UNIVERSITY OF GHANA

COLLEGE OF BASIC AND APPLIED SCIENCES

INVESTIGATING THE ROLE OF MIR-4725 AND ITS NOVEL SNP IN TRIPLE NEGATIVE BREAST CANCER

BY

EMMANUEL LANTE LAMPTEY

(10570421)

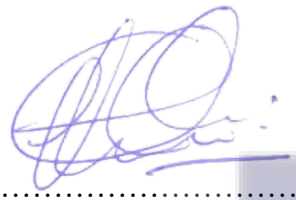
THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF MPHIL IN MOLECULAR BIOLOGY DEGREE

DEPARTMENT OF BIOCHEMISTRY, CELL AND MOLECULAR BIOLOGY

April, 2023

DECLARATION

I, Emmanuel Lante Lamptey, hereby declare that this thesis is the outcome of my own research project under the supervision of Dr Lily Paemka of the Department of Biochemistry, Cell and Molecular Biology, and Dr Kwabena Owusu Danquah of Noguchi Memorial Institute for Biomedical Research, University of Ghana. To the best of our knowledge, this thesis has not been presented for the award of any degree or published elsewhere. Any mention of other authors' work has been duly acknowledged and properly referenced.



15th November 2023

……

Emmanuel Lante Lamptey (10570421) Date

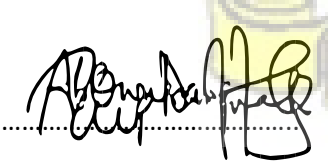
Student

20th November 2023

……

Dr Lily Paemka Date

Supervisor



22th November 2023

Dr. Kwabena Owusu Danquah Date

Co-supervisor

DEDICATION

I dedicate this work to my mother, Miss Theresa Lartey, for her immeasurable support during my education. I am especially grateful to her for her encouragement, motivation, and prayers. I also dedicate this work to all women with breast cancer and survivors of the disease.

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to my supervisor, Dr. Lily Paemka, for her guidance and mentorship throughout this research and Dr. Kwabena Owusu Danquah for the time and effort to ensure the success of this work.

I also appreciate the West African Genetic Medicine Centre (WAGMC), a World Bank ACE Impact Project at the University of Ghana, for funding through my MPhil fellowship to do this work. I would also like to thank WACCBIP and the Department of Biochemistry, Cell, and Molecular Biology for providing the laboratory space for conducting this research.

Special thanks to the members of Paemka lab for the support and intellectual input to fine-tune this work, as well as the welcoming environment they provided to enhance my learning process and my MPhil colleagues for their contributions and assistance with this work.

Finally, the biggest thanks to the Almighty God for His Grace and Favour throughout this journey.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Acute myeloblastic leukaemia	AML
Deoxyribonucleic acid	DNA
Dulbecco's modified eagle medium	DMEM
Epithelial-mesenchymal transition	EMT
Estrogen receptor	ER
Glioblastoma multiforme	GBM
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH
Human epidermal growth factor 2	HER2
Luria-Bertani	LB
Matrix metalloproteinase	MMP
MircoRNA	miRNA
Minimum free energy	MFE
Mitogen-activated protein kinase	MAPK
Nucleotides	nt
Phosphate buffered saline	PBS
Phosphoinositide-3 kinase	PI3K
Polymerase chain reaction	PCR
Primary miRNAs	pri-miRNA
Research electronic data capture	REDCap
Restriction fragment length polymorphism	RFLP
Reverse transcription quantitative PCR	RT-qPCR
Ribonucleic Acid	RNA
RNA-induced silencing complex	RISC
Single nucleotide polymorphism	SNP

Sub-Saharan Africa	SSA
Triple-negative breast cancer	TNBC
Tris-EDTA	TE
Tumour microenvironment	TME
Untranslated region	UTR
World health organization	WHO

ABSTRACT

Breast cancer is the primary cause of female cancer-related deaths across the world. One specific subtype of breast cancer, known as triple-negative breast cancer (TNBC), responsible for ~45% of breast cancer frequency in West Africa and is distinguished by the null expression of estrogen, progesterone, and human epidermal growth factor 2 receptors. It is the most aggressive subtype and has a poor prognosis, especially in women of African ancestry. TNBC is a complex and heterogeneous disease with diverse molecular subtypes and its incidence is rapidly increasing in Africa. Genetic and epigenetic factors, including microRNAs (miRs), play a significant role in TNBC pathogenesis. miRs are short non-coding RNA sequences that are involved post-transcriptional regulation and their dysregulation is a key epigenetic factor in TNBC. Mutations (such as SNPs) in miRNA genes impact their processing and expression, and have been linked to TNBC pathophysiology. A novel SNP, rs73991220 was discovered in miR-4725, and has been associated with the risk of estrogen receptor-negative breast cancer in women of African ancestry. In a preliminary study of a Ghanaian breast cancer cohort, the SNP rs73991220 was identified in two out of six TNBC tumour samples. However, the function of miR-4725 and the impact of the SNP rs73991220 on miR-4725 expression have not been established in breast cancer. This study sought to identify the role of miR-4725 in TNBC and how the SNP affects miR-4725 expression.

Twenty-four breast tumour samples and their matched normal adjacent tissues were screened for SNP rs73991220 (G) using PCR-RFLP. Two of the twenty-four (8%) tumour samples were found to have the heterozygous AG genotype and were both TNBC subtypes. The effect of the SNP on the stability of the secondary structure of the primary miR-4725 transcript was determined using the RNAfold online tool. Subsequently, the expression of miR-4725 was determined in the 2 AG and 2 AA TNBC tumour samples. It was observed that the SNP (G) conferred more stability to pri-miR-4725 compared to A; there was however, no significant

difference in expression between the AA and AG genotypes of the TNBC tumour tissues. Furthermore, the putative targets of the mature miR (miR-4725-5p) were determined using online bioinformatics tools, and the target validation was performed by overexpressing miR-4725-5p in MDA-MB-468, a basal-like 1 subtype of the TNBC cell line obtained from an African-American woman. *KIF2C* was found to be significantly repressed with exogenous expression of miR-4725 and hence predicted to be the most likely target of miR-4725-5p in the MDA-MB-468 cell line out of five validated putative targets. Additionally, other tumour-promoting markers (*WISP1, SNAIL, MYC,* and *VEGFA*) were upregulated following miR-4725-5p overexpression. These findings suggest that miR-4725-5p may have a tumour-promoting function in MDA-MB468 cells, which may be independent of *KIF2C* suppression and the SNP rs73991220 may favour miR-4725 expression thus contributing to the increased risk of TNBC progression.

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background

Breast cancer is a complex genetic disease with high incidence and mortality rates among women worldwide. It has surpassed lung cancer in cancer incidence, with an estimated 2.3 million newly diagnosed cases in 2020 (Sung *et al*., 2021). According to recent reports breast cancer incidence and mortality in the African population are increasing significantly (Adeloye *et al.*, 2018) with an estimated 32% of Ghanaian women diagnosed with the disease (GLOBOCAN, 2020), with further increase expected.

Breast cancer is heterogeneous, and manifests with several clinical, histological, and molecular characteristics; thus, the disease is classified into subtypes. Hormone receptors; estrogen -ER, progesterone -PR, and human epidermal growth factor 2 -HER2 have been used as molecular markers to subtype the disease into luminal (A and B), HER2 enriched and TNBC. TNBC is the most aggressive of these subtypes, has limited treatment options owing to the lack of expression of hormone receptors, which is commonly targeted in therapy. Approximately 15% of breast cancer cases are TNBC globally (Stanisławek, 2021). Furthermore, the TNBC subtype is more common in West Africans (53.2%) than in African Americans (29.8%) and White Americans (15.5%), implying that a genetic component may be responsible for this difference (Jiagge *et al*., 2016).

Many factors contribute to breast cancer pathogenesis; however, genetic abnormalities in protein-coding genes present a major risk to development of breast cancers. These include mutations in *BRCA* 1 and *BRCA* 2, as well as *TP53* tumour suppressor genes. However, epigenetic and noncoding components of the genome, such as microRNAs (miRNAs), have

also been implicated in the pathogenesis of breast cancer as well as other types of tumours (Qian *et al*., 2016; Ellsworth *et al*., 2019).

The microRNA (miR) network is a non-protein-coding component composed of short single-stranded RNA molecules of approximately 22 nucleotides (nt) long and are responsible for post-transcriptional regulation. miRs are typically transcribed by RNA polymerase II or III from the genome as part of introns (mirtrons) or as primary miRNAs (pri-miRNA) which are processed into matured 22nt miRNAs. miRNAs are involved in regulating many biological processes, including cell cycle regulation, proliferation, and apoptosis (Gong *et al*., 2015). They have been shown to regulate the temporal and spatial expression of genes by promoting mRNA degradation or suppressing mRNA translation (Hu *et al*., 2013). Dysregulation and/or dysfunction of miRNAs have been linked to tumorigenesis and cancer progression. Factors such as single-nucleotide polymorphisms (SNPs) within primary (pri)-miRNAs play important roles in miRNA dysregulation and dysfunction.

SNPs have been linked to various cancers, including colon, breast, and bladder cancers (Jurj *et al*, 2020). SNPs in protein-coding genes may have a significant effect on protein structure, function, and/or abundance. Similarly, SNPs within miR genes have been shown to affect miR processing, resulting in aberrant expression that has been linked to various cancers (Jurj *et al*., 2020). Changes in miR expression and processing can affect many targeted mRNAs, resulting in altered cellular respond. Therefore, miRs and non-functional SNPs should be investigated as potential biomarker for breast cancer (Iorio & Croce, 2012).

1.2 Problem statement

Breast cancer is multifaceted and complex disease and still a global health problem and a major cause of death among women. Although significant progress has been made to understand the molecular basis of breast cancer pathogenesis, and the quest to identify more potent biomarkers and targets for therapy for the disease has yielded many useful discoveries, there are still many gaps, particularly within the African population, which is the world's most genetically diverse. Moreover, as the disease evolves with the emergence of drug resistance, more potent and precise molecular targets are required to improve diagnosis and treatment, particularly in more aggressive forms such as TNBC. Currently, there is a dearth of research in Ghanaian populations regarding the molecular genetics of the development and progression of Triple-Negative Breast Cancer (TNBC). This subtype of breast cancer is particularly prevalent among women of African descent under the age of 40 and tends to progress rapidly, leading to late detection (Plasilova *et al.*, 2016; Stanisławek, 2021). and a limited range of treatment options, primarily surgery and some forms of chemotherapy (Hercules *et al.*, 2022). However, in recently, a new targeted chemotherapy drug, Sacituzumab govitecan-hziyno, has been introduced for the treatment of TNBC in many countries (So *et al.*, 2022). Nonetheless, given the heterogeneity of TNBC and the high genetic diversity including SNPs of the Ghanaian and African populations, there is a significant risk of drug resistance and reduced efficacy with this new therapy. The purpose of this study is to understand unique molecular signatures of TNBC and explore the potential of miR-4725 as a biomarker or therapeutic target for people of African descent, in order to provide more personalized diagnosis and treatment for this diverse population.

1.3 Significance of Study

miRNAs are essential components of normal cells. Their dysregulation has been implicated in breast cancer pathogenesis. The chaotic nature of tumour cells causes miRNA dysregulation. Additionally, random mutations in the miRNA gene that may be caused by this chaos have been shown to affect miRNA expression or binding to its target.

In genome-wide association studies, Qian *et al*. (2016) discovered that SNPs within the pri-miRNA, as well as genes involved in miRNA biosynthesis, are associated with breast cancer risk in women of African ancestry. The study also discovered that risk differed depending on hormone receptor status and that the novel SNP rs73991220 (A/G) in mir-4725 is associated with ER-negative breast cancer (Qian *et al*., 2016).

A preliminary study found the SNP rs73991220 in two out of six TNBC tumour samples but not in their matched adjacent non-neoplastic or five matched non-TNBC sample controls in Ghanaian breast cancer patients using Sanger sequencing (Mbirbah, unpublished). This indicated the presence of the SNP in the population; thus, further screening is required to ascertain the frequency of the SNP in the breast cancer cohort. Furthermore, the precise function of miR-4725 has not been defined as well as its role in breast cancer. This study provided data on the association of the SNP rs73991220 and miR-4725 expression and the potential role of miR-4725-5p in TNBC pathophysiology. This information provides basis for the further exploration of miR-4725 as a potential biomarker or therapeutic target for TNBC in people of with African ancestry.

1.4 Research questions

1. What is the function of miR-4725?
2. What is the relationship between SNP rs73991220 and miR-4725 expression?
3. What role does miR-4725 have in the pathophysiology of triple-negative breast cancer.

1.5 Aim and specific objectives

1.5.1 Aim

This study aimed to identify the role of miR-4725 and its SNP rs73991220, in the pathophysiology of TNBC.

1.5.2 Specific objectives

The specific objectives were;

1. To determine the effect of SNP rs73991220 and miR-4725 expression.
2. To determine putative targets of miR-4725 using *in silico* analysis and validate predicted genes using gene expression analysis.
3. To determine the effect of the miR-4725 expression on selected oncogenes.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Cancer pathobiology

Cancer is a complex genetic disease characterized by uncontrolled growth and the spread of aberrant cells in the body (Malumbres, 2020). Cancer pathobiology is a complex interaction between genetic and environmental factors that contributes to the development and progression of cancer (Patierno, 2020). Tumorigenesis is usually characterized by genetic and epigenetic alterations, as well as chronic "insults" to cells that drive neoplastic transformation (Cullen & Breen, 2016).

Accumulation of mutations in essential regulatory genes involved in cell growth and replication plays a significant role in tumour development (Costa *et al*., 2020). Tumour protein p53 (*TP53*), a tumour suppressor gene, arrests cells with damaged DNA from the G1 phase to the S phase of the cell cycle, allow reversal of the damage, and induce apoptosis when the repair of defective cells is unsuccessful (Cullen & Breen, 2016). Mutations within these types of genes could lead to loss of regulatory function, which allows cells with the compromised genetic makeup to thrive, consequently leading to neoplastic transformation of the cell lineage (Costa *et al*., 2020).

Gain-of-function mutations that result in the perpetual activation of oncogenic functions, consequently lead to cancer development (Galluzzi *et al*., 2020). An important example of such oncogenes is the *RAS* gene family, which codes for proteins involved in promoting cell growth and division (Hanrahan *et al*., 2020). Another example of genes that can undergo gain-of-function mutations is *BCL2*, which encodes an anti-apoptotic protein. (Cullen & Breen, 2016; Pylayeva-Gupta *et al*., 2011).

Epigenetic modifications that regulate gene expression also contribute to neoplastic transformation and cancer progression (Kanwal *et al*. 2015). Epigenetic changes are heritable and reversible alterations affecting gene expression even in the absence of genomic mutations (Cullen & Breen, 2016). These processes include DNA methylation, histone acetylation, and small non-coding RNA expression, particularly miRNA (Kanwal *et al*., 2015). Histone modification patterns, global alterations in DNA methylation, and altered expression profiles of chromatin-modifying enzymes as well as miRNAs, encompass the cancer epigenome (Kanwal *et al*., 2015). These changes lead to a global aberrant gene expression profile and the resultant neoplastic state (Sharma *et al*., 2009).

The immune system also plays a significant role in tumorigenesis and tumour progression; however, its involvement is complex and a "double-edged sword" (Hanahan & Coussens, 2012). Immune responses mounted against tumour cells can equitably suppress and promote tumour development and progression (Pandya *et al*., 2016). Although the immune system is responsible for the clearance of abnormal cells, the mechanisms of response contribute immensely to tumour development and progression owing to chronic stimulation of both the innate and adaptive systems by tumour antigens (Disis, 2010; Gonzalez *et al*., 2018; Gregory & Houghton, 2011). Additionally, cancer cells can circumvent the immune system by downregulating receptors that make them targets for immune destruction. Alternatively, they upregulate the expression of molecules such as *PD-L1* (Topalian *et al*., 2015) that suppress the function of immune cells such as T cells, which are responsible for the identification and destruction of these abnormal cells thereby permitting cancer cells to continue to grow and spread undetected (Gregory & Houghton, 2011; Pandya *et al*., 2016)

In addition to genetic and epigenetic factors, chromosomal translocations and exposure to certain environmental conditions such as cigarette smoke, ultraviolet radiation, and certain microorganisms (viruses and bacteria) can increase the risk of developing cancer (Anand *et al*., 2008). For instance, smoking is the primary cause of lung cancer, accounting for approximately 80-90 per cent of lung cancer fatalities (American Cancer Society, 2021).

2.2 Cancer hallmarks

Cancer hallmarks, as established by Hanahan and Weinberg in 2000, are characteristics exhibited by neoplastic cells during their evolution from normal cells. These hallmarks of cancer represent the genetic and cellular changes that contribute to carcinogenesis and tumour progression (Hanahan & Weinberg, 2000, 2011). To unravel the complexities of cancer development and malignant progression, Hanahan and Weinberg initially proposed six major cellular phenotypes: evading growth suppressors, sustaining proliferative signalling, enabling replicative immortality, resisting cell death, inducing angiogenesis, and activating invasion and metastasis (Hanahan, 2022; Hanahan & Weinberg, 2011). However, with increasing data on the mechanisms of tumorigenesis and progression, a myriad of molecular and cellular phenotypes displayed by transforming cells has been described (Galluzzi *et al*., 2020). These characteristics are based on the fundamental knowledge that cancer development and malignant progression are multistep processes involving many contributing factors (Carmen, 2017; Hanahan, 2022; Hanahan & Weinberg, 2011). These factors include epigenetic alterations, chromosomal aberrations, oncogenic microorganisms, as well as cellular activity of the stroma within the tumour microenvironment (TME) (Figure 2.1) (Fouad & Carmen, 2017; Hanahan, 2022; Hanahan & Weinberg, 2011).

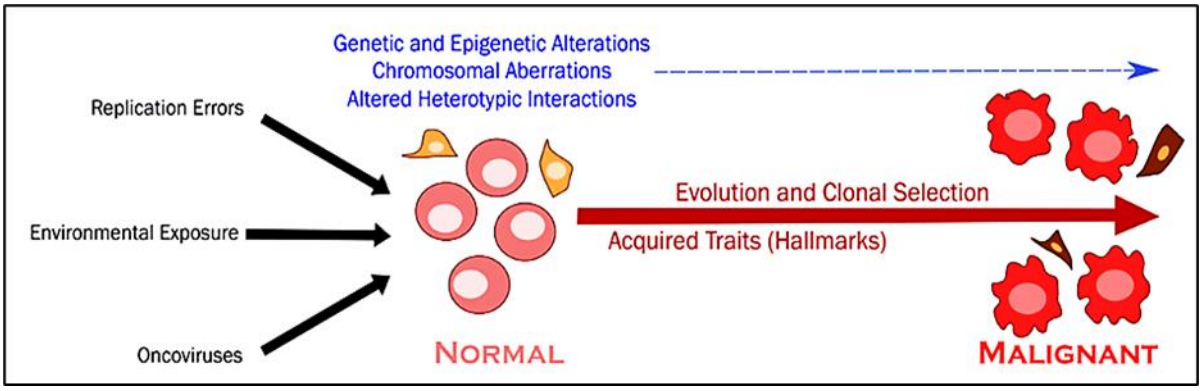


Figure 2.1 Transformation of normal cells. A variety of chronic insults leads to the transformation of these cells from normal to malignant (Fouad & Carmen, 2017).

Cancer cells defy most biological systems employed to regulate cell division and growth through a variety of mechanisms to maintain proliferative signalling (Cheng *et al*., 2008; Bhowmick *et al*., 2004). Cancer cells utilize two classic mechanisms to evade cell proliferation checkpoints: they generate mitotic signals and induce normal cells within the tumour surrounding stroma to release growth factors (Hanrahan *et al*., 2020). In response to these signals, cancer cells express cognate ligands that permit their interminable proliferation. Oncogenes, such as those in the *RAS* signalling pathway, are involved in the sustained proliferation of cancer cells (Hanrahan *et al*., 2020). Proteins encoded by these genes promote cell growth through the mitogen-activated protein kinase (MAPK) and phosphoinositide-3 kinase (PI3K)/Akt signalling pathways (Pylayeva-Gupta *et al*., 2011). Mutations in *RAS* family genes, such as KRAS, NRAS, as well as in *PI3K*, are among the most prevalent genetic abnormalities reported in human malignancies, including breast, lung, and endometrial cancers (Hanrahan *et al*., 2020). These mutations continuously activate these proteins, resulting in uncontrolled cell growth and division (Pylayeva-Gupta *et al*., 2011).

In addition, cancer cells can evade the typical processes of growth regulation and apoptosis (Hanrahan *et al*., 2020). This process of evasion is achieved through the inactivation of tumour suppressor genes, such as *TP53*. *TP53* encodes the tumour protein p53, which is a DNA damage sensor (Galluzzi *et al*., 2020) which triggers apoptosis in cells with significant

DNA damage (Junttila and Evan, 2009). Mutations in the *TP53* gene, which lead to loss of function and uncontrolled cell proliferation and division, have been found in a wide range of human malignancies (Galluzzi *et al*., 2020). Alternatively, cancer cells can circumvent apoptosis by promoting anti-apoptotic genes including *BCL-2* and suppressing caspases activity making them resistant to apoptotic signals and thus promoting their survival (Adams & Cory, 2007; Fouad & Carmen, 2017; Hanahan & Weinberg, 2011).

Another distinct characteristic of cancer cells is their ability to indefinitely proliferate, a condition known as replicative immortality. Telomerase, an enzyme that is normally dormant in normal cells, is commonly reactivated in cancer cells to maintain the length of telomeres, which are protective endcaps of chromosomes; thus, telomeric shortening does not occur. Eventually, cells do not become senescent or die but are rather immortalized (Blasco, 2005; Hanahan & Weinberg, 2011).

