate of breast cancer increases with age. Hormonal level is one of the major factors that triggers breast cancer it is suggested that women who hits puberty at earlier age have higher level of estrogen and progesterone which ultimately leads to higher chance of developing breast carcinoma and similarly for those who reach their menopause after 55 years (Key *et al.*, 2002).

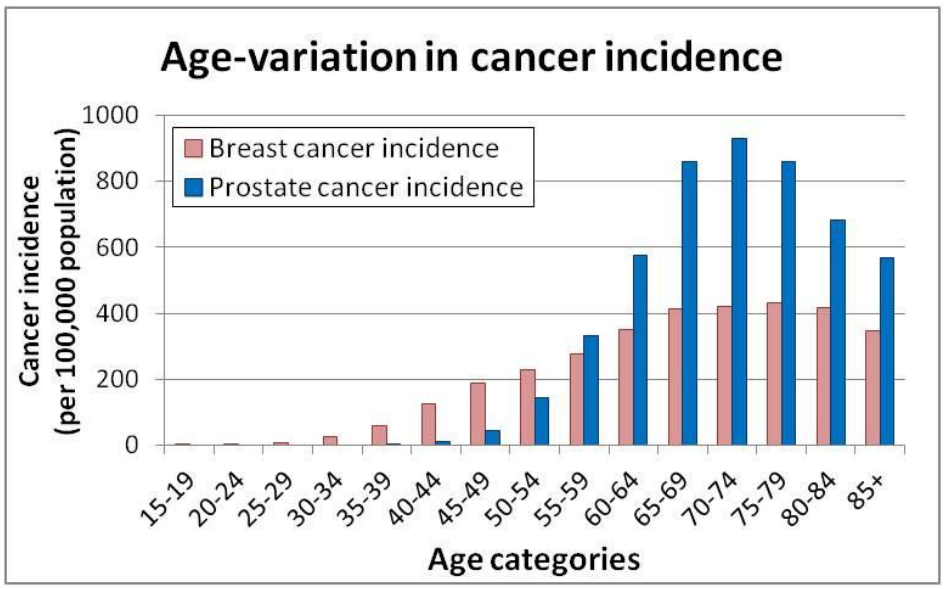

Figure 1.3: Association of breast cancer with age (Katatrepsis, 2012).

There are some oncogenic viruses that are suspected as causative agents of cancer e.g. Mouse memory tumor virus (MMTV), Human Papilloma Virus (HPV), Epstein Barr Virus (EBV), Human T-cell Lymphoma Virus (H-TLV).There are studies that reported association between viral sequence presence and development of breast cancer in which they compared normal tissues with tumorous one and found that 38% cases have 660- Base Pair (bp) of MMTV ENV gene like sequences that are not present in normal tissues (Wang *et al.,* 1998; Wang *et al*., 2001). In some studies they screened Long Terminal Repeat (LTR) sequences and full length proviral genome by Polymerase chain reaction (PCR) analysis. This association varies across the globe (Wang *et al*., 2001).

Mouse mammary Tumor Virus is milk transmitted filterable virus that belongs to retroviridea family of beta-retrovirus genus which develops oncogenesis in mice. Its genome consists of two copies of Single stranded Ribonucleic Acid (ssRNA) and reverse transcriptase and integrase enzymes that help in replication through cellular Transfer RNA (tRNA) (Maitra

et al., 2006) and integration of their genome within host respectively (Büchen-Osmond, 2006). It comprises two kinds of protein i) proteins that are common to all retroviruses Envelope (Env), Pol and group specific antigen (Gag) ii) accessory proteins (Buchen-Osmund, 2006). Long terminal repeats are present at the flanking regions of their genome. Three kind of regulatory elements are present in LTR that regulate the transcription of both MMTV genes and downstream genes of their integration sites (Bramblett *et al.,* 1995) their transcription is also regulated by steroid hormone through Hormone Responsive Element (HRE) (Majors and Varmus, 1983).

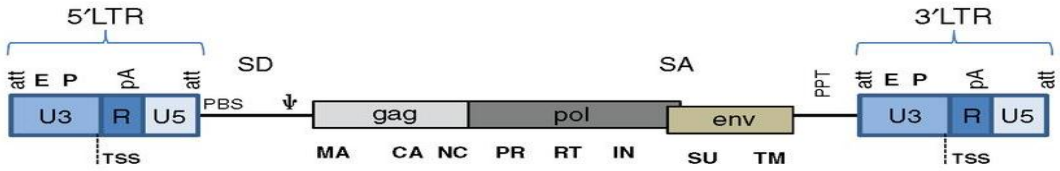

Figure 1.4: Genome of Mouse Mammary tumor Virus. (Ranzani *et al.,* 2013)

Lipid bilayer is major component of viral envelope it also contain some glycoproteins that linked through disulfide bond. Viral core is protected by icosahedral capsid protein and matrix proteins are located between envelope and capsid protein.

MMTV induce mammary cancer through hyperplastic outgrowth of pre malignant mammary gland**.** Insertion site 1 and 2 (Int-1 and 2) are two major integration sites that play important role in their pathogenesis. Insertional mutagenesis is a random process which leads to increase in stemness (Lamb *et al.,* 2015). Patterning and cell fate determination during embryogenesis are regulated by the genes of Wingless type (WNT) family that are located on q arm of chromosome 12 (Callahan *et al.,* 2008). Some important signaling pathways that play important role in oncogenesis like mitogen-activated protein kinase (MAPK),

Phosphoinositide phospholipase Cγ (PLCγ) and Phosphoinositide 3-kinase (PI3K)-AKT are controlled by Fibroblast growth factor 3 (FGF3), one of the integration site of MMTV**,** belongs to FGF family.

It is stated in some studies that expression of WNT-1 gene is higher in specimen that are positive for MMTV ENV and LTR when compared with negative sample (Pegram *et al.,* 2005). FGF 18 and 20 are downstream genes of canonical WNT pathway so during oncogenesis FGF signaling will also enhance. β-catenin dependent WNT pathway that induce carcinogenesis. In mammary gland leads to the enhanced expression of cellular myelocytomatosis (c-myc) and cyclin D1, i.e. known as key oncogene. Their overexpression is also reported in mice screened with MMTV. So they can act as potential biomarker for detecting breast cancer in humans (Musgrove, 2003).

MMTV integration leads to the activation of many downstream genes of WNT pathway including Matrix Mettalo Proteinase 7 (MMP7) involve in invasion**,** cyclin D1 and myc (proliferation), Map3k8 (high expressed in breast tumors), Wrch1 (migration), GSK3b (key component of PI3K-AKT pathway), Transcription factor (tcf) and WISP1 (angiogenesis), these genes can act as potential biomarkers for detecting breast cancer at early stages (Theodorou *et al.,* 2007; Wang *et al.,* 2014). R-spondin (Rspo) and FGF gene family are also characterized as important integration sites for MMTV. Overexpression of Rspo3 has been reported in epithelial mammary cell lines that are deficient in p-53 (Theodorou *et al.,* 2007).

Expression of MMTV target genes was also deregulated in 5-43% cases when they screened 295 tumors through microarray (Theodorou *et al.,* 2007). In Pakistani population 20 and 26% samples were screened with LTR and ENV sequences respectively (Naushad *et al.,*

2014). There are some studies that also reported zero prevalence of MMTV (Zangen *et al.*, 2002; Mant *et al.,* 2004).

EBV and HPV have been screened in both normal and tumorous tissues. Prevalence of HPV Genotype 18, 16 and 33 have been reported in various studies. It is also stated that women suffering from cervical and breast cancer were screened with same HPV genotype (Wang *et al.,* 2012).

International Agency for Research on Cancer (IARC) classified HPV as an oncovirus it shows 0-86% association with breast cancer, it is a Deoxyribo Nucliec Acid (DNA) virus that integrate its genome and then uses host machinery for viral protein production. Viral proteins (E6, 7 and 5) leads to carcinogenesis through inactivation of pRb and p53 (tumor suppressor genes) and tyrosine kinase respectively.

Human papilloma virus is Non-enveloped virus protected with icosahedral capsid (derived from 72capsomeres) and contains almost 8 kilo base (kb) of ds DNA. Three kinds of viral protein are present major and minor capsid protein and late protein 1(L1). They contain two types of Open Reading Frame (ORF) late and early that are categorized on the basis of their genome location. Structural (6) and non-structural (2) Genes are responsible for replication and assembly of HPV genome they contain 7 early genes (E1-7) and two late genes (L1 and 2) (Kajitani *et al.,* 2012). C-myc that is a potential biomarker for breast cancer also shows significant relation with HPV 16, 6 and 7 as the elevated expression of c-myc was reported in the cells infected with HPV (Alibek *et al.,* 2013).

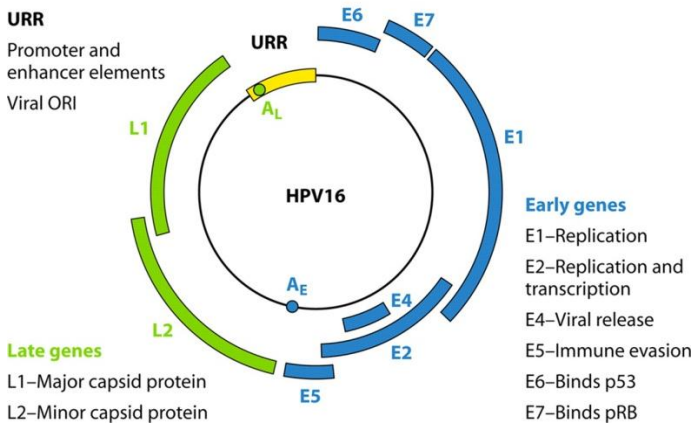

Figure 1.5: Genome of HPV. (Stanley, 2012*)*

Epstein Bar virus is also characterized as oncovirus that contributes in carcinogenesis through modifying phenotype via metastasizing rather than directly causing it. In tumor cells it leads to various alterations which ultimately results in cancer progression. It is reported in various studies that in case of breast cancer one in every 1000 cells are infected with EBV. They also play important role in cancer treatment by affecting chemotherapy through chemotherapeutic drug resistance.

Epstein Bar Virus is a double standard DNA (dsDNA) (with the size of 172 bp approximately) virus surrounded by octahedral capsid which is further enveloped by cell membrane of host. Its genome is divided into short and long sequences known as internal and terminal direct repeats, and can encode 100 genes. They mostly infect humans and can be transmitted through body fluid mostly through saliva (Glaser *et al.,* 2004). In various studies association between EBV and breast cancer was conformed through PCR analysis. High expression of Immunoglobin A (IgA) antibody is associated with presence of viral capsid of EBV (Alibek *et al.,* 2013).

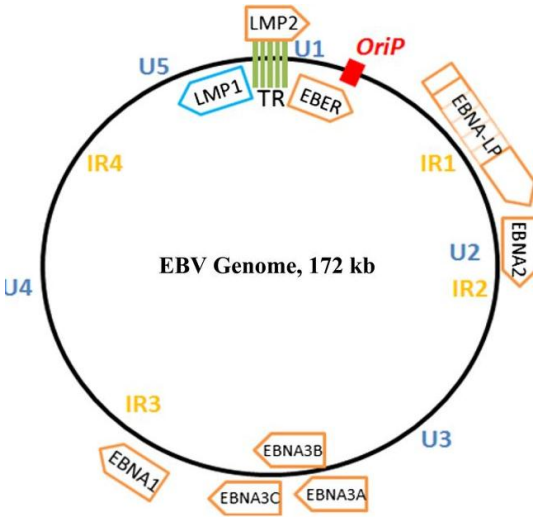

Figure 1.6: Genome of EBV. (Chi Young *et al.,* 2013)

Exact pathways that are involved in infection are still not known. So, it is important to find out the correlation of these viruses with breast cancer because already present data is not sufficient for studying the oncogenic ability of these viruses that leads to breast cancer.

Very little work has been done in Asian countries while in Pakistan data is almost negligible. To develop a proper connection it is important that data should be collected from all over the world. So it is very important to report the exact prevalence of these viruses (MMTV, EBV, and HPV) in Pakistani population because in this region death rate due to breast cancer and infectious diseases is higher. So to avoid this we should take proper measures and awareness shall be raised. This study was therefore designed by keeping this in mind and therefore both tissue and blood samples of the patients were analyzed for the presence of the aforementioned viruses, to ensure credibility of the results.

Aims of the study include;

I. Screening of MMTV LTR, MMTVENV, EBV and HPV viral regions and their comparative analysis in both tissue biopsies and blood samples.

II. To study the association of MMTV, HPV and EBV with breast cancer.

III. To study the association of these viruses and their coinfection with lymph node involvement, types of breast cancer, tumor grade and age of patients.

CHAPTER 2

REVIEW OF LITERATURE

Breast cancer is most common cause of death among female globally. In Pakistan prevalence of breast cancer is high in urban as well as rural areas. It is either transmitted sporadically or inherited. There are no proper screening centers in Pakistan so the disease cannot be diagnosed at early stage due to which mortality rate is becoming high (Menhas and Umer, 2015). After Jews, Karachi have high incidence rate of breast cancer and the reason behind this is not clear. Diet and reproductive factors also play important role in carcinogenesis e.g. multiple births, high fertility rate, prolonged effect of reproductive hormone and breast feeding early pregnancy and late menopause. However they did not report the role of BRCA1 and 2 in their study (Bhurgri *et al.*, 2000).

