

EXPRESSION ANALYSIS OF B4GALT5 & B4GALT6 IN BREAST CANCER

Submitted in partial fulfilment for the award of the
Bachelor of Technology Degree in Biotechnology

by

S NITISH KUMAR (40230032)



**DEPARTMENT OF BIOTECHNOLOGY
SCHOOL OF BIO & CHEMICAL ENGINEERING**

**SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY**

(DEEMED TO BE UNIVERSITY)

Accredited with Grade “A++” by NAAC

JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI - 600 119

APRIL - 2024



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

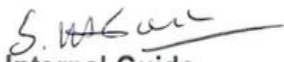
Accredited A++ Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

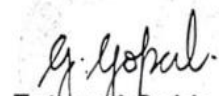
DEPARTMENT OF BIOTECHNOLOGY

BONAFIDE CERTIFICATE

This is to certify that this project report is the bonafide work of **S NITISH KUMAR** (40230032). Who has done the project work entitled "**EXPRESSION ANALYSIS OF B4GALT5 & B4GALT6 IN BREAST CANCER**" submitted for the degree of **BACHELOR OF TECHNOLOGY IN BIOTECHNOLOGY (BTech)**, under our supervision from October 2023 to march 2024.


Internal Guide

Dr.S. USHA NANDHINI


External Guide

Dr.G. GOPAL



Head of the Department
Dr. V. RAMESH KUMAR

Submitted for Viva voce Examination held on

Internal Examiner

External Examiner

DECLARATION

I hereby declare that the BTech thesis entitled “**EXPRESSION ANALYSIS OF B4GALT5 & B4GALT6 IN BREAST CANCER**” is the record of project work carried out by me, during the period from October 2023 to march 2024 under the guidance and supervision of **Dr. R.P. OVIYA**, Assistant professor, Department of biotechnology, Sathyabama institute of science and technology and **Dr.S.USHA NANDHINI**, Assistant professor, Department of biotechnology, Sathyabama institute of science and technology, is submitted in partial fulfillment of the requirement for the award of Bachelor of Technology degree in Biotechnology.

Place: Chennai

Date:

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I take this moment to bow to the supreme power, who is the meaning of the Vedas, the source of light, for the everything in my life.

I am pleased to acknowledge my sincere thanks to Board of Management of **SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. RAMESH KUMAR M.Phil.,Ph.D.**, Dean, School of Bio and chemical Engineering and Head of the Department, Dept. of Biotechnology for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr. OVIYA R.P**, Assistant professor, Department of biotechnology and **Dr. S.USHA NANDHINI**, Assistant professor, Department of biotechnology, Sathyabama institute of science and technology for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I would also like to acknowledge my sincere gratitude to Cancer Institute WIA, Molecular Oncology department for allowing me to participate and working in their laboratory facilities for my final year project and providing me their guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I am forever grateful to my mentor **Dr. G. GOPAL**, Professor, Department of Molecular Oncology, Cancer Institute WIA, without whom this work would never have seen fruition. He gave me the freedom to work at my own pace and allowed me to experiment with various ideas. He has always been an approachable guide and encouraged me to discuss research questions. These discussions have added considerable value to me as a future researcher.

I wish to express my thanks to all Teaching and Non-teaching staff members of the Department of Biotechnology who were helpful in many ways for the completion of the project.

ABSTRACT

Breast cancer is one of the most prevalent malignancies among women worldwide, and understanding the molecular mechanisms underlying its development and progression is crucial for improving diagnostic and therapeutic strategies. The β 1,4-galactosyltransferase (β 4GALT) family of enzymes plays a vital role in the biosynthesis of glycoconjugates, which are involved in various cellular processes, including cell adhesion, signalling, and proliferation.

This study aimed to investigate the expression patterns of B4GALT5 and B4GALT6 in breast cancer samples and cell lines, and to evaluate their potential as biomarkers or therapeutic targets. Quantitative real-time PCR (qRT-PCR) were performed on a cohort of breast cancer patient samples and adjacent non-cancerous tissues to assess gene expression levels of B4GALT5 and B4GALT6, respectively. Additionally, a panel of breast cancer cell lines representing different molecular subtypes was examined for their B4GALT5 and B4GALT6 expression levels.

Our results revealed a significant downregulation of β 4GALT5 and β 4GALT6 in breast cancer tissue samples compared to their normal counterparts. In additionally, lower expression levels of these genes were also observed in aggressive breast cancer cell lines, suggesting their potential involvement in tumor progression and metastasis. It was also confirmed with the bioinformatics analysis with help of various database.

This study provides valuable insights into the expression profiles of B4GALT5 and B4GALT6 in breast cancer and highlights their potential as biomarkers or therapeutic targets. Further investigations are warranted to elucidate the underlying mechanisms and explore the therapeutic potential of targeting these enzymes in breast cancer.

KEY WORDS: Breast cancer, β 4GALT5, β 4GALT6, Quantitative real-time PCR

TABLE OF CONTENTS

CHAPTER NO.	CONTENTS	PAGE NO.
	ABSTRACT	vi
	LIST OF TABLES	ix
	LIST OF FIGURES	x
	LIST OF ABBREVIATION	xiii
1	INTRODUCTION	1
	1.1 BREAST CANCER BIOLOGY AND THEIR MECHANISMS	7
	1.2 BREAST CANCER AND THEIR TYPES	9
	1.3 MOLECULAR SUBTYPES OF BC	11
	1.4 STAGING OF THE BC	13
	1.5 PREDECTIVE MARKERS OF BC	15
2	REVIEW OF LITERATURE	20
3	AIM AND OBJECTIVE	24
4	MATERIALS AND METHODS	25
	IN VITRO STUDY USING CELL LINE	
	4.1 CELL LINES AND TYPES OF CELL LINES	25
	4.2 CULTURING OF CELL LINES	27
	4.2.1 Mammalian cell culture	27
	4.2.2 Total Rna isolation	28
	4.2.3 Agarose gel electrophoresis	28
	4.2.4 Cdna synthesis	28
	4.2.5 Quantitative real time – polymerase chain reaction (QRT-PCR)	39
	STUDY IN TISSUE SAMPLES	
	4.3 BC SAMPLE COLLECTION	30
	4.4 BC SAMPLE PREPARATION	32
	4.4.1 Total Rna isolation	32
	4.4.2 Agarose gel electrophoresis	33
	4.4.3 Cdna synthesis	33
	4.4.4 Designing primers	34

	4.4.5 Quantitative real time – polymerase chain reaction (QRT-PCR)	34
	BIOINFORMATICS ANALYSIS	
	4.5 EXPRESSION ANALYSIS OF GENES	36
	4.5.1 GEPIA DATABASE	36
	4.5.2 GOBO DATABASE	36
	4.5.3 C-BIOPORTAL DATABASE	37
	4.5.4 COSOMIC DATABASE	37
5	RESULTS AND DISCUSSION	39
	5.1 IN VITRO STUDY USING CELL LINE	39
	5.2 STUDY IN TISSUE SAMPLES	44
	5.3 BIOINFORMATICS ANALYSIS	49
6	STATISTICAL ANALYSIS	67
7	SUMMARY AND CONCLUSION	71
	REFERENCES	72
	ANNEXURE	77

LIST OF TABLES

TABLE NUMBER	TITLE	PAGE NO
1.1	BC staging according to AJCC 7 th edition	14
4.1	List of cell lines used	27
4.2	Lists of primers used for QRT-PCR	34
5.1	Quantification value for cell lines	41
5.2	Fold change value for B4GALT5 & B4GALT6 GENES in Cell lines	43
5.3	Quantification value for breast cancer tissue samples (tumor & paired normal)	44
5.4	Fold change value for B4GALT5 & B4GALT6 GENES	46
6.1	CT value of B4GALT5	67
6.2	ANOVA test – B4GALT5	68
6.3	CT value of B4GALT6	69
6.4	ANOVA test – B4GALT6	70