Cancer cells require a blood supply for development and propagation (Rankin & Giaccia, 2020). Angiogenesis, a term that describes the formation of new blood vessels. This process is normally "switched off" after development but is frequently enhanced in cancers (Madu *et al*., 2020). Vascular endothelial growth factor A (*VEGFA*), an angiogenic growth factors, facilitates this process and are secreted by neoplastic cells (Rankin & Giaccia, 2020). VEGF stimulates the formation of new blood vessels. High levels of VEGF have been reported in many human malignancies, including breast, lung, and colorectal cancer (Ferrara, 2010).

Cancer cells' ability to invade surrounding tissues and colonize distant sites to form secondary growths is a prominent feature of malignancy known as metastasis (Jain *et al*., 2020). This is a complicated process that results in epithelial cells losing their cell-to-

cell/cell-to-extracellular matrix adhesion capability and polarity (Das *et al*., 2019). Epithelial cells transform into mesenchymal cells in a process known as epithelial-mesenchymal transition (EMT) (Das *et al*., 2019). Genes encoding matrix metalloproteinases (*MMPs*) encode a family of enzymes that contribute to cancer cell invasion (Kapoor *et al*., 2016). Some MMPs (such as *MMP 2* and *MMP 9*) are known to be upregulated in cancers, including breast carcinomas (Kapoor *et al*., 2016). They break down the basement membrane and extracellular matrix, which allows cancer cells to invade surrounding tissue (Brown & Murray, 2015). Other genes, such as *TWIST* and *SNAIL*, participate in the EMT process, allowing cancer cells to migrate and metastasize to distal organs (Benson *et al*., 2013; Fouad & Carmen, 2017; Hanahan & Weinberg, 2011; Micalizzi *et al*., 2010).

Further understanding of the mechanisms of neoplastic changes and malignant progression has led to the development of eight proposed hallmarks and two enabling processes (Hanahan, 2022). The two processes that enable neoplastic changes include genome instability and tumour-promoting inflammation (Hanahan & Weinberg, 2016). A new review by Hanahan proposed four prospective characteristics that provide new insights into neoplastic development and malignant progression including unlocking phenotypic plasticity, non-mutational epigenetic reprogramming, polymorphic microbiomes, and senescent cells (Figure 2.2) (Hanahan, 2022).

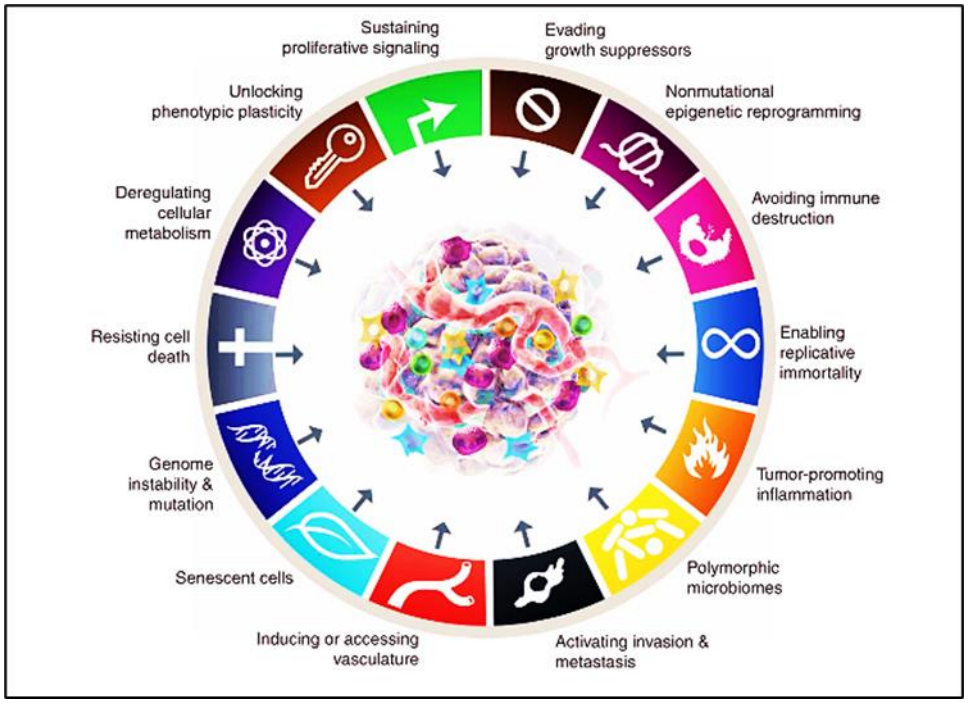


Figure 2.2 Prospective Cancer Hallmarks. The ten core characteristics include the four newly proposed features of tumorigenesis. (Hanahan, 2022).

2.3 Cancer epidemiology

Cancer incidence and mortality are rapidly increasing worldwide, and as a result, the global impact of the disease is progressively worsening (Mattiuzzi & Lippi, 2019). The increase in cancer incidence and mortality can be attributed to alterations in the prevalence and distribution of major cancer risk factors that are characteristic of varying degrees of socioeconomic development (Mattiuzzi & Lippi, 2019).

Cancer is rapidly becoming the predominant factor causing increased mortality in several continental sub regions, including sub-Saharan Africa (SSA) (Bray *et al*., 2018; Ngwa *et al*., 2022). The actual burden of cancer, particularly breast and prostate cancers, in SSA, is significantly underestimated despite an increase in cancer mortality (Anyigba *et al*., 2021; Sung *et al*., 2021).

2.4 Breast cancer in sub-Saharan Africa

The global incidence of breast cancer is 11.7%, whereas that of lung cancer is 11.4%, which makes breast cancer the commonest cancer. It is rank fifth with regards to cancer mortality worldwide, according to the GLOBOCAN 2020 report. In women globally, it is, regrettably, the foremost cause of cancer-related death, affecting approximately 2.3 million women and resulting in approximately 685,000 deaths in 2020 (GLOBOCAN, 2020)

Although breast cancer incidence in SSA appears to be lower than in first-world countries, the mortality rate is disproportionately high. (Ngwa *et al*., 2022). This disparity has been attributed to the absence of infrastructure that permits early detection and effective treatment; additionally, the actual disease burden is statistically underrepresented in the SSA region as reviewed by Anyigba *et al*. (2021).

Breast cancer accounts for 32% of newly reported cancer cases in the Ghanaian female population (GLOBOCAN, 2021). Regrettably, breast cancer in Ghanaian women is often diagnosed late after the onset (Martei *et al*., 2018), resulting in high frequency of advanced stage cancers and, therefore, significantly reduces the five-year survival rate to less than 40% compared to 86% in developed countries (Akuoko *et al*., 2022; Anyigba *et al*., 2021; Sighoko *et al*., 2013). The distinguishing factors that create this disproportion include late detection of breast cancer symptoms, socio-cultural and economic factors as well as absence of screening programs (Asoogo & Duma, 2015; Bonsu & Ncama, 2019; Martei *et al*., 2018). To improve breast cancer survival rates in SSA, efforts have been made to sensitize the population and provide more efficient healthcare systems to combat issues of poor detection rates, late detection, and substandard treatment of the disease (Kohler *et al*., 2017).

2.5 Breast cancer pathogenesis

The pathogenesis of breast cancer does not significantly differ from that of other cancers. Similar to most cancers, this disease is complicated and has many contributing factors that ultimately result in the accumulation of genetic mutations and cellular alterations at various phases of development and progression (Low *et al*., 2018; Vodicka *et al*., 2021). The pathophysiology of breast cancer can be broadly classified into three stages: initiation (tumorigenesis), promotion, and progression (Basu, 2018).

Tumorigenesis results from the accumulation of genetic alterations in healthy breasts (Nielsen *et al*., 2016). The most commonly mutated genes, also classified as high-penetrance genes in breast cancer, include the tumour suppressor genes *BRCA1* and *BRCA2* (Ali *et al*., 2021). *BRCA1* and *BRCA2* encode regulatory molecules that ensure DNA damage repair; therefore, alterations in these genes lead to inadequate and inefficient repair, resulting in the accumulation of mutations and high susceptibility to breast cancer (Pilger *et al*., 2021; Sadeghi *et al*., 2020). Other genes commonly mutated in breast cancer include *p53*, *PTEN*, and *APC* (Sheikh *et al*., 2015).

The accumulation of essential genetic mutations and epigenetic alterations in transforming cells characterizes the promotion stage (Feng *et al*, 2018). These modifications result in the silencing of tumour suppressor genes such as *TP53* while activating of oncogenes such as RAS and *HER2* (Feng *et al*, 2018). Consequently, cancer cells acquire a new quality that confers immortality and their proliferation proceeds unregulated (Feng *et al*., 2018)

The progression stage is characterized by the development of a malignant neoplasm and the dissemination of cancer cells to distant sites in the body (metastasis) away from the original

start site (Al-Mahmood *et al*., 2018). The process of metastasis is complicated and requires the activation of numerous signalling pathways such as the PI3K/Akt and MAPK pathways (Wee & Wang, 2017; Xu *et al*., 2015).

Furthermore, the epithelial-mesenchymal transition (EMT) program plays a pivotal role in the development of breast cancer malignancy (Bill & Christofori, 2015). The EMT program is a distinct set of genetic and cellular alterations that enables cancer cells to separate from the initial tumour and migrate to distant sites in the body (Bill & Christofori, 2015; Xu *et al*., 2015).

Accurate and precise diagnosis of breast cancer requires comprehensive knowledge of its pathogenesis at the cellular and molecular levels (Rakha & Green, 2017). Additionally, targeted therapies developed based on an understanding of the molecular pathways and cellular characteristics of breast cancer have resulted in enhanced patient prognosis (Tang *et al*., 2016; Yin *et al*., 2020).

2.6 Breast cancer classification and molecular subtypes

The categorization of breast cancer into various clinical classifications and molecular subtypes is a significant factor in determining the treatment and disease management options (Tang *et al*., 2016). Breast cancer can be classified into two principal categories: invasive and non-invasive (*in situ*) carcinomas (Araújo *et al*., 2017). Invasive carcinomas can infiltrate the surrounding tissue and possibly more remote sites from the tumour origin, whereas non-invasive carcinomas exclusively remain in the ducts or lobules of the breast (Araújo *et al*., 2017).

Recently, however, it has become crucial to classify breast cancer according to the various molecular features of the tumour to direct the development of more effective treatment options (Garrido-Castro *et al*., 2019). Four main molecular subtypes of breast cancer are distinguished based on immunohistochemistry-based categorizations: luminal A, luminal B, HER2-enriched, and triple-negative (Tsang & Tse, 2020).

2.6.1 Luminal A

Luminal A subtype of breast cancer is characterized by the expression of estrogen receptors and progesterone receptors and the absence of HER2 (Tsang & Tse, 2020). Approximately 65% of all breast cancers fall into this category (Zhang *et al*., 2022). They are low-grade tumours, develop slowly and tend to respond better to hormone therapy, thus, have a better prognosis (Stanisławek, 2021).

2.6.2 Luminal B

The Luminal B subtype has ER and PR expression and a relatively high expression of HER2 expression (Tsang & Tse, 2020). Approximately 20% of breast cancers are classified as Luminal B (Feng, *et al*., 2018). These tumours have high expression of proliferation-related genes (such as MKi67); are less responsive to hormone therapy and higher risk of recurrence (Rakha & Green, 2017; Stanisławek, 2021).

2.6.3 HER2-enriched

HER2-enriched subtype represents tumours that lack *ER* and *PR* but the *HER2* gene is amplified together with other proliferation gene clusters such as *GRB7* (Feng *et al*., 2018; Stanisławek, 2021). This subtype constitutes 10 to 15% of breast cancer cases (Stanisławek, 2021). HER2-positive breast cancer is more aggressive and has a higher risk of recurrence

(Tsang & Tse, 2020). This subtype is treated with targeted therapies such as Herceptin or trastuzumab, a monoclonal antibody that binds to the HER2 protein (Feng *et al*., 2018).

2.6.4 TNBC

TNBC is a diverse group of breast cancers that are ER-negative, PR-negative, and HER2-negative. TNBC accounts for around 20% of all breast cancers and is more prevalent in women under the age of 40 and in African-American women (Plasilova *et al*., 2016; Stanisławek, 2021). These tumours are more aggressive and highly metastatic with the highest risk of recurrence. Metastasis in TNBC is complicated and poorly understood (Al-Mahmood *et al*., 2018; Rankin & Giaccia, 2020); however, various factors that have been identified as key contributors to the process of metastasis include angiogenesis, tumour-stroma associations, survival in circulation, intravasation into circulation via basement membranes, extravasation into foreign tissues, and genetic and epigenetic modifications (Al-Mahmood *et al*., 2018; Rankin & Giaccia, 2020). Treatment options for triple-negative breast cancer are limited to chemotherapy and surgical excisions. This is because, TNBCs cannot be managed with endocrine therapy or treatments designed to target HER2 (Al-Mahmood *et al*., 2018; Dai *et al*., 2015; Lehmann *et al*., 2016). The triple negative characteristic with absence of hormone receptors accounts for the poor sensitivity to standard hormone treatment options (Dai *et al*., 2015).

TNBCs are heterogenous and may be classified into six subtypes based on gene expression profiling: basal-like (BL1 and BL2), mesenchymal (M), mesenchymal stem-like (MSL), immunomodulatory (IM), and luminal androgen receptor (LAR), as well as an unidentified group (UNS)] However, the clinical significance of the subtyping is yet unknown, and

additional study is needed to determine its influence on TNBC therapy options (Stanisławek, 2021).

TNBC accounts for approximately 80% of breast cancers caused by a BRCA1 germline mutation, whereas 11-16% of all TNBCs have *BRCA1* or *BRCA2* germline mutations. This indicates a major role of genetic factors playing a significant role in TNBC development. However, epigenetic factors including miRNAs initiate the sub clonal evolution of TNBC and shape its progression (Zolota *et al*., 2021).

2.7 MicroRNA

MiRNAs (miRs) are small non-coding RNA sequences of 19 to 24 nucleotides long and play a significant role in post-transcriptional regulation (Bail *et al*., 2010). miRs are found in the genome, mainly within introns or intergenic domains, and are transcribed as primary miRNAs (pri-miRs) by RNA polymerase II and III, respectively (Bail *et al*., 2010). The pri-miR produced after initial transcription is a lengthier sequence of approximately 200 nucleotides, which is then processed into a mature miRNA (Yates *et al*., 2013).

Regulation of gene expression is achieved by miRs through the ability to interact with mRNA and inhibit translation or facilitate the degradation of mRNA transcripts (Correia de Sousa, M *et al*., 2019). In most cases, miRNAs interact with the 3'-UTR (untranslated region) of target mRNAs to suppress their expression (Correia de Sousa *et al*., 2019). However, the interaction of miRNAs with other regions, including the 5′-UTR, coding sequences, and gene promoters, has also been reported (Correia de Sousa *et al*., 2019; O'Brien *et al*., 2018). Developmental processes in living organisms are critically regulated by miRNA action, which also affects health or fosters disease development (Gebert *et al*., 2019).

2.8. MiRNA biosynthesis

The biosynthesis of miRNAs is an intricate process that involves the transcription of miRNA genes, subsequent processing of the primary miRNA transcript (pri-miRNA), and maturation of the resulting miRNA (Gebert *et al*., 2019). MiRNA biosynthesis begins with the transcription of miRNA genes by RNA polymerase II (Yates *et al*., 2013). MiRNA genes are typically situated in intergenic regions or within introns of protein-coding genes (Yates *et al*., 2013). The initial transcript produced by RNA polymerase II is the primary miRNA transcript (pri-miRNA) (Yates *et al*., 2013).

Following the production of pri-miRNA, the subsequent phase of miRNA biosynthesis involves the processing of the initial transcript into a mature transcript (Yates *et al*., 2013). This function is executed by the microprocessor complex, which comprises endonuclease Drosha and its cofactor DGCR8 (Ha & Kim, 2014; Kian *et al*., 2018). The microprocessor complex cleaves pri-miRNAs to produce a stem-loop structure known as the precursor miRNA (pre-miRNA) (Ha & Kim, 2014; Kian *et al*., 2018).

Subsequently, the exportin-5 protein aids in relocating the modified pre-miRNA from the nucleus to the cytoplasm. Once in the cytoplasm, pre-miRNA undergoes additional processing by the endonuclease Dicer and its cofactor TRBP. This step results in the production of mature miRNAs, which are short single-stranded RNA molecules.

After processing, the mature miRNA adheres to the RNA-induced silencing complex (RISC) and guides the RISC to effectively locate a targeted mRNA sequence through complementary base-pairing (Ha & Kim, 2014; Kian *et al*., 2018; Yates *et al*., 2013). Silencing of the target gene is then accomplished by miRNA, either by facilitating mRNA cleavage or by interfering

with translation pathways (Gebert *et al*., 2019). A schematic of the process is shown in figure 2.3.

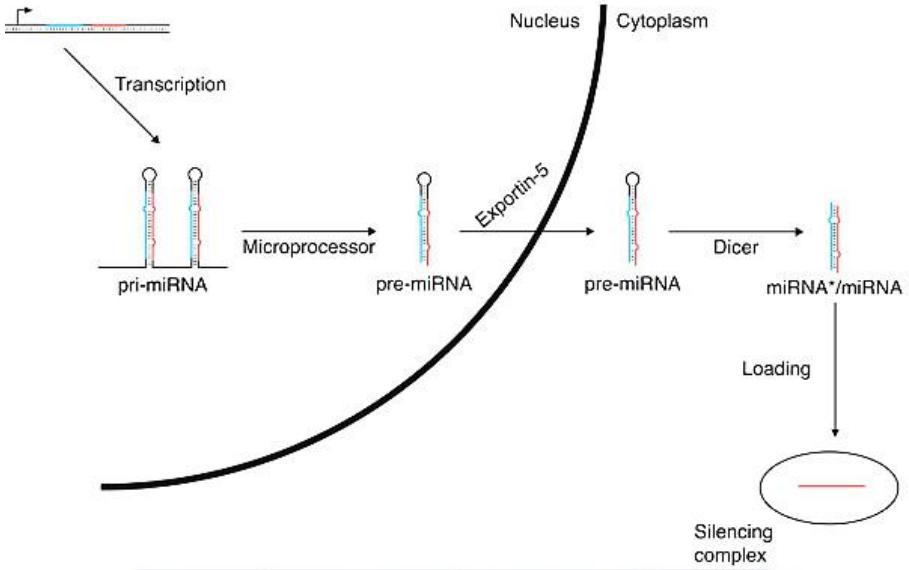


Figure 2.3. The Canonical miRNA biogenesis pathway. Adapted from Blelloch (2009).

Several factors can influence the biogenesis process, resulting in abnormal miRNA regulatory function. These include the effect of other aberrantly expressed transcription factors of miRNA in an abnormal cellular state, as well as mutations in the miRNA genes or genes involved in their synthesis (Qian *et al*., 2016). miRNA dysregulation/function can promote tumorigenesis or contribute to tumour progression.

2.9 Mutations in miRNA

Mutations in miRNAs can significantly affect their stability, secondary structure, and expression (Ha & Kim, 2014). Mutations in or in the vicinity of a pre-miRNA hairpin structure may have a greater impact on miRNA biosynthesis and the function than mutations in other regions pri-miRNA or precursor (Rice *et al*, 2020). Additionally, mutations found in the seed region of miRNAs can affect their ability to bind target mRNAs. For example, a mutation in the seed region of miR-96-5p was shown to disrupt the base pairing between

miR-96-5p and its target mRNA; as a result, repression of the target gene significantly decreased (Olavarrieta *et al*., 2009).

According to one study, the SNP rs2910164 in miR-146a was associated with a higher risk of papillary thyroid carcinoma and reduced levels of pre-miR-146a and mature miR-146a-5p by nearly twofold, resulting in a less efficient downregulation of miR-146a-5p targets (Jazdzewski *et al*, 2008). Mutation in the loop region of the miR-30e-3p precursor decreased its processing by the Drosha-DGCR8 complex, consequently, levels of mature miR-30e-3p were reduced in the cells (Wang *et al*., 2019).

Several factors influence the effect of a mutation on the structure of pri or pre-miRNA, such as the type of mutation, its localization, and the thermodynamic stability of the precursor. Mutations have been shown to affect miRNA expression by altering its stability. In a study, a mutation in the stem region of miR-128 altered the thermodynamic stability of its secondary structure, resulting in decreased expression of mature miR-128 (Cho, 2010; Machowska *et al*., 2022).

2.10 MiRNA secondary structure stability effect on expression

The integrity of a miRNA's secondary structure can affect its expression and processing (Linbo *et al*., 2011). miRNAs with more stable secondary structures have higher levels of expression in cells which have attributed to the fact that stable miRNAs are ultimately more resistant to degradation (Wu *et al*. 2008).

Other publications have indicated that alterations to a miRNA's secondary structure stability can affect the processing of the miRNA by ribonucleases that cleave it from its precursor

(Gebert *et al*., 2019; Ha & Kim, 2014). This implies that miRNAs with more stable secondary structures are less efficiently processed, suggesting that the stability of the secondary structure can affect the accessibility of the miRNA to ribonucleases (Ha & Kim, 2014).

It is important to acknowledge that the influence of miRNA stability on its processing and subsequent expression is not fully understood, as multiple factors are involved in this relationship. These factors include cell type, miRNA family, and the specific conditions of the experiment, which can result in varying observations when studying miRNA (Wu *et al*., 2008).

2.11 The role of miRNAs in TNBC

Recent research has shown that miRNAs play a crucial role in the development and progression of breast cancer, thus, TNBC (Kian *et al*., 2018; Shen *et al*., 2009; Wu *et al*., 2008). miRNAs regulate the expression of regulatory genes that play tumour suppressor or oncogenic roles and shape the molecular landscape of TNBC. miRNAs that are dysregulated in TNBC have been identified and they include miR-10b, miR-21, miR-155, and miR-221/222. These miRNAs are involved in various aspects of TNBC progression, including cell proliferation, invasion and metastasis. (Santana *et al*., 2023; Yang *et al*., 2015).