Breast Cancer (BC) Sub Type	Age %	Area	Age Range	Reference
Infiltrating ductal carcinoma	90%	Hyderabad, Jamshoro, Sindh	33-45	(Malik et al., 2010)
	81%	Karachi	48	(Siddiqui et al., 2000)
	82.60%	Peshawar	40-59	(Naeem et al., 2008)
	78%	Karachi	15-80	(Baloch et al., 2014)
	81%	Rawalpindi, ISB, NWFP, upper Punjab	36-60	(Mamoon et al., 2009a)
	91%	Karachi	30-66	(Malik, 2002)
	37%	Karachi	40-49	(Siddiqui et al., 2003)
Invasive intraductal carcinoma	94%	Karachi	31-53	(Afridi and Ahmed, 2012)
Ductal carcinoma in situ	16.25%	Karachi	48	(Siddiqui et al., 2000)
	2.40%	Karachi	15-80	(Baloch et al., 2014)
	1%	Karachi	48-95	(Bhurgri, 2005)
	90%	Hyderabad, Jamshoro, Sindh	33-45	(Malik et al., 2010)
Mucinous carcinoma	2.17%	Peshawar	40-59	(Naeem et al., 2008)
	12%	Karachi	15-80	(Baloch et al., 2014)
	0.52%	Karachi	48	(Siddiqui et al., 2000)
Infiltering lobular carcinoma	6.50%	Peshawar	40-59	(Naeem et al., 2008)
	0.34%	Karachi	48	(Siddiqui et al., 2000)
	1.20%	Karachi	15-80	(Baloch et al., 2014)
Papillary carcinoma	4.35%	Peshawar	40-59	(Naeem et al., 2008)
	0.17%	Karachi	48	(Siddiqui et al., 2000)
Invasive lobular carcinoma	6.50%	Peshawar	40-59	(Naeem et al., 2008)
Medullary carcinoma	2.17%	Peshawar	40-59	(Naeem et al., 2008)
	6%	Karachi	15-80	(Baloch et al., 2014)
Benin lumps	39.70%	Karachi	15-80	(Baloch et al., 2014)
	30.91%	Rawalpindi, , ISB, NWFP, upper Punjab	36-60	(Mamoon et al., 2009a)
	92.10%	Karachi	48- 95	(Bhurgri, 2005)
Total BC	19%	Baluchistan	31-50	(Doutani H et al., 2012)
	45.41%	Lahore	>18	(2014)

Figure 2.1: Subtypes of breast cancer in various regions of Pakistan (Basra *et al.*, 2016).

Incidence of breast cancer increase with age especially after menopause and also early mensuration increases the chances of developing cancer (McPherson *et al.,* 2000). Rise in hormones result into earlier breast development and their interaction with hormones which ultimately leads to carcinogenesis. Girls are developing cancer at very young age because of early puberty (Key *et al.,* 2002).Invasive Ductal carcinoma is one of the prevalent types of breast cancer followed by insitu ductal carcinoma in Pakistan and globally (Badar *et al*., 2015).

There are some oncogenic viruses that are suspected as causative agents of cancer e.g. MMTV, HPV, EBV, and HTLV. There are studies that reported association between viral sequence presence and development of breast cancer (Alibek *et al.,* 2013).

MMTV is a milk transmitted filterable virus that belong Retroviridea family of beta-retrovirus genus which develops oncogenesis in mice. It is transmitted either vertically from mother to child or lactating pups gets infected through infected mother. There are studies in which they compared normal tissues with tumorous one and found that 38% cases have 660-bp of MMTV ENV gene like sequences that are not present in normal tissues (Wang *et al.*, 1998: Wang *et al.,* 2001). Some studies also reported that they screen LTR sequences and full length proviral genome by PCR analysis. This association varies across the globe. There are many hypotheses that are presented but their proper role in oncogenesis is still under the consideration (Salmons and Gunzburg, 2013). Long terminal repeats are present at the flanking regions of their genome. Three kind of regulatory elements are present in LTR that regulate the transcription of both MMTV genes and downstream genes of their integration sites (Bramblett *et al.,* 1995) their transcription is also regulated by steroid hormone through HRE (Majors and Varmus, 1983).

MMTV is a milk transmitted filterable virus that belong Retroviridea family of beta-retrovirus genus which develops oncogenesis in mice. Its genome consists of two copies of ssRNA and reverse transcriptase and integrase enzymes that help in replication through cellular tRNA (Maitra *et al.,* 2006) and integration of their genome within host respectively (Büchen-osmond, 2006). 5' end of their RNA is protected by methylguanosine cap and poly A tail is present at 3' end they also contain unique regions at 3' end U3 (1200 base pair in length), U5 located at 5' end (120 base pair long), packaging signals. Downstream of U5 there are Primer binding site (18 base pair in length) that is complementary to tRNA-lys, and leads to the production of – strand of DNA. 5'LTR that is present in U3 region contain enhancers and promoters. Transcription starts at 5' U3/R region and at R/U5 junction (i.e. located in 3-LTR) it terminates. Three kind of polyprotein precursors are produced by ribosomal frameshift during translation Pr160Gag-Pro-Pol, Pr110Gag-Pro, Pr77Gag (Jacks *et al.,* 1987).

MMTV enter into the host through interaction of transferrin 1(i.e. cellular receptor) with Envelope glycoprotein. Modified form of virion known as Reverse Transcription Complex (RTC), a ribonucleoprotein, located in cytoplasm is site of viral DNA synthesis (Mougel *et al*., 2009).

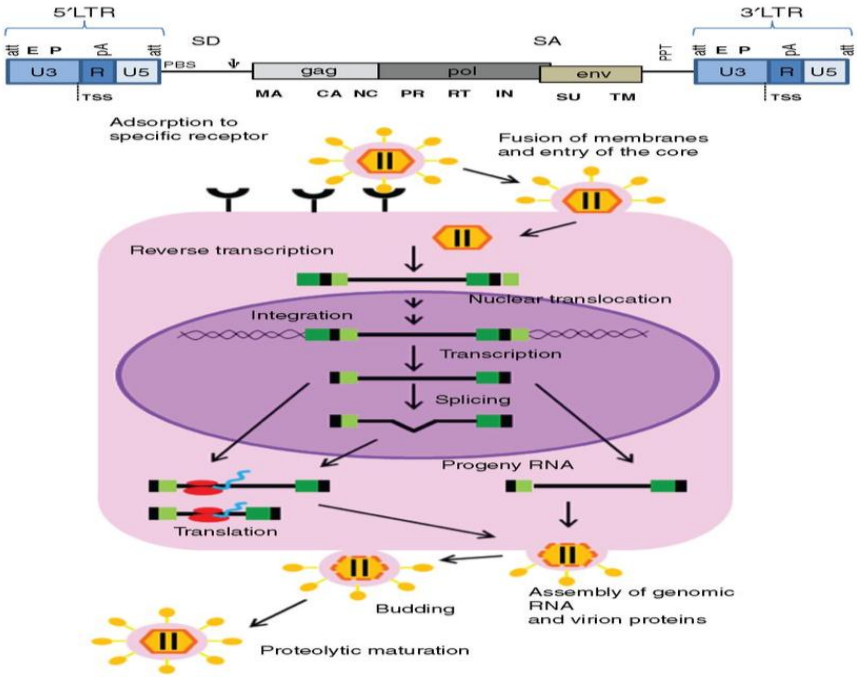

Figure 2.2: Illustrating Genome and life cycle of MMTV. (Ranzani *et al.*, 2013)

Through the process of Reverse transcription viral RNA is converted into dsDNA, in this process tRNA act as a primer and RNA-DNA hybrid is formed with the help of reverse transcriptase and their ribonuclease activity degrade most of the virion RNA and the resulting product is known as provirus. Pre-integration complex consist of provirus that is associated with nucleoprotein complex, this complex is integrated into host genome through integrase i.e. viral protein. Actively dividing cells are prone to viral infection specially MMTV, because during replication there is no nuclear membrane so that they can easily gain access to nucleus. This complex will be integrated and transcribe into messenger RNA (mRNA) with the help of cellular RNA Polymerase II (RNA pol II) and that is further translated into viral proteins with the help of cellular machinery. These proteins assembled to form immature virion; prior to budding they are associated with cell membrane. Immature virion is transformed into mature one through PR-mediated processing. There are two modes of infection either endogenous or

exogenous. In 90% cases exogenous mode of transmission is observed (mostly occurred at 6 months) and in 40% cases endogenous infection (age of onset is 20 months) (Bentvelzen, 1974; Nandi and McGrath, 1973).

MMTV spread through immune host system, after consuming infected breast viruses travel to payer's patches with the help of toll like receptor 4 (present in alimentary), where they are processed and presented by B Cells to T cells which leads to the activation and release of Interleukin 4 and 2 (IL-4 and IL-2) (Rassa *et al*., 2002). So, viral load will be increases in blood through cell division because lymphoid cells containing integrated viral particles will also increases. After this they viral particles move to the mammary gland with the help mesenteric lymph nodes and then finally enter through Transferrin 1 (Tfr1) receptors present in mammary gland of mice (Karapetian *et al*., 1994). In humans hormonal changes during lactation and pregnancy also trigger the expression of MMTV (Cato *et al*., 1987). Some studies stated that adult mice are not affected by MMTV because in their intestinal tract specific receptors for MMTV are down regulated. Nasal route of infection is primary from where this virus spread to whole system (Velin *et al*., 1996).

Two major hypotheses are reported for their role in cancer. Insertional mutagenesis is the first mechanism through which MMTV affect the expression of downstream genes and second one is that MMTV also encode oncoproteins. MMTV induce mammary cancer through hyperplastic outgrowth of pre malignant mammary gland. Int-1 and 2 are two major integration sites that play important role in their pathogenesis. Insertional mutagenesis is a random process which leads to increase in stemness (Lamb *et al.,* 2015). Patterning and cell fate determination during embryogenesis are regulated by the genes of WNT family that are located on q arm of chromosome 12 (Callahan and Smith, 2008). Some important signaling

pathways that play important role in oncogenesis like MAPK, PLCγ and PI3K-AKT are controlled by FGF3**,** one of the integration site of MMTV**,** belongs to FGF family.

MMTV causes pathogenesis through Insertional mutagenesis that is a random process in which MMTV integrate its genome near proto-oncogenes. HREs (hormonal responsive elements) located in LTR region of MMTV, also play role in carcinogenesis by up regulating the downstream genes in response to steroid hormones (Callahan and Smith, 2008).

Theodoru *et al*., (2007) reported 17 new insertion sites of MMTV that play important role in breast cancer development. Sites present adjacent to WNT and FGF genes are major insertion sites. In respective tumors it is observed that their expression is enhanced which confirms the enhancer mediated activation (Theodorou *et al.,* 2007). Activation of various oncogenes leads to tumor genesis in mammary gland (MacArthur *et al.,* 1995).

MMTV integration leads to the activation of many downstream genes of WNT pathway including MMP7(involve in invasion) **,** cyclin D1 and myc (proliferation), and Map3k8(high expressed in breast tumors) , these genes can act as potential biomarkers for detecting breast cancer at early stages (Theodorou *et al.,* 2007; Wang *et al.,* 2014). Rspo and FGF gene family are also characterized as important integration sites for MMTV. Overexpression of Rspo3 has been reported in epithelial mammary cell lines that are deficient in p-53 (Theodorou *et al.,* 2007).

MMTV proteins Env and SAg have high oncogenic potential. They reported that cells expressing ENV protein show low expression level of E- cadherin, they are also sensitive to apoptosis that is triggered by Tumor Necrosis Factor-α (TNF-α). They also investigate the oncogenic potential of Super Antigens (SAg) for this they transplant SAg expressing cell lines

in fat pads which resulted into the tumor development thus confirming their role as a carcinogen (Mukhopadhyay *et al.,* 1995; Katz *et al.,* 2005). Phosphorylation pattern of Signal peptide of ENV was also demonstrated that weather they are anti-oncogenic or carcinogen. This study found that phosphorylation of this protein play vital role in Erb-B signal pathways and in ribosomal protein of nucleolus L5 (Feldman *et al.,* 2012).

Hormones trigger the development of breast cancer in both mice and humans. In mice MMTV expression is high during pregnancy while in case of humans it has been reported that 62% cases of gestational breast cancer were screened positive for MMTV (Wang *et al.,* 2003). It is also stated that oncogenes are also somehow associated with MMTV Insertional mutagenesis (Johal *et al*., 2007).

In various studies they amplified viral sequences of Env and LTR regions from tumor tissues. In some studies they screen complete proviral sequences while in other they found that these sequences are of human origins upon sequencing. Zangen *et al.,* (2002) and Bindra *et al*., (2007) studied expression of ENV and found no results. Researchers also performed expression profiling of samples that were screened positive for env and they report the up regulation of 27 genes that involve genes from TNF-α, interferon and gene from Tumor Growth Factor β (TGFβ). In mice MMTV uses transferrin receptor to gain entry into the cell lines while in humans their proper mode of entry is still unknown. It has been suggested that sequencing of integration sites help in understanding that weather the mechanism of carcinogenesis is same in mice and humans or not (Fernandez-Cobo *et al*., 2006).

Risk of cancer can be increased by the presence of various viruses. EBV is one of the oncovirus that help in carcinogenesis as it acts as cofactor. It has been stated in various studies

that EBV and HPV co-infection is quite common in case of cervical cancer (Corbex *et al*., 2014; Aromseree *et al*., 2015). EBV is resident of B and T cells so it is easily attracted by pro inflammatory components towards the site of carcinogenesis. If we understand the role of EBV as helper of HPV it will help us in developing advanced therapies and prevention.

In some studies they proposed the hypothesis that EBV play important role in breast cancer as they may play vital role in early oncogenesis and then help in oncogenic property development i.e. angiogenesis, metastasis and invasiveness (Yasui *et al.*, 2001; Amarante and Watanabe, 2009). Main site for EBV infection and development are B, T cells and epithelial cells. Various factors that helps in EBV binding and their entry into the cells include, gp110 (help them to enter into both epithelial and B cells) Glyco protein 350 (gp350) (that help them to enter into the B cells through CD21), BMRF2 (viral receptor) and beta-1integrase (present on epithelial cells) helps them to enter into epithelial cells (Draborg *et al.*, 2013).