LIST OF FIGURES

FIGURE NUMBER	TITLE	PAGE NO
1.1	The four main stages of cancer cell process	9
1.2	The types of breast cancer	11
1.3	Molecular types of breast cancer	13
1.4	The predictive marker of breast cancer	19
4.1	Tumor & paired normal breast sample collected from FS	32
5.1	HEK 293 cell lines	39
5.2	ZR751 cell lines	39
5.3	MCF7 cell lines	40
5.4	Cell culture technique	40
5.5	RNA isolation	40
5.6	HEK 293 RNA band	41
5.7	ZR751 RNA band	41
5.8	MCF7 RNA band	42
5.9	Raw data extracted from the QRT-PCR instrument	42
5.10	Calculation of fold change	42
5.11	Fold change comparison between the B4GALT5 & B4GALT 6 genes	43
5.12	RNA isolation	44
5.13	CDNA SYNTHESIS	45
5.14	Thermal cycle instrument	45
5.15	B4GALT5 & B4GALT6 band shows in the range of 350- 500 BP	46
5.16	Quantstudio 12 th K Realtime PCR	46
5.17	Fold change comparison between the B4GALT5 & B4GALT 6 genes	47
5.18	Comparison of B4GALT5 between tumor & paired normal	48

5.19	Comparison of B4GALT6 between tumor & paired normal	48
5.20	Box plot of breast cancer comparison between tumor & paired normal With B4GALT5 gene	49
5.21	Survival analysis of breast cancer comparison between tumor & paired normal with B4GALT5 gene	49
5.22	Box plot of breast cancer comparison between tumor & paired normal With B4GALT6 GENE	50
5.23	Survival analysis of breast cancer comparison between tumor & paired normal with B4GALT6 gene	50
5.24	Gene set analysis of B4GALT 5 gene in all tumors	51
5.25	Gene set analysis of B4GALT 5 gene in normal with tumor	52
5.26	Comparison of B4GALT 5 GENE in across all types of breast cancer	53
5.27	Comparison of B4GALT 5 gene in across all subtypes of breast cancer	53
5.28	Box plot of B4GALT 5 gene in across all subtypes of breast cancer	54
5.29	KaplanMeier Summary of B4GALT 5 gene for each cell lines	55
5.30	Gene set analysis of B4GALT 6 gene in all tumors	55
5.31	Gene set analysis of B4GALT 6 gene in normal with tumor	56
5.32	Comparison of B4GALT 6 GENE in across all types of breast cancer	57
5.33	Comparison of B4GALT 6 gene in across all subtypes of breast cancer	57
5.34	Box plot of B4GALT 6 gene in across all subtypes of breast cancer	58
5.35	KaplanMeier Summary of B4GALT 6 gene for each cell lines	59
5.36	Alteration of B4GALT 5 in breast cancer cell	59
5.37	Graph between the altered & unaltered group – Overall	60

	survival rate	
5.38	Proteogenomic Status of samples	60
5.39	Alteration of B4GALT 6 in breast cancer cell	61
5.40	Graph between the altered & unaltered group – Overall survival rate	62
5.41	Proteogenomic Status of samples	62
5.42	Gene view for B4GALT 5	63
5.43	Tissue distribution of mutations across the primary tissue types	64
5.44	Substitution mutations	64
5.45	Gene view for B4GALT 6	65
5.46	Tissue distribution of mutations across the primary tissue types	65
5.47	Substitution mutations	60
6.1	Linear regression line- B4GALT5	68
6.2	Linear regression line- B4GALT6	70

LIST OF ABBREVIATIONS

BC	- Breast cancer
B4GALT5	- beta-1,4-galactosyltransferase 5
B4GALT6	- beta-1,4-galactosyltransferase 6
SN2	- Substitution Nucleophilic Bimolecular
CSCs	- cancer stem cells
Cdna	- Complementary DNA
Gb3	- globotriaosylceramide
TNBC	- triple-negative breast cancer
ER	- hormone receptors- Estrogen
PR	- progesterone receptors
HER2	- human epidermal growth factor receptor 2
EGFR	- epidermal growth factor receptor
FISH	- fluorescent in-situ hybridization
CISH	- chromogenic in situ hybridization

CHAPTER 1

INTRODUCTION

Breast cancer remains one of the most prevalent and deadly malignancies among women worldwide. Despite advances in early detection and treatment strategies, the molecular mechanisms underlying breast cancer initiation, progression, and metastasis are not fully understood. Emerging evidence suggests that alterations in glycosylation pathways, which govern the synthesis and modification of glycoconjugates on cell surfaces and secreted proteins, play a crucial role in cancer development and progression.

Among the key enzymes involved in glycosylation are the beta-1,4-galactosyltransferase (B4GALT) family members, responsible for the biosynthesis of complex glycosphingolipids. Two members of this family, B4GALT5 and B4GALT6, have garnered significant attention due to their potential implications in breast cancer biology. These enzymes catalyze the transfer of galactose residues to specific glycosphingolipid precursors, thereby modulating the composition and structure of these essential membrane components.(Hennet, 2002)

Galactose is commonly found in several classes of glycoconjugates in both prokaryotes and eukaryotes. Various galactosyltransferase (GalT) enzymes catalyse the addition of galactose (Gal) in two anomeric configurations through α 1-2, α 1-3, α 1-4, α 1-6 or β 1-3, β 1-4 linkages following the standard reaction: $\text{UDP-Gal} + \text{acceptor} \rightarrow \text{Gal-acceptor} + \text{UDP}$ (Hennet, 2002)

The variety of galactosylation reactions significantly contributes to the tremendous diversity of oligosaccharide structures expressed by living organisms. In prokaryotes, GalTs are involved in the biosynthesis of the lipopolysaccharide O-antigen, capsular polysaccharides and pilin glycans.(Sato et al., 2001). In the yeast *Schizosaccharomyces pombe*, a GalT participates in the formation of cell wall galactomannan, and in plants several GalTs synthesize plastid galactolipids, galactosylated flavonols, galactomannan as well as complex asparagine-linked glycans. In keeping with the overall complexity of

higher eukaryotes, the vast diversity of galactosylated structures is paralleled by several GalT gene duplication events that give rise to several groups of enzymes with different acceptor specificities and distinct patterns of tissue expression.(Chen et al., 2006)

B4GALT FAMILY OF PROTEINS

The B4GALT enzymes belong to the glycosyltransferase family and share a common structural fold consisting of two $\beta/\alpha/\beta$ domains. The catalytic domain contains a conserved DXD motif that binds and activates the donor sugar nucleotide (UDP-galactose), while the stem region, which varies among different B4GALT enzymes, determines the specificity for the acceptor molecule. The catalytic mechanism involves an SN₂-like displacement reaction, where the activated galactose is transferred to the acceptor molecule. Each member of the B4GALT family exhibits a distinct substrate specificity, allowing for the synthesis of a diverse range of glycoconjugates, such as N-linked glycoproteins (B4GALT1), glycosphingolipids like lactosylceramide and Gb3 (B4GALT5 and B4GALT6, respectively), glycolipids (B4GALT3), and proteoglycans (B4GALT4).(Lo, 1998)