In one study, miR-10b was shown to promote metastasis in TNBC by targeting the *HOXD10* gene, which suppresses the expression of the pro-metastatic gene *RhoC* (Ma *et al*., 2007, Yang *et al*., 2015). Likewise, miR-221/222 promotes TNBC cell growth and invasion by repressing the tumour suppressor gene *PTEN* (Li *et al*., 2020; Zhao *et al*., 2008). Another study showed that miR-155 regulates the expression of various proinflammatory genes in

TNBC cells, suggesting that it may contribute to the development of TNBC by promoting inflammation (Tili *et al*., 2011).

In addition to their roles in TNBC pathophysiology, miRNAs have emerged as potential diagnostic and prognostic biomarkers for TNBC. The expression signature of a particular miRNA or a combination of miRNAs can be identified as a prognostic determinant or biomarker in TNBC. In a study by De Rinaldis *et al*. (2013), seven miRNAs were found to be associated with TNBC clinical prognosis. Moreover, upregulation of miR-16-2* and miR-766 have been associated with favourable distant metastases-free survival (DMFS), while miR-381 and miR-409-5p have been associated with poor DMFS. miR-409-5p, miR-376b, miR-410, and miR-193a-3p together have also been negatively associated with breast cancer-specific survival (De Rinaldis *et al*., 2013; Yang *et al*., 2015).

Overall, these findings suggest that dysregulation of miRNAs plays a crucial role in TNBC pathophysiology and that targeting these miRNAs may represent a promising therapeutic strategy for TNBC treatment.

2.12 MiR-4725 in human disease

There is a paucity of data on the miR-4725 and it function and the role it plays in cancer, specifically in breast cancer has not been elucidated till date. In one study, miR-4725-3p was found to be significantly upregulated in response to xanthohumol in glioblastoma multiforme (GBM). MiR-4725-3p in this study was found to inhibited cell invasion in GBM by downregulation stromal interacting molecule 1 (STIM1) which promoted glioma cell invasion (Ho *et al*., 2018). In another study, miR-4725-3p was identified as a core regulatory miRNA in RNA networks involving ncRNAs, miRNAs and mRNAs when esophageal

squamous cell carcinomas (ESCC) were treated with neoadjuvant chemoradiotherapy (Shao & Li, 2019).

Furthermore, in study to identify an extracellular biomarker for glaucoma, miR-4725-3p has also been found to be upregulated in plasma and aqueous humour samples of glaucoma patients (Hindle *et al.*, 2019). The miR-4725-5p isomir has also been found to be downregulated in patients with acute myeloblastic leukaemia (AML) compared to healthy controls in a study using miRNA microarray followed by RT-qPCR validation (Ozdogan *et al.*, 2017).

Clearly, miR-4725 is involved in different diseases and have varied expression pattern in different tissue and disease condition, however, there is an unmet understanding the exact role the gene plays in diseases.

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Study design

A cross-sectional study was used to determine the frequency of the SNP rs73991220 and miR-4725 expression pattern using archival breast cancer tissue samples collected from treatment naïve patients during a mastectomy procedure at the Korle-Bu Teaching Hospital. An experimental design was also used to determine the role of the miR-4725 using the triple-negative breast cancer cell line MDA-MB468.

3.2 Sample selection.

Thirty matched (tumour and adjacent-normal) breast tissue samples were selected randomly from the archival breast tissue samples stored at -80°C. Immunohistochemistry (IHC) data was later retrieved for the patient data of the selected samples from the REDCap (Research electronic data capture) database to classify tissue samples by hormone receptor status. Six of the matched pairs had no IHC data and hence were excluded from the study. The remaining 24 matched breast tissue samples were grouped into TNBC and non-TNBC.

3.3 Ethical clearance

Samples used in this study were a part of collected breast tissue samples for a previous breast cancer study for which ethical approval had been sort from the Korle-Bu Teaching Hospital Protocol with IRB number KBTH-IRB /0009/2018. Additionally, the study participants recruited for the study signed a written informed consent before the tissue samples were collected.

3.4 DNA extraction

The Quick-DNA™ Miniprep plus Kit (Zymo Research) was used to extract genomic DNA from tissue samples and cell lines, following the manufacturer's protocol. Tissue samples were removed from the -80 °C freezer and allowed to thaw on ice before cutting. Approximately 25 mg of breast tissue was cut using a sterile surgical blade and pair of forceps and transferred into a sterile 2 ml Eppendorf tube. A mixture containing 95µl of solid tissue buffer, 95µl of molecular grade water and 10µl of proteinase K was added to the cut tissue samples, vortexed for 15 s, and incubated at 55 °C overnight. The solubilized tissue was centrifuged at 14,000 × g for 1 min to separate debris, and the supernatant was transferred to a 1.5 ml Eppendorf tube.

Genomic binding buffer was added to the supernatant and vortexed for 15 s, and then the whole mixture was transferred into a Zymo-Spin™ IIC-XLR Column in a collection tube and centrifuged at 14,000 × g for 1 min. The flow-through in the collection tube was discarded and the column was subjected to three successive washing with 400µl of pre-wash buffer, 700µl and 400µl of g-DNA buffer. DNA was then eluted in a 60µl elution buffer, and the purity and concentration of DNA were determined using a Nanodrop spectrophotometer (Thermo Fisher Scientific).

3.5 MiR-4725 genotyping.

MiR-4725 was genotyped for the SNP rs73991220 in breast tissue and cell lines (MDA-MB468 and MDA-MB231) using restriction digest following amplification of the gene by Polymerase Chain Reaction (PCR). The SNP was found to be a restriction site for NlaIII (Hin1II). Hence, specific primers were designed to amplify 280 bp, including flanking regions of the gene.

Table 3.1 **Primer sequence for miR-4725 amplification**

	Sequence	**Amplicon size**
Forward	5'-*GGCTTTAGGTGGAACCCAGG*-3'	280bp
Reverse	5'-*ACAGCTGCCCCTGAAAGTC*-3'	

3.5.1 Gene amplification (polymerase chain reaction-PCR)

The final volume of 20µl PCR mix contained 2µl of gDNA, 1µl each of forward and reverse primers, 10µl of Q5 High Fidelity HotStart Master Mix 2X (New England Biolabs) and 6µl of nuclease-free water. The reaction was done using the conditions in Table 3.2.

Table 3.2 **PCR conditions**

Step	**Temperature (°C)**	**Time (seconds)**	**Cycles**
Initial denaturation	98	30	1
Denaturation	98	10	30
Annealing	58	15	30
Extension	72	15	30
Final extension	72	180	1

3.5.2 Restriction digest

PCR amplicons were then digested with NlaIII purchased from Thermo Fisher Scientific in a final reaction mix containing 5µl of amplicons, 1µl of 10X buffer G, 0.5 µl of NlaIII restriction and 8.5µl of nuclease-free water to a final reaction volume of 15 µl. Digestion was performed at 37 °C for 2 h, followed by 20 min of enzyme inactivation at 65 °C. The digested products were resolved by electrophoresis on a 2% agarose gel for 90 min, and the bands were visualized using a gel imager. The homozygous wild type shows a single band with

lower molecular weight (~140bp). However, the SNP allele is resistant to digestion and hence a homozygous mutant (GG) has band size same as the PCR product (280bp). A heterozygous genotype (AG) showed a double band at 140bp and 280bp whereas a homozygous (AA) wild-type showed a single digested band at 140bp marker (figure 3.1).

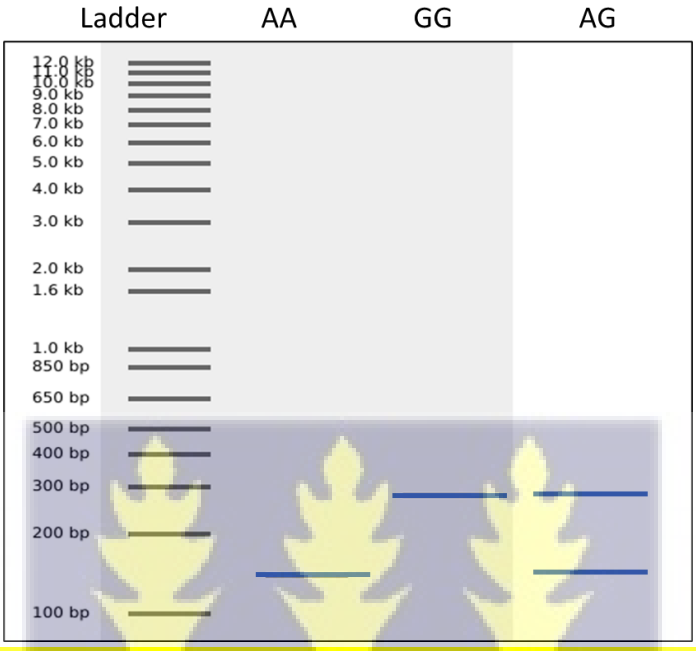


Figure 3.1 Virtual digest showing of PCR amplicon of miR-4725: AA is homozygous wild-type (band size =140bp), GG is homozygous mutant (band size = 280bp) and AG heterozygous genotypes (band sizes =140bp and 280bp).

3.6 RNA extraction

3.6.1 Breast tissue samples

Breast tissue samples were stored in TRIzol reagent, and Direct-zol™ RNA Miniprep Plus (Zymo Research) was used to extract total RNA from the samples. Briefly, tissue samples were thawed on ice, and approximately 25 mg of tissue was cut with a sterile surgical blade and placed into 2 ml Eppendorf tubes. Eight-hundred microlitres of Tri-zol was added to the cut tissue sample, and bashing beads were added and homogenized for 15 min. The samples were centrifuged after homogenization at 14,000 × g for 1 min, and the supernatant was

carefully transferred to a 1.5 ml Eppendorf tube. Eight-hundred microlitres of absolute ethanol were added to the supernatant and thoroughly mixed by pipetting up and down.

The mixture was then transferred to a Zymo-Spin™ IIICG Column in a collection tube and centrifuged at 14,000 × g for 30 s. The flow-through was discarded. The column was washed and underwent a DNA digestion step with DNase1 enzyme for 15 minutes. The column was washed twice with 400µl of Direct-zol™ RNA PreWash buffer and 700µl of RNA wash buffer respectively. Subsequently, RNA was eluted in 60µl of 1X Tris EDTA (TE) buffer (pH=8.0) and stored at -80℃ until use.

3.7 Plasmid expansion

3.7.1 Bacteria transformation

Human microRNA (miR4725 (MI0017362) expression Plasmid under control of a *CMV* promotor was purchased from Origene Technologies (USA). Plasmid came in a 2-D bar-coded matrix tube as dried plasmid DNA. A volume of 100µl of nuclease-free water was used to reconstitute the plasmid DNA to a final concentration of 100ng/µl of plasmid DNA. The pCMV-Mir backbone of the plasmid contains a kanamycin resistance marker for selection in *Escherichia coli (E. coli).*

Chemically competent *DH5a E. coli* strain, a high-efficiency transformation strain, was transformed with the plasmid using the heat-shock method optimized in the lab. Briefly, the chemically competent DH5a was thawed on ice and 50µl was pipetted in a pre-chilled round bottom 2ml Eppendorf tube on ice. One tube for the miR-4725 plasmid and another for a no-plasmid control. Five microlitres of a 10ng/µl dilution of miR-4725 plasmid was added to tube labelled miR-4725 and 5µl of nuclease-free water was added to the no plasmid control tube. The tubes were mixed gently by flicking the bottom and incubated on ice for 30

minutes. The tubes were transferred into a water bath at 42℃ for 45 seconds and then placed back on the ice for about 5 minutes during which 500µl of Luria-Bertani (LB) broth was added to the transformation mix and incubated at 37℃ for 1 hour with shaking at 180rpm.

After incubation, 250µl of each culture was plated on an appropriate LB agar plate containing kanamycin for selection. The non transformed bacterial culture was plated on a kanamycin plate (negative control) and on a plate free of antibiotics as a positive control (bacteria viability control). The plates were incubated at 37℃ overnight.

Six colonies from the transformed bacteria culture on kanamycin plate were selected and transferred into 5 ml of LB broth containing kanamycin in six separate 15 ml falcon tubes and incubated overnight at 37℃ with shaking at 180rpm. Tube caps were left slightly unscrewed to allow for aeration.

3.7.2 Plasmid purification

Plasmid was extracted from bacterial cells using PureYield™ Plasmid Miniprep System (Promega). The 5ml broth culture was centrifuged at 14,000g for 5 minutes to pellet the bacterial cells. The supernatant was poured off and the cells were resuspended in 600µl of Tris EDTA (TE) buffer (pH=8.0) and transferred into a 2ml round bottom Eppendorf tube. Hundred microlitres of cell lysis buffer was added and thoroughly mixed. Subsequently, 350µl of neutralization solution (at 4℃) was added and mixed by pipetting until the mixture was completely neutralized. Cell debris was separated from the solution by centrifugation at 14,000g for 3 minutes. The supernatant was carefully transferred into a PureYield™ minicolumn on a collection tube. The column was centrifuged at 14,000g for 30 seconds and the flowthrough was discarded. The column was then washed with 200µl of endotoxin removal buffer and 400µl of column wash solution sequentially. Plasmid was eluted into a

1.5ml Eppendorf tube with 60µl of 1X TE buffer (pH=8.0). Plasmid purity and quantity were determined with a nanodrop (Thermofisher Scientific) and stored at –20°C until ready for use.

3.8 Cell cultures

MDA-MB468 cell line was cultured in DMEM (Dulbecco's Modified Eagle Medium – United States) supplemented with 10% foetal bovine serum and penicillin/streptomycin antibiotics in a T 75 culture flask. All experiments were done when cells were at 60-80% confluence. Before experiments, cells were trypsinized with 2ml of 1X trypsin-EDTA for 5 minutes at 37°C and 5% CO_2. The reaction was stopped with 5ml of DMEM supplemented with FBS and cells were transferred into a 15ml falcon tube for centrifugation. Cells were centrifuged at 1000rpm for 5 mins to obtain cell pellets. Pellets were resuspended in 1ml of media and then cells were counted before seeding the required cells for experiments.

3.8.1 RNA extraction from cell lines

RNA was extracted from MDA-MB-468 using Quick-RNA™ Miniprep plus Kit. Cells were lysed in wells after media was removed post-transfection using 200µl of RNA lysis buffer per well for 5 minutes. The lysate was transferred into a Spin-Away™ Filter[1] (yellow) in a collection tube and centrifuged at 14,000g for 1 minute to remove the majority of genomic DNA which bound to the column. An equal volume of absolute ethanol was added to the flowthrough and mixed, then transferred to a Zymo-Spin™ IIICG Column[1] (green) in a collection tube and centrifuged. The column was washed and treated with DNase1 for 15 minutes to further remove genomic DNA. Subsequently, 400µl of RNA Prep Buffer was added to the column and centrifuged at 14,000g for 30 seconds. The column was washed twice using 700µl and 400µl of RNA Wash Buffer in succession. The RNA was then eluted

in 60µl of 1X TE Buffer (pH=8.0) and stored at -80°C until ready for use. The purity and quantity of RNA were measured using a nanodrop spectrophotometer after extraction and prior to use in subsequent experiments. (Thermofisher Scientific).

3.9 Transfection of cell lines

MDA-MB4-68 cell line was transfected with *pCMV-MIR4725* plasmid in increasing concentration of plasmid DNA in a 12-well plate (0.5µg, 1µg and 1.5µg per well) with Polyfect transfection reagent (Qiagen, Germany) using optimized manufacturer's protocol. Approximately 300,000 cells were seeded in each well of a 12-well plate in 1ml of DMEM supplemented with 1% penicillin/streptomycin and 10% foetal bovine serum a day before transfection. On the day of transfection, media was pipetted off and cells were rinsed with 1x phosphate-buffered saline (PBS) and 1ml of fresh media was added. The transfection mix contained specific concentration of plasmid DNA (0.5µg, 1µg and 1.5µg per well) to a final volume of 20µl with incomplete DMEM (free of antibiotics and FBS). Three microlitres of polyfect reagent was added to the mix, vortexed, spun down and incubated for 15 minutes at room temperature. A 100µl of complete media was then added after incubation and the mix was pipetted into each well. Cells were allowed to incubate for 48 hours at 37°C in 5% CO_2 after transfection before RNA extraction for expression analysis.

3.10 MiR-4725 expression analysis

Both mature miR-4725-5p (sequence; *AGACCCUGCAGCCUUCCCACC*) and miR-4725-3p (sequence; *UGGGGAAGGCGUCAGUGUCGGG*) TaqMan Assay were purchased from Thermofisher Scientific as well as a small nuclear RNA U6 as endogenous control (RNU6). Quantification using the TaqMan assay involved a two-step reverse transcription quantitative PCR (RT-qPCR) kit following the manufacturer's protocol with modifications.

3.10.1 cDNA synthesis

cDNA was obtained through reverse transcription from 0.5µg of total RNA samples using specific miR-4725-3p, 5p and RNU6 stem-loop primers from the TaqMan MicroRNA Assays and reverse transcription reagents from the TaqMan® MicroRNA Reverse Transcription Kit (Thermofisher scientific). The stem-loop primer binds to the 3' end of the mature miRNA giving the reverse transcriptase a starting material for the extension of the cDNA. Briefly, a reaction mix containing 0.2µl RNase inhibitor, 0.3µl dNTPs, 1.5µl of RT buffer, 2µl of stem-loop primer, 2µl of multi-scribe reverse transcriptase, 0.5µg of total RNA and nuclease-free water to a 15µl final volume in a PCR tube. Tube was placed in a thermocycler with conditions set as seen in Table 3.3.

Table 3.3 Reverse transcription reaction condition

Step	Temperature (°C)	Times (min)
1	16	30
2	43	30
3	85	5
4	4	hold

3.10.2 qPCR

cDNA product from the reaction described in 3.10.1 was used as a template in addition to TaqMan MicroRNA Assays (containing a primer-probe of miR-4725-3p, 5p and RNU6), TaqMan™ Universal Master Mix II, with no UNG used in quantification of the miRNA expression. The assay consists of a probe which has miRNA-specific primers, a 5' reporter FAM dye and a 3' non-fluorescent quencher which inhibits the fluorescence of the dye. The quencher is cleaved by the polymerase before the extension, after specific binding of the

primers and allowing the FAM dye to emit a fluorescence which is detected. Briefly, the reaction mix consisted of 10µl of master mix, 1µl of the primer-probe, 8.5µl of nuclease-free water and 0.5µl of cDNA template. The PCR was run using the conditions in Table 3.4 below.

The comparative ΔΔCt method was used to determine the fold change expression after the appropriate threshold was set and samples normalized with snRNA U6 assay as an endogenous control in each sample.

Table 3.4 Quantitative PCR reaction conditions

Step	Temperature (°C)	Time	Cycles
Enzyme activation	95	10 min	
Denaturation	95	15 sec	40
Annealing/Extension	60	60 sec (read)	40

3.11 *In silico* prediction of miR-4725-5p targets and gene ontology

The isomir miR-4725-5p is the mature miRNA that is expressed and its putative targets were determined using various online tools. First, the gene expression profile in TNBC of samples in The Cancer Genome Atlas (TCGA) was accessed using the University of Alabama at Birmingham Cancer (UALCAN) database analysis portal (https://ualcan.path.uab.edu/) (Chandrashekar *et al*, 2022). The top 250 significantly upregulated genes were obtained. Subsequently, the predicted targets of miR-4725-5p were also obtained from TargetScan database version 8.0 (https://www.targetscan.org/vert_80/). The putative targets were determined by finding the intersection between the two gene setlist. Gene ontology and Reactome was determined using Metascape v3.5.20230101 (https://metascape.org) and graphs plotted with SRplot software (https://www.bioinformatics.com.cn/en).

3.12 Target gene expression

The expression of miR-4725-5p putative target genes determined using bioinformatics, as well as selected genes involved in endothelial-mesenchymal transmission (EMT), angiogenesis and apoptosis were evaluated using a Luna® Universal One-Step RT-qPCR Kit using specific primers designed for each gene (Table 3.5). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) expression was used as an endogenous control for the normalization of gene expression in samples. RT-qPCR was performed following the manufacturer's instructions. Briefly, a final reaction mix containing 0.5µg of total RNA, 10µl of Luna Universal One-Step Reaction Mix (2X), 1µl of Luna WarmStart RT Enzyme Mix (20X), 0.8µl each of both forward and reverse primers and nuclease-free water to make up a 20µl final volume in a 96 well optical plate. RT-qPCR was run using cycling conditions in Table 3.5 on a QuantStudio 5 thermocycler (Thermofisher Scientific). CT values were retrieved and the comparative ΔΔCT was used to calculate the fold change of expression using RNA from untransfected MDA-MB468 cell line as normal control.

Table 3.5 Reverse transcription qPCR run conditions

Step	Temperature (°C)	Time	cycles
Reverse transcription	55	15 min	1
Initial denaturation	95	1 min	1
Denaturation	95	15 sec	40
Annealing	Dependent on the primer sequence	15 sec	
Extension	60	1 min (Read)	

Table 3. 6 Primer sequences for gene expression analysis

Gene Name	Forward 5'>3'	Reverse 5'>3'
Putative Gene Targets		
ANLN	*TCGAAGATGGTGTGTTCTTTCT*	*TTCTCTGTTGGCTGGTTCTATC*
MKi67	*GACCTCAAACTGGCTCCTAATC*	*GCTGCCAGATAGAGTCAGAAAG*
ASFIB	*CCATGGTTGGAAGATGGTATCAG*	*TCCTAGCAGAGGACAGGTTAAG*
KIF2C	*TTCAGTGGAATGGGCAGAAG*	*CCTTCGGATGTAAGGGAAGAAG*
NUSAP1	*GCAGTCTTCTGCTAGCCAAT*	*GGCCTTTCTATCCCAGCTTAC*
Other genes		
VEGF	*CAGGACATTGCTGTGCTTT*	*CTCAGAAGCAGGTGAGAGTAAG*
WISP1	*ACTCATTAAGGCAGGGAAGAAG*	*CCACAGTACTTGGGTTGATAGG*
C-MYC	*AAGCTGAGGCACACAAAGA*	*GCTTGGACAGGTTAGGAGTAAA*
SNAIL1	*CCACGAGGTGTGACTAACTATG*	*ACCAAACAGGAGGCTGAAATA*
TWIST1	*AGGCATCACTATGGACTTTCTC*	*GGCCAGTTTGATCCCAGTAT*

Anillin (ANLN), Marker of proliferation Ki67 (MKi67), Anti-silencing function 1 B (ASFIB), Kinesin family member 2C (KIF2C), Nucleolar and spindle-associated protein 1 (NUSAP1), Vascular endothelial growth factor A (VEGFA), WNT-induced signalling protein 1 (WISP1),

3.13 Alamar blue assay

Alamar Blue is a cell viability assay reagent containing resazurin, a cell-permeable, non-toxic, and faintly fluorescent blue indicator dye. Resazurin is a redox (oxidation-reduction) indicator that changes colour in response to cellular metabolic reduction. The reduced form, resorufin, is pink and strongly fluorescent, with the intensity of fluorescence proportional to the amount of respiring live cells. Alamar Blue is a direct biomarker of cell health, detecting the quantity of oxidation during respiration and quantifying cell viability and cytotoxicity.