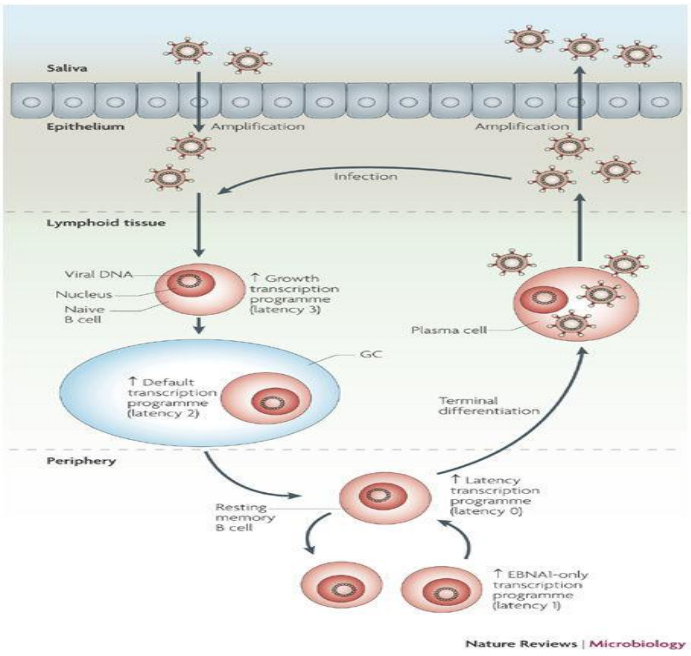

Figure 2.3: EBV Mechanism of Infection. (Thorley *et al.*, 2008)

There mode of infection is divided into lytic and latent phase. Latent phase is a condition in which they do not produce any virions, infection of epithelial cells and B lymphocytes are latent. Various malignancies, tumors and B cell immortalization caused by EBV latency because they can hide from immune system. They down regulate some of the target genes of Cytotoxic T cells (CTLs) in this way they escape from host immune system.

Epstein Bar Virus a dsDNA (with the size of 172 bp approx.) virus surrounded by octahedral capsid which is further enveloped by cell membrane of host. Its genome is divided into short and long sequences known as internal and terminal direct repeats, and can encode 100 genes. They mostly infect humans and can be transmitted through body fluid mostly through saliva (Glaser *et al.,* 2004). In various studies association between EBV and breast cancer was conformed through PCR analysis. High expression of IgA antibody is associated with presence of viral capsid of EBV (Alibek *et al.,* 2013).

Connection between EBV and breast cancer is confirmed in a study in which they screened EBV in both tissues as well as in peripheral blood (Perkins *et al.,* 2006). Infectious role of EBV is also investigated by transfecting breast cancer cell lines (BT474and MCF7) and as result they show tumorigenic activity this study confirm the oncogenic role of EBV. Data reported in various studies also shows that EBV is not involved in breast cancer because viral load detected in biopsies was very low (Arbach *et al.,* 2006).

Out of thirty two twenty five studies showed positive association of EBV with breast carcinoma and all of them are PCR based. In five studies they screen the presence of EBV through Immuno histochemistry (IHC) and in situ hybridization and find zero prevalence (Glenn *et al.,* 2012). Some of the normal breast tissues and breast milk also screened positive for EBV. Oncogenes expressed by EBV are Latent membrane Protein 1 an 2 (LMP1 and LMP

2) they leads to the transformation of fibroblast and lymphocytes and through their Ig receptors they also affect signaling pathways of lymphocytes respectively, but their positive association with breast cancer cannot be recognized yet (Lawson and Heng, 2010).

EBV also resulted into Immortalization of epithelial cells through a protein called BARF1 (Lawson and Heng, 2010). Oncogene activates some of the signaling pathways that play important role in carcinogenesis like HER2 and HER 3 that further activate Extracellular signal regulated kinases (ERK) and AKT (Glenn *et al*., 2012).

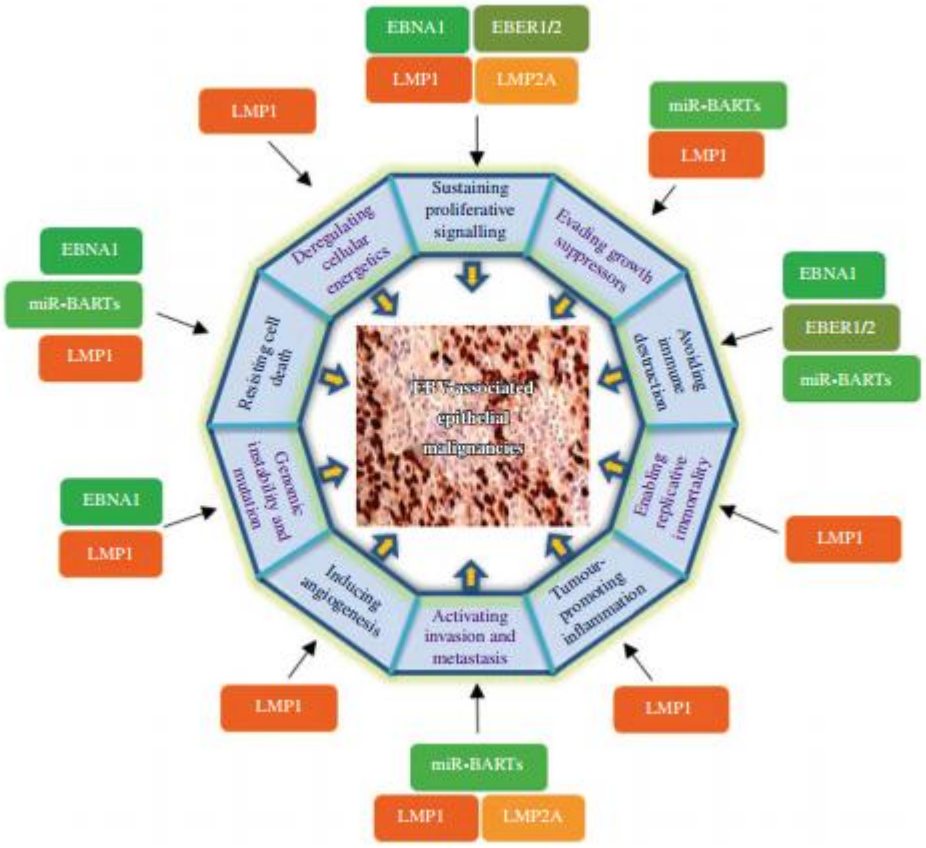

Figure 2.4: Oncogenic potential of different EBV latent genes (Tsao *et al.*, 2015).

.

HPV is Non-enveloped virus protected with icosahedral capsid (derived from 72capsomeres) and contains almost 8kb of ds DNA. Three kinds of viral protein are present major and minor capsid protein and late protein 1(L1). They contain two types of ORF late and early that are categorized on the basis of their genome location. Structural (6) and non-structural (2) genes are responsible for replication and assembly of HPV genome. They contain 7 early genes (E1-7) and two late genes(L1 and 2) (Kajitani *et al.,* 2012).c-myc that is a potential biomarker for breast cancer also shows significant relation with HPV 16, 6 and 7 as the elevated expression of c-myc was reported in the cells infected with HPV (Alibek, 2013).

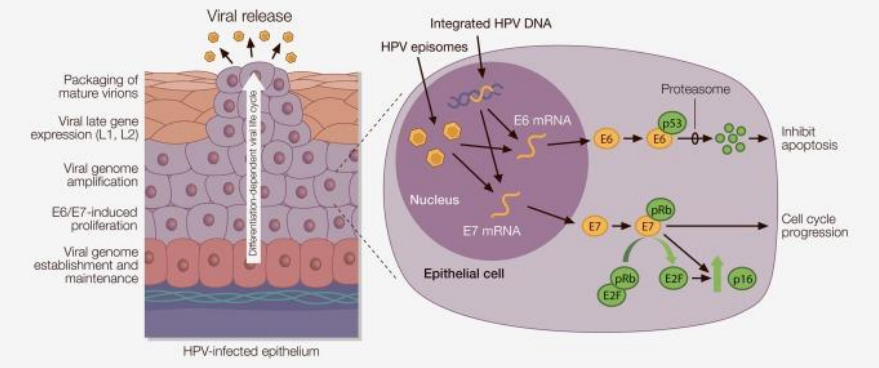

Figure 2.5: Mechanism of HPV infection. (Andersen *et al.,* 2014)

Studies shows that 29 % cases of breast cancer are screened positive for HPV while some studies show no association of HPV with breast cancer. Scientist are interested to study the association of HPV with breast cancer because risk factors involve in carcinogenesis are not completely known, risk of HPV infection in breast tissues of sexually active population is very high, mammary ducts are prone to infections because of their exposition to external environment (Wang *et al.,* 2012).Young women are sexually active then older ones so prevalence of HPV is high in young patients suffering from breast cancer. Presence of HPV

is also reported in patients suffering from both cervical and breast cancer (Amarante and Watanabe, 2009).

Ethnicity and genetic background also affect the prevalence of HPV as it varies from 0-86%, 13 % cases in Europe and 43% cases in America and Australia. Serotype 16, 18 and 33 are characterized as high risk HPV. Studies from China and Japan reported that rare subtypes of HPV (56 and 36) were screened positive in tumor tissues (Paoli and Carbone, 2013).

HPV can cause cancer in various vital organs because of their ability to transport themselves from one site to another through either hematological route or lymphatic system (Hennig *et al.,* 1999). It is reported in a study that 44% cases were positive for HPV subtype 33 and none of them were positive for high risk serotypes that are 16 and 18. But it is still unclear that they either infect before or after the development of cancer. It has been reported that HPV was screened positive in 4% samples of breast milk out of 223 samples regardless of cervical status, oral HPV status of mother, ethnicity and geographical data. Presence of HPV in breast milk is positively associated with oral HPV status (Sarkola *et al.,* 2008). More than one subtype can also involve in carcinogenesis. Invasive carcinoma is caused by E6 oncogene of HPV that leads to overexpression of Id-1(Akil *et al.,* 2006). It is stated that HPV holds almost similar oncogenic properties in both cervical and breast cancer so the use of HPV vaccine has been suggested to avoid HPV pathogenesis (Heng *et al.,* 2009).

It is quite difficult to detect HPV DNA because it is incorporated into host genome after various cycles, but HPV specific primers can be used for their detection. HPV leads to deactivation of tumor suppressor genes (p53 and RB1) through their early proteins E6 and E7 but alone they are not enough for the development of cancer (Narisawa-Saito and Kiyono,

2007). So it is concluded that integration sites of HPV are screened in breast cancer biopsies but their proper mechanism of carcinogenesis is still not known.

MMTV has a significant role as it is speculated to be a causative agent to breast cancer. On the contrary, HPV and EBV are not known to be directly causing the cancer. However, these viruses, upon integration, result in the down regulation of tumor suppressor genes whereas in the up regulation of proto oncogenes, thus causing breast carcinoma. Moreover, presence of one of these viruses creates a favorable microenvironment that encourages the infection and pathogenesis of another virus. Role of co infection in the development of breast carcinoma is still not clear.

Cervical cancer and Burkett's lymphoma these are the major diseases caused by HPV and EBV respectively. So, it is hypothesized that EBV and HPV infection in combination with other risk factors like hormonal level, age and diet may lead to the development of breast carcinoma. Patients suffering from breast carcinoma may get immuno-compromised because of the chemo and radio therapy such patients are more prone to viral infections which ultimately lead to invasiveness of breast carcinoma (Naushad *et al.*, 2017).

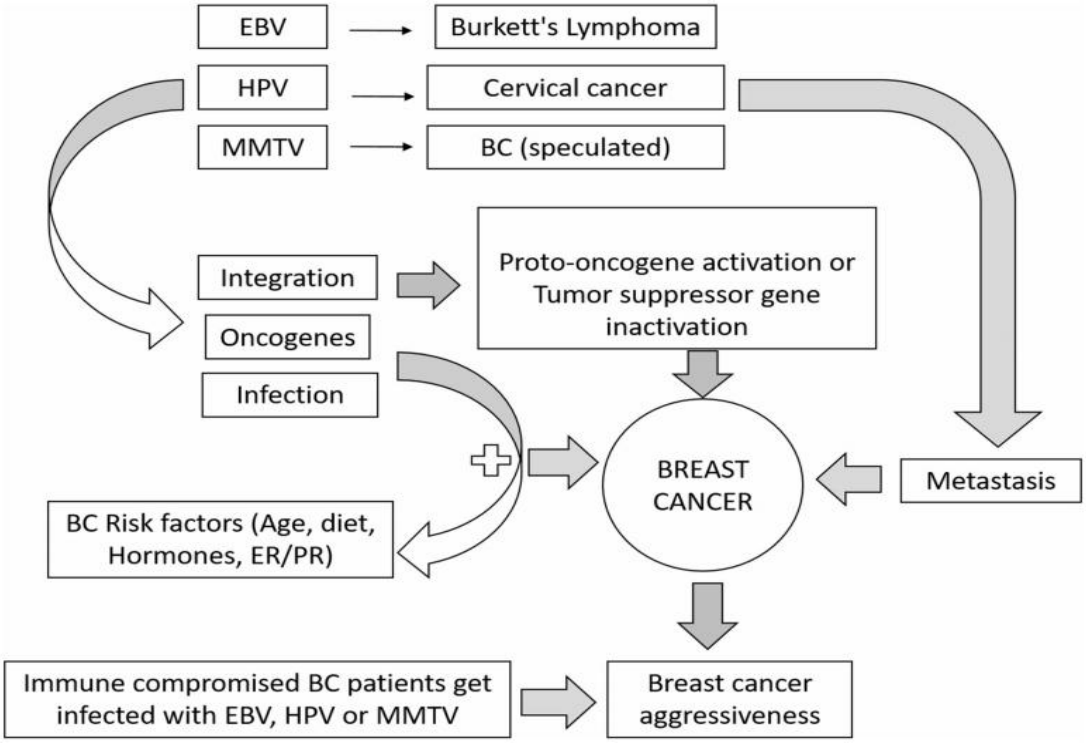

Figure 2.6: Role of viral etiology in breast cancer metastasis (Naushad *et al.*, 2017).

EBV help in oncogenesis of cervical cancer by acting as a cofactor same can be applied for breast cancer but the available data is still not sufficient to build a proper relationship between these viruses and breast cancer so understanding the role of these oncogenic viruses will help us to design biomarkers for early detection and cocktail chemotherapy for better treatment (Khenchouche *et al.*, 2013).

CHAPTER 3

MATERIALS AND METHODS

3.1 Sample collection

Blood (n=15) and tumorous tissue (n=13) samples of breast cancer patients were collected from oncology department of Holy family Hospital Rawalpindi through proper consent. Inclusion criteria include breast cancer patients and patients suffering from any other infection were excluded from the study. Blood samples of control (n=28) were also collected from healthy individuals of ASAB, NUST.