B4GALT enzymes are widely expressed in various tissues, but some members exhibit tissue-specific or developmental stage-specific expression patterns, like B4GALT3 being predominantly expressed in the brain and having a role in neuronal development and function. Most B4GALT enzymes are localized in the Golgi apparatus, participating in the later stages of glycoconjugate biosynthesis, but some are also found in the endoplasmic reticulum or plasma membrane. Proper glycosylation catalyzed by B4GALT enzymes is crucial for the folding, stability, and function of many glycoproteins and glycolipids involved in various biological processes, including cell adhesion, migration, signaling, immune recognition, and embryonic development. Dysregulation or mutations in B4GALT genes can lead to genetic disorders like congenital disorders of glycosylation (CDGs), affecting multiple organ systems and causing developmental delays, neurological abnormalities, and immune dysfunction.(Miyazaki et al., 2019)

ROLE OF B4GALT FAMILY OF PROTEIN IN CANCER CELL METABOLISM

The B4GALT family of proteins significantly impacts cancer cell metabolism and progression through their involvement in the biosynthesis and regulation of glycosphingolipids and glycoproteins. Enzymes like B4GALT5 and B4GALT6 catalyze the synthesis of specific glycosphingolipids, such as lactosylceramide and globotriaosylceramide (Gb3), which are frequently overexpressed in various cancers and contribute to enhanced cancer cell proliferation, survival, metastasis, and drug resistance, with Gb3 particularly promoting cancer cell invasion and metastasis by modulating signaling pathways involved in cell migration and epithelial-mesenchymal transition (EMT)(Pouncey et al., 1991). Additionally, B4GALT1, the major enzyme responsible for the biosynthesis of N-linked glycoproteins, plays a crucial role in the aberrant glycosylation of cancer-associated glycoproteins like mucins, receptor tyrosine kinases, and adhesion molecules, which can modulate their activity, localization, and interactions, promoting cancer progression and metastasis. Moreover, B4GALT enzymes contribute to the metabolic reprogramming of cancer cells by modulating the biosynthesis of glycoconjugates involved in various metabolic pathways, such as glycolysis, pentose phosphate pathway, and amino acid metabolism, while also participating in the crosstalk between glycosylation and metabolic pathways, where metabolic intermediates can influence glycosylation patterns, and vice versa. Consequently, B4GALT enzymes and their associated glycosylation pathways have emerged as promising therapeutic targets in cancer, with inhibition or modulation of their activity offering potential strategies for disrupting cancer cell growth, metastasis, and drug resistance.(Youakim & Shur, 1993)

B4GALT5

The molecular cloning of the b4GalT5 cDNA relied on polymerase chain reaction (PCR) amplification using degenerate primers designed after regions conserved between b4GalT1 and a structurally related b1-4 GlcNAc-transferase from the snail *Lymnaea stagnalis*. The b4GalT5 gene maps to chromosome 20q13. Unlike b4GalT1 to 4, the b4GalT5 gene contains 9 exons, suggesting a

more distant evolutionary relationship to the previous members of the b4GalT gene family(Kumagai et al., 2010). Also, the b4GalT5 protein shows only about 35% identity to b4GalT1 to 4. The human b4GalT5 cDNA was expressed as a soluble protein A-fusion construct. The activity toward GlcNAc was quite low. The apparent Km for the GlcNAc acceptor was 33 mM compared with a Km of 0.6 mM for b4GalT1.(Sato & Furukawa, 2007) The addition of α -lactalbumin had no inhibitory effect on the N-acetyllactosamine synthase activity of b4GalT5. Further characterizations have shown that b4GalT5 participates in the biosynthesis of N-glycans, whereas a preference toward O-linked acceptors has been noted by others. The biological relevance of b4GalT5 is currently unknown.(Shirane et al., 1999)

B4GALT5, the enzyme responsible for catalyzing the biosynthesis of lactosylceramide, a precursor for more complex glycosphingolipids, plays a critical pathological role in various diseases, particularly cancer. Its upregulation leads to overproduction of lactosylceramide, which has been associated with enhanced cancer cell proliferation, survival, metastasis, and drug resistance in breast, colon, ovarian, and lung cancers.(Sato & Furukawa, 2004) Moreover, B4GALT5 expression is elevated in cancer stem cells (CSCs), maintaining their stem cell-like properties and contributing to tumor recurrence and metastasis. This enzyme and its product, lactosylceramide, regulate signaling pathways involved in cancer progression, such as PI3K/Akt, MAPK, and NF- κ B, and influence the tumor microenvironment, promoting angiogenesis and immune evasion.(Chatterjee et al., 2019) Consequently, B4GALT5 has emerged as an attractive therapeutic target, with inhibition strategies showing promising results in reducing cancer cell proliferation, metastasis, and CSC properties in preclinical studies, highlighting its potential for combination therapies. Additionally, dysregulation of B4GALT5 contributes to insulin resistance and metabolic disorders.(Sato et al., 2000)

B4GALT6

The synthesis of the glycolipid lactosylceramide requires the action of a b4GalT activity. Recombinant b4GalT1 can transfer Gal to Glc-ceramide, though

to low efficiency compared with its activity toward GlcNAc acceptors. To identify a better catalyst for the lactosylceramide synthase reaction, Nomura et al. have purified the activity from rat brain, which led to the cloning of a novel cDNA called b4GalT6.(Sato et al., 2001) The corresponding gene is localized on chromosome 18 at position q11. Its genomic organization, including 9 exons, is like that of b4GalT5. The identity at the amino acid level was also highest with b4GalT5, reaching 68%. A comparison of the lactosylceramide synthase activity between the known b4GalT enzymes was not included in the primary description of b4GalT6, but the subsequent description of the Pro5Lec20 Chinese hamster ovary (CHO) cell deficient for b4GalT6 showed that the loss of this enzyme has no effect on glycolipid profiles, thus indicating that other b4GalTs can also function as lactosylceramide synthase in cells. The characterization of N-glycans in Pro5Lec20 CHO cells demonstrated that b4GalT6 also accounts for N-glycan biosynthesis, a finding that has been confirmed by the recent characterization of recombinant b4GalT6 activity. No data exist about the roles of b4GalT6 in animal development and physiology.(Guo et al., 2001)

B4GALT6, the enzyme catalyzing the biosynthesis of globotriaosylceramide (Gb3), plays a critical pathological role in cancer progression and certain infectious diseases. Its overexpression leads to Gb3 accumulation, which has been associated with increased cancer cell proliferation, survival, invasion, and metastasis in breast, colon, ovarian, and kidney cancers by modulating signaling pathways like Src and PI3K/Akt, promoting cancer cell migration and epithelial-mesenchymal transition (EMT).(Hennet, 2002) Additionally, Gb3 serves as a receptor for Shiga toxins produced by pathogenic bacteria like Shiga toxin-producing Escherichia coli, leading to inhibition of protein synthesis and cell death, causing diseases like hemolytic uremic syndrome (HUS) and hemorrhagic colitis. B4GALT6 and Gb3 also regulate cellular processes like proliferation, apoptosis, and cell cycle, contributing to various diseases' pathogenesis. Consequently, B4GALT6 has emerged as a potential therapeutic target, with inhibition strategies showing promising results in reducing cancer progression and mitigating Shiga toxin effects in preclinical studies, highlighting its potential for combination therapies.