Three hundred thousand (300,000) cells of MDA-MB468 were seeded in a 12-well plate in 500ul of DMEM (supplemented with 1% penstrep and 10% FBS) and incubated at 37°C and 5% CO_2 overnight. Cells were then transfected with increasing concentrations of pCMV-mir-4725 (0.5, 1.0 and 1.5 µg) and 0.5 µg of control plasmid. The transfected cells were

incubated for 24 hours at 37°C. After the 24-hour lapsed, the media was changed and 50ul of 5mg/ml of Alamar blue reagent was added and incubated for 5 hours. The supernatant was pipetted into a 96-well plate and the fluorescence was read at 590nm emission (excitation=545nm).

3. 14 Statistical analyses

All data generated from this study were captured and calculated using Microsoft Excel. Graph Pad Prism 8.4.3 Software was used for further data analysis. Categorial data were represented with graphs and pie chart. Differential gene expression (fold change) data was calculated using the ΔΔCT method and represented as bar graphs. Significant differences between differential gene expressions were determined using ANOVA and compared relative to the control. The significant level was set at 5% (p-value < 0.05) for all statistical analyses.

CHAPTER FOUR

4.0 RESULTS

4.1 Genotype of SNP rs73991220 in selected breast cancer samples.

Forty-eight (48) breast cancer samples (24 tumour and 24 normal-adjacent) were screened for the 50A>G mutation (rs73991220) in the miR-4725 gene. Four out of the 24 samples (17%) were TNBC (Figure 4.1), 12 (50%) were non-TNBC and 8 (33%) had no immunohistochemistry data (Figure 4.1). Two out of 24 (8%) of the tumour samples were found to have a heterozygous (AG) genotype (Figure 4.2). These 2 samples were TNBC. Samples and account for 50% of the TNBC samples. All the normal-adjacent samples had a homozygous (AA) genotype.

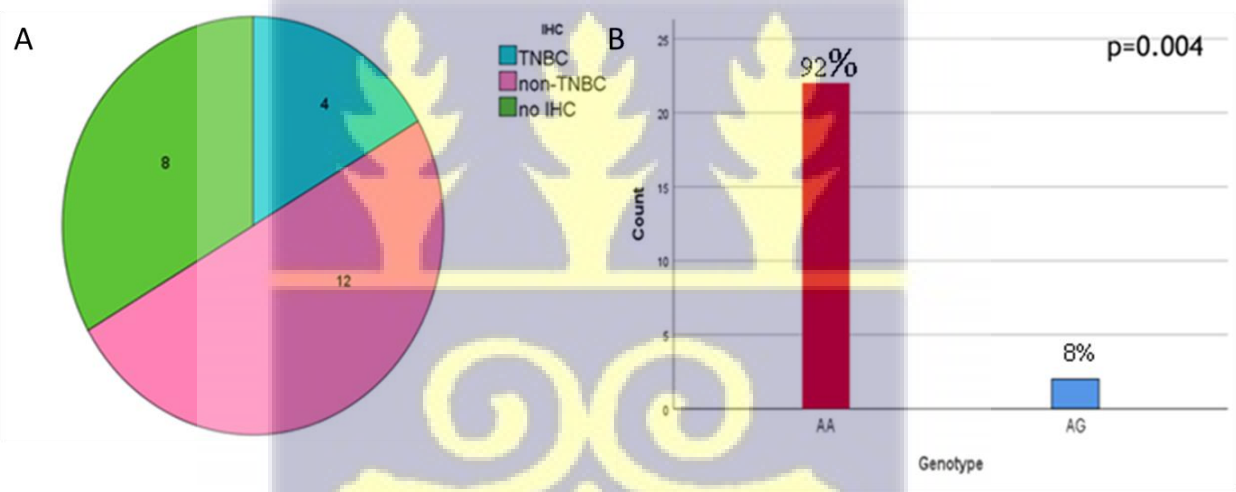


Figure 4.1 Screening the SNP rs73991220 in a breast tissue sample (24 tumour and 24 normal-adjacent): (A) Classification of the tumour sample. (B) Allele frequency distribution. TNBC-triple-negative breast cancer, IHC- immunohistochemistry

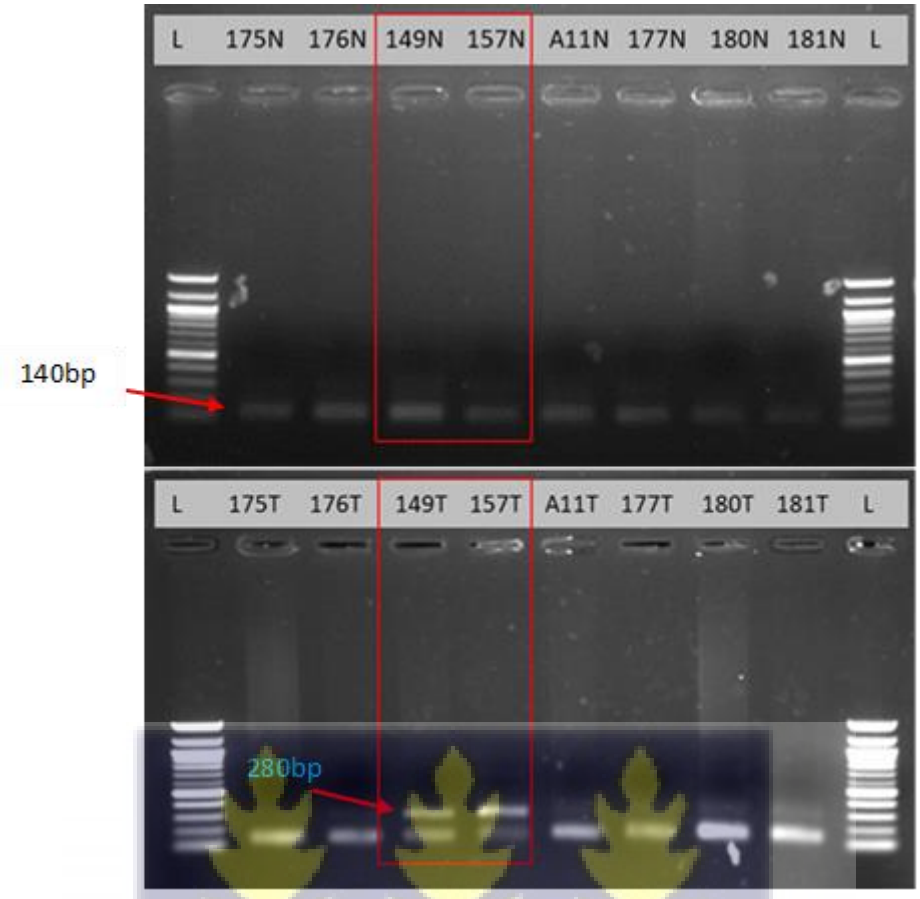


Figure 4.2 Representative gel of Restriction digest product. A single band at 140bp indicates a homozygous Wild Type (AA) genotype. A double band at 280bp and 140bp represents a heterozygous (AG) genotype.

4.2 G mutant allele confers higher stability to miR-4725 secondary folding structure

The secondary structure of the primary (pri) miR-4725 as well as the minimum free energy (MFE) of the wild type (A) and mutant (G) gene was determined using the RNAfold webserver (http://rna.tbi.univie.ac.at/). The SNP rs73991220 (G mutant) did not show any effect on the secondary structure of the pri-miR-4725. However, the A>G mutation reduced the entropy at its position (50) within the RNA and the base pairs within its vicinity in the RNA sequence resulting in a lower total minimum free energy (MFE) (-47.24kcal/mol) of the mutant compared to the wild type (-44.85kcal/mol). A lower MFE indicates a more thermodynamically stable structure of the RNA molecule. Thus, the *in-silico* analysis

suggests that the G mutant confers higher stability to the miR-4725 secondary folding structure (Figure 4.3).

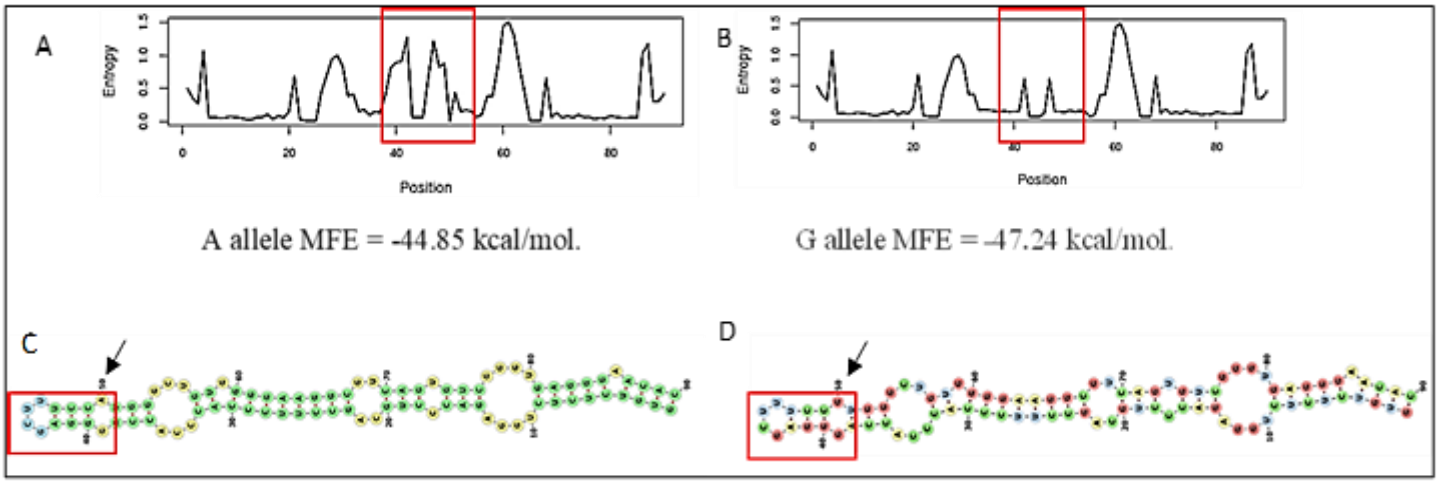


Figure 4. 3 G mutant allele confers higher stability to miR-4725 secondary structure. (A) And (B) show the entropy within the sequences of the pri-miR-4724. (C) and (D) are predicted folded structures of the pri-miR-4725. MFE- minimum free energy

Furthermore, the expression miR-4725-5p was quantified and compared between TNBC tumour samples with the heterozygous genotype AG (n=2), and homozygous wild type AA (n=2). Expression in the homozygous wild type sample was observed to be reduced compared to the heterozygous sample (mean Fold change of 0.02 vs 0.06). However, the observed difference was not statistically significant (p= 0.5029) (Figure 4.4).

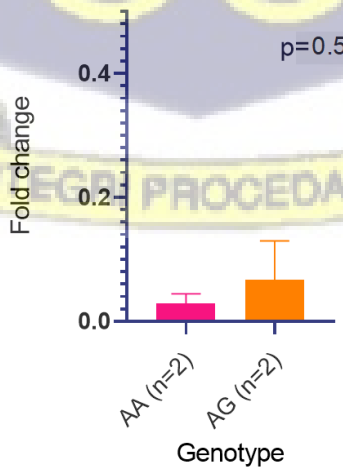


Figure 4.4 Expression of miR-4725-5p in 2 TNBC tumour sample with homogenous (AA) and the 2 TNBC heterogonous (AG) genotypes.

4.5 MiR4725 expression in breast cancer

Expression of miR-4725 was determined from RNA sequence data of TCGA invasive breast carcinoma patients' using UALCAN analysis software. There were 749 primary invasive breast carcinoma samples and 76 normal breast samples. Expression data showed miR-4725 has general low reads per million (RPM) (upper limit of 3.6RPM). However, expression was relatively higher in tumour samples compared to normal (figure 4.5).

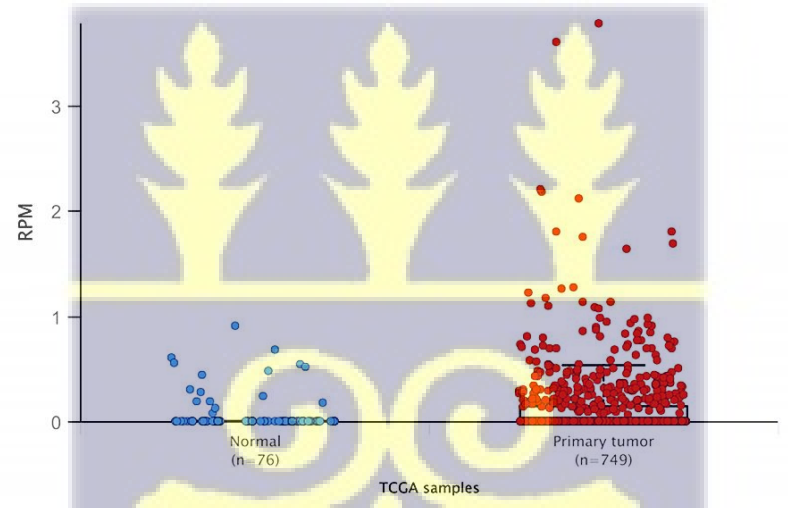


Figure 4.5 mir-4724-5p Expression in TCGA samples. Jitter plot of miR-4725 expression in the primary tumour of invasive breast carcinoma and normal samples from TCGA database. RPM-reads per million, TCGA- The Cancer Genome Atlas

4.3 MiR-4725-5p putative targets

MiR-4725 is downregulated in TNBC breast tumours and more so in the normal breast tissues (appendix), hence, significantly upregulated genes in TNBC were determined using UALCAN online analysis software (figure 4.6). Additionally, the target list consisting of 3479 possible genes with the potential complementary sequence for the mature miRNA

(miR-4725-5p) was obtained from TargetScan. A Venn diagram was used to depict the distribution of the potential targets of miR-4725-5p in TNBC. A total of 36 genes were obtained as potential targets (figure 4.7).

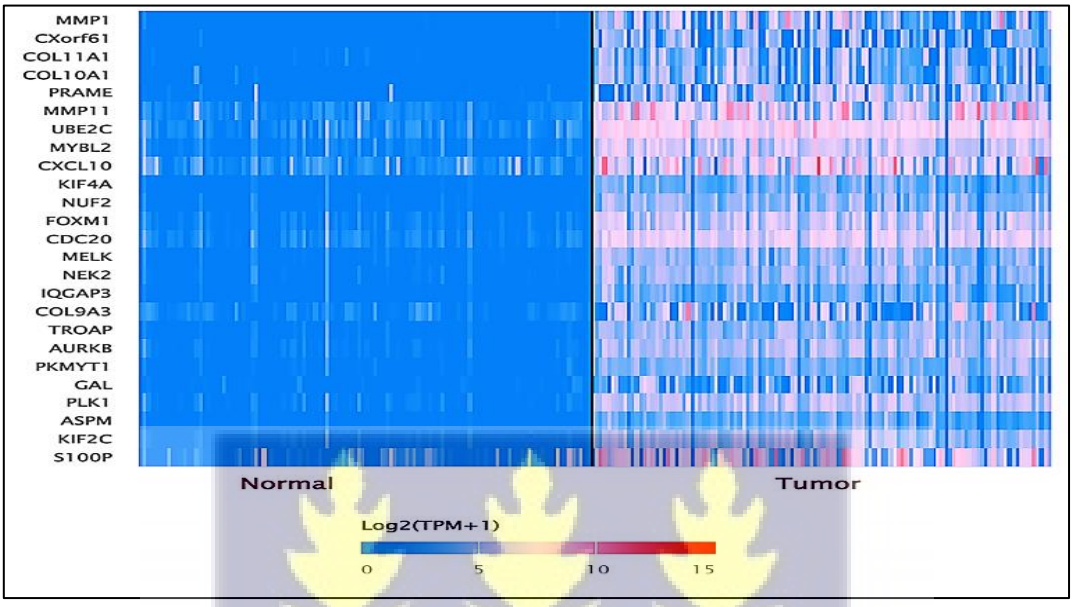


Figure 4.6 Heatmap of differentially upregulated genes in TNBC from TCGA expression database generated using UALCAN online analysis software. 250 significantly upregulated gene were identified.

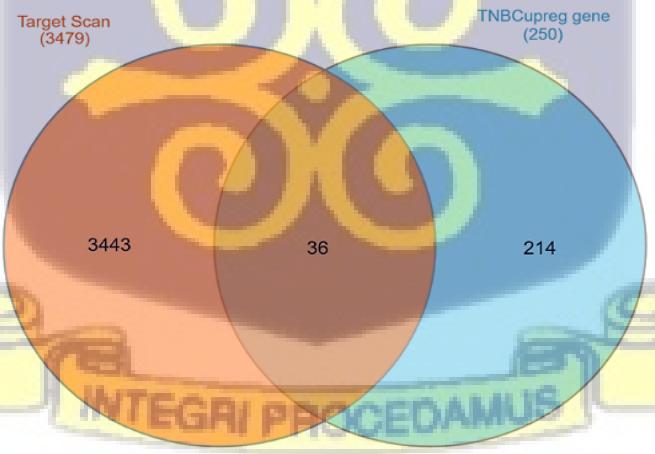


Figure 4.7 Predicted miR-4725-5p targets in TNBC. Venn diagram indicates the intersection of genes between significantly upregulated genes in TNBC and TargetScan gene list of miR-4725-5p Targets

4.4 Gene ontology of target genes

The gene ontology, as well as Reactome pathway analysis of the 36 potential targets of miR-4725-5p, was done using Metascape online software and the graphs were generated using the SRplot online tool. Generally, the 36 genes were found to be involved in mitotic processes and regulation (figure 4.8). Highly enriched genes such as kinesin family member 2 C (*KIF2C*) also known as mitotic-centromere associated kinesin (*MCAK*) and centromere protein A (CENPA) were observed to be involved in many Reactome pathways including amplification of signal form kinetochore, mitotic prometaphase and sister chromatid segregation (figure 4.9).

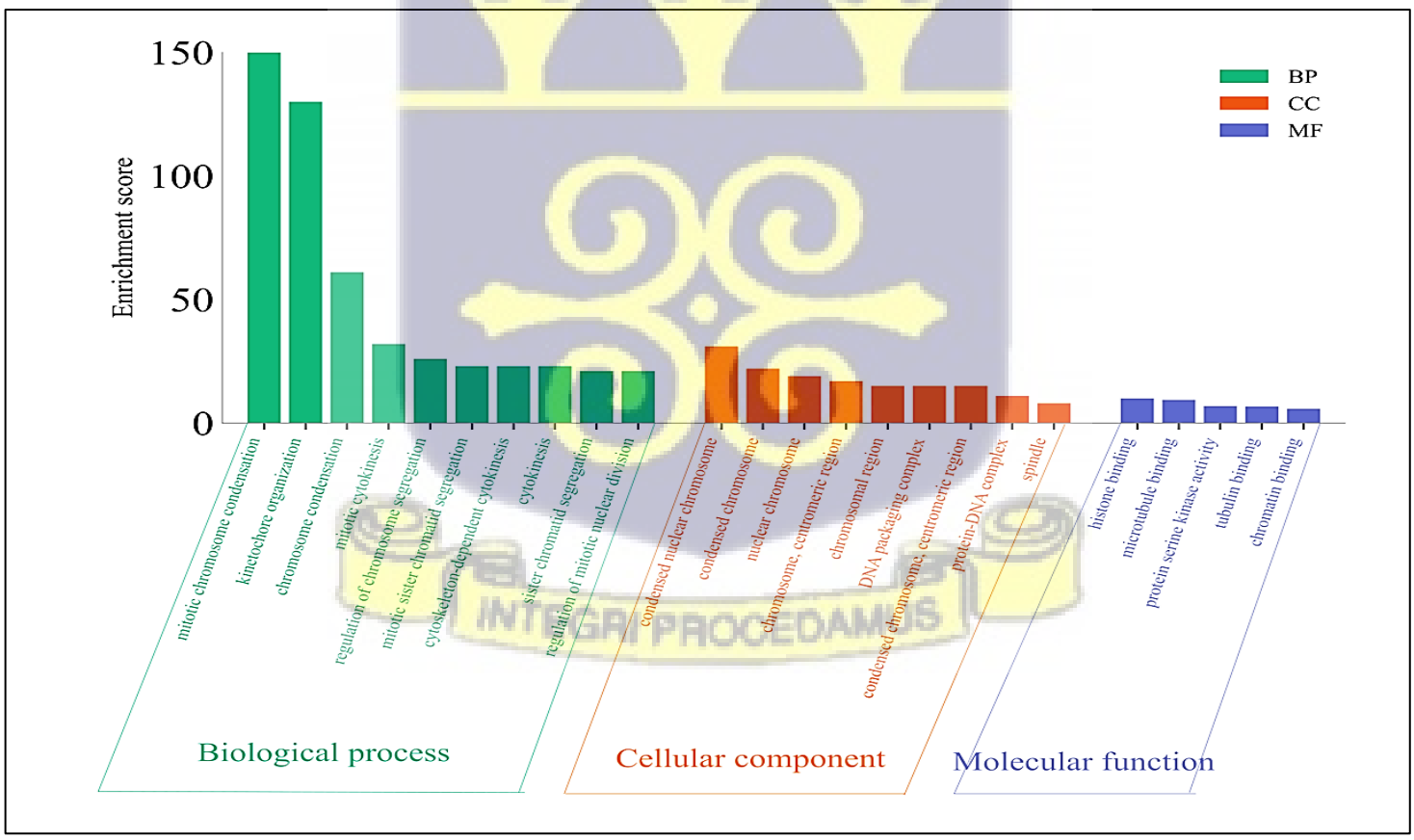


Figure 4.8 Gene ontology of the 36 identified Targets of miR-4725-5p. The putative target genes are enriched in biological processes encompassing chromosomal movement and packing and cytokinesis.