Labeled Sterile Ethylenediamine Tetraacetic Acid (EDTA) containing vacationers were used for collections of blood sample, tissue biopsies were collected in 15 mL falcon tubes labeled with patient Identity (ID). Collected samples were store at -20 °C.

Standard protocols were used to extract whole genomic DNA from blood as well as tissue samples.

3.2 Genomic DNA extraction from blood

In 1.5militer (mL) of autoclaved eppendorf tube 750 micro liter (uL) of blood and Solution A (Soln A) , i.e. lysis buffer (0.32 Molar (M) sucrose, 1M Tris Hydrochloride (Tris-HCl), Triton X100, 1M Magnesium Chloride ($MgCl_2$), was added and homogenized by inverting microfuge tube 3-4 times and then incubated for 10 minutes (mins) at room temperature. Pellet was washed by adding 400 uL of solution A followed by 2mins centrifugation at 13000 rpm. 5uL of proteinase K , 12uL of 20% Sodium dodecyl Sulfate

(SDS) and 400uL of Soln B (20 mili Molar (mM) Tris-HCl, 4mM EDTA disodium salt dehydrate (Na2EDTA) ,100mM Sodium Chloride (NaCl) at 7.4 Power of H+ (pH) was added in nuclear pellet and then it was incubated at 37 degree Celsius (°C) for overnight.

Fresh mixture containing equal volume of solution D (chloroform 24 Volume (vol): isoamylalcohol vol1) and C (Phenol) was prepared on next day and 500uL of this mixture was added in microfuge tube and then centrifuged for 10 mins at 13000 rpm. As a result two layers were formed, i) aqueous phase (containing nucleic acid) and second one is organic phase contain phenol. Through careful pipetting aqueous layer was transferred into a new sterile microfuge tube without disturbing the protein layer present between the aqueous and organic phase. Thread shaped DNA was precipitated by adding $1/10^{th}$ vol (approx.55uL) of 3M sodium acetate (pH6) and 2vol of 100% ice chilled ethanol followed by rapid inversion. DNA was pelleted out by 10 mins centrifugation at 13000 rpm and supernatant was discarded. 70% ethanol was used for further washing through centrifugation (13000 rpm for 7 min). Pellet was dried at room temp for 15-25 mins and then dissolved in 50 to 100uL of Tris-EDTA (TE) buffer either at room temperature or by incubating it for 1Hour (hr) at 37°C. Samples were stored at -20°C for long term

3.2.1 Genomic DNA extraction from tissue specimen

Sample preparation:

Samples were either processed within 2hours of collection or stored at -20°C for long term. Tissue was minced into tiny pieces (approx. 1-2 millimeter (mm)) in mortar with the help of liquid nitrogen and blade. Frozen powder of tissue is transferred into a sterile microfuge tube followed by the addition of 920 uL of DNA digestion buffer.

Cell lysis and digestion:

10% SDS (50uL) and proteinase K (30uL) was added in digestion buffer which resulted into the formation of viscous solution, solution was properly mixed through vortex. Samples were incubated either at 37°C for overnight or at 55°C for 3h which resulted into clear and viscous solution.

Extraction of DNA:

Half of the solution was transferred into new microfuge tube and then equilibrated phenol was added in equal amount gently followed by centrifugation at 1700g for 10mins. This resulted into the formation of two layers, i) aqueous phase (containing nucleic acid) and ii) organic phase contain phenol. Aqueous layer was transferred carefully into new sterile microfuge tube without disturbing protein layer. Centrifugation process was repeated by adding equal volume of chloroform.

Purification and precipitation of DNA:

DNA was precipitated by adding $1/10^{th}$ vol of 3M sodium acetate (pH 5.2) and 2vol of 100% ice chilled ethanol to aqueous phase followed by rapid inversion. If the DNA is not precipitated then incubate the mixture at -20°C for half an hour. After this samples were centrifuged at full speed (approx. 1300 rpm) for 10mins supernatant was discarded and pellet was washed again with 70% ethanol through centrifugation (13000 rpm for 7 min). DNA pellets was dried at room temp for 15-25 mins and then dissolved into 50 to 100uL of TE buffer either at room temperature or by incubating for 1 hour at 37°C. Samples were either stored at 4°C or for long term storage at -20°C. Only 25uL of TE buffer (for dissolving DNA pellet) was added if the yield of DNA was low.

Quantification of DNA:

Concentration of DNA was calculated in two ways.

I. Agarose Gel electrophoresis

II. DNA Quantification

Agarose gel electrophoresis:

Agarose gel was used to determine the quantity of DNA. 2uL of DNA sample mixed with 3uL of loading dye was loaded in 1% agarose gel. After running the gel at 90volts for 15mins it was assessed visually and also in gel documentation system for taking proper photograph.

DNA qualification:

Nanodrop-2000 was used to analyze the extracted DNA for this purpose 1uL of DNA sample was used. Value at 260 nano meter (nm) as well as the ratio of 260/280 nm was checked to analyze the quality of DNA. Value less than 1.8 showed that samples are contaminated either by protein or phenol.

3.3 Amplification of housekeeping gene (β-Globin)

Beta globin characterized as a house keeping gene was amplified using specific primers to check the integrity of DNA. For this purpose 12.5 uL of reaction mixture was prepared that contain 1X Thermus aquaticus (Taq) buffer, 25mM $MgCl_2$, 2mM Deoxyribonucleotide triphosphate (DNTPs), 10 Pico molar (p mol) forward and reverse primer, 1Unit (U) of thermo stable Taq Polymerase and 50 nanogram (ng) i.e almost 1uL of

DNA template and then final volume of 12.5uL was made through PCR water. Sequences of primers used for Beta globin amplification are mentioned below in table.

During master mix preparation proper measures were taken to avoid contamination and then PCR was performed using standard profile for the amplification of 110 bp of beta globin. PCR profile that was used for amplification; initial denaturation for 5 minutes at 95°C, denaturation for 30 sec at 95°C, primer annealing at 54 °C for 30 Seconds (sec) , extension for 30 sec at 72°C for 35 cycles , final extension for 10mins at 72 °C.

Table 3.1: Primer Sequences for amplification of β-globin gene.

Serial no.	**Primer**	**sequence 5'—> 3'**	**Melting Temperature**
1	β-globin Forward	ACACAACTGTGTTCACTAGC	58.4^0C
2	β-globin Reverse	CAACTTCATCCACGTTCACC	60.4^0C

Analysis of PCR product by gel electrophoresis:

PCR products were analyzed through 2% agrose gel that was prepared in 1X Tris/Borate/EDTA (TBE) buffer and same buffer was used in gel tank to run the gel. 5uL of PCR product mixed with 4uL of loading dye was loaded along with the 2 uL of 100bp DNA marker or ladder. After running the gel at 100volts for 30mins it was visualized and photographed in gel documentation system.

3.4 Detection of HPV L1 Gene using GP5+/GP6+ primers

Table 3.2 includes the primer sequences used for the amplification of L1 gene of HPV.

Table 3.2: Primer Sequence for amplification of GP5/GP6+

Serial no.	**Primer**	**sequence 5'—> 3'**	**Melting Temperature**
1	GP5+	TTTGTTACTGTGGTAGATACTAC	49.9 °C
2	GP6+	GAAAAATAAACTGTAAATCATATTC	46.2 °C

Optimization of HPV GP5+/GP6+ primers and screening of samples:

For optimization of these primers DNA sample of positive control was used. Gradient PCR was set up at 43, 45, and 49°C annealing temperature and HPV DNA fragment of 150 bp was amplified success fully at 43°C annealing temperature. After successful optimization of HPV primers DNA extracted from blood and tissue samples were screened for the presence of HPV DNA fragment. For this purpose Reaction mixture of 12.5 uL was prepared that contain Nuclease free water, 1X Taq buffer , 25mM $MgCl_2$, 2mM DNTPs , 10pmol forward and reverse primer, 1U of thermo stable Taq pol and 50ng(1uL) of DNA template. PCR profile that was used for amplification; initial denaturation for 5 minutes at 95°C, denaturation for 40 sec at 95°C, primer annealing at 43 °C for 45 sec , extension for 40 sec at 72°C for 35 cycles , final extension for 10mins at 72 °C.

Analysis of PCR product by agarose gel electrophoresis:

To analyze the 150 bp of PCR product 2% gel, 2 grams (g) of agrose was dissolved in 100mL of 1XTBE buffer followed by boiling in microwave oven upon slightest cooling 10uL of EtBr is mixed in agrose solution and then poured in gel caster. 5uL of PCR product mixed with 4uL of loading dye and loaded into the gel along with 100bp DNA ladder. After running the gel for 30 mins at 100volts gel was assessed under ultra violet light and Gel documentation system was used for taking photograph.

3.5 Detection of EBV DNA using EBER 2 primers

Screening of EBV DNA in breast cancer samples was performed by using EBV specific primers that are mentioned below.

Table 3.3: Primer Sequence for Screening EBV

Serial no.	Primer	sequence 5'—> 3'	Melting Temperature
1.1	EBER2F	AGG ACA GCC GTT GCC CTA GTG	58.3 °C
2.2	EBER2R	TAG CGG ACA AGC CGA ATA CCC T	56.7 °C

Optimization of EBV using EBER2F/EBER2R primers and screening of samples:

For optimization of these primers DNA sample of positive control was used. Gradient PCR was set up at 51, 56 and 60°C annealing temperature and EBV DNA fragment of 170 bp was amplified success fully at 51°C annealing temperature. After successful optimization of

EBER primers DNA extracted from blood and tissue samples were screened for the presence of EBV DNA fragment. For this purpose Reaction mixture of 12.5 uL was prepared that contain Nuclease free water, 1X Taq buffer , 25mM $MgCl_2$, 2mM DNTPs , 10pmol forward and reverse primer, 1U of thermo stable Taq pol and 50ng(1uL) of DNA template. PCR profile that was used for amplification; initial denaturation for 5 minutes at 95°C, denaturation for 40 sec at 95°C, primer annealing at 51 °C for 45 sec extension for 40 sec at 72°C for 35 cycles , final extension for 10mins at 72 °C.

Analysis of PCR product by agarose gel electrophoresis:

To analyze the 150 bp of PCR product 2% gel, 2g of agrose was dissolved in 100mL of 1XTBE buffer followed by boiling in microwave oven upon slightest cooling 10uL of EtBr is mixed in agrose solution and then poured in gel caster, 5uL of PCR product mixed with 4uL of loading dye and loaded into the gel along with 100bp DNA ladder. After running the gel for 30 mins at 100volt gel was assessed under UV light and Gel documentation system was used for taking photograph.

3.6 Detection of MMTV

Table 3.4 includes the primer sequences for the MMTV LTR regions.

Table 3.4: Primer Sequence for Screening of LTR Region.

Serial no.	**Primer**	**sequence 5'—> 3'**	**Melting Temperature**
1	LTR3	CGTGTGTTTGTGTCTGTTCG	54.5^0C
2	LTR5	GGTGGCAACCAGGGACTTAT	57.3^0C

Optimization of MMTV LTR region using LTR3/LTR5 primers and screening of samples:

For optimization of these primers Mice DNA sample was used. Gradient PCR was set up at 48, 49 and 50°C annealing temperature and LTR gene of 663 bp was amplified success fully at 50C and 49°C annealing temperature. After successful optimization of LTR primers DNA extracted from blood and tissue samples were screened for the presence of LTR region. For this purpose Reaction mixture of 12.5 uL was prepared that contain Nuclease free water, 1X Taq buffer , 25mM $MgCl_2$, 2mM DNTPs , 10pmol forward and reverse primer, 1U of thermo stable Taq pol and 50ng(1uL) of DNA template. PCR profile that was used for amplification; initial denaturation for 5 minutes at 95°C, denaturation for 30 sec at 95°C, primer annealing at 49 °C for 40 sec , extension for 40 sec at 72°C for 35 cycles , final extension for 10mins at 72 °C.

Detection of MMTV ENV region:

Table 3.5 includes the primer sequences used for the amplification of LTR region of MMTV.

Table 3.5: Primer Sequence for Screening ENV Region of MMTV

Serial no.	**Primer**	**sequence 5'—> 3'**	**Melting Temperature**
1	ENV-F	CCTCACTGCCAGATC	48.7^0C
2	ENV-R	CTATCTGTGGCATACCT	45.8^0C

Optimization of MMTV ENV region using ENV-R/ENV-F primers and screening of samples:

For optimization of these primers Mice DNA sample was used. Gradient PCR was set up at 48, 49 and 50°C annealing temperature and ENV gene of 663 bp was amplified success fully at 50°C annealing temperature. After successful optimization of ENV primers DNA extracted from blood and tissue samples were screened for the presence of ENV region of MMTV. For this purpose Reaction mixture of 12.5 uL was prepared that contain Nuclease free water, 1X Taq buffer, 25mM $MgCl_2$, 2mM DNTPs , 10pmol forward and reverse primer, 1U of thermo stable Taq pol and 50ng(1uL) of DNA template. PCR profile that was used for amplification; initial denaturation for 5 minutes at 95°C, denaturation for 30 sec at 95°C,

primer annealing at 50 °C for 40 sec , extension for 40 sec at 72°C for 35 cycles , final extension for 10mins at 72 °C.

Analysis of PCR product by agarose gel electrophoresis:

To analyze the PCR product 2% gel, 2g of agrose was dissolved in 100mL of 1XTBE buffer followed by boiling in microwave oven upon slightest cooling 10uL of EtBr is mixed in agrose solution and then poured in gel caster, 5uL of PCR product mixed with 4uL of loading dye and loaded into the gel along with 100bp DNA ladder. After running the gel for 30 mins at 100volt gel was assessed under UV light and Gel documentation system was used for taking photograph.

3.7 Statistical analysis

Statistical analysis was performed using SPSS (Social Science Software) and Graph pad prism to find out the association of viral presence with breast cancer as well as lymphnode involvement. Prevalence of viral sequence and their coinfection was also studied in various factors including types of breast carcinoma, age, and tumor grade.