However, further research is needed to develop selective and effective inhibitors, as dysregulation of Gb3 biosynthesis is also implicated in metabolic disorders and neurodegenerative diseases.(Nihei et al., 2023)

THE EXPRESSION B4GALT 5 & B4GALT 6 STUDY IN BREAST CANCER AS A TOOL TO IDENTIFY

The expression levels of B4GALT5 and B4GALT6 in breast cancer have emerged as valuable tools for identifying and characterizing breast cancer subtypes, predicting prognosis and therapeutic response, and serving as potential therapeutic targets. These enzymes exhibit varying expression patterns among different breast cancer subtypes, with higher levels observed in more aggressive subtypes like HER2-positive and triple-negative breast cancer (TNBC), which are associated with poorer prognosis and increased metastatic potential(Sato et al., 2001). Importantly, B4GALT5 and B4GALT6 play crucial roles in the maintenance and regulation of breast cancer stem cells (BCSCs), contributing to their self-renewal capacity and tumorigenic potential. High expression levels of these enzymes in breast cancer tumors have been linked to poorer prognosis, shorter disease-free survival, increased risk of metastasis, and resistance to certain chemotherapeutic agents and targeted therapies, suggesting their potential as prognostic and predictive biomarkers.(Borsig et al., 1997) Consequently, B4GALT5 and B4GALT6 have emerged as attractive therapeutic targets, with inhibition strategies showing promising results in reducing breast cancer cell proliferation, invasion, and metastasis in preclinical studies, highlighting their potential for combination therapies.(Ma et al., 1997) Furthermore, the expression levels of these enzymes, in combination with other molecular markers, have been proposed as part of diagnostic and prognostic panels for breast cancer, aiming to improve subtyping accuracy, risk stratification, and personalized treatment selection

THE POTENTIAL OF B4GALT 5 & B4GALT 6 AS A NOVAL DIAGNOSTIC MARKER IN BREAST CANCER

B4GALT5 and B4GALT6 exhibit differential expression patterns among breast cancer subtypes, with higher levels observed in aggressive subtypes like

HER2-positive and triple-negative breast cancer (TNBC), aiding in accurate subtyping and guiding treatment decisions. Their elevated expression correlates with poor prognostic features such as higher tumor grade, advanced stage, lymph node metastasis, and reduced overall survival, suggesting potential as valuable prognostic markers for identifying high-risk patients (Paone et al., 1980). Additionally, their expression levels are elevated in early breast cancer stages, making them candidates for early detection and screening, while monitoring their levels during treatment could assess response and detect recurrence or metastasis. (Echeverri & Orozco, 2022) Incorporating B4GALT5 and B4GALT6 expression data into diagnostic panels or multimarker approaches, combined with established biomarkers like hormone receptor status and gene expression profiles, could improve diagnostic accuracy and enable personalized strategies. Moreover, detecting their expression or associated glycosylation patterns in circulating tumor cells or exosomes through liquid biopsy could facilitate non-invasive, real-time monitoring of breast cancer progression and personalized treatment adjustments. (Odunsi et al., 2002)

1.1 BREAST CANCER BIOLOGY AND THEIR MECHANISMS

Breast cancer is a multifaceted disease characterized by the uncontrolled proliferation of abnormal cells within the breast tissue. It is one of the most prevalent and devastating forms of cancer affecting women worldwide. The development and progression of breast cancer involves intricate biological mechanisms that span various cellular and molecular levels. At the genetic level, alterations in critical genes, such as tumor suppressor genes (e.g., BRCA1, BRCA2, and p53) and oncogenes (e.g., HER2 and c-Myc), play a pivotal role in driving the malignant transformation of cells. These genetic aberrations can lead to dysregulation of key cellular processes, including uncontrolled cell growth, evasion of apoptosis, and sustained proliferative signaling. Hormonal influences, particularly those mediated by estrogen and progesterone receptors, are critically involved in the development and progression of certain types of breast cancer. Hormonal signaling pathways can promote cell proliferation, survival, and tumor growth, making them attractive therapeutic targets.

Growth factor signaling pathways, such as those involving epidermal growth factor receptor (EGFR) and human epidermal growth factor receptor 2 (HER2), are often dysregulated in breast cancer. Aberrant activation of these pathways can lead to sustained proliferative signaling and enhanced cell survival, contributing to cancer progression. Angiogenesis, the process of forming new blood vessels, plays a crucial role in supporting the growth and metastasis of breast tumors. Vascular endothelial growth factor (VEGF) and other angiogenic factors promote the formation of new blood vessels, facilitating the delivery of nutrients and oxygen to the rapidly dividing cancer cells. The epithelial-mesenchymal transition (EMT) is a pivotal process that enables cancer cells to acquire invasive and migratory properties, facilitating metastasis. Transcription factors (e.g., SNAIL and TWIST) and signaling pathways (e.g., TGF- β and Wnt) regulate the EMT process, allowing cancer cells to disseminate to distant sites. The tumor microenvironment, comprising various cell types such as stromal cells, immune cells, and the extracellular matrix, plays a crucial role in shaping cancer progression and metastasis.

1.1.1 Here is an Overview of the Four Main Stages

Stage 0 (In Situ): Cancer cells are limited to the place of origin at this early stage and have not spread to neighboring tissues. Because the aberrant cells are present but not yet invasive, this is sometimes referred to as carcinoma in situ.

Stages I and II (Localized): The cancer is regarded as localized during these phases. The tumor is small and limited to its original location in Stage I. Stage II refers to a bigger tumor that has not yet migrated to distant places but may include adjacent tissues or lymph nodes.

Stage III (Regional): By the time the cancer reaches this intermediate stage, it has usually progressed to adjacent lymph nodes and may even touch adjacent tissues. There is a higher chance of more metastases because the tumor is larger.

The most advanced stage, known as **Stage IV (Distant Metastasis)**, occurs when cancer has spread to distant organs or tissues. Since stage IV tumors are frequently more difficult to cure, palliative care may become more important to control.

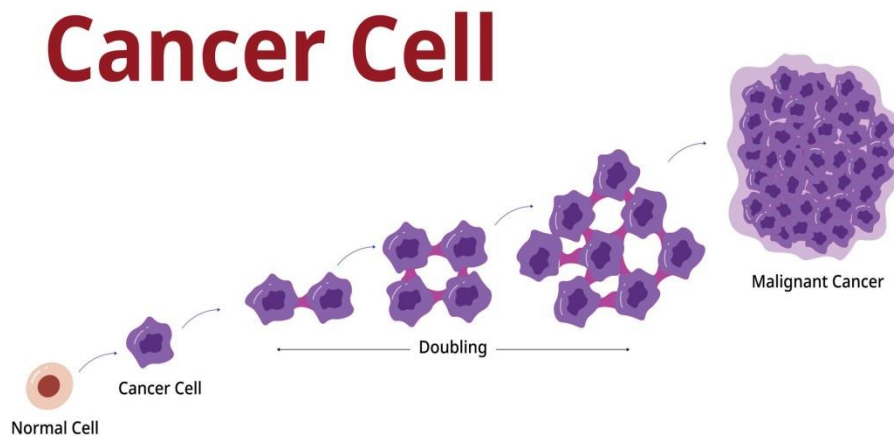


Fig. 1.1: The Four Main Stages of Cancer Cell Process

1.2 BREAST CANCER AND THEIR TYPES

Breast cancer is a common and complex type of cancer that primarily affects women but can also affect men. It has now surpassed lung cancer as the leading cause of global cancer incidence in 2020, with an estimated 2.3 million new cases, representing 11.7% of all cancer cases. Epidemiological studies have shown that the global burden of BC is expected to cross almost 2 million by the year 2030. In India, the incidence has increased significantly, almost by 50%, between 1965 and 1985. It starts with irregular cell growth in breast tissue, leading to the formation of tumors. These tumors can be benign (non-cancerous) or malignant (cancerous). Breast cancers are classified into different molecular subtypes, including Luminal A, Luminal B, HER2-Enriched, and Triple-Negative Breast Cancer (TNBC). Each subtype has unique features that influence treatment decisions. Early detection methods, like screening, are crucial, along

with newer targeted therapies and ongoing research, to improve the outcomes for individuals battling breast cancer.