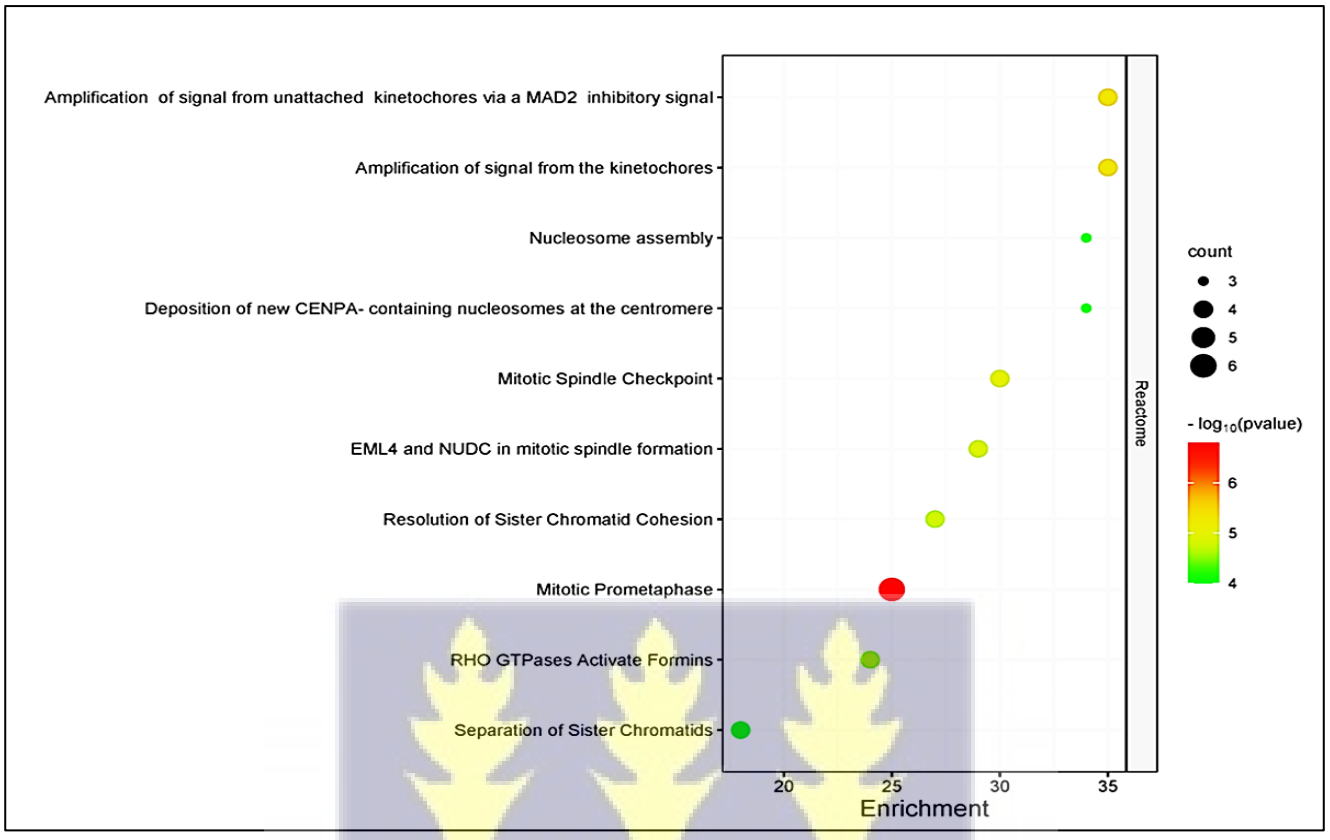


Figure 4.9 Reactome pathway of 36 identified genes from enrichment analysis. Significantly enrich gene (red) are involved in mitotic prometaphase and highly enriched gene in kinetochore signalling.

4.6 Baseline expression of miR-4725-5p in MDA-MB468 and MDA-MB231 cell lines.

The baseline expression of mir-4725-5p was determined within two triple-negative breast cancer cell lines MDA-MB468 and MDA-MB231. MiR-4725 expression was significantly low with the fold change of 0.00164 and 0.000185 in MDA-MB-468 and MDA-MB-231 respectively (figure 4.10).

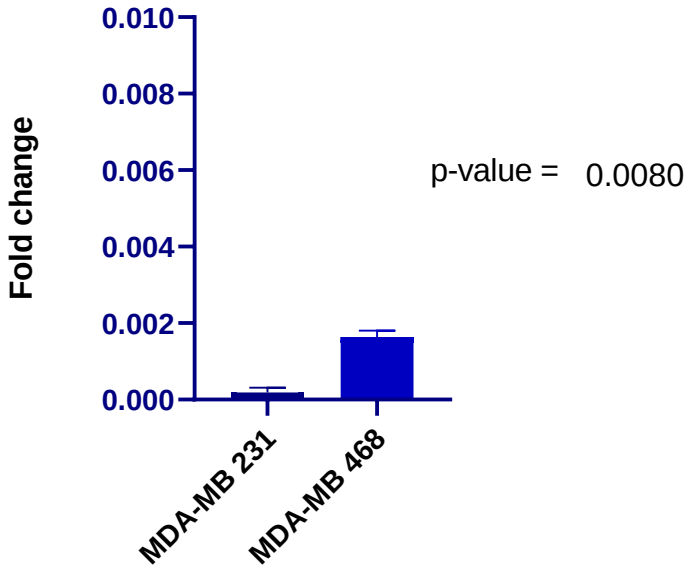


Figure 4.10 MiR-4725-5p expression in TNBC cell lines. Mature miR-4724-5p expression in MDA-MB231 and MDA-MB468 determined through RT-qPCR

With the observed low endogenous expression of miR-4725 in MDA-MB 468 cell lines, the cells were transfected with varying concentrations of pCMV-miR-4725 (CAT#: SC401868) plasmid which expresses the mature miR-4725-5p. Transfection was validated with the quantification of the mature miR-4725-5p expression. The result showed a significant increase in the relative fold change (RFC) of the miR-4725-5p (to the increasing plasmid DNA (pDNA) concentration (0.5µg, 1.0µg and 1.5µg) in MDA-MB-468 cells (figure 4.11).

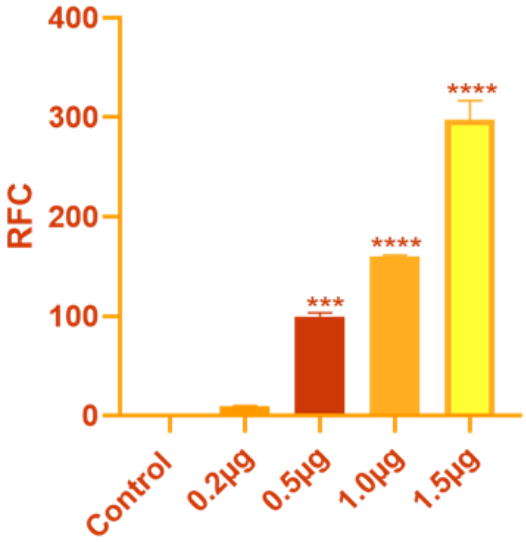


Figure 4.11 MiR-4725-5p expression validation. Increasing levels of miR-4725-5p expression after transfecting with increasing plasmid concentration ns- no significance, *-p<0.05, **-p<0.01.***- p<0.001, ****- p<0.0001

4.7 MiR-4725-5p putative target gene expression analysis

Five putative target genes were selected for use in this study based on a preliminary analysis done using gene expression data from 37 TNBC patients of African descent. Expression levels of the 5 putative targets (*ANLN*- anillin, *MKi67*- marker of proliferation ki67, *ASF1B*- anti-silencing function 1B chaperone protein, *NUSAP1*- nucleolar and spindle associated protein 1 and *KIF2C*- kinesin family member 2C) was determined after transfection of MDA-MB468 cell lines with miR-4725-5p expression plasmid. KIF2C expression was significantly reduced in response to the increased expression of the miR-4725-5p (RFC= 0.323, 0.157, 0.128 for 0.5 µg, 1.0 µg and 1.5µg of pDNA). *NUSAP1* and *ANLN* showed a dose-dependent decrease in expression, however, a significant reduction was observed in cells transfected with the 1.5µg concentration of pDNA (RFC=0.733 and 0.310 respectively). A dose-dependent increase in *ASF1B* expression was observed although it wasn't statistically significant. Likewise, MKi67 had no significant change in expression in the cells transfected with 0.5 µg and 1.0 µg of pDNA but not in the 1.5 µg transfection (figure 4.12).

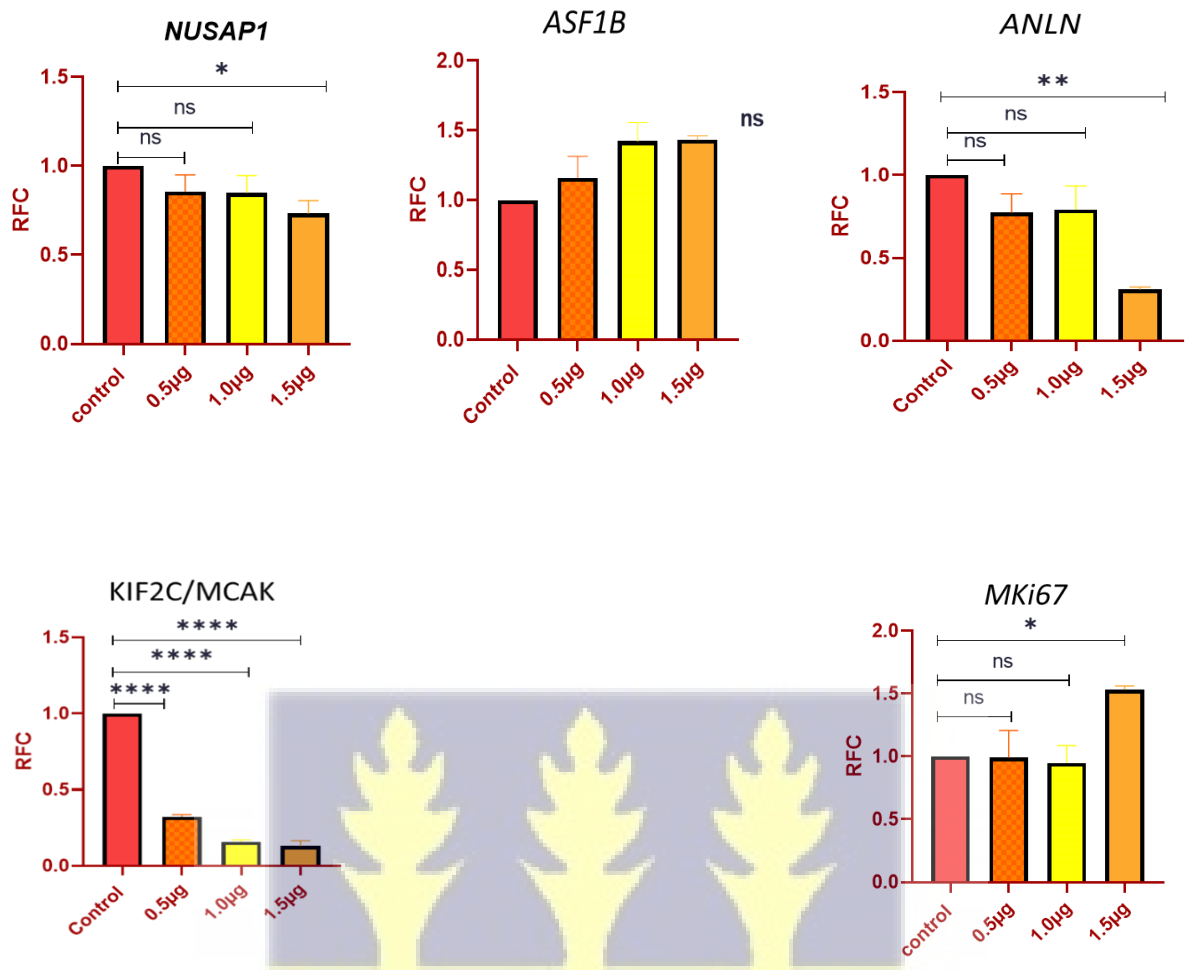


Figure 4.12 Gene expression levels of Putative targets. *NUSAP1*-nucleolar and spindle associated protein 1, *ASF1B*- ant-silencing function 1B chaperone protein, *ANLN*- anillin, *KIF2C*-kinesin family member 2, *MKi67*- marker of proliferation Ki 67. ns- No significance, *-p<0.05, **-p<0.01.***- p<0.001, ****- p<0.0001. *KIF2C* expression was most significantly reduced in response to exogenous miR-4725-5p

4.8 Expression of *KIF2C* in breast cancer

The pattern of expression of *KIF2C* was determined in breast cancer and the subclasses of breast cancer (Luminal, HER2-enriched and TNBC) using GEPIA online tool (https://gepia.cancer-pku.cn/) and UALCAN online analysis software (https://ualcan.path.uab.edu/). *KIF2C* was observed to be generally significantly upregulated in breast tumours than in normal breast tissues. Within the molecular subclasses of breast

cancer, TNBC had a higher expression of *KIF2C* than in HER2-enriched and luminal subtypes (figure 4.13).

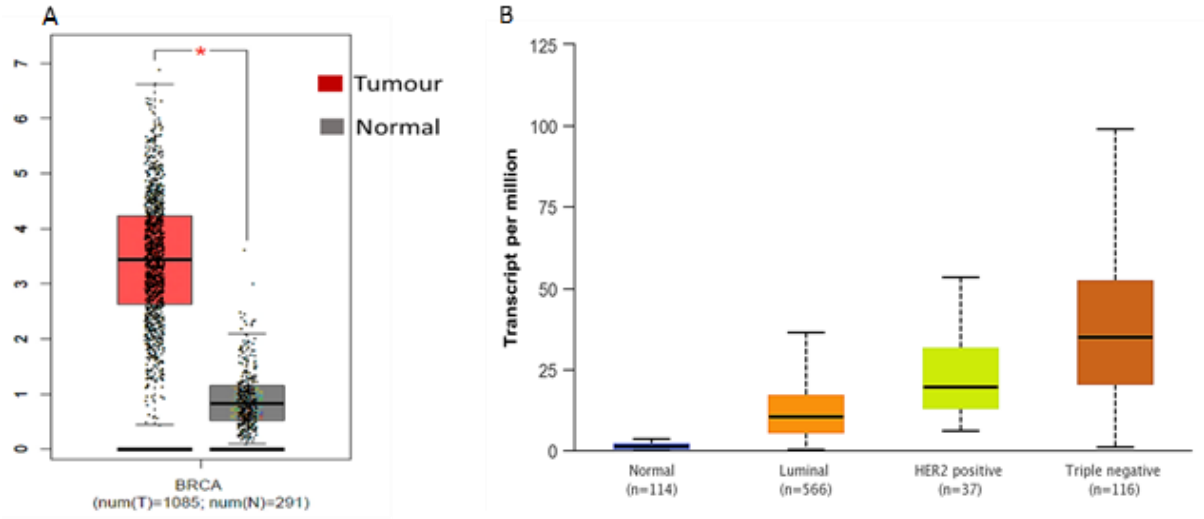


Figure 4.13 Differential expressions of KIF2C in Breast cancer subtypes. in normal and breast tumour samples from TCGA breast cancer samples. Data generated with GEPIA online software (A) and UALCAN (B). KIF2C is highly expressed in TNBC subtype than in other subtypes of breast cancer.

4.8 miR-4725-5p effect on cancer hallmarks

The expression of some genes known to play significant roles in breast cancer progressions such as proliferation, metastasis, endothelial-mesenchymal transition (EMT) and vascularization was determined in transfected cells.

Alamar blue assay was used to determine the viability of the cells after transfection. There were no significant changes observed between cell only, and transfected cells (control and increasing concentration of pDNA). Two gene markers *c-MYC* and *TP53* were used to assess proliferation and apoptosis, respectively. The expression of *c-MYC* was observed to increase with increasing pDNA. However, there were no significant changes observed with *TP53* expression (Figure 4.14).

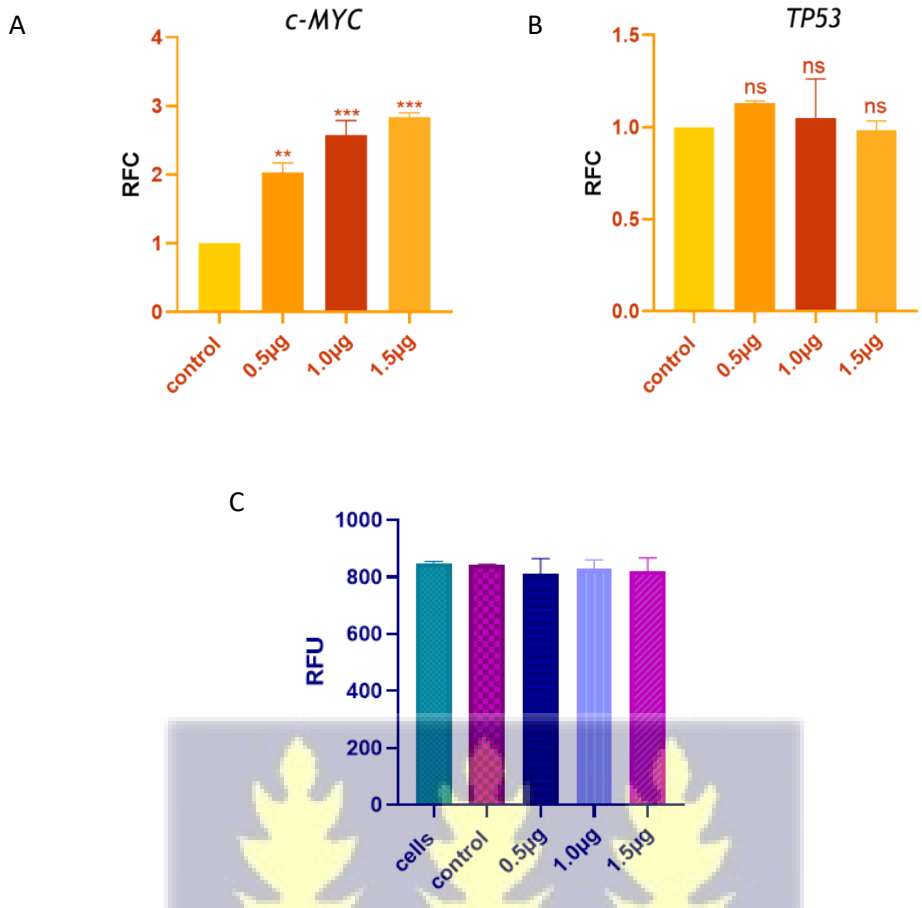


Figure 4.14 MiR-4725-5p effects on proliferation and apoptosis in MDA-MB468 cell line. (A) And (B) are *c-MYC* and *TP53* gene expression respectively. C is an Alamar blue assay. Expression of c-MYC was significantly increased with increasing concentration of plasmid DNA whereas TP53 expression as well as the cell viability (Alamar blue assay) was not affected by the increasing concentration of plasmid DNA. ns- no significance, *-p<0.05, **-p<0.01.***- p<0.001, ****- p<0.0001. RCF; relative fold change. RFU: relative fluorescent units

Expression of markers of endothelial-mesenchymal transition/metastasis and angiogenesis were also determined in the transfected MDA-MB468 cells. WNT-induced signalling pathway protein 1 (*WISP1*) and *SNAI1* also known as *SNAIL* were also observed to be significantly upregulated. Vascular endothelial growth factor (*VEGFA*), a marker of angiogenesis showed significant upregulation (Figure 4.15).

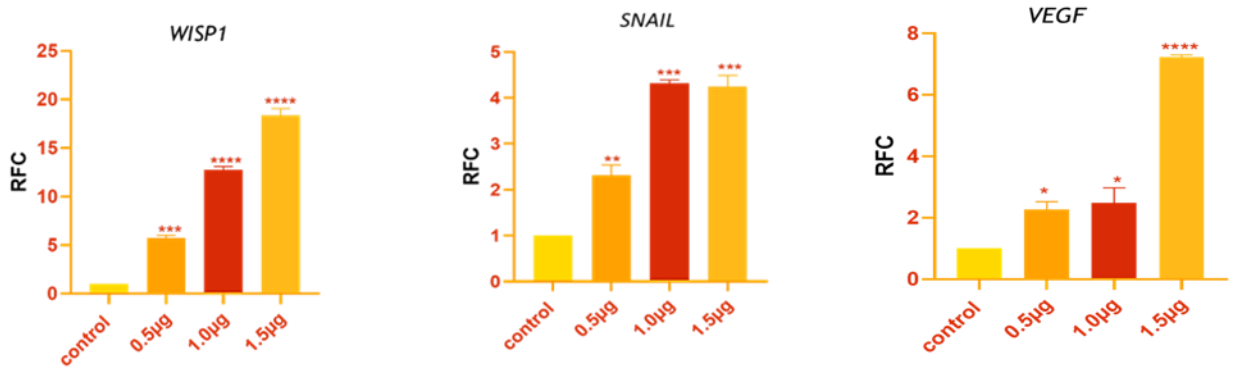


Figure 4.15 **M**iR-4725-5p promotes endothelial-mesenchymal transition (EMT) and angiogenesis. *WISP1*, *SNAIL* and *VEGF* which are tumour promoting markers are upregulated in responses to exogenous miR-4725-5p ns- no significance, *-p<0.05, **-p<0.01.***- p<0.001, ****- p<0.0001. RCF; relative fold change. RFU: relative fluorescent units

CHAPTER FIVE

5.0 DISCUSSION AND CONCLUSION

5.1 Discussion

TNBC is a highly aggressive form of breast cancer with complex and heterogeneous molecular characteristics. Unlike other breast cancer subtypes, TNBC expresses none of the characteristic hormonal receptors, estrogen receptor (ER), progesterone receptor (PR), or human epidermal growth factor 2 receptor (HER2); thus, therapies targeting these hormone receptors are unsuitable. Recently, Sacituzumab govitecan-hziy, a targeted chemotherapy which targets TROP-2-expressing cancer cells, has been approved for use in TNBC treatment (Hasanzad *et al*., 2023; So *et al*., 2022). However, due to its complex molecular characteristics and interindividual variations, the potential of Sacituzumab govitecan-hziy may soon face challenges and drug resistance may develop. The need for highly efficient therapeutic targets to manage the disease has led to an increased interest in miRNAs, which have been found to significantly contribute to tumorigenesis and cancer progression. Generally, miRNAs are involved in post-transcriptional gene regulation. However, dysregulation of miRNAs leads to tumour development and progression. Dysregulation of miRNAs has been associated with various factors, including SNPs within miRNA genes and genes involved in their biosynthesis. This can result in altered expression of miRNAs (Arun *et al*., 2022; Hu *et al*., 2008).