CHAPTER 4

RESULTS

4.1 Data collection

In this study 28 blood (n=15) and tumor tissue (n=13) samples of breast cancer patients and 28 controls taken from healthy females were included. Clinical and pathological data was also gathered. Age group of the patients ranged from 20 to above 60. Figure 4.1shows the no of breast cancer patients falling in different age categories. According to this bar chart 27% patient fall in the range of 41-49 and 50-60years, 20% fall in the range of 30-40 and above 60years while only 6% cases fall in the range of 21-29years.

Figure 4.2 depicts the percentage distribution of type of carcinoma and according to this data invasive ductal carcinoma is a dominant type of breast cancer as it is observed in 60% cases. Tumor grade II (67%) and grade III (27%) were found to be common among patients (fig 4.3). According to Fig 4.4 47% patients were suffering from stage II, 40% stage I while only 13% were suffering from stage III of breast cancer.Further detail which includes Estrogen/Progesterone (ER/PR) status, metastasis status, and involvement of lymph node is listed in table 4.1.

Table 4.1: Clinical and pathological data of breast cancer patients.

Age Group	
21-29	1
30-40	3
41-49	4
50-59	4
> 60	3
Type of Breast Carcinoma	
Invasive Ductal Carcinoma(IDC)	9
Ductal Carcinoma in situ (DCIS)	4

Spindle Cell Carcinoma	1
Papillary Carcinoma	1
Tumor Grade	
Grade I	1
Grade II	10
Grade III	4
ER/PR Status	
Absent	12
Present	3
Lymph Node Involvement	
Yes	7
No	8

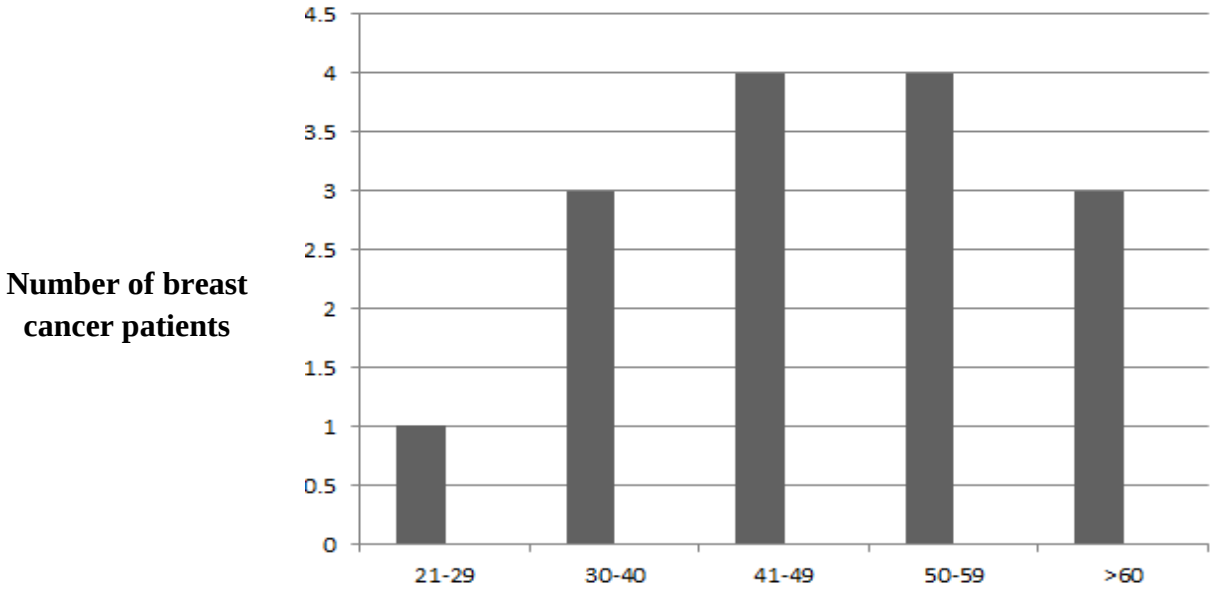

Figure 4.1: Age distribution of breast cancer patients. This bar chart shows that 27 % patient fall in the range of 41-49 and 50-60years, 20% fall in the range of 30-40 and above 60years while only 6% cases fall in the range of 21-29years in studied samples.

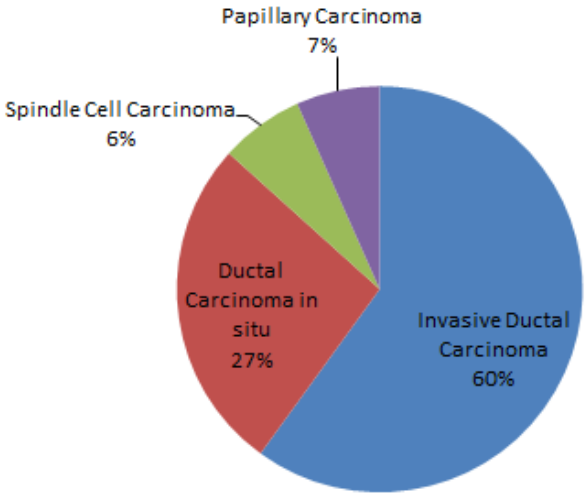

Figure 4.2: Percentage distribution of breast cancer patients falling in different categories of breast carcinoma. According to this 60 % cases fall in IDC while 27% in DCIS.

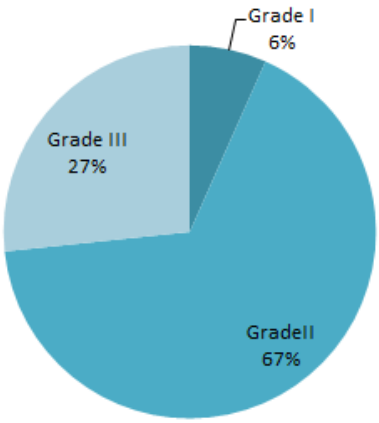

Figure 4.3: Percentage distribution of breast cancer patients falling in different categories of tumor grade. This shows that 67% cases fall in tumor grade II.

4.2 DNA extraction

DNA was extracted from tissue and blood samples of breast cancer patients and blood samples of healthy control through phenol chloroform method (detail description is given in chapter 3). Extracted genomic DNA is shown in figure 4.5.

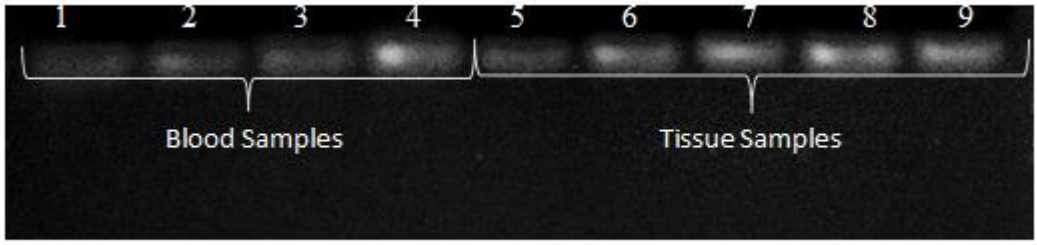

Figure 4.4: Representative gel of Genomic DNA. Lane 1-4 depicts DNA extracted from tissue samples and lane 5-9 depicts DNA extracted from blood samples.

4.3 Amplification of housekeeping gene

Beta- globin (an internal control) was used to check the integrity of extracted DNA. 110 bp of beta globin was successfully amplified in all DNA samples. Some of them are shown in figure 4.2.

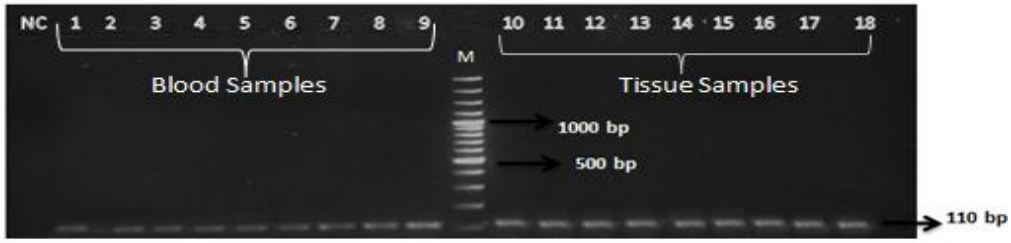

Figure 4.5: Representative gel of amplified 110bp of β-globin gene fragment. Lane 1-9 depicts beta globin amplified from blood samples. Lane 10-18 depicts beta globin amplified from tissue samples of breast cancer patients loaded along with 100 bp DNA marker.

4.4 Optimization of GP5+/GP6+ primers for HPV L1 gene

Positive control was used to optimize GP5+/GP6+ primers that target L1 gene of HPV. For this gradient PCR was set up at 43, 45 and 48°C. Figure 4.3 shows that primers are successfully optimized at 43°C.

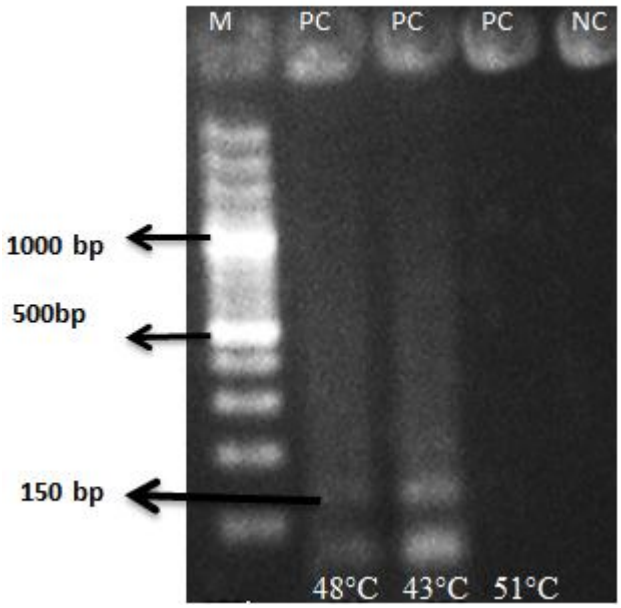

Figure 4.6: Optimization of GP5+/GP6+ Primers. Lane 2 contains 150 bp of PCR product amplified from positive control at 48°C lane 3 contains positive control amplified at 43°C annealing temperature. While lane 4 contains no PCR product at 51°C annealing temperature and lane 5 contain reagent control.

4.4.1 HPV detection in breast cancer samples and controls:

Breast cancer samples were screened for the presence of L1 gene, out of 28, 4 tissue and 4 blood samples were screened positive for the presence of L1 gene of HPV while controls were screened negative for the presence of L1 gene. PCR products of positive samples were loaded in 2% agarose along with 100bp DNA ladder.

Figure 4.7: Representative gel of amplification of HPV L1 gene in control samples. Lane M contains 100bp DNA marker and lane PC and NC contain positive control of HPV and reagent control respectively. Lane 1-17 contains control samples with negative results.

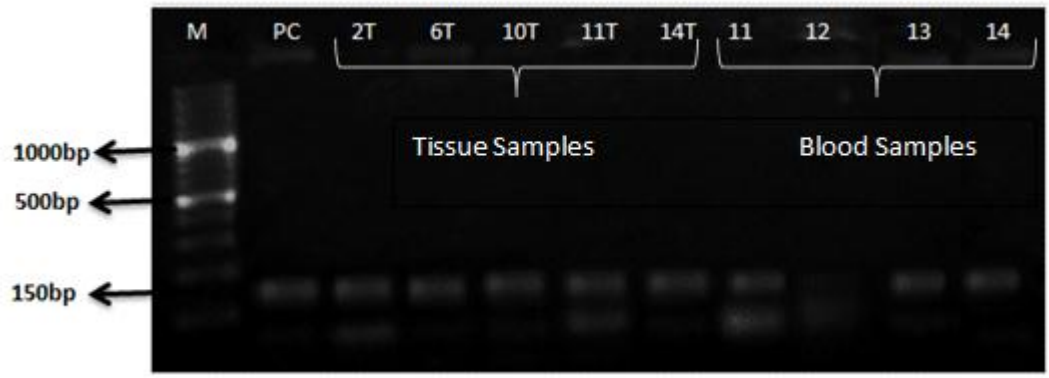

Figure 4.8: Amplification of HPV L1 gene in breast cancer samples. Lane 2 contains PCR product amplified from positive control. Lane 2T to 14 T show PCR product amplified from tissue samples, while lane 11- 14 show PCR products amplified from blood samples of breast cancer patients.

4.5 Optimization of EBER2R/EBER2F primers for EBV

Positive control was used to optimize EBER2R/EBER2F primers that are specific for EBV DNA detection. For this gradient PCR was set up at 51, 56 and 60°C. Figure 4.5 shows that primers are successfully optimized at 51°C.

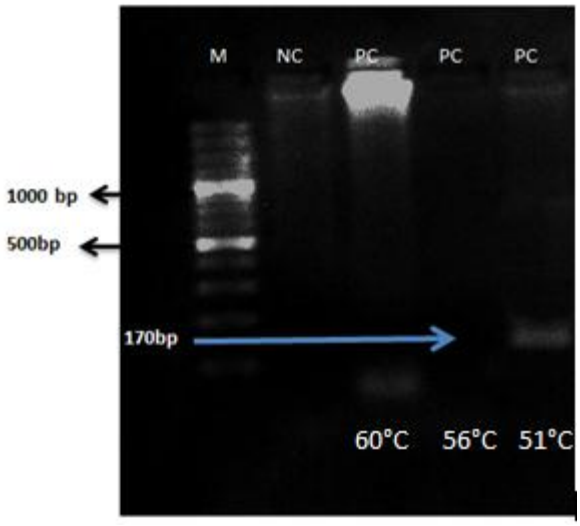

Figure 4.9: Optimization of EBER2R/EBER2F primers. Lane 2 contains negative control lane 3 and 4 contain no PCR product at 60°C and 56°C annealing temperature respectively while Lane 5 contains 170 bp of PCR product that is successfully optimized at 51°C.

4.5.1 EBV detection in breast cancer samples and controls:

Breast cancer samples were screened for the presence of EBV DNA fragment, out of 28 tissue and 6 Blood samples and 7 tissue samples were screened positive while controls were screened negative for the presence of EBV DNA fragment. PCR products of positive samples were loaded in 2% agarose along with 100bp DNA ladder.