Breast cancer encompasses various types, each characterized by distinct features and molecular profiles. The main types include

1 *Ductal Carcinoma In- Situ (DCIS):*

Ductal Carcinoma In- Situ (DCIS) is a type of breast cancer that does not penetrate the basement membrane of tubules. During the early phase, malignant cells are located solely inside the duct and do not invade the adjacent tissue or disseminate to remote parts of the organism. DCIS often presents itself in mammographic findings for breast screening purposes; it is pre-cancerous. Though DCIS itself is noninvasive, if left untreated, it can eventually develop into invasive breast cancer. Treatment for DCIS may involve lumpectomy or mastectomy surgery as well as radiation therapy. The purpose is to prevent invasion and make sure there are no negative effects on the person who has it.

2 *Invasive Ductal Carcinoma (IDC):*

The most occurring sort of breast cancer is Invasive Ductal Carcinoma (IDC) which initiates in the milk ducts and breaks out from them into neighboring tissues. IDC has a propensity to spread to other parts of the body, unlike non-invasive conditions such as Ductal Carcinoma in Situ (DCIS). It is often diagnosed through physical examination and medical imaging. The treatment options usually include surgery like lumpectomy or mastectomy followed by additional therapies including radiation, chemotherapy, and targeted therapy based on specific molecular features. Those with invasive ductal carcinoma need early detection for successful management and improved outcomes.

3 *Invasive Lobular Carcinoma (ILC):*

Invasive Lobular Carcinoma (ILC) is a specific form of breast cancer that starts in the milk-producing glands. Unlike the more common type, Invasive Ductal Carcinoma, ILC frequently doesn't form a clear lump and might present as a subtle thickening of breast tissue. A unique feature of ILC is its single-file

growth pattern, making it difficult to detect using typical imaging techniques. Treatment typically involves surgery, radiation therapy, chemotherapy, and hormone therapy. Early detection of ILC is crucial for successful management and a favorable outlook for individuals diagnosed with this type of breast cancer.

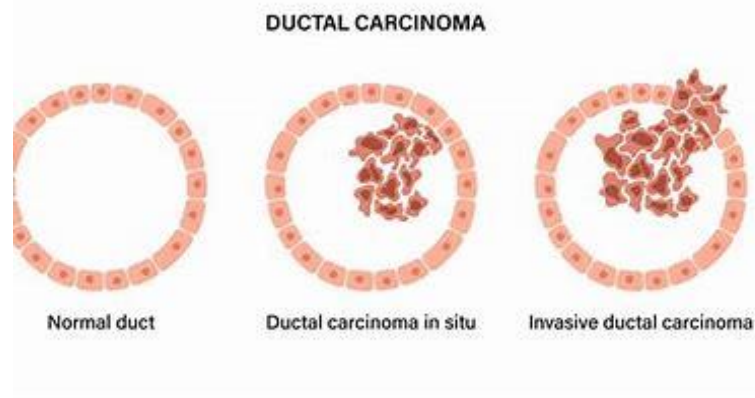


Fig. 1.2: The Types of Breast Cancer

1.3 MOLECULAR SUBTYPES OF BC

BC is a heterogeneous disease both at the clinical as well as molecular levels. This molecular subtype of breast cancer (BC) are classifications based on the distinct molecular characteristics and gene expression patterns observed in tumors. The classification of BC based on molecular subtypes are constantly evolving and the exact definition of it is still an area of active research.

1.3.1 LUMINAL SUBTYPE

These form most BC cases, more than 60% of the cases. They are named so because of their molecular profile being like luminal breast epithelium. The hormone receptors- estrogen and progesterone receptors (ER and PR) are overexpressed and represent the most diverse subtype. It is further divided into luminal A and luminal B subtypes; The Luminal A subtype of breast cancer is characterized by the presence of hormone receptors (ER and/or PR) and the absence of HER2 overexpression. These tumors exhibit low proliferative activity, contributing to a favorable prognosis, and are often diagnosed in

postmenopausal women. The Luminal B subtype of breast cancer is characterized by the presence of hormone receptors (ER and/or PR), with variability in HER2 status (positive or negative). Luminal B tumors exhibit higher proliferative activity compared to Luminal A, influencing their prognosis, which may be less favorable. Treatment strategies for Luminal B involve a combination of endocrine therapy, like Luminal A, and HER2-targeted therapies if HER2-positive.

1.3.2 HER2 ENRICHED SUBTYPE

The HER2-Enriched subtype of breast cancer is characterized by its negativity for hormone receptors (ER and PR) and positive status for the human epidermal growth factor receptor 2 (HER2), resulting in high proliferative activity and aggressive behaviour. Known for its potential for rapid tumor growth, HER2-Enriched breast cancer has a poorer prognosis compared to some subtypes. Treatment strategies primarily involve HER2-targeted therapies, such as trastuzumab, pertuzumab, and ado-trastuzumab emtansine, aiming to inhibit the overactive HER2 signalling pathway.

1.3.3 TRIPLE-NEGATIVE BREAST CANCER (TNBC):

Triple-Negative Breast Cancer (TNBC) is a type of breast cancer that lacks the presence of Estrogen Receptor (ER), Progesterone Receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression, causing it to be more aggressive and grow rapidly. TNBC does not respond to hormone-blocking or HER2-targeted therapies, making chemotherapy the main treatment. It has different subtypes, such as basal-like, immunomodulatory, mesenchymal, and luminal androgen receptor, contributing to its heterogeneity. Overall, due to limited targeted therapies, TNBC comes with a poorer prognosis compared to other types of breast cancer.

Molecular types of breast cancer

	ER	PR	HER2
Luminal A	+	+	-
Luminal B	+	-	+
HER2 positive	-	-	+
Triple negative	-	-	-

October is breast cancer awareness month

08 Oct 2017 @kegill

Fig. 1.3: Molecular Types of Breast Cancer

1.4 STAGING OF THE BC

THE TNM STAGING OF BC

BC is categorized into TNM stages, refers to the primary breast tumor dimensions, regional nodal status and presence or absence of distant metastasis. TNM staging is done based on the American joint committee on cancer (AJCC) system. Anatomically, BC is grouped into stages 0 to IV.

1.4.1 LOCALLY ADVANCED BC

Locally advanced breast cancer (LABC) is the most advanced, non-metastatic stage of the disease. It is heterogenous in clinical presentation and biological behavior. In India, almost 60% of the cases are locally advanced at presentation. LABC includes stage IIB and III BC. LABC patients are usually presented with large breast mass, pain, oedema, redness, nipple retraction, skin dimpling (peau d'orange), diffuse induration, mostly with clinically palpable axillary and/ or supraclavicular adenopathy and occasionally ulceration of the overlying skin. LABC has a higher risk for recurrence and lower overall survival (OS) in comparison with the early BC. According to Surveillance, Epidemiology and End Results (SEER) prediction, the 5-y OS of LABC is around 85%, against the early BC where it is around 95 %. The first step in the management of the LABC involves histologic diagnosis and determination of hormone receptor status.

TABLE1.1: BC Staging According to AJCC 7th Edition.