In this study, the novel SNP rs73991220 (A/G) in mir4725 identified by Qian *et al*. (2016) to be associated with the risk of estrogen receptor-negative breast cancer was found to affect the stability of the mir-4725 secondary structure. Secondary structure stability has been shown to affect miRNAs processing by Drosha complex and dicer downstream (Gebert *et al*., 2019; Ha & Kim, 2014). More stable miRNA secondary structures affect further processing and, hence,

decreased expression of mature miRNAs. The G mutant allele of mir-4725 was shown to have lower entropy and MFE, and hence, a more stable secondary structure than the reference A allele. Furthermore, a comparison of the expression of mir-4725 in TNBC tissue samples with the heterozygous AG genotype and homozygous AA revealed a nonsignificant difference in expression. It is worth mentioning however that the expression miR-4725-5p in the tumour samples was observed to be generally reduced and hence a heterozygous genotype may not have any significant influence in the expression of mature mir4725.

The 5p arm of miR-4725 was investigated in this study as expression analysis of breast cancer showed that mir4725-5p is the isomiR expressed in the breast (Appendix 4). Gene ontology of the thirty-six putative targets showed that miR-4725-5p could have a conserved mitotic function, specifically involved in chromosomal movement and packaging. The identified 36 targets including *ANLN, MKI67, KIF2C, NUSAP1* and *ASF1B* have been implicated in oncogenic roles in many cancers including TNBC (Gao *et al*., 2021; Han *et al*., 2018; Maryam & Chin, 2021; Sun *et al*., 2020).

ANLN, *MKI67*, *NUSAP1*, *KIF2C*, and *ASF1B* are upregulated in TNBC and involved in disease progression (Han *et al*., 2018; Jiang *et al*., 2021; Maryam & Chin, 2021; Sun *et al*., 2020). The expression of these genes was determined in MDA-MB-468 cells 48 h post-transfection with increasing concentrations of the miR-4725 expression plasmid. *KIF2C* was found to be significantly downregulated with increasing concentrations of mir-4725 plasmid. *KIF2C* or mitotic centromere-associated kinesin (*MCAK*) is important for mitosis in eukaryotic cells. It encodes a kinesin-like protein that can depolymerize microtubules at the plus-end, thereby promoting mitotic chromosome segregation. (Barr & Gergely, 2008). KIF2C forms a complex with KIF18B to execute its function. However, this complex is

negatively regulated by AURORA kinases by the phosphorylation of KIF2C (Tanenbaum *et al*., 2011). It has also recently been shown to be a crucial component of DNA double-strand break repair (Zhu *et al*., 2020).

In many cancers, including breast cancer, KIF2C is overexpressed (Appendix 5), and has been shown to play various roles in tumour progression, including increased proliferation, metastasis, maintenance of stemness, and drug resistance (Jiang *et al.*, 2021) and consequently, predicted to be a potential biomarker or therapeutic target. It was therefore hypothesized that miR-4725-5p will reduce tumorigenesis following KIF2C downregulation. In contrast, gene associated with tumorigenesis and cancer progression were observed to be upregulated in the MDA-MB468 cell line following exogenous expression of miR-4725-5p. Tumorigenic phenotypes such as metastasis, EMT, angiogenesis, proliferation, and apoptosis were assessed using P53, MYC, WISP1, VEGF, and SNAIL The genes *MYC*, *WISP1*, *VEGF*, and *SNAIL* were significantly upregulated in response to mir-4725-5p expression in MDA-MB468 cells. These genes are upregulated in breast cancers and, therefore, in TNBC, where they play tumour-promoting roles (Casciano *et al*., 2020; Ferrara, 2010; Liu *et al*., 2019; Zhang *et al*., 2013).

c-*MYC* is a member of the *MYC* family of proto-oncogenes, which are important signalling hubs for many cellular processes that maintain the growth, proliferation, and survival of cancer cells. MYC is amplified in approximately 25% of breast cancers and occurs more frequently in TNBC (~50%), and is associated with the risk of relapse, poor prognosis, and death (Casciano *et al*., 2020).

WISP1 belongs to the CCN growth factor family and plays an important role in tumorigenesis and progression of a wide range of human malignancies. WISP proteins appear to be important for cancer cell proliferation, apoptosis, invasion, and metastasis (Liu *et al.*, 2019). Upregulation of WISP1 in breast cancer promotes EMT and expression of N-cadherin, Snail, and β-catenin, and represses E-cadherin (Chaing *et al.*, 2015). In addition to these processes, WISP1 has also been shown to suppress antitumor immunity via the downregulation of interleukin 12 (IL-12) in breast cancer (Klinke, 2014).

SNAI1 or Snail is a conserved zinc finger transcription factor, which, together with SNAI2/SLUG, forms a small family of EMT inducers during embryogenesis. SNAIL maintains the stemness characteristics of TNBC cells and drives mesenchymal stem cell differentiation, fuelling tumour invasion, metastasis, and chemoresistance (Tsirigoti *et al.*, 2022).

The upregulation of the angiogenic marker *VEGFA* in MDA-MB468 cells suggests the potential of mir4725-5p to promote blood vessel formation for tumour progression. *VEGFA* encodes ligands that are crucial for coordinating new blood vessel formation during embryonic and postnatal development (Ferrara, 2010).

Together, these genes (*MYC, WISP1, VEGF, and SNAIL*) may be involved in interconnecting pathways that lead to tumour progression. Although these genes on their own have specific pathways through which they are activated, WNT signalling and mitogen-activated protein kinase (*MAPK*) activation are pathways common to these genes. These two pathways are chiefly implicated in cancer development and progression. These pathways are involved in many cellular processes including the activation of the expression of transcription factors and

genes that promoting cell growth, differentiation and proliferation. Aberrant regulation of WNT and MAPK pathways results in development and progression of many cancers including breast cancer (Shorning *et al*., 2020). It can therefore be inferred that miR-4725-5p may be involved in these pathways by repressing the regulator gene(s) within these pathways, leading to the upregulation of SNAIL, WISP1, VEGFA, and c-MYC, thus facilitating tumour progression.

5.2 Conclusion

In conclusion, SNP rs73881220 is associated with TNBC and confers greater stability to the secondary structure of mir4725, which may favour the expression of mature mir-4725-5p. Furthermore, the downregulation of *KIF2C* by the overexpression of miR-4725-5p indicates that KIF2C is a significant target of miR-4725-5p in MDA-MB468 cells. Overexpression of miR-4725-5p resulted in the upregulation of *SNAIL*, *WISP1*, *c-MYC*, and *VEGFA*, which are pro-tumour gene products involved in WNT and MAPK signalling pathways. Therefore, miR-4725-5p may be an oncomir that contributes to cancer progression by upregulating oncogenes including SNAIL, WISP1, VEGFA, and c-MYC. Although KIF2C, a tumour promoting gene, is a significant target of miR-4725-5p in this study, its repression did not influence the expression of other tumour promoting genes in the MDA-MB468 cell line. Thus, the repression of KIF2C by miR-4725-5p may be independent of miR-4725-5p function as an oncomir (figure 5.1).

5.2.1 Graphical summary

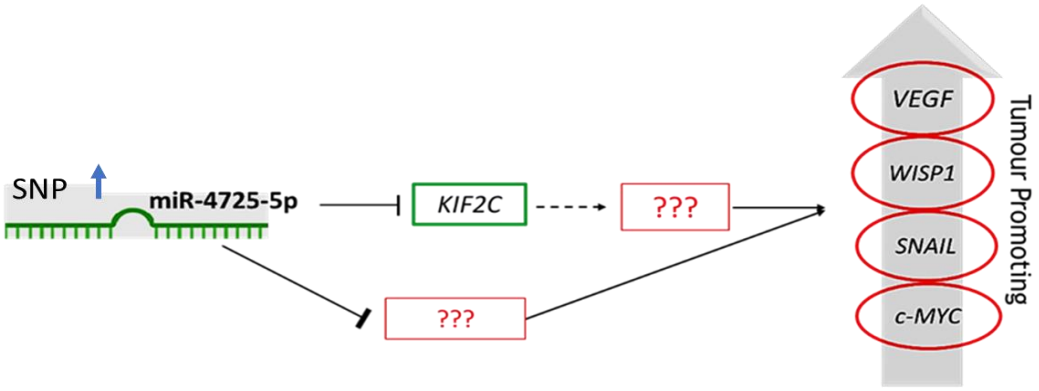


Figure 5.1 Possible interaction of miR-4725-5p in TNBC progression. miR-4725-5p may promote tumour by repressing KIF2C or independent of KIF2C

5.3 Limitations and recommendations

This study was unable to address the effect of the homozygous genotype of SNP rs73991220 (G) on miR-4725 expression. Moreover, only two tumours were identified as having the heterozygous genotype; hence, comparative expression showed no significant differences. Hence, a larger sample size should be investigated for the frequency of the SNP and to determine the association of the SNP to the expression of miR-4725 within the tissue samples. Additionally, an *in vitro* investigation of how the homozygous genotype of the SNP rs73991220 affects the expression of miR-4725 should would be necessary to further validate the effect of the SNP on mir4725 expression.

Only five of the 36 identified putative targets were investigated. However, a global expression profile, when miR-4725-5p is stably expressed, should be investigated to understand the mechanism of its gene regulation in TNBC and normal breast. A specific focus on miR-4725-5p downregulation of KIF2C and other downstream effect should be investigated to determine the possible mechanism involved in the tumour-promoting phenotype observed in this study.

REFERENCES

Adams, J. M., & Cory, S. (2007). The Bcl-2 apoptotic switch in cancer development and therapy. Oncogene, 26(9), 1324–1337.

Adeloye, D., Sowunmi, O. Y., Jacobs, W., David, R. A., Adeosun, A. A., Amuta, A. O., ... & Chan, K. Y. (2018). Estimating the incidence of breast cancer in Africa: a systematic review and meta-analysis. Journal of global health, 8(1).

Akuoko, C. P., Chambers, S., & Yates, P. (2022). Healthcare providers' perspectives of the supportive care needs of women with advanced breast cancer in Ghana. BMC Women's Health, 1–16.

Ali, R. M., McIntosh, S. A., & Savage, K. I. (2021). Homologous recombination deficiency in breast cancer: Implications for risk, cancer development, and therapy. Genes, Chromosomes and Cancer, 60(5), 358-372.

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

Al-Mahmood, S., Sapiezynski, J., Garbuzenko, O. B., & Minko, T. (2018). Metastatic and triple-negative breast cancer: challenges and treatment options. Drug delivery and translational research, 8, 1483-1507.

Anand, P., Kunnumakara, A. B., Sundaram, C., Harikumar, K. B., Tharakan, S. T., Lai, O. S., Sung, B., & Aggarwal, B. B. (2008). Cancer is a preventable disease that requires major lifestyle changes. Pharmaceutical research, 25, 2097-2116.

Anyigba, C. A., Awandare, G. A., & Paemka, L. (2021). Breast cancer in sub-Saharan Africa: The current state and uncertain future. Experimental Biology and Medicine, 246(12), 1377-1387.

Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., Polónia, A., & Campilho, A. (2017). Classification of breast cancer histology images using convolutional neural networks. PloS one, 12(6), e0177544.

Arun, R. P., Cahill, H. F., & Marcato, P. (2022). Breast Cancer Subtype-Specific miRNAs: Networks, Impacts, and the Potential for Intervention. 1–14.

Asoogo, C., & Duma, S. E. (2015). Factors contributing to late breast cancer presentation for health care amongst women in Kumasi, Ghana. Curationis, 38(1).

Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L. A., Hart, R. P., & Kiledjian, M. (2010). Differential regulation of microRNA stability. Rna, 16(5), 1032-1039.

Basu, A. K. (2018). DNA damage, mutagenesis and cancer. International journal of molecular sciences, 19(4), 970.

Bazzini, A. A., Xu, P., Calarco, J. A., & Jacobsen, S. E. (2008). MicroRNA stability and expression through the plant life cycle. Nucleic Acids Research, 36(17), 5635-5644.

Benson, C. S., Babu, S. D., Radhakrishna, S., Selvamurugan, N., & Sankar, B. R. (2013). Expression of matrix metalloproteinases in human breast cancer tissues. Disease Markers, 34(6), 395–405.

Bhowmick, N. A., Neilson, E. G., & Moses, H. L. (2004). Stromal fibroblasts in cancer initiation and progression. Nature, 432(7015), 332-337.

Bill, R., & Christofori, G. (2015). The relevance of EMT in breast cancer metastasis: Correlation or causality? FEBS letters, 589(14), 1577-1587.

Blasco, M. A. (2005). Telomeres and human disease: Ageing, cancer and beyond. Nature Reviews Genetics, 6(8), 611–622.

Blelloch. (2009). Small RNAs – their biogenesis, regulation and function in embryonic stem cells. StemBook, 1–16.

Bonsu, A. B., & Ncama, B. P. (2019). Recognizing and appraising symptoms of breast cancer as a reason for delayed presentation in Ghanaian women: A qualitative study. PLoS ONE, 14(1).

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

Brown, G. T., & Murray, G. I. (2015). Current mechanistic insights into the roles of matrix metalloproteinases in tumour invasion and metastasis. The Journal of pathology, 237(3), 273-281.

Carey, L. A., Perou, C. M., Livasy, C. A., Dressler, L. G., Cowan, D., Conway, K., Karaca, G., Troester, M. A., Tse, C. K., & Edmiston, S. (2006). Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. Jama, 295(21), 2492-2502.

Casciano, J. C., Perry, C., Cohen-Nowak, A. J., Miller, K. D., Vande Voorde, J., Zhang, Q., Chalmers, S., Sandison, M. E., Liu, Q., Hedley, A., McBryan, T., Tang, H. Y., Gorman, N., Beer, T., Speicher, D. W., Adams, P. D., Liu, X., Schlegel, R., McCarron, J. G., … Schug, Z. T. (2020). MYC regulates fatty acid metabolism through a multigenic program in claudin-low triple negative breast cancer. British Journal of Cancer, 122(6), 868–884.

Chandrashekar, D. S., Karthikeyan, S. K., Korla, P. K., Patel, H., Shovon, A. R., Athar, M., ... & Varambally, S. (2022). UALCAN: An update to the integrated cancer data analysis platform. Neoplasia, 25, 18-27.

Chen, X., *et al*. (2010). MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and activates Akt signaling in human ovarian cancer. Journal of Biological Chemistry, 285(8), 5761-5770.

Cheng, N., Chytil, A., Shyr, Y., Joly, A., & Moses, H. L. (2008). Transforming growth factor-β signaling–deficient fibroblasts enhance hepatocyte growth factor signaling in mammary carcinoma cells to promote scattering and invasion. Molecular Cancer Research, 6(10), 1521-1533.

Chiang, K. C., Yeh, C. N., Chung, L. C., Feng, T. H., Sun, C. C., Chen, M. F., ... & Juang, H. H. (2015). WNT-1 inducible signaling pathway protein-1 enhances growth and tumorigenesis in human breast cancer. Scientific reports, 5(1), 1-12.

Cho, W. C. (2010). MicroRNAs: potential biomarkers for cancer diagnosis, prognosis and targets for therapy. The international journal of biochemistry & cell biology, 42(8), 1273-1281.

Ciszewski, T., & Jopacka-Szatan, K. (2015). Breast cancer risk factors. Prz Menopauzalny, 14, 196-202.

Colditz, G. A., Kaphingst, K. A., Hankinson, S. E., & Rosner, B. (2012). Family history and risk of breast cancer: nurses' health study. Breast cancer research and treatment, 133, 1097-1104.

Correia de Sousa, M., Gjorgjieva, M., Dolicka, D., Sobolewski, C., & Foti, M. (2019). Deciphering miRNAs' action through miRNA editing. International journal of molecular sciences, 20(24), 6249.

Costa, M. W., Hasham, M. G., & Rosenthal, N. (2020). Molecular Tools in Cancer Research. In Abeloff's Clinical Oncology, 2-23. Elsevier.

Cullen, J. M., & Breen, M. (2016). An Overview of Molecular Cancer Pathogenesis, Prognosis, and Diagnosis. Tumors in Domestic Animals, 1–26.

Dai, X., Li, T., Bai, Z., Yang, Y., Liu, X., Zhan, J., & Shi, B. (2015). Breast cancer intrinsic subtype classification, clinical use and future trends. American journal of cancer research, 5(10), 2929.

Das, V., Bhattacharya, S., Chikkaputtaiah, C., Hazra, S., & Pal, M. (2019). The basics of epithelial–mesenchymal transition (EMT): A study from a structure, dynamics, and functional perspective. Journal of Cellular Physiology, 234(9), 14535-14555.

Davis, D. L., Bradlow, H. L., Wolff, M., Woodruff, T., Hoel, D. G., & Anton-Culver, H. (1993). Medical hypothesis: xenoestrogens as preventable causes of breast cancer. Environmental health perspectives, 101(5), 372-377.

de Rinaldis, E., Gazinska, P., Mera, A., Modrusan, Z., Fedorowicz, G. M., Burford, B., ... & Tutt, A. (2013). Integrated genomic analysis of triple-negative breast cancers reveals novel microRNAs associated with clinical and molecular phenotypes and sheds light on the pathways they control. BMC genomics, 14(1), 1-19.

Dharambir, K., Deeksha, P., Riya, S., Vivek Kumar, G., Neelam, G., Deepika, K., Atef, Z., Atef, Z., Shubham, K., & Assaye Belay, G. (2022). Global Increase in Breast Cancer Incidence: Risk Factors and Preventive Measures. BioMed Research International.

Disis, M. L. (2010). Immune regulation of cancer. In Journal of Clinical Oncology 28(29) 4531–4538.

Ellsworth, D. L., Turner, C. E., Ellsworth, R. E., & Li, C. J. (2019). A Review of the Hereditary Component of Triple Negative Breast Cancer: High- and Moderate-Penetrance Breast Cancer Genes, Low-Penetrance Loci, and the Role of Nontraditional Genetic Elements. Journal of Oncology, 2019.

Rakha, E. A., & Green, A. R. (2017). Molecular classification of breast cancer: what the pathologist needs to know. Pathology, 49(2), 111-119.

Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., Ji, X., Liu, W., Huang, B., Luo, W., Liu, B., Lei, Y., Du, S., Vuppalapati, A., Luu, H. H., Haydon, R. C., He, T.-C., & Ren, G. (2018). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes & Diseases, 5(2), 77-106.

Ferrara, N. (2010). Pathways mediating VEGF-independent tumor angiogenesis. Cytokine and Growth Factor Reviews, 21(1), 21–26.

Ferrara, N. (2010). Pathways mediating VEGF-independent tumor angiogenesis. Cytokine and Growth Factor Reviews, 21(1), 21–26.

Fouad, Y. A., & Carmen, A. (2017). Revisiting the hallmarks of cancer Yousef. Am J Cancer Res, 7(5), 1016–1036.

Galluzzi, L., Linkermann, A., Kepp, O., & Kroemer, G. (2020). Pathophysiology of cancer cell death. In Abeloff's Clinical Oncology. 74-83. Elsevier.

Gao, L. B., Bai, P., Pan, X. M., Jia, J., Li, L. J., Liang, W. B., ... & Zhang, L. (2011). The association between two polymorphisms in pre-miRNAs and breast cancer risk: a meta-analysis. Breast cancer research and treatment, 125, 571-574.

Gao, Z., Jia, H., Yu, F., Guo, H., & Li, B. (2021). KIF2C promotes the proliferation of hepatocellular carcinoma cells in vitro and in vivo. Experimental and Therapeutic Medicine, 22(4), 1–9

Garrido-Castro, A. C., Lin, N. U., & Polyak, K. (2019). Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment Heterogeneity of Triple-Negative Breast Cancer. Cancer discovery, 9(2), 176-198.

Gearhart-Serna, L. M., Davis, J. B., Jolly, M. K., Jayasundara, N., Sauer, S. J., Di Giulio, R. T., & Devi, G. R. (2020). A polycyclic aromatic hydrocarbon-enriched environmental chemical mixture enhances AhR, antiapoptotic signaling and a proliferative phenotype in breast cancer cells. Carcinogenesis, 41(12), 1648-1659.

Gebert, L. F., & MacRae, I. J. (2019). Regulation of microRNA function in animals. Nature reviews Molecular cell biology, 20(1), 21-37. Global Burden of Disease Cancer Collaboration; Sung H; Ferlay J, *et al*. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: Cancer J. Clin. 2021;71(3):209-249.