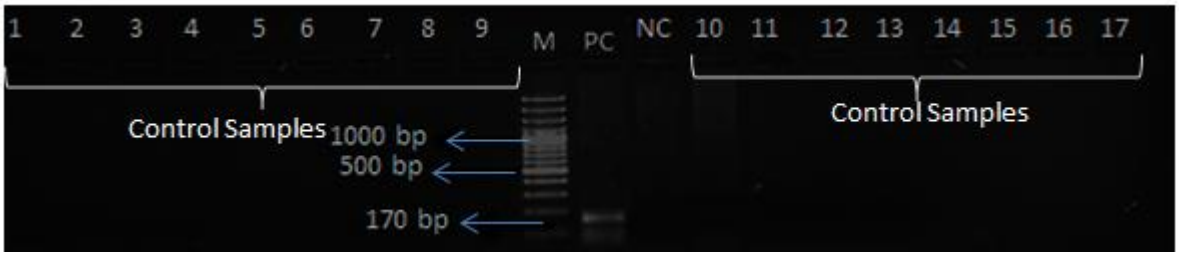

Figure 4.10: Representative gel of amplification of EBV DNA in control samples. Lane M contains 100bp DNA marker and lane PC and NC contain positive control of EBV and reagent control respectively. Lane 1-9 and 10-17 contains control samples with negative results.

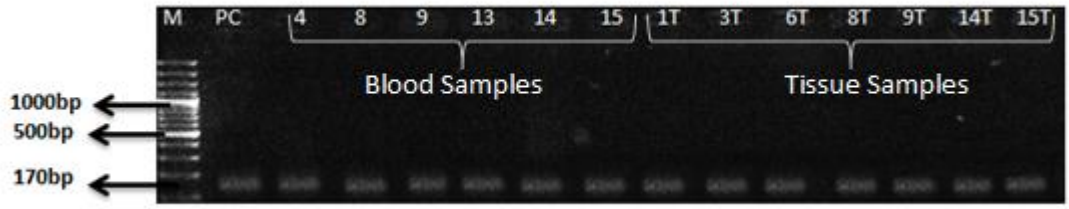

Figure 4.11: Detection of EBV DNA in breast cancer samples. Lane 2 contains PCR product amplified from positive control. Lane 4-15 contains PCR product amplified from blood samples. Lane 1T to 15T contains PCR products amplified from tissue samples.

4.6 Optimization of LTR3/LTR5 primers for MMTV LTR

DNA extracted from mice samples were used as Positive control to optimize LTR3/LTR5 primers that are specific for the detection of LTR region. Primers were successfully optimized at 49°C annealing temperature as shown in figure 4.5.

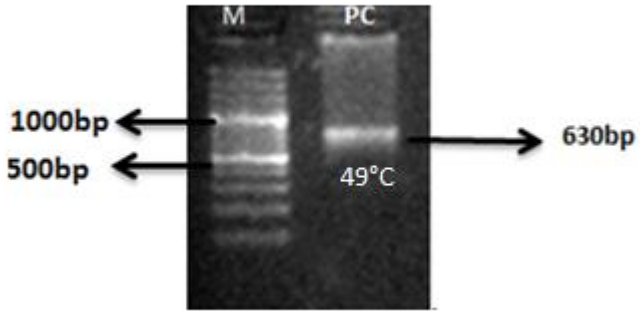

Figure 4.12: Optimization of LTR 3/LTR 5 primers. Lane 2 contains 630 bp of MMTV LTR that was successfully optimized at 49°C annealing temperature from the DNA extracted from mice who were suffering from mammary tumor.

4.6.1 Screening of breast cancer and control for MMTV LTR region:

Breast cancer samples were screened for the presence of LTR region, out of 28, 4 tissues and 3 blood samples were screened positive while controls were screened negative for the presence of LTR region of MMTV. PCR products of positive samples were loaded in 2% agarose along with 100bp DNA ladder.

Figure 4.13: Representative gel of amplification of MMTV LTR region in control samples. Lane M contains 100bp DNA marker and lane PC and NC contain positive control of MMTV LTR and reagent control respectively. Lane 1-12 contains control samples with negative results.

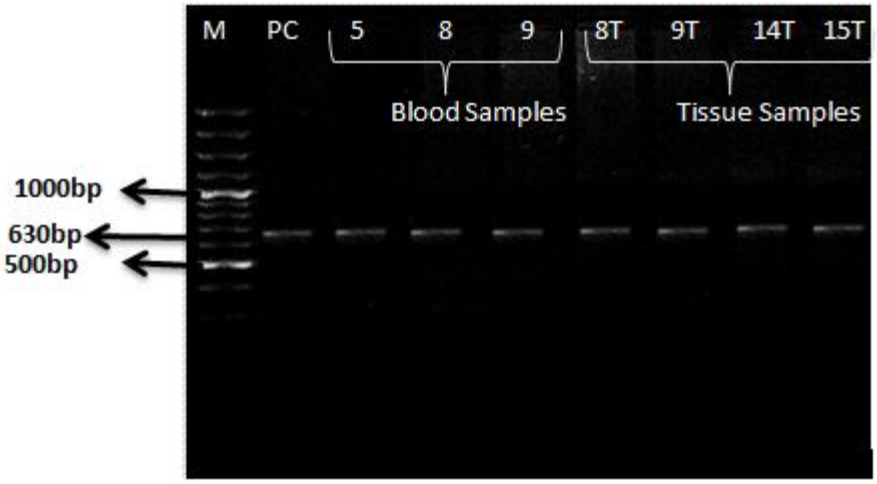

Figure 4.14: Detection of MMTV LTR region in breast cancer samples. Lane 2 contains PCR product amplified from positive control. Lane 5-8 contains PCR product amplified from blood samples and lane 8T-15T contains PCR product amplified from tissue samples of breast cancer patients.

4.7 Optimization of ENV-R/ENV-F primers for MMTV ENV

DNA extracted form MMTV infected mice were used as positive control to optimize ENV-R/ENV-F primers that are specific for the detection of ENV region. Primers were successfully amplified at 50°C annealing temperature.

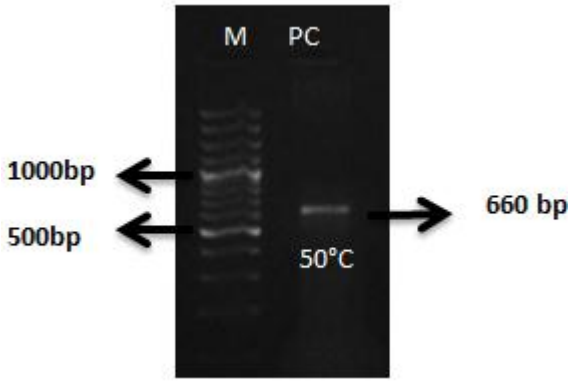

Figure 4.15: Optimization of ENV-R/ENV-F primers. Lane 2 contains 660 bp of MMTV ENV that was optimized at 50°C annealing temperature, in DNA sample extracted from mice suffering from mammary tumor.

4.7.1 Screening of breast cancer and control for MMTV ENV region:

Breast cancer samples were screened for the presence of ENV region, out of 28, 3 tissue and none of the blood samples were screened positive while controls are negative for the presence of ENV region of MMTV. PCR products of positive samples were loaded in 2% agarose along with 100bp of DNA ladder.

Figure 4.16: Representative gel of amplification of MMTV ENV region in control samples. Lane M contains 100bp DNA marker and lane PC and NC contain positive control of MMTV ENV and reagent control respectively. Lane 1-12 contains control samples with negative results.

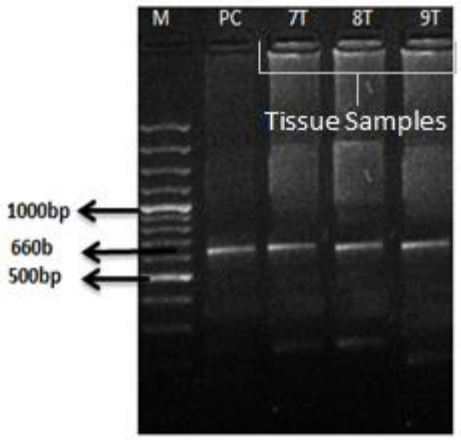

Figure 4.17: Detection of MMTV ENV region in breast cancer samples. Lane 2 contains PCR product amplified from positive control. Lane 7T-9T contains 660 bp of MMTV ENV region that was amplified from breast cancer tissue samples.

4.8 STATISTICAL ANALYSIS

4.8.1 Prevalence of mono and co-infection in breast cancer patients:

Fig 4.15 shows viral prevalence in breast cancer patients. This data suggest that in 28.57% cases no viruses were detected while mean mono and co infection is observed in 39 and 32% cases respectively.

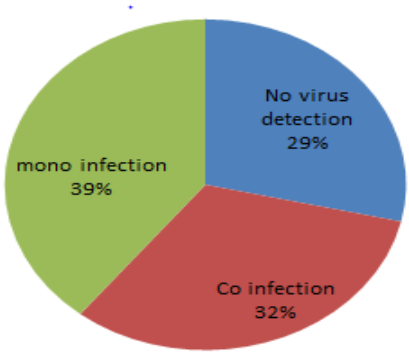

Figure 4.18: Mean mono-infection (39%) and co-infection (32%) of EBV, MMTV LTR, MMTV ENV and HPV in studied breast cancer samples.

4.8.2 Comparative analysis of tissue and blood samples:

The comparative analysis showed positive results for both blood and tissue samples in case of MMTV LTR, EBV and HPV. In case of MMTV LTR, EBV, HPV no significant difference can be observed in viral expression within tissue and blood samples of breast cancer, however, in case of MMTV ENV no positive results were generated in case of blood samples.

Table 4.2: Comparative analysis of viral expression in blood and tissue samples of breast cancer patients.

	MMTV LTR		**MMTV ENV**		**EBV**		**HPV**	
	Blood Samples	**Tissue Samples**	**Blood samples**	**Tissue samples**	**Blood samples**	**Tissue samples**	**Blood Samples**	**Tissue Samples**
Positive	20%	33.3%	0%	25%	33.3%	58.3%	26.67%	41.6%
Negative	80%	66.6%	100%	75%	66.%	41.6%	73.33%	58.3%

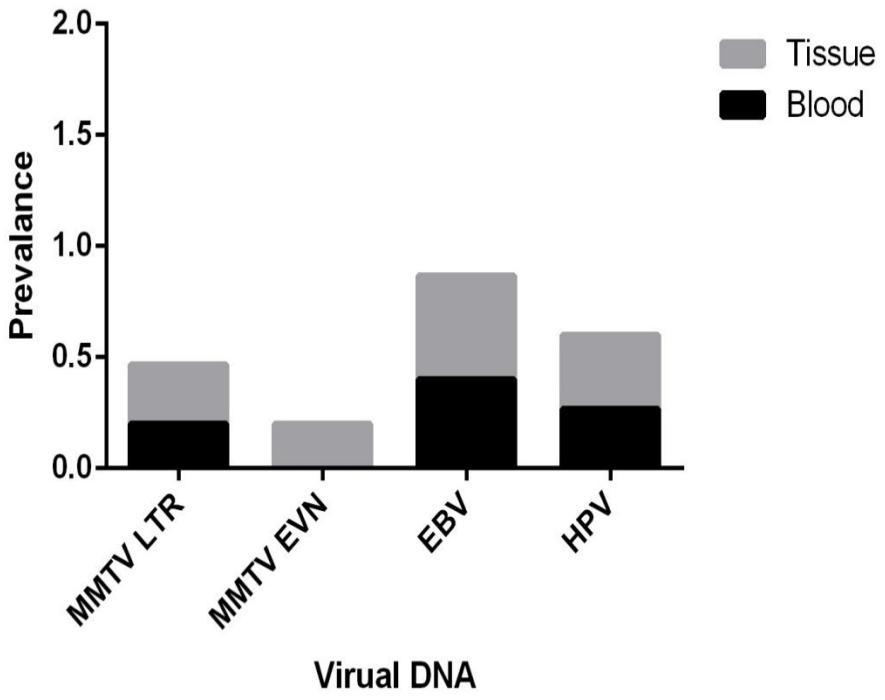

Figure 4.19: Graphical representation of breast cancer samples that were screened positive for viral presence in blood and tumorous tissue of breast cancer.

4.8.3 Association of viruses with breast cancer:

Fisher exact test was performed separately on each virus to find out the association of these viruses with breast cancer development. Results can be seen in figure 4.16. Results showed that EBV, HPV and MMTV LTR are significantly associated with breast cancer development in studied samples while there was no significant association between MMTV ENV and breast cancer

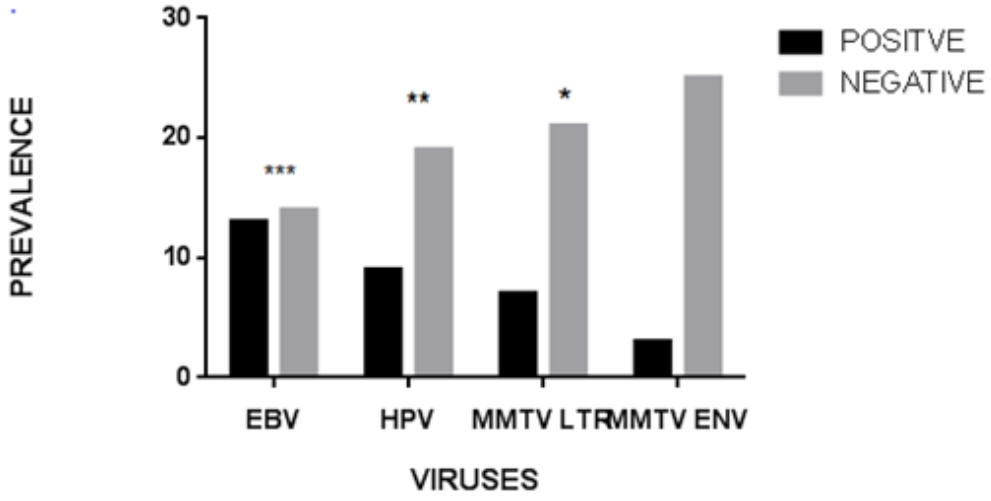

Figure 4.20: Association of MMTV LTR, ENV, EBV and HPV in Breast cancer patients. Fisher exact test showed that EBV **(***= 0.0001**), HPV **(**=0.0018**) and MMTV LTR **(*=0.0102**) are significantly associated with breast cancer development. While MMTV ENV (**P value= 0.2364**) showed negative association with breast cancer development in studied samples of breast cancer.