Anatomical stage	TNM	Description
Stage 0	tisN0M0	Carcinoma in-situ
Stage I A	T1N0M0	Tumor ≤20 mm and no regional lymph node metastases
Stage I B	T0N1miM0 T1N1miM0	Tumor ≤20 mm and cancer cells found in lymph nodes
Stage II A	T0N1M0 T1N1M0 T2N0M0	Tumor < 20 mm and metastases to the movable ipsilateral level I, II axillary lymph node(s). Tumor 20-50 mm and no regional lymph node metastases
Stage II B	T2N1M0 T3N0M0	Tumor 20-50 mm and metastases to the movable ipsilateral level I, II axillary lymph node(s). Tumor> 50 mm and no regional lymph node metastases
Stage IIIA	T0N2M0 T1N2M0 T2N2M0 T3N1M0 T3N2M0	Tumor < 50 mm and metastases in ipsilateral level I, II axillary lymph nodes that are clinically fixed or matted; or in clinically detected* ipsilateral internal mammary nodes in the absence of clinically evident axillary lymph node metastases. Tumor> 50 mm and spread to either one of the above
Stage IIIB	T4N0M0 T4N1M0 T4N2M0	Tumor of any size with direct extension to the chest wall and/or to the skin (ulceration or skin nodules) and metastases in ipsilateral level I, II axillary lymph nodes that are clinically fixed or matted; or in clinically detected* ipsilateral internal mammary nodes in the absence of clinically evident axillary lymph node metastases
Stage IIIC	Any T	Tumor of any size and metastases in ipsilateral

	N3M0	infraclavicular (level III axillary) lymph node(s) with or without level I, II axillary lymph node involvement; or in clinically detected* ipsilateral internal mammary lymph node(s) with clinically evident level I, II axillary lymph node metastases; or metastases in ipsilateral supraclavicular lymph node(s) with or without axillary or internal mammary lymph node involvement
Stage IV	Any T any N M0	Tumor of any size, with distant metastasis to organs like lung, liver, brain or bones.

1.5 PREDECTIVE MARKERS OF BC

According to the WHO, "a biomarker is any substance, structure or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease. A predictive marker is a biomarker that predicts the response to therapeutic interventions. Identification of a reliable marker for predicting the PCR helps in determining the probabilities of patients responding to the therapy and hence a step towards personalized treatment. Apart from that, predictive markers also aid in identifying optimum drug dose and envisaging toxicity and thus greatly improving the patients' quality of life, Patient's age, menstrual status, tumor size, type and grade. lymph node status is some of the existing predictive markers, which are part of the studies with varying conclusions. Hence it is not very clear as to whether these factors serve as individual predictive markers for PCR.

1.5.1 HORMONE RECEPTOR STATUS (ER, PR)

Hormonal receptors, ER and PR are intracellular receptors. Estrogen and progesterone, being steroid hormones easily enter the cell and bind their respective receptors. ER consists of two subtypes- $ER\alpha$ and $ER\beta$, former is the primary ER in mammary glands and overexpression is seen in almost 70% of the BC cases. $ER\alpha$ is a ligand- dependent transcription factor, upon binding to

estrogen, it translocates into the nucleus and brings about the transcription of genes like cyclin D1 and c-myc.

Like the ER, PR also has two isoforms- PRA and PRB, but with almost equal expression in normal breasts. Breast tumorigenesis involves a lack of the coordinated expression of PRA and PRB, with a high PRA/PRB ratio associated with poor prognosis. The mechanism of progesterone signaling through PR is similar to that of ER. It is a mediator of stemness and expansion in mammary glands. ER, and PR status act as both prognostic as well as predictive markers. They predict the response of the tumor to endocrine therapy (selective estrogen modulators and aromatase inhibitors). As per the latest the American Society of Clinical Oncology/College of American Pathologists guidelines, it is recommended that ER and PR are tested by standard IHC. ER-negative tumors are found to be more chemosensitive and positive ER/PR results in poor response to NAT. The inter-laboratory reproducibility of IHC based evaluation of ER PR status is less with variability of almost 30%. There also seems to be inconsistency between the ER/PR status between core needle biopsy and surgical specimen in some cases. This can be attributed to the heterogeneity of the tumor or the experimental

1.5.2 HER2 STATUS

HER2 (encoded by Her2 or c-erbB-2 or neu) is a tyrosine kinase of the epidermal growth factor receptor (EGFR) family. HER2 is found to be overexpressed in 15-25% of breast cancer patients, predominantly due to gene amplification. HER2 is commonly evaluated using IHC and fluorescent in-situ hybridization (FISH) or chromogenic in situ hybridization (CISH). IHC semi-quantitatively measures the protein levels, whereas the FISH and CISH quantify HER2 gene copy number. As with the ER measurement, HER2 IHC also has less reproducibility (24). A meta-analysis revealed that 13% of cancers that were HER2 positive in the primary tumor were found to be negative at the metastatic site. HER2 is more successful as a predictive marker than a prognostic marker. HER2 positive tumors are generally of high grade, high proliferation rate, ER-negative and poor prognosis. HER2 overexpression seems to influence the

response to chemotherapeutics (anthracycline-based and a study reported 5-FU, methotrexate, cyclophosphamide response), anti-estrogen (tamoxifen) and anti-HER2 therapy (trastuzumab, lapatinib, pertuzumab and trastuzumab emtansine). HER2 positive BC showed comparative resistance to tamoxifen, though ER/PR positive, and in such cases aromatase inhibitors are preferred. HER2 positive BC, showed 64% PCR when treated with a combination of trastuzumab and chemotherapy. Tumors with activating mutations in phosphatidylinositol 3-kinase catalytic subunit (PIK3CA) and Her2 showed better PCR rates for NAT with docetaxel, carboplatin, trastuzumab, lapatinib.

1.5.3 Ki-67 EXPRESSION

Proliferation markers like Ki67, topoisomerase IIa (Topo-II), proliferating cell nuclear antigen, cyclins and mitosins have been studied for their role as predictive markers. Ki67 is a non-histone protein of the nucleus and is expressed during the cell cycle, with DNA binding functions. Ki67, which is used in routine clinical practice as a prognostic marker has been shown to have failed as a predictive factor, owing to the heterogeneity of breast tumors. Assessment of Ki67 by immunohistochemistry (IHC) in the nucleus is taken as the proliferation index. Though limited, pieces of evidence revealed high Ki67 to be associated with clinical and PCR to NAT. Similar results are also indicated by meta-analysis where high Ki67 correlated with higher PCR rates, but poor prognosis. Low Ki67% associated with PCR when treated with pre-operative tamoxifen.

Topoisomerases are involved in the modification of the DNA topology, facilitating replication and transcription. Topo-II generates a transient double-strand break in DNA to release the torsional hindrance during replication. Owing to its importance in cellular processes, targeting them is used as a strategy to treat cancer, for example, anthracyclines. Generally, the Topo-II expression negatively correlated with prognosis. Co-amplification of Her2 and Topo-II positively correlated with favorable response to anthracycline-based NAT in the primary tumor. Two studies reported contradictory results in Topo-II association with therapeutic response. One reports Topo-II levels to be negatively correlated

with PCR in NAT with docetaxel and epirubicin. The other study with TNBC also treated neoadjuvant taxane and anthracycline, the Topo-II overexpression independently predicts PCR and a favorable DFS.

1.5.4 CYTOKERATIN 5/6

Cytokeratin's (CKs) are a group of intermediate filament proteins found in the cytoskeleton of epithelial cells. They play a structural role, providing support and maintaining the integrity of epithelial tissues, particularly its role in identifying the basal-like subtype, commonly known as triple-negative breast cancer. While not considered a predictive marker, CK 5/6 expression serves as a valuable tool for categorizing tumors based on their molecular characteristics. The thesis delves into the implications of CK 5/6 as a diagnostic marker, shedding light on its association with the aggressive basal-like subtype, which lacks hormone receptor expression and HER2 overexpression. Understanding CK 5/6's significance contributes to the nuanced classification of breast cancer subtypes, guiding therapeutic decisions and offering insights into the prognosis of affected individuals.

1.5.5 APOPTOSIS MARKERS

Apoptosis markers, pivotal indicators of programmed cell death essential for maintaining tissue homeostasis and preventing diseases such as cancer. The study delves into prominent apoptosis markers, including caspases (such as Caspase-3, Caspase-8, and Caspase-9), PARP (Poly (ADP-ribose) polymerase), and Bcl-2 family proteins, which collectively regulate the intricate process of apoptosis. Annexin V and the TUNEL assay are explored for their roles in detecting apoptotic cells, particularly in flow cytometry and histological analyses.