Gonzalez, H., Hagerling, C., & Werb, Z. (2018). Roles of the immune system in cancer: from tumor initiation to metastatic progression. Genes & development, 32(19-20), 1267-1284.

Gregory, A. D., & Houghton, A. M. G. (2011). Tumor-associated neutrophils: new targets for cancer therapy. Cancer Research, 71(7), 2411–2416.

Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. Nature Reviews Molecular Cell Biology, 15(8), 509-524.

Han, G., Zhang, X., Liu, P., Yu, Q., Li, Z., Yu, Q., & Wei, X. (2018). Knockdown of anti-silencing function 1B histone chaperone induces cell apoptosis via repressing PI3K/Akt pathway in prostate cancer. International Journal of Oncology, 53(5), 2056–2066.

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. In Cancer Discover. American Association for Cancer Research Inc. 12(1) 31–46.

Hanahan, D., & Coussens, L. M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer cell, 21(3), 309-322.

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. cell, 100(1), 57-70.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. cell, 144(5), 646-674.

Hanahan, D., & Weinberg, R. A. (2016). Biological hallmarks of cancer. Holland-Frei Cancer Medicine, 1-10.

Hanahan, A. J., Iyer, G., & Solit, D. B. (2020). 2 - Intracellular Signaling. In J. E. Niederhuber, J. O. Armitage, M. B. Kastan, J. H. Doroshow, & J. E. Tepper (Eds.), Abeloff's Clinical Oncology 12 (6), 24-46.

Hasanzad, M., Sarhangi, N., Aghaei Meybodi, H. R., & Nikfar, S. (2023). Pharmacogenetics and toxicology. In Reference Module in Biomedical Sciences. Elsevier.

Hercules, S. M., Alnajar, M., Chen, C., Mladjenovic, S. M., Shipeolu, B. A., Perkovic, O., Pond, G. R., Mbuagbaw, L., Blenman, K. R., & Daniel, J. M. (2022). Triple-negative breast cancer prevalence in Africa: a systematic review and meta-analysis. BMJ Open, 12(5), e055735.

Hindle, A. G., Thoonen, R., Jasien, J. V., Grange, R. M. H., Amin, K., Wise, J., Ozaki, M., Ritch, R., Malhotra, R., & Buys, E. S. (2019). Identification of candidate miRNA biomarkers for glaucoma. Investigative Ophthalmology and Visual Science, 60(1), 134–146.

Ho, K. H., Chang, C. K., Chen, P. H., Wang, Y. J., Chang, W. C., & Chen, K. C. (2018). miR-4725-3p targeting stromal interacting molecule 1 signaling is involved in xanthohumol inhibition of glioma cell invasion. Journal of Neurochemistry, 146(3), 269–288.

Howell, A., Anderson, A. S., Clarke, R. B., Duffy, S. W., Evans, D. G., Garcia-Closas, M., Gescher, A. J., Key, T. J., Saxton, J. M., & Harvie, M. N. (2014). Risk determination and prevention of breast cancer. Breast Cancer Research, 16(5), 1-19.

Ho-Yen, C., Bowen, R. L., & Jones, J. (2012). Characterization of basal-like breast cancer: an update. Diagnostic Histopathology, 18(3), 104-111.

Hu, Y., Qin, X., Yan, D., Cao, H., Zhou, L., Fan, F., ... & Feng, J. (2017). Genome-wide profiling of micro-RNA expression in gefitinib-resistant human lung adenocarcinoma using microarray for the identification of miR-149-5p modulation. Tumor Biology, 39(3), 1010428317691659.

Hu, Z., Chen, J., Tian, T., Zhou, X., Gu, H., Xu, L., Zeng, Y., Miao, R., Jin, G., Ma, H., Chen, Y., & Shen, H. (2008). Genetic variants of miRNA sequences and non-small cell lung cancer survival. Journal of Clinical Investigation, 118(7), 2600–2608.

Iorio, M. V., & Croce, C. M. (2012). MicroRNA dysregulation in cancer: Diagnostics, monitoring and therapeutics. A comprehensive review. EMBO Molecular Medicine, 4(3), 143–159.

Jain, R. K., Martin, J. D., Chauhan, V. P., & Duda, D. G. (2020). Tumor microenvironment: vascular and extravascular compartment. In Abeloff's Clinical Oncology 108-126. Elsevier.

Jazdzewski, K., Murray, E. L., Franssila, K., Jarzab, B., Schoenberg, D. R., & de la Chapelle, A. (2008). Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. Proceedings of the National Academy of Sciences, 105(20), 7269-7274.

Jiang, C. F., Xie, Y. X., Qian, Y. C., Wang, M., Liu, L. Z., & Shu, Y. Q. (2021). TBX15 / miR - 152 / KIF2C pathway regulates breast cancer doxorubicin resistance via promoting PKM2 ubiquitination. Cancer Cell International, 1–18.

Junttila, M. R., & Evan, G. I. (2009). P53 a Jack of all trades but master of none. Nature Reviews Cancer, 9(11), 821–829.

Jurj, M. A., Buse, M., Zimta, A. A., Paradiso, A., Korban, S. S., Pop, L. A., & Berindan-Neagoe, I. (2020). Critical analysis of genome-wide association studies: Triple negative breast cancer quae exempli causa. International Journal of Molecular Sciences, 21(16), 1–22.

Kanwal, R., Gupta, K., & Gupta, S. (2015). Cancer epigenetics: an introduction. Cancer Epigenetics: Risk Assessment, Diagnosis, Treatment, and Prognosis, 3-25.

Kapoor, C., Vaidya, S., Wadhwan, V., Kaur, G., & Pathak, A. (2016). Seesaw of matrix metalloproteinases (MMPs). Journal of cancer research and therapeutics, 12(1), 28-35.

Kian, R., Moradi, S., & Ghorbian, S. (2018). Role of components of microRNA machinery in carcinogenesis. Experimental oncology.

Klinke II, D. J. (2014). Induction of Wnt-inducible signaling protein-1 correlates with invasive breast cancer oncogenesis and reduced type 1 cell-mediated cytotoxic immunity: a retrospective study. PLoS computational biology, 10(1), e1003409.

Kohler, R. E., Gopal, S., Miller, A. R., Lee, C. N., Reeve, B. B., Weiner, B. J., & Wheeler, S. B. (2017). A framework for improving early detection of breast cancer in sub-Saharan Africa: A qualitative study of help-seeking behaviors among Malawian women. Patient education and counseling, 100(1), 167-173.

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. The Journal of clinical investigation, 121(7), 2750-2767.

Lehmann, B. D., Jovanović, B., Chen, X. I., Estrada, M. V., Johnson, K. N., Shyr, Y., ... & Pietenpol, J. A. (2016). Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. PloS one, 11(6), e0157368.

Li, S., Le, T. N. Y., Nguyen, T. D., Trinh, T. A., & Nguyen, T. A. (2021). Bulges control pri-miRNA processing in a position and strand-dependent manner. RNA biology, 18(11), 1716-1726.

Li, S., Li, Q., Lü, J., Zhao, Q., Li, D., Shen, L., ... & Yu, Z. (2020). Targeted inhibition of miR-221/222 promotes cell sensitivity to cisplatin in triple-negative breast cancer MDA-MB-231 cells. Frontiers in genetics, 10, 1278.

Ling, H., Fabbri, M., Calin, G. A., & Liu, C. G. (2013). Mutant p53 can induce miR-200c expression via the miR-200c/141 promoter. The international journal of biochemistry & cell biology, 45(11), 2507-2517.

Liu, C., *et al*. (2011). MicroRNA-155 promotes breast cancer cell proliferation and invasion by targeting the tumor suppressor gene programmed cell death 4. Cancer Research, 71(5), 1818-1826.

Liu, Y., Song, Y., Ye, M., Hu, X., Wang, Z. P., & Zhu, X. (2019). The emerging role of WISP proteins in tumorigenesis and cancer therapy. Journal of Translational Medicine, 17(1), 1–14.

Low, S. K., Zembutsu, H., & Nakamura, Y. (2018). Breast cancer: The translation of big genomic data to cancer precision medicine. Cancer science, 109(3), 497-506.

Ma, L., Teruya-Feldstein, J., & Weinberg, R. A. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature, 449(7163), 682-688.

Machowska, M., Galka-marciniak, P., & Kozlowski, P. (2022). Consequences of genetic variants in miRNA genes. Computational and Structural Biotechnology Journal, 20, 6443–6457.

Madu, C. O., Wang, S., Madu, C. O., & Lu, Y. (2020). Angiogenesis in Breast Cancer Progression, Diagnosis, and Treatment [Review]. Journal of Cancer, 11(15), 4474-4494.

Malumbres, M. (2020). Control of the cell cycle. In Abeloff's Clinical Oncology (pp. 56-73). Elsevier.

Martei, Y. M., Vanderpuye, V., & Jones, B. A. (2018). Fear of Mastectomy Associated with Delayed Breast Cancer Presentation Among Ghanaian Women. The Oncologist, 23(12), 1446–1452.

Maryam, A., & Chin, Y. R. (2021). ANLN Enhances Triple-Negative Breast Cancer Stemness Through TWIST1 and BMP2 and Promotes its Spheroid Growth. 8(July), 1–10.

Mattiuzzi, C., & Lippi, G. (2019). Current cancer epidemiology. Journal of epidemiology and global health, 9(4), 217.

Micalizzi, D. S., Farabaugh, S. M., & Ford, H. L. (2010). Epithelial-mesenchymal transition in cancer: Parallels between normal development and tumor progression. Journal of Mammary Gland Biology and Neoplasia, 15(2), 117–134.

Ngwa, W., Addai, B. W., Adewole, I., Ainsworth, V., Alaro, J., Alatise, O. I., Ali, Z., Anderson, B. O., Anorlu, R., & Avery, S. (2022). Cancer in sub-Saharan Africa: a lancet oncology commission. The Lancet Oncology.

Nielsen, F. C., van Overeem Hansen, T., & Sørensen, C. S. (2016). Hereditary breast and ovarian cancer: new genes in confined pathways. Nature Reviews Cancer, 16(9), 599-612.

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of microRNA biogenesis, mechanisms of actions, and circulation. Frontiers in endocrinology, 9, 402.

Olavarrieta, L., Aguirre, L. A., Castillo, I., Steel, K. P., Dalmay, T., Moreno, F., Miguel, A., & America, N. (2009). Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. 41(5), 609–613.

Ozdogan, H., Gur Dedeoglu, B., Oztemur Islakoglu, Y., Aydos, A., Kose, S., Atalay, A., Yegin, Z. A., Avcu, F., Ocean Cetinkaya, D., & Ilhan, O. (2017). DICER1 gene and miRNA dysregulation in mesenchymal stem cells of patients with myelodysplastic syndrome and acute myeloblastic leukemia. Leukemia Research, 63, 62–71.

Pandya, P. H., Murray, M. E., Pollok, K. E., & Renbarger, J. L. (2016). The Immune System in Cancer Pathogenesis: Potential Therapeutic Approaches. Journal of Immunology Research, 2016.

Patierno, S. R. (2020). 10 - Environmental Factors. In J. E. Niederhuber, J. O. Armitage, M. B. Kastan, J. H. Doroshow, & J. E. Tepper (Eds.), Abeloff's Clinical Oncology 6, 139-153.

Pilger, D., Seymour, L. W., & Jackson, S. P. (2021). Interfaces between cellular responses to DNA damage and cancer immunotherapy. Genes & development, 35(9-10), 602-618.

Plasilova, M. L., Hayse, B., Killelea, B. K., Horowitz, N. R., Chagpar, A. B., & Lannin, D. R. (2016). Features of triple-negative breast cancer Analysis of 38,813 cases from the national cancer database. Medicine (United States), 95(35).

Pylayeva-Gupta, Y., Grabocka, E., & Bar-Sagi, D. (2011). RAS oncogenes: Weaving a tumorigenic web. In Nature Reviews Cancer 11(11), 761–774.

Qian, F., Feng, Y., Zheng, Y., Ogundiran, T. O., Ojengbede, O., Zheng, W., Blot, W., Ambrosone, C. B., John, E. M., Bernstein, L., Hu, J. J., Ziegler, R. G., Nyante, S., Bandera, E. V., Ingles, S. A., Press, M. F., Nathanson, K. L., Hennis, A., Nemesure, B., … Huo, D. (2016). Genetic variants in microRNA and microRNA biogenesis pathway genes and breast cancer risk among women of African ancestry. Human Genetics, 135(10), 1145–1159.

Rahman, M. M., Brane, A. C., & Tollefsbol, T. O. (2019). MicroRNAs and epigenetics strategies to reverse breast cancer. Cells, 8(10), 1214.

Rakha, E. A., & Green, A. R. (2017). Molecular classification of breast cancer: what the pathologist needs to know. Pathology, 49(2), 111-119.

Rankin, E. B., & Giaccia, A. J. (2020). 3 - Cellular Microenvironment and Metastases. In J. E. Niederhuber, J. O. Armitage, M. B. Kastan, J. H. Doroshow, & J. E. Tepper (Eds.), Abeloff's Clinical Oncology (Sixth Edition) (pp. 47-55. e43). Elsevier.

Reis-Filho, J. S., & Pusztai, L. (2011). Gene expression profiling in breast cancer: classification, prognostication, and prediction. The Lancet, 378(9805), 1812-1823.

Rice, G. M., Shivashankar, V., Ma, E. J., Baryza, J. L., & Nutiu, R. (2020). Functional atlas of primary miRNA maturation by the microprocessor. Molecular cell, 80(5), 892-902.

Sadeghi, F., Asgari, M., Matloubi, M., Ranjbar, M., Karkhaneh Yousefi, N., Azari, T., & Zaki-Dizaji, M. (2020). Molecular contribution of BRCA1 and BRCA2 to genome instability in breast cancer patients: Review of radiosensitivity assays. Biological Procedures Online, 22, 1-28.

Santana, T. A. B. d. S., de Oliveira Passamai, L., de Miranda, F. S., Borin, T. F., Borges, G. F., Luiz, W. B., & Campos, L. C. G. (2023). The Role of miRNAs in the Prognosis of Triple-Negative Breast Cancer: A Systematic Review and Meta-Analysis. Diagnostics, 13(1).

Shao, M., & Li, W. (2019). Transcriptional factor regulation network and competitive endogenous RNA (ceRNA) network determining response of esophageal squamous cell carcinomas to neoadjuvant chemoradiotherapy. PeerJ, 2019(3).

Sharma, S., Kelly, T. K., & Jones, P. A. (2009). Epigenetics in cancer. Carcinogenesis, 31(1), 27–36. https://doi.org/10.1093/carcin/bgp220

Sheikh, A., Hussain, S. A., Ghori, Q., Naeem, N., Fazil, A., Giri, S., Sathian, B., Mainali, P., & Al Tamimi, D. M. (2015). The spectrum of genetic mutations in breast cancer. Asian Pacific Journal of Cancer Prevention, 16(6), 2177-2185.

Shen, J., Ambrosone, C. B., & Zhao, H. (2009). Novel genetic variants in microRNA genes and familial breast cancer. International journal of cancer, 124(5), 1178-1182.

Shen, X., *et al*. (2013). MicroRNA-21 and breast cancer: a possible biomarker and therapeutic target. Cancer Epidemiology, 37(6), 755-760.

Shorning, B. Y., Dass, M. S., Smalley, M. J., & Pearson, H. B. (2020). The PI3K-AKT-mTOR pathway and prostate cancer: At the crossroads of AR, MAPK, and WNT signaling. In International Journal of Molecular Sciences (Vol. 21, Issue 12, pp. 1–47). MDPI AG. https://doi.org/10.3390/ijms21124507

Sighoko, D., Kamaté, B., Traore, C., Mallé, B., Coulibaly, B., Karidiatou, A., Diallo, C., Bah, E., Mccormack, V., Muwonge, R., Bourgeois, D., Gormally, E., Paula, M., Bayo, S., & Hainaut, P. (2013). Breast cancer in pre-menopausal women in West Africa: Analysis of temporal trends and evaluation of risk factors associated with reproductive life q. The Breast, 22(5), 828–835.

Singletary, S. E. (2003). Rating the risk factors for breast cancer. Annals of surgery, 237(4), 474.

Sirohi, B., Chalkidou, K., Pramesh, C., Anderson, B. O., Loeher, P., El Dewachi, O., Shamieh, O., Shrikhande, S. V., Venkataramanan, R., & Parham, G. (2018). Developing institutions for cancer care in low-income and middle-income countries: from cancer units to comprehensive cancer centres. The Lancet Oncology, 19(8), e395-e406.

So, J. Y., Ohm, J., Lipkowitz, S., & Yang, L. (2022). Triple negative breast cancer (TNBC): Non-genetic tumor heterogeneity and immune microenvironment: Emerging treatment options. In Pharmacology and Therapeutics (Vol. 237). Elsevier Inc. https://doi.org/10.1016/j.pharmthera.2022.108253

Stanisławek, A. (2021). Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies— An Updated Review. 1–30.

Sun, L., Shi, C., Liu, S., Zhang, E., Yan, L., Ji, C., & Zhao, Y. (2020). Overexpression of NuSAP1 is predictive of an unfavourable prognosis and promotes proliferation and invasion of triple-negative breast cancer cells via the Wnt/β-catenin/EMT signalling axis. Gene, 747.

Sun, Y.-S., Zhao, Z., Yang, Z.-N., Xu, F., Lu, H.-J., Zhu, Z.-Y., Shi, W., Jiang, J., Yao, P.-P., & Zhu, H.-P. (2017). Risk factors and preventions of breast cancer. International journal of biological sciences, 13(11), 1387.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality

Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians, 71(3), 209–249.

Tanenbaum, M. E., Macurek, L., van der Vaart, B., Galli, M., Akhmanova, A., & Medema, R. H. (2011). A complex of Kif18b and MCAK promotes microtubule depolymerization and is negatively regulated by Aurora kinases. Current Biology, 21(16), 1356-1365.

Tang, Y., Wang, Y., Kiani, M. F., & Wang, B. (2016). Classification, treatment strategy, and associated drug resistance in breast cancer. Clinical breast cancer, 16(5), 335-343.

Tili, E., Michaille, J. J., Wernicke, D., Alder, H., Costinean, S., Volinia, S., & Croce, C. M. (2011). Mutator activity induced by microRNA-155 (miR-155) links inflammation and cancer. Proceedings of the National Academy of Sciences of the United States of America, 108(12), 4908–4913.

Topalian, S. L., Drake, C. G., & Pardoll, D. M. (2015). Immune checkpoint blockade: A common denominator approach to cancer therapy. In Cancer Cell (Vol. 27, Issue 4, pp. 450–461). Cell Press.

Tsang, J., & Tse, G. M. (2020). Molecular classification of breast cancer. Advances in anatomic pathology, 27(1), 27-35.

Tsirigoti, C., Ali, M. M., Maturi, V., Heldin, C. H., & Moustakas, A. (2022). Loss of SNAI1 induces cellular plasticity in invasive triple-negative breast cancer cells. Cell Death and Disease, 13(9).

Vodicka, P., Andera, L., Opattova, A., & Vodickova, L. (2021). The interactions of DNA repair, telomere homeostasis, and p53 mutational status in solid cancers: Risk, prognosis, and prediction. Cancers, 13(3), 479.

Wang, Q., Xu, M., Sun, Y., Chen, J., Chen, C., Qian, C., Chen, Y., Cao, L., Xu, Q., & Du, X. (2019). Gene expression profiling for diagnosis of triple-negative breast cancer: a multicenter, retrospective cohort study. Frontiers in Oncology, 9, 354.

Wang, X., *et al*. (2018). MicroRNA-21 promotes immune evasion in breast cancer by targeting PD-L1. Oncogene, 37(31), 4174-4183.

Wang, Z., Chen, J., Sun, Y., Li, C., & Li, Z. (2019). A single nucleotide polymorphism in the miR-30e-3p precursor is associated with susceptibility to hepatocellular carcinoma. Bioscience reports, 39(9), BSR20191113.

Wee, P., & Wang, Z. (2017). Epidermal growth factor receptor cell proliferation signaling pathways. Cancers, 9(5), 52.

Wu, M., Jolicoeur, N., Li, Z., Zhang, L., Fortin, Y., L'Abbe, D., Yu, Z., & Shen, S.-H. (2008). Genetic variations of microRNAs in human cancer and their effects on the expression of miRNAs. Carcinogenesis, 29(9), 1710-1716.

Xu, W., Yang, Z., & Lu, N. (2015). A new role for the PI3K/Akt signaling pathway in the epithelial-mesenchymal transition. Cell adhesion & migration, 9(4), 317-324.

Yang, F., Zhang, W., Shen, Y., & Guan, X. (2015). Identification of dysregulated microRNAs in triple-negative breast cancer (review). International Journal of Oncology, 46(3), 927–932.

Yates, L. A., Norbury, C. J., & Gilbert, R. J. (2013). The long and short of microRNA. cell, 153(3), 516-519.

Yin, L., Duan, J.-J., Bian, X.-W., & Yu, S.-c. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. Breast Cancer Research, 22, 1-13.

Zhang, H., Katerji, H., Turner, B. M., Audeh, W., & Hicks, D. G. (2022). HER2-low breast cancers: incidence, HER2 staining patterns, clinicopathologic features, MammaPrint and BluePrint genomic profiles. Modern Pathology, 35(8), 1075-1082.

Zhang, X., *et al*. (2010). A functional variant in miRNA-146a is associated with susceptibility to breast cancer. Cancer Research, 70(17), 6707-6711.

Zhang, X., *et al*. (2011). A functional variant in miRNA-499 is associated with susceptibility to coronary artery disease. Human Molecular Genetics, 20(24), 4958-4964

Zhao, J. J., Lin, J., Yang, H., Kong, W., He, L., Ma, X., Coppola, D., & Cheng, J. Q. (2008). MicroRNA-221/222 negatively regulates estrogen receptor α and is associated with tamoxifen resistance in breast cancer. Journal of Biological Chemistry, 283(45), 31079–31086.