4.8.4 Frequency of MMTV, HPV, EBV and their coinfection in lymph node involvement:

Results obtained shows that EBV, MMTV LTR ,MMTV ENV , HPV and their viral co-infection is significantly associated with Lymph node involvement as the p-value (0.035) obtained is less than 0.05.

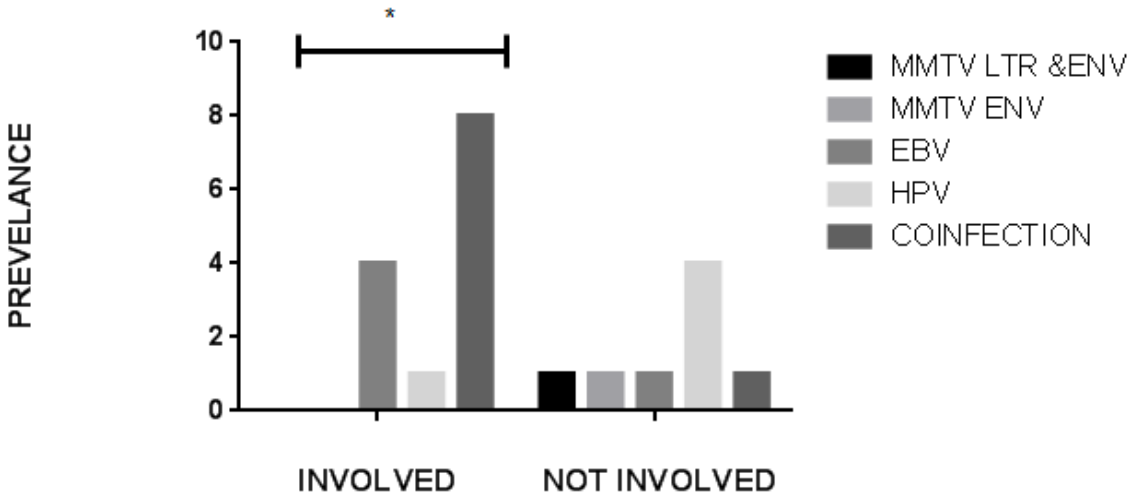

Figure 4.21: Frequency of MMTV LTR, ENV, EBV, HPV and their co-infection in lymph node involvement. P value (*=0.035) for these viruses and their co-infection show statistically significant results in studied samples of breast cancer.

4.8.5 Frequency of MMTV, EBV, HPV and their co-infection in different types of breast carcinoma:

Data obtained was graphically represented. It was observed that viruses and their co-infection are mostly involved in Invasive ductal carcinoma (IDC) and Ductal carcinoma in situ (DCIS) that are the common types of breast cancer.

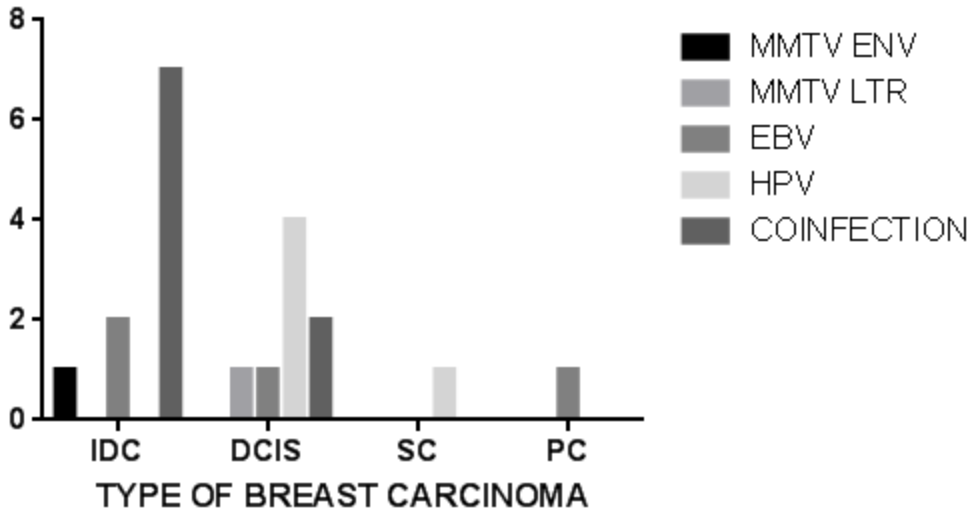

Figure 4.22: Frequency of EBV, MMTV LTR, MMTV ENV, HPV and their co-infection in different types of breast carcinoma**.** This shows that in studied samples EBV and its co-infection is mostly involved in IDC while HPV is commonly observed in DCIS.

4.8.6 Frequency of MMTV, HPV, EBV and their co-infection in different categories of tumor grade:

Frequency of MMTV LTR, MMTV ENV, EBV and HPV individually and synergistically was represented graphically. Data shows that viruses and their coinfection are mostly involved in tumor grade 2.

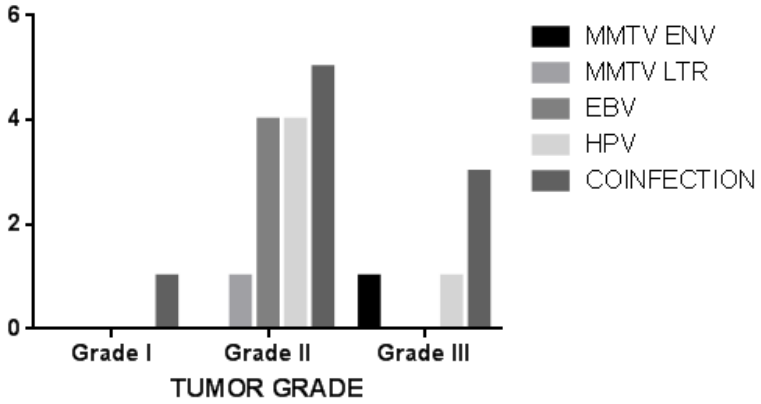

Figure 4.23: Frequency of MMTV, EBV, HPV like DNA and their co-infection in tumor grade. Viral regions are mostly screened in the samples that fall in the category of tumor grade II.

4.8.7 Frequency of MMTV, HPV, EBV and their co-infection in below and above 40 age categories:

Occurrence of these viruses and their confection was observed in age group above and below 40 years. Data is represented graphically. This data suggest that patients above 40 years showed greater viral prevalence.

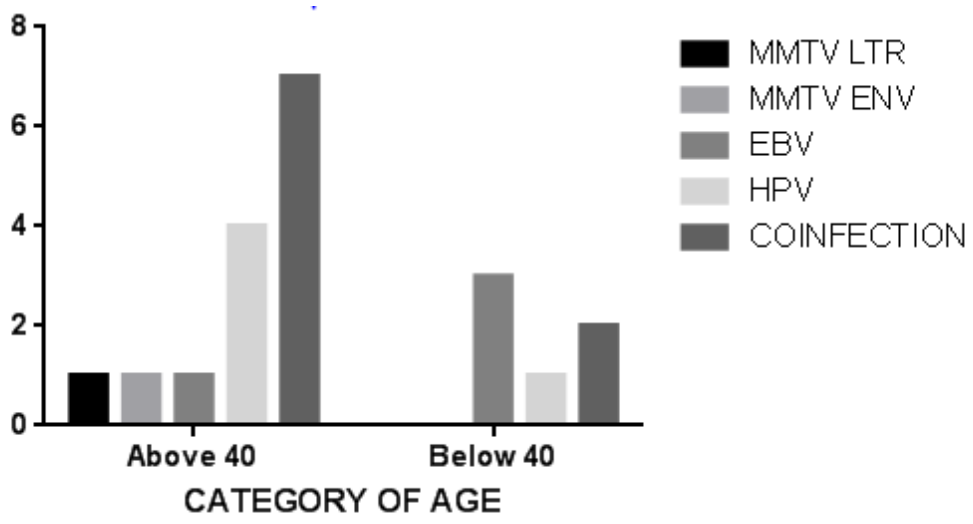

Figure 4.24: Frequency of EBV, HPV, MMTV and their co-infection in age group above and below 40. This chart depicts that viral regions are mostly screened in the patients above 40 years.

CHAPTER 5

DISCUSSION

Data obtained in this study showed mono-infection (39%) as well as coinfection (32%) of MMTV, HPV and EBV viruses in breast cancer patients. Statistical analysis showed that breast carcinoma is significantly associated with MMTV LTR, EBV and HPV while there was no significant association for MMTV ENV.

A number of studies have reported successful amplification of ENV and LTR regions of MMTV from tumorous tissues of breast cancer patients. Ford *et al.,* (2004) reported that 32% samples of breast cancer were screened positive for MMTV. In a study conducted by Pogo and Holland in 1997 MMTV env gene was screened positive in 39%. Wang *et al.,* (2001) reported the presence of MMTV LTR regions in 20% cases of Pakistani breast cancer patients and 26% cases were screened positive for MMTV ENV region in a study conducted by Naushad *et al.,* (2014). Some studies even reported complete proviral sequence screened from the patients of breast cancer and these studies supported our findings.

However, variation exists as studies have reported as high as 78% prevalence in Tunisian population (Hachana *et al.,* 2008) to almost no prevalence of MMTV at all (Fukuoka *et al.,* 2008; Witt *et al.,* 2003; Zangen *et al.,* 2002). These variations can be justified by different geographical regions or may be due to different strategies used for detection.

Our data suggested that MMTV LTR (*=0.0102) is significantly associated with breast cancer development. While MMTV ENV (P value= 0.2364) showed negative association with breast cancer development. However, number of samples included in this study was not enough for suggesting and understanding the role of MMTV in oncogenesis.

EBV may not have direct oncogenic properties in breast cancer development but it might act as a helper virus or cofactor in collaboration with other viruses such as MMTV and HPV. Out of thirty two, twenty five studies showed positive association of EBV with breast carcinoma and all of them are PCR based. In five studies they screen the presence of EBV through IHC and in situ hybridization and find zero prevalence (Glenn *et al.,* 2012).

EBV also resulted into immortalization of epithelial cells through a protein called BARF1 (Lawson and Heng, 2010). BARF1 is also characterized as an oncogene because it activates some of the signaling pathways that play important role in carcinogenesis like HER2 and HER 3 that further activate ERK and AKT (Glenn *et al*., 2012).

In vivo studies have reported interaction between the oncoproteins LMP1 of EBV and E6 & E7 of HPV, reporting a reduced level of apoptosis (Hagensee *et al*., 2011). These findings support the idea of virus collaboration in breast cancer development. In the recent studies it was proposed that EBV might gain the oncogenic potential when present in a chronic inflammatory environment. Thus EBV is assumed to be linked with aggressive forms of breast cancers (Aguayo *et al*., 2011).

We observed coinfection in 32% cases and these findings are supported by the study of Shi and his co-workers (Shi *et al*., 2016) where they reported higher viral load of EBV in case of co-infection with HPV. Coinfection occurs when inflammatory cytokines are released by the cells that are infected by HPV thus providing suitable conditions for EBV infection.

In this study we screened tissue and blood samples of breast cancer patients. Statistical analysis (Fisher exact test) of our data showed that EBV (*******= 0.0001) is significantly associated with breast cancer development.

HPV leads to deactivation of tumor suppressor genes (p53 and RB1) through their early proteins E6 and E7 but alone they are not enough for the development of cancer (Narisawa-Saito and Kiyono, 2007). So it is concluded that integration sites of HPV are screened in breast cancer biopsies but their proper mechanism of carcinogenesis is still not known.

Studies show that 29 % cases of breast cancer are screened positive for HPV while some studies showed no association of HPV with breast cancer. Ethnicity and genetic background also affect the prevalence of HPV varying from 0-86%. For instance, 13 % cases are reported in Europe and 43% cases in America and Australia. Serotype 16, 18 and 33 are characterized as high risk HPV. Studies from China and Japan reported that rare subtypes of HPV (56 and 36) were also screened positive in tumor tissues (Paoli and Carbone, 2013).

In this study presence of 150bp of HPV DNA was screened through GP5+ and GP6+ primers and it has been found that HPV is significantly associated (**=0.0018) with breast carcinoma. Coinfection of viruses and their correlation with factors like lymph node involvement, tumor grade types of breast carcinoma and age group of patients is also studied in this project.

Lymph node involvement is mostly observed (*=0.035) in samples that were screened positive for EBV alone along with its coinfection with other viruses. No noteworthy results were observed for those that were screened positive for MMTV and HPV. So it is hypothesized that coinfection of EBV play significant role in cancer metastasis and invasiveness through lymph node.

It has been observed that EBV and HPV positive samples mostly fall in the category of tumor grade II. This might be explained by the fact that E6 protein of HPV leads to

inactivation of tumor suppressor gene (p53) by forming a complex with it that ultimately leads to oncogenesis (Thomas *et al.,* 2000).

IDC and DCIS are commonly observed types of breast cancer in Pakistani patients. Samples that were included in this study were 27% DCIS and 60% IDC. HPV positive samples were mostly DCIS this might be because HPV is normally associated with metastatic but noninvasive type of tumors (Jiang *et al*., 2015).

It has been stated that EBV and its co-infection play vital role in metastasis and invasiveness of breast carcinoma so that is why prevalence of EBV is high in IDC then DCIS that is a more violent form of a cancer (Romain *et al*., 2000). In this study we didn't find any significant association between EBV and DCIS due to small sample size but it has been observed that samples that were positive for EBV along with its coinfection mostly fall in the category of IDC.

Our data suggested that patients above 40 years show higher prevalence of these viruses (MMTV, HPV and EBV) than those below 40 years of age. This is because longer period of incubation will increase the viral titer in cells as it is observed in case of EBV because EBER is characterized as a gene with late expression.