1.5.6 OTHER MARKERS WITH PREDICTIVE POTENTIAL

BRCA1 is involved in the repair of DNA double-strand breaks and is also associated with nucleotide excision repair. Downregulation of breast and ovarian cancer susceptibility gene (BRCA1/2) increased the sensitivity of tumor to DNA damaging chemotherapy (example, PARP inhibitor Olaparib; cisplatin and

etoposide), whereas intact BRCA1/2 increased the susceptibility to anti-microtubule agents. Two classes of microtubule interfering agents are used- vinca alkaloids and taxanes, the former destabilize, and the latter stabilize the microtubules. Further, studies have shown that BRCA1 mutation carriers were more sensitive to anthracycline-based regimens than BRCA2 and non-carriers. DNA recombinase RAD51, a component of homologous recombinant repair has shown the potential to predict PCR. Low RAD51 score associated with high grade, high proliferative index, TNBC and PCR.

Multi-analyte tests like urokinase plasminogen activator (uPA)/PAI-1, Oncotype DX and MammaPrint though show prognostic prediction potential, are not reliable in predicting the response to a specific form of chemotherapy. Two randomized trials, NSABP B-20 and SWOG-8814 explored the role of Oncotype DX as a predictive signature. It is suggested that the recurrence score (RS) was associated with response to adjuvant chemotherapy. High CD4+ and CD8+ tumor-infiltrating lymphocytes were associated with PCR followed by NAT and a better prognosis.

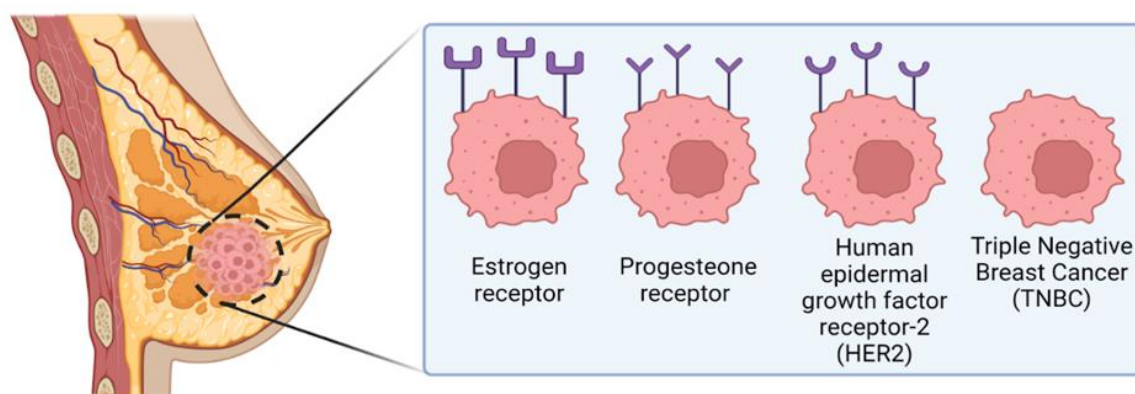


Fig.1.4: The Predictive Marker of Breast Cancer

CHAPTER 2

LITERATURE SURVEY

THE ROLE B4GALT 5 & 6 IN HUMAN CANCER

These studies exploring the roles of the B4GALT family, the hedgehog pathway, P-glycoprotein, and multidrug resistance-associated protein 1 (MRP1) in multidrug resistance mechanisms in cancer cells, particularly leukaemia. Relevant articles discuss the interaction between glycosylation and multidrug resistance, the involvement of hedgehog signaling in cancer drug resistance, and the functions of P-glycoprotein and MRP1 in limiting drug penetration. Additionally, reviews on strategies for targeting multidrug resistance in leukemia provide context for understanding the importance of the proposed research in combating drug resistance in leukemia treatment. **Zhou et al. (2013)**

The roles of UGCG, B4GalT5, mitochondrial oxidative stress, and the ERK signaling pathway in heart hypertrophy and cardiac remodelling. Relevant articles discuss the involvement of glycosphingolipids in heart disease, the role of B4GalT5 in cell adhesion signaling and mitochondrial function, and the contribution of ERK signaling to cardiac hypertrophy. This comprehensive survey provides background information crucial for understanding the proposed research's focus on the intricate connections between UGCG, B4GalT5, mitochondrial oxidative stress, and ERK signaling in the modulation of heart hypertrophy. **(Cui et al., 2023)**

The mechanisms of aging in *Caenorhabditis elegans*, changes in glycosylation patterns during aging, the role of *Caenorhabditis elegans* as a model organism for aging research, and the impact of glycosylation on aging and age-related diseases. Relevant articles discuss the utility of *Caenorhabditis elegans* as a model system for studying aging, genetic mechanisms underlying aging, and the implications of altered galactosyltransferase expression and glycosylation patterns in aging-related processes. This comprehensive survey provides background information crucial for understanding the proposed

research's focus on β -1,4-galactosyltransferase expression during aging in *Caenorhabditis elegans*. **(Kremer et al., 2020)**

The roles of B4GALT5, lumican, and the Akt/GSK-3 β / β -catenin pathway in cardiac fibrosis. Relevant articles discuss the involvement of glycosylation in cardiac hypertrophy and fibrosis, the multifunctional roles of lumican in regulating the extracellular matrix, and the significance of Akt signaling in cardiac hypertrophy and heart failure. Additionally, studies on the role of GSK-3 β in mitochondrial dysfunction provide insights into potential mechanisms underlying cardiac fibrosis via the Akt/GSK-3 β / β -catenin pathway. This comprehensive survey provides background information crucial for understanding the proposed research's focus on attenuating cardiac fibrosis through B4GALT5 downregulation and its interactions with lumican and the Akt/GSK-3 β / β -catenin pathway. **(X. Zhang et al., 2024)**

THE ROLE B4GALT 5 & 6 IN BREAST CANCER

The immunological aspects of porcine reproductive and respiratory syndrome virus (PRRSV) infection, the role of glycosyltransferases such as B4GALT5 in immune cell development and function, and the innate and adaptive immune responses to PRRSV infection. Relevant articles discuss immune modulation by virus-encoded proteins, the impact of glycosylation on immune cell differentiation and activation, and the interactions between PRRSV and the host immune system. This comprehensive survey provides background information crucial for understanding the proposed research's focus on elucidating the immunological regulation roles of B4GALT5 during PRRSV infection. **(L. Zhang et al., 2018)**

The role of cell surface galactosyltransferase in various cellular processes, including cell interactions, adhesion, fertilization, migration, and cancer progression. Relevant articles discuss the dynamic control of cell adhesion and migration mediated by cell surface galactosyltransferase, its impact on cancer metastasis, glycan-mediated cell interactions in stem cell differentiation and organ regeneration, and its potential interactions with adhesion molecules such as integrins. This comprehensive survey provides

background information crucial for understanding the expression and function of cell surface β 1,4-galactosyltransferase and its implications in diverse biological processes. **(Cooke & Shur, 1994)**

The roles of glycosphingolipids in neuronal development and myelin formation, the importance of B4galt5 and B4galt6 genes in myelination processes, and the interplay between glycosphingolipid metabolism and neural development. Relevant articles discuss the essential roles of glycosphingolipids in cellular processes underlying brain development, including neuronal generation and differentiation, as well as their involvement in myelin stability and synaptic plasticity. This comprehensive survey provides background information crucial for understanding the pivotal roles of lactosylceramide synthases encoded by B4galt5 and B4galt6 genes in neuronal generation and myelin formation in mice. **(Yoshihara et al., 2018)**