Zheng, R., Sun, K., Zhang, S., Zeng, H., Zou, X., Chen, R., Gu, X., Wei, W., & He, J. (2019). Report of cancer epidemiology in China, 2015. Zhonghua zhong liu za zhi [Chinese journal of oncology], 41(1), 19-28.

Zhu, S., Paydar, M., Wang, F., Li, Y., Wang, L., Barrette, B., Bessho, T., Kwok, B. H., & Peng, A. (2020). Kinesin Kif2C in regulation of DNA double strand break dynamics and repair. ELife, 9, 1–22.

Zolota, V., Tzelepi, V., Piperigkou, Z., Kourea, H., Papakonstantinou, E., Argentou, M. I., & Karamanos, N. K. (2021). Epigenetic alterations in triple-negative breast cancer—the critical role of extracellular matrix. Cancers, 13(4), 1–17.

APPENDICES

Appendix I: List of 36 putative target genes of miR-4725

Gene	Ensembl Gene ID	Description
MFAP2	ENSG00000117122	microfibril associated protein 2
KIF2C	ENSG00000142945	kinesin family member 2C
MEX3A	ENSG00000254726	mex-3 RNA binding family member A
SLAMF8	ENSG00000158714	SLAM family member 8
CENPL	ENSG00000120334	centromere protein L
LAD1	ENSG00000159166	ladinin 1
SDC1	ENSG00000115884	syndecan 1
CENPA	ENSG00000115163	centromere protein A
QPCT	ENSG00000115828	glutaminyl-peptide cyclotransferase
NCAPH	ENSG00000121152	non-SMC condensin I complex subunit H
POC1A	ENSG00000164087	POC1 centriolar protein A
MCM2	ENSG00000073111	minichromosome maintenance complex component 2
SMC4	ENSG00000113810	structural maintenance of chromosomes 4
LAMP3	ENSG00000078081	lysosomal associated membrane protein 3
LRRC15	ENSG00000172061	leucine rich repeat containing 15
KCNK5	ENSG00000164626	potassium two pore domain channel subfamily K member 5
ANLN	ENSG00000011426	anillin actin binding protein
EIF4EBP1	ENSG00000187840	eukaryotic translation initiation factor 4E binding protein 1
RDH10	ENSG00000121039	retinol dehydrogenase 10
MELK	ENSG00000165304	maternal embryonic leucine zipper kinase
MKI67	ENSG00000148773	marker of proliferation Ki-67
CHEK1	ENSG00000149554	checkpoint kinase 1
NUSAP1	ENSG00000137804	nucleolar and spindle associated protein 1
SBK1	ENSG00000188322	SH3 domain binding kinase 1
CENPN	ENSG00000166451	centromere protein N
GINS2	ENSG00000131153	GINS complex subunit 2
FANCA	ENSG00000187741	FA complementation group A
DBNDD1	ENSG00000003249	dysbindin domain containing 1
LIMD2	ENSG00000136490	LIM domain containing 2
TK1	ENSG00000167900	thymidine kinase 1
SECTM1	ENSG00000141574	secreted and transmembrane 1
FAM83D	ENSG00000101447	family with sequence similarity 83 member D
CHAF1B	ENSG00000159259	chromatin assembly factor 1 subunit B
TMSB15B	ENSG00000158427	thymosin beta 15B
ASF1B	ENSG00000105011	anti-silencing function 1B histone chaperone
TCF19	ENSG00000137310	transcription factor 19

Appendix 2

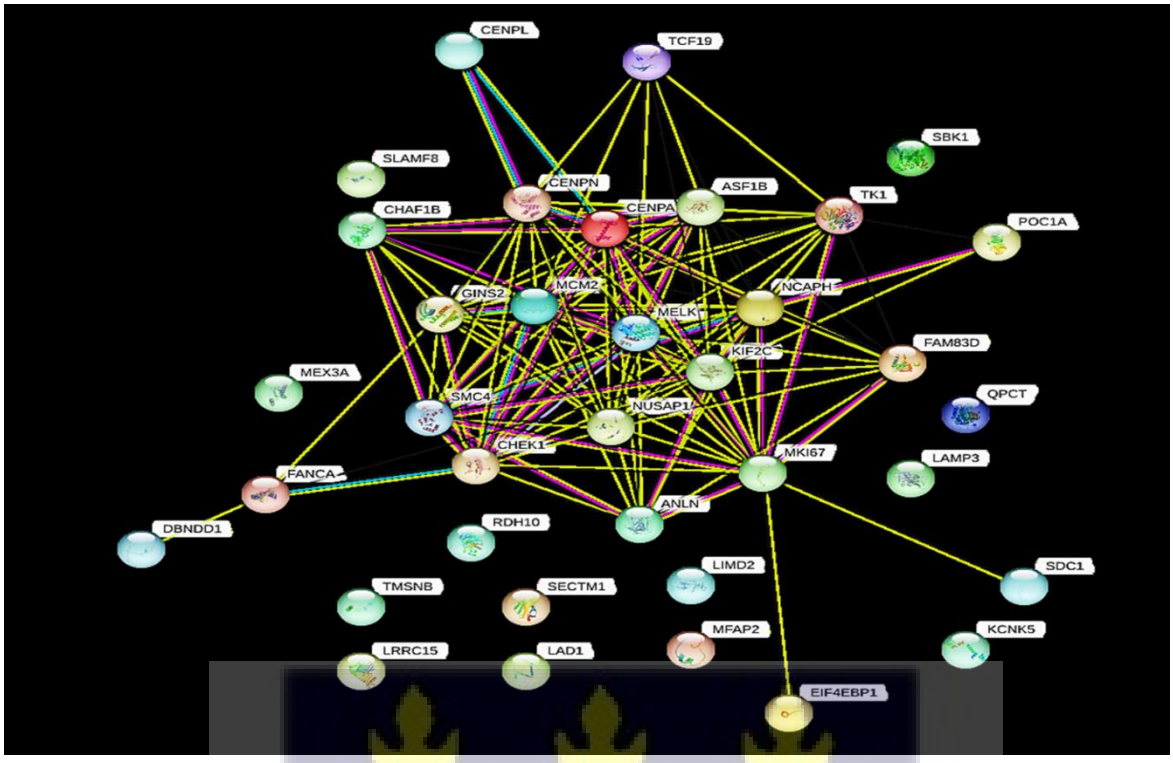


Figure 1: Network interactions between putative target genes. Most of the identified genes are found in to interact with each other. Generated with cytoscape.

Apendix 3: Expression of 5 selected gene target in breast cancer

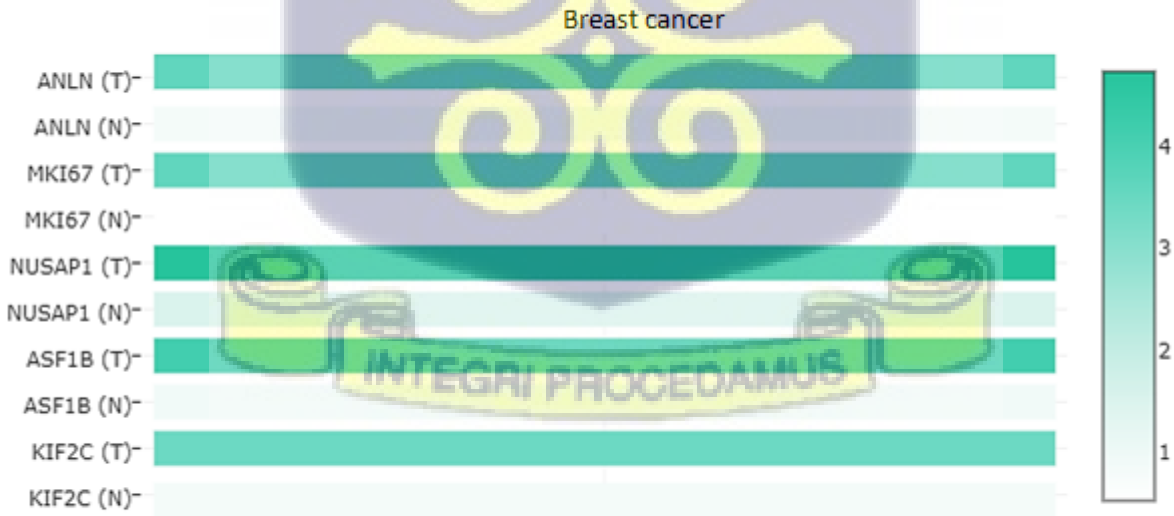


Figure 2 : differential expression 5 putative targets gene in tumour and normal match samples from TCGA database. All genes are upregulated in tumour but not in normal matched samples. Generated with GEPIA.

Appendix 4: Expression of miR-4725 in tissue samples and cell lines

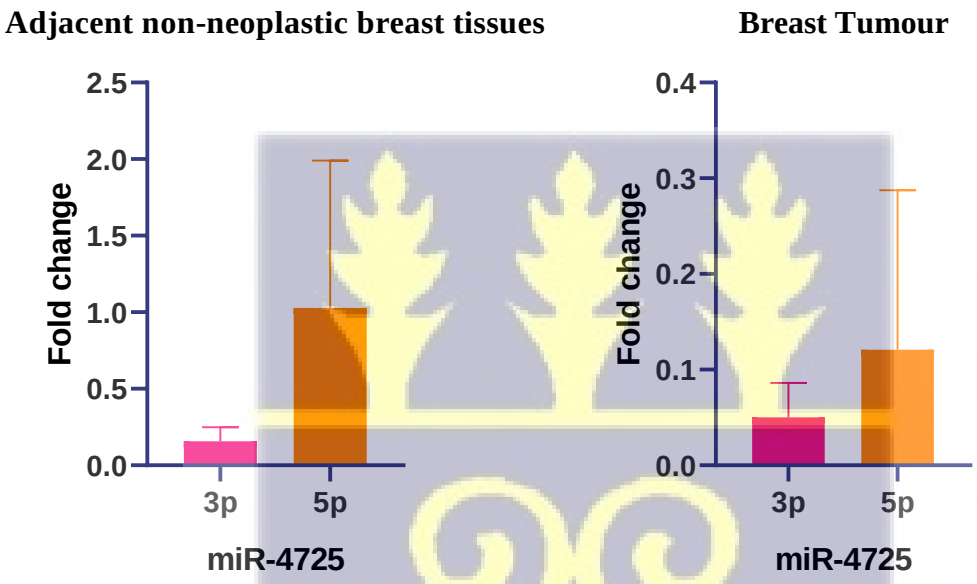


Figure 3a: expression analysis of miR4725 isotypes in Ghanaian tumour and adjacent non-neoplastic tissues (n=6).

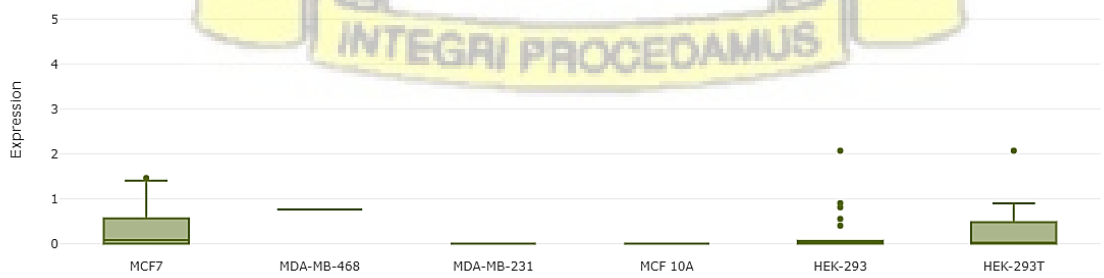


Figure 3b. Expression of miR-4725 in cell lines. Expression is generally low although variable. (generated with DIANA Tools-miTED)

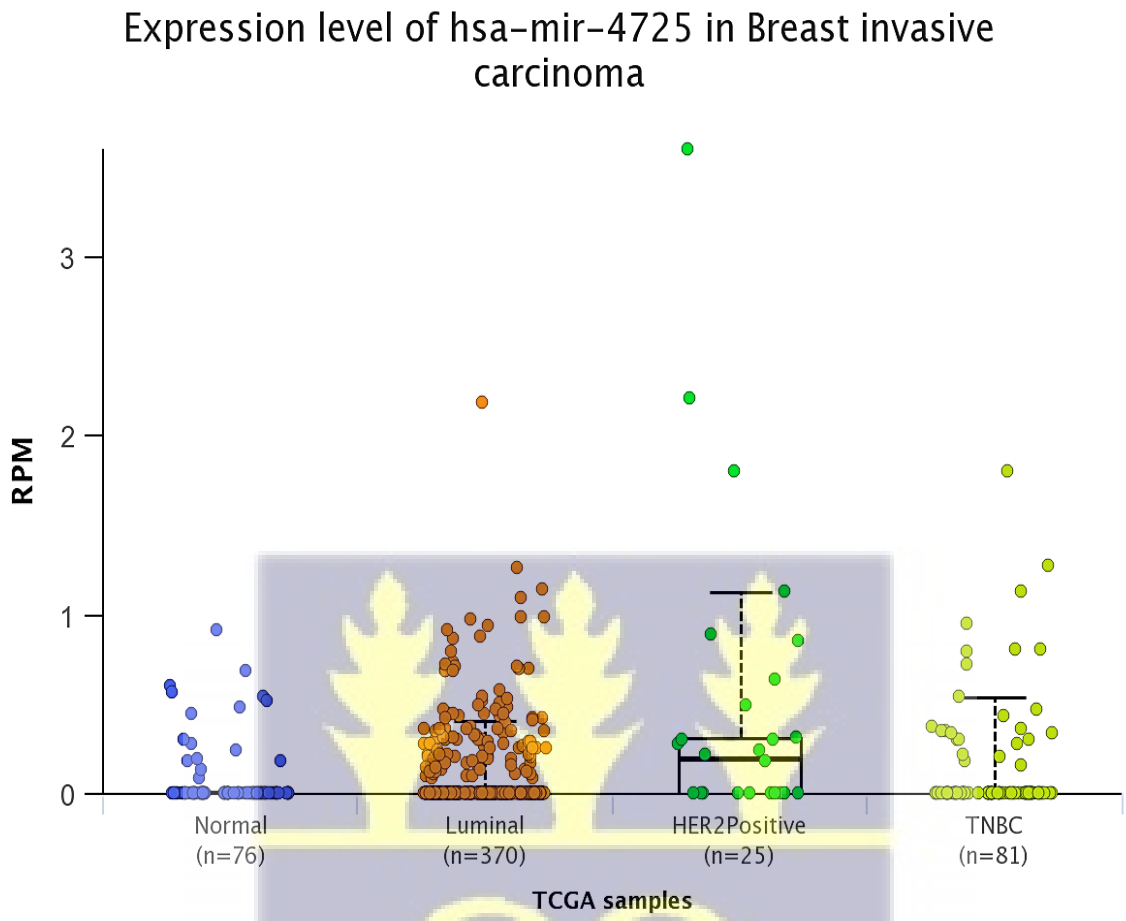


Figure 3c: Expression of miR4725 in breast cancer subtypes. Luminal and TNBC subtypes have a relatively lower mean expression. (Generated with UALCAN analysis software)

Appendix 5

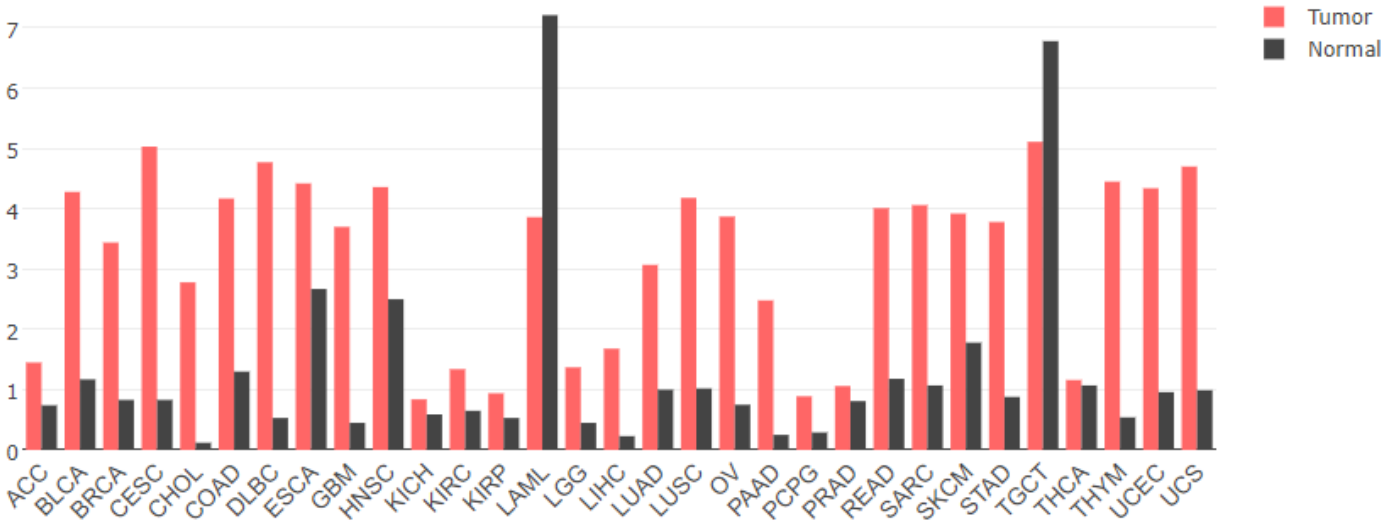


Figure 4: Expression of KIF2C in cancers. A comparison between matched normal and tumour samples. KIF2C is generally upregulated in all tumour samples.

Appendix 6 survival analysis of miR4725 expression in breast cancer

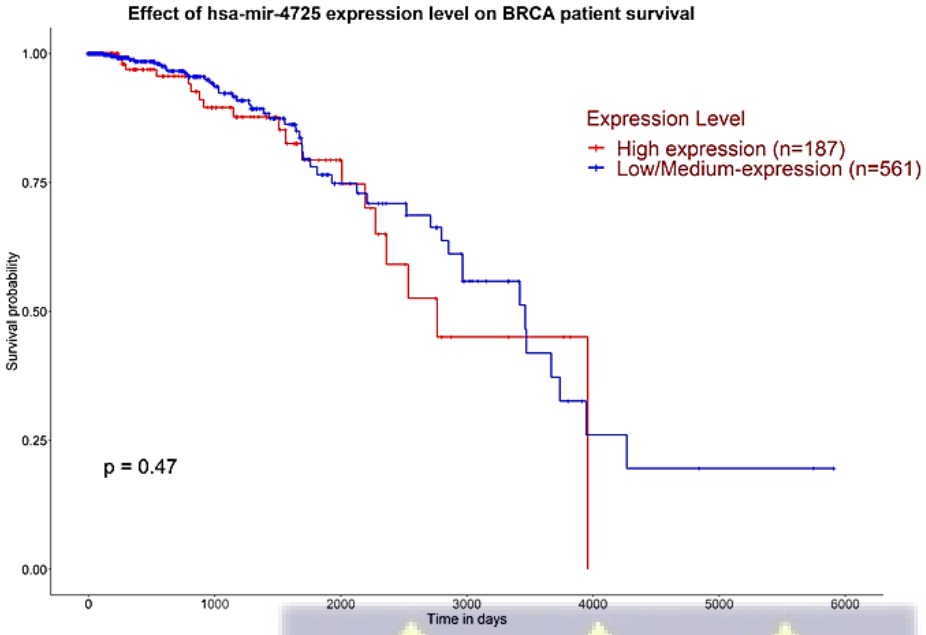


Figure 5a: Survival analysis of miR-4725 expression in breast cancer. High expression is associated with low survival although not statistically significant.

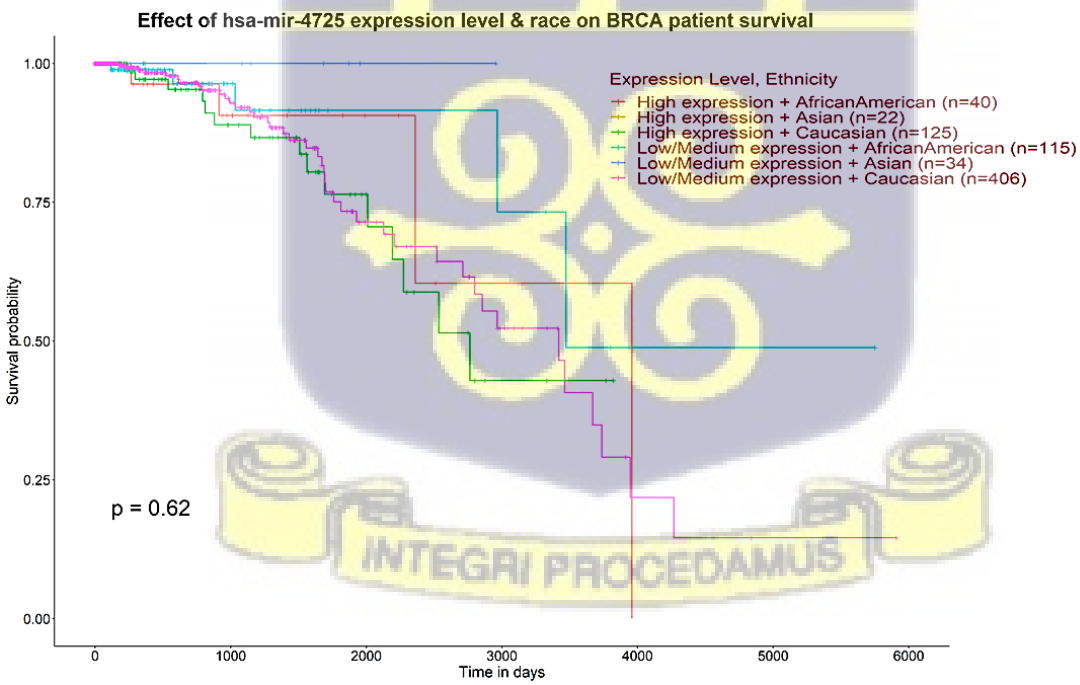


Figure 5b: Survival analysis of miR-4725 expression in breast cancer subtypes. High expression is associated with low survival in TNBC although not statistically significant.