In this study we didn't find any significant association of mono and co infection of these viruses with tumor grade, types of breast cancer and age because of the small sample size but further studies with large sample size may help us to understand the oncogenic properties of these viruses in detail that will ultimately lead to the development of potential biomarker for better diagnosis and also in treatment of breast cancer through improving chemotherapy.

CONCLUSION & FUTURE PROSPECTS

DNA samples extracted from tissue and blood samples of breast cancer patients were screened for the presence of EBV, HPV and MMTV LTR and ENV regions. According to our data breast cancer is significantly associated with EBV, HPV and MMTV LTR. While, there was no significant relationship of MMTV ENV with breast carcinoma. It is observed in studied samples that mean mono (39%) and co-infection (32%) of EBV , MMTV LTR and HPV play significantly important role in the development of breast cancer .

Statistical analysis of the data showed that EBV and its co-infection are significantly associated with lymph node involvement. This data suggested that Epstein bar virus along with Mouse mammary tumor virus and Human papilloma virus play a significant role in invasiveness and metastasis of breast cancer. However, MMTV and HPV alone do not showed any significant association with lymph node involvement. The results showed no significant association between the prevalence of these viruses with tumor grade, types of breast cancers and age of patients.

Further studies with large sample size may help in understanding the role of these oncogenic viruses in progression and development of breast cancer in Pakistan through identifying the important receptors and protein targets of these oncogenic viruses, which will ultimately help us in development of potential biomarkers for the early diagnosis. Further research in this area will also help us in designing new strategies for cancer treatment through improving chemotherapy. In this way this study will ultimately benefit the healthcare sector of Pakistan.

REFERENCES

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), 1.

Salmons, B., & Gunzburg,W. H. (2013). Revisiting a role for a mammary tumor retrovirus in human breast cancer. *International Journal of Cancer*, 133(7), 1530-1535.

Callahan, R., & Smith, G. H. (2008). Common integration sites for MMTV in viral induced mouse mammary tumors. *Journal of mammary gland biology and neoplasia*, 13(3), 309-321.

Musgrove, E. A. (2003). Wnt signalling via the epidermal growth factor receptor: a role in breast cancer. *Breast Cancer Research*, 6(2).

Pegram, M. D., Pietras, R., Bajamonde, A., Klein, P., & Fyfe, G. (2005). Targeted therapy: wave of the future. *Journal of clinical oncology*, 23(8), 1776-1781.

Lamb, R., Bonuccelli, G., Ozsvári, B., Peiris-Pagès, M., Fiorillo, M., Smith, D. L., & Chiu, M. (2015). Mitochondrial mass, a new metabolic biomarker for stem-like cancer cells: understanding WNT/FGF-driven anabolic signaling. *Oncotarget*, 6(31), 30453.

Theodorou, V., Kimm, M. A., Boer, M., Wessels, L., Theelen, W., Jonkers, J., & Hilkens, J. (2007). MMTV insertional mutagenesis identifies genes, gene families and pathways involved in mammary cancer. *Nature Genetics*, 39(6), 759-769.

Wang, J. M., Huang, F. C., Kuo, M. H. J., Wang, Z. F., Tseng, T. Y., Chang, L. C., & Lin, J. J. (2014). Inhibition of cancer cell migration and invasion through suppressing the

Wnt1-mediating signal pathway by G-quadruplex structure stabilizers. *Journal of Biological Chemistry*, 289(21), 14612-14623.

Qu, B., Liu, B. R., Du, Y. J., Chen, J., Cheng, Y. Q., Xu, W., & Wang, X. H. (2014). Wnt/β-catenin signaling pathway may regulate the expression of angiogenic growth factors in hepatocellular carcinoma. *Oncology letters*, 7(4), 1175-1178.

Menhas, R., & Umer, S. (2015). Breast Cancer among Pakistani Women. Iranian journal of public health, 44(4), 586-7.

Bhurgri, Y., Bhurgri, A., Hassan, S. H., Zaidi, S. H. M., Rahim, A., Sankaranarayanan, R., & Parkin, D. M. (2000). Cancer incidence in Karachi, Pakistan: first results from Karachi cancer registry. *International Journal Of Cancer*, 85(3), 325-329.

Ranzani, M., Annunziato, S., Adams, D. J., & Montini, E. (2013). Cancer gene discovery: exploiting insertional mutagenesis. *Molecular Cancer Research*, 11(10), 1141-1158.

Thorley-Lawson, D. A., & Allday, M. J. (2008). The curious case of the tumour virus: 50 years of Burkitt's lymphoma. *Nature Reviews. Microbiology*, 6(12), 913.

Stanley, M. A. (2012). Epithelial cell responses to infection with human papillomavirus. *Clinical Microbiology Reviews*, 25(2), 215-222.

Andersen, A. S., Sølling, K., Sophie, A., Ovesen, T., & Rusan, M. (2014). The interplay between HPV and host immunity in head and neck squamous cell carcinoma. *International Journal Of Cancer*, 134(12), 2755-2763.

Basra, M. A., Saher, M., Athar, M. M., & Raza, M. H. (2016). Breast Cancer in Pakistan a Critical Appraisal of the Situation Regarding Female Health and Where the Nation Stands. *Asian Pacific journal of cancer prevention: APJCP*, 17(7), 3035-3041.

Damin, A. P., Karam, R., Zettler, C. G., Caleffi, M., & Alexandre, C. O. (2004). Evidence for an association of human papillomavirus and breast carcinomas. *Breast Cancer Research and Treatment*, 84(2), 131-137.

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

De Villiers, E. M., Sandstrom, R. E., zur Hausen, H., & Buck, C. E. (2005). Presence of papillomavirus sequences in condylomatous lesions of the mamillae and in invasive carcinoma of the breast. *Breast Cancer Research*, 7(1), R1-11.

De Villiers, E. M., Schmidt, R., Delius, H., & zur Hausen, H. (2002). Heterogeneity of TT virus related sequences isolated from human tumour biopsy specimens. *Journal of Molecular Medicine*, 80(1), 44-50.

Deshpande, C. G., 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), 1193-1199.

Indik, S., Günzburg, W. H., Salmons, B., & Rouault, F. (2005a). Mouse mammary tumor virus infects human cells. Cancer research, 65(15): 6651 6659.

Indik, S., Günzburg, W. H., Salmons, B., & Rouault, F. (2005b). A novel, mouse mammary tumor virus encoded protein with Rev-like properties. *Virology*, 337(1), 1-6.

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 Biomarkers & Prevention*, 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 tumor virus as multiple viruses in breast cancer. *PloS One*, 7(11), e48788.

Jiang, R., Ekshyyan, O., Moore-Medlin, T., Rong, X., Nathan, S., Gu, X., Guidry, J. T. (2015). Association between human papilloma virus/Epstein–Barr virus co infection and oral carcinogenesis. *Journal of Oral Pathology & Medicine*, 44(1), 28-36.

Perkins, R. S., Sahm, K., Marando, C., Dickson-Witmer, D., Pahnke, G. R., Mitchell, M., & Dunn, S. P. (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), R70.

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.

Richardson, A. K., Currie, M. J., Robinson, B. A., Morrin, H., Phung, Y., Pearson, J. F., & Walker, L. C. (2015). Cytomegalovirus and Epstein - Barr virus in Breast Cancer. *PloS One*, 10(2), e0118989.

Romain, S., Spyratos, F., Descotes, F., Daver, A., Rostaing Puissant, B., Bougnoux, P., Martin, P. M. (2000). Prognostic of DNA-synthesizing enzyme activities (thymidine kinase and thymidylate synthase) in 908 T1–T2, N0–N1, M0 breast cancers: A retrospective multicenter study. *International Journal of Cancer*, 87(6), 860-868.

Saha, A., Kaul, R., Murakami, M., & Robertson, E. S. (2010). Tumor viruses and cancer biology: Modulating signaling pathways for therapeutic intervention. *CancerBiology & Biology & Therapy*, 10(10), 961-978.

Cato, A., Henderson, D., & Ponta, H. (1987). The hormone response element of the mouse mammary tumour virus DNA mediates the progestin and androgen induction of transcription in the proviral long terminal repeat region. *The EMBO journal*, 6(2), 363.

Dickson, C., & Atterwill, M. (1980). Structure and processing of the mouse mammary tumor virus glycoprotein precursor pr73env. *Journal Of Virology*, 35(2), 349-361.

Dickson, C., & Skehel, J. J. (1974). The polypeptide composition of mouse mammary tumor virus. *Virology*, 58(2), 387-395.

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.

Etkind, P. R., Stewart, A. F., Dorai, T., Purcell, D. J., & Wiernik, P. H. (2004). Clonal isolation of different strains of mouse mammary tumor virus-like DNA sequences from both the breast tumors and non-Hodgkin's lymphomas of individual patients diagnosed with both malignancies. *Clinical Cancer Research*, 10(17), 5656-5664.

Faedo, M., Ford, C. E., Mehta, R., Blazek, K., & Rawlinson, W. D. (2004). Mouse mammary tumor-like virus is associated with p53 nuclear accumulation and progesterone receptor positivity but not estrogen positivity in human female breast cancer. *Clinical Cancer Research*, 10(13), 4417-4419.

Faschinger, A., Rouault, F., Sollner, J., Lukas, A., Salmons, B., Günzburg, W. H., & Indik, S. (2008). Mouse mammary tumor virus integration site selection in human and mouse genomes. *Journal Of Virology*, 82(3), 1360-1367.

Fernandez-Cobo, M., Melana, S., Holland, J., & Pogo, B. (2006). Transcription profile of a human breast cancer cell line expressing MMTV-like sequences. *Infectious Agents and Cancer,* 1(1), 7.

Stanley, M. A. (2012). Epithelial cell responses to infection with human papillomavirus. *Clinical Microbiology Reviews*, 25(2), 215-222.

Thorley-Lawson, D. A. (2001). Epstein-Barr virus: exploiting the immune system. *Nature Reviews Immunology*, 1(1), 75-82.

Thomas, M., Pim, D., & Banks, L. (1999). The role of the E6-p53 interaction in the molecular pathogenesis of HPV. *Oncogene,* 18(53), 7690-7700.

Torabizadeh, Z., Nadji, A., Naghshvar, F., Nosrati, A., & Parsa, M. (2014). Association between Epstein - Barr virus (EBV) and Breast Cancer. *Research in Molecular Medicine*, 2(4), 24-29.

Trabelsi, A., Rammeh, S., Stita, W., Mokni, M., Mourou, A., & Korbi, S. (2008). Detection of Epstein-Barr virus in breast cancers with lymphoid stroma. *In Annales de Biologie Clinique,* 66 (1), 59-62.

Akil, N., Yasmeen, A., Kassab, A., Ghabreau, L., Darnel, A. D., & Al Moustafa, A. E.(2008). High-risk human papillomavirus infections in breast cancer in Syrian women and their association with Id-1 expression: a tissue microarray study. *British Journal of Cancer*, 99(3), 404-407.

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.

Zangen, R., Harden, S., Cohen, D., Parrella, P., & Sidransky, D. (2002). Mouse mammary tumor-like env gene as a molecular marker for breast cancer. *International Journal of Cancer*, 102(3), 304-307.

Choi, Y., Kappler, J. W., & Marrack, P. (1991). A superantigen encoded in the open reading frame of the 3′ long terminal repeat of mouse mammary tumour virus. *Nature*, 350(6315), 203-207.

Zapata-Benavides, P., Saavedra-Alonso, S., Zamora-Avila, D., Vargas-Rodarte, C., Barrera-Rodriguez, R., Salinas-Silva, J., Trejo-Avila, L. (2007). Mouse mammary tumor virus-like gene sequences in breast cancer samples of Mexican women. *Intervirology*, 50(6), 402-407.

Wu, X., Li, Y., Crise, B., & Burgess, S. M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science*, 300(5626), 1749-1751.

Yanagawa, S., Kakimi, K., Tanaka, H., Murakami, A., Nakagawa, Y., Kubo, Y., . . . Masuda, T. (1993). Mouse mammary tumor virus with rearranged long terminal repeats causes murine lymphomas. *Journal Of Virology*, 67(1), 112-118.

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., Pelisson, I., Melana, S. M., Holland, J. F., & Pogo, B. G.-T. (2001). Detection of MMTV-like LTR and LTR-env gene sequences in human breast cancer. *International Journal Of Oncology*, 18(5), 1041-1044.

Witt, A., Hartmann, B., Marton, E., Zeillinger, R., Schreiber, M., & Kubista, E. (2003). The mouse mammary tumor virus-like env gene sequence is not detectable in breast cancer tissue of Austrian patients. *Oncology reports*, 10(4), 1025-1029.

Arbach, H., Viglasky, V., Lefeu, F., Guinebretiere, J. M., Ramirez, V., Bride, N., & Mourah, S. (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.

Sadia, H., Gomez,M. K., Rahat, T. B., & Naushad, W. (2015). 96PMMTV and breast cancer: Integration sites of MMTV in breast cancer patients and their potential role in cancer progression. *Annals of Oncology*, 26(9), ix28-ix28.

Aromseree, S., Pientong, C., Swangphon, P., Chaiwongkot, A., Patarapadungkit, N.,Kleebkaow, P., & Ekalaksananan, T. (2015). Possible contributing role of Epstein-Barr virus (EBV) as a cofactor in human papillomavirus (HPV)-associated cervical carcinogenesis. *Journal of Clinical Virology*, 73, 70-76.

Badar, F., Mahmood, S., Faraz, R., Yousaf, A., ul Quader, A., Asif, H., & Yousaf, A.(2015). Epidemiology of Breast Cancer at the Shaukat Khanum Memorial Cancer Hospital and Research Center, Lahore, Pakistan. *Journal of the College of Physicians and Surgeons--Pakistan: JCPSP*, 25(10), 738-742.

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).

Majors, J., & Varmus, H. E. (1983). A small region of the mouse mammary tumor virus long terminal repeat confers glucocorticoid hormone regulation on a linked heterologous gene. *Proceedings of the National Academy of Sciences*, 80(19), 5866-5870.