The role of glycosphingolipids, including lactosylceramide, in embryonic development, the importance of B4GALT5 in lactosylceramide synthesis and embryonic development, and the functional significance of extra-embryonic tissues in mouse development and viability. Relevant articles provide insights into the regulatory mechanisms underlying glycosphingolipid synthesis, the genetic interactions between B4GALT5 and other glycosyltransferases, and the essential functions of extra-embryonic tissues in supporting embryonic growth and viability. This comprehensive survey offers background information crucial for understanding the pivotal role of B4GALT5 as a lactosylceramide synthase essential for mouse extra-embryonic development. **(Nishie et al., 2010)**

The functional differences between β 4GalT6 and β 4GalT5, and the regulation of lactosylceramide synthesis. Relevant articles provide insights into substrate recognition by glycosyltransferases, differential recognition of sorting signals by adaptor protein complexes, and developmental changes in glycosphingolipid synthesis. This comprehensive survey offers background information crucial for understanding the involvement of β 4GalT6 in lactosylceramide synthesis with less intensity compared to β 4GalT5. **(Tokuda et al., 2013)**

the role of beta-1,4-galactosyltransferase V (B4GALT5) in mouse development, growth regulation, and viability. Relevant articles provide insights into the function of B4GALT5 in glycosylation processes, its involvement in glycosphingolipid synthesis, and the consequences of B4GALT5 deficiency on mouse growth and viability. Additionally, studies on the lethality of mutant mice with deficiencies in other glycosyltransferases offer context for understanding the impact of B4GALT5 mutations on mouse development and survival. This comprehensive survey offers background information crucial for understanding the early lethality observed in beta-1,4-galactosyltransferase V-mutant mice due to growth retardation. **(Kumagai et al., 2009)**

CHAPTER 3

AIM AND OBJECTIVES

AIM

To study the gene expression of B4GALT5 & B4GALT6 in breast cancer using real-time polymerase chain reaction.

OBJECTIVES

1.Expression Profiling: To determine the levels of B4GALT5 and B4GALT6 expression in various breast cancer subtypes compared to adjacent normal tissues(PSO1,PSO2).

2.Correlation with Clinical Parameters: To assess the potential associations between the expression levels of B4GALT5 and B4GALT6 and relevant clinical parameters (PSO1, PSO2).

3.Glycosylation Analysis: To investigate the impact of altered expression of B4GALT5 and B4GALT6 on glycosylation patterns in breast cancer cells (PSO1, PSO2).

4.Biomarker Discovery: To evaluate the feasibility of B4GALT5 and B4GALT6 as diagnostic or prognostic biomarkers for breast cancer (PO3, PO6).

CHAPTER 4

MATERIALS AND METHODS

IN VITRO STUDY USING CELL LINE

4.1 CELL LINES AND TYPES OF CELL LINES

Cell lines are groups of cells that originally came from animal or human tissues. They have been modified to grow continuously in controlled lab conditions. These cell lines are valuable tools for medical and biological research. They help scientists understand disease mechanisms, test drugs, and study cell function and biology. Each cell line is different based on where it came from. In general, they need the right amounts of nutrients, acidity levels, and temperature to grow and reproduce. Over time, cell lines can change genetically. To prevent this, scientists keep frozen stocks of cells from early stages to ensure their accuracy.

4.1.1 CONTROL CELL

The HEK293 cell line originated from embryonic human kidney cells that were immortalized through transformation with sheared fragments of adenovirus 5 DNA, as indicated by the '293' label representing the specific culture established. HEK293 has emerged as one of the most widely used cell lines in academic research and biotechnology applications owing to several desirable properties – it has adherent epithelial-like morphology growing in monolayers suitable for cell culture vessels, and unmatched efficiency at transient gene expression enabling reliable exogenous protein production. Its extraordinarily rapid division rate of 18-24 hours has supported high yield generation of recombinant proteins, viral vectors for vaccines, gene therapy as well as basic research.

4.1.2 BREAST CANCER CELL LINES

1) MCF-7

The MCF7 cell line was established in 1970 from the pleural effusion of a female patient diagnosed with metastatic breast adenocarcinoma. MCF7 cells exhibit an epithelial morphology and form adherent monolayers with doubling times ranging from 38-60 hours. A defining property of this widely studied line is the preservation of estrogen receptors as well as expression of luminal epithelial markers like cytokeratins. Notably, MCF7 cells can form spheroid structures in 3D cultures and produce tumors when xenografted in mice, providing excellent in vivo breast cancer models. Although the production of milk proteins or lipids is compromised, the retention of a certain degree of differentiation enables MCF7 to serve as an important tool in breast cancer research on processes like proliferation, cell signaling, metastasis and angiogenesis. the MCF7 human breast adenocarcinoma cell line has become indispensable for breast cancer studies owing to its origins in metastatic epithelial cells coupled with a high degree of physiological.

2)ZR-751

The ZR-75-1 or ZR751 cell line was established in 1970 from the malignant ascites of a 63-year-old Caucasian female patient diagnosed with invasive ductal breast carcinoma. The cell line is named after the Zachariae and Ramos researchers who developed this continuing cell culture. ZR751 exhibits typical epithelial morphology, growing as adherent monolayers with a doubling rate of approximately 60 hours. As a luminally-derived cell line, ZR751 cells express estrogen and progesterone receptors in addition to luminal biomarkers like cytokeratin 19. Notably, ZR751 cultures display specialized structures called domes and acini-like spheroids that are indicative of functional polarity, not commonly retained in cancer lines. However, the cells can be highly aggressive and invasive if xenografted in immunocompromised mice. The conservation of hormone receptors and partial differentiation attributes make the ZR751 cell line well-suited for evaluating hormone response, screening anti-estrogenic compounds, elucidating signaling pathways driving abnormal proliferation as well

as identifying molecular factors involved in metastasis and angiogenesis. The ZR751 line serves as an important luminal subtype breast cancer model owing to maintained estrogen sensitivity pathways and invasive disease characteristics mirroring the original clinical case.

TABLE 4.1 : List of Cell Lines Used

Type of Cell Line	Cell Line	Medium Used	Source of Cell Lines
Control cell	HEK293	DMEM with 10% FBS	CANCER INSTITUTE
Breast cancer	MCF7 & ZR-751	DMEM with 10% FBS	CANCER INSTITUTE

4.2 CULTURING OF CELL LINES

4.2.1 Mammalian cell culture

Human derived cell lines (Table 2.1) were maintained in T-25 cm² flasks (Corning Inc, USA), in Dulbecco's Modified Eagle Medium (DMEM) (HiMedia, India) or Roswell Park Memorial Institute -1640 (RPM1-1640) (HiMedia, India), supplemented with 10% Fetal bovine serum (FBS) (GIBCO, USA) containing 100 mg/L penicillin and 100 mg/L streptomycin (HiMedia, India) at 37°C and 5% CO₂ in a humidified incubator. For sub-culturing, 80 to 90% confluent cells were washed in sterile phosphate-buffered saline (PBS) and then 0.5 to 2 mL Trypsin-EDTA solution was added [0.25% w/v Trypsin (Sigma Aldrich, India) and ethylene diamine tetraacetate (EDTA) (MB biomedical, India) 0.001% w/v dissolved in PBS and 0.2 µm filter sterilized]. Then the cells were centrifuged at 800 rpm for 2 min and the pellet was suspended in a growth medium containing 10% FBS to inactivate trypsin. 10 µL of cell suspension was stained with 10 µL of 0.4% trypan-blue solution and the cell numbers were estimated using Countess II FL Automated Cell Counter (Thermo Scientific, USA). Cell line stocks were frozen in liquid nitrogen using a freezing mixture of 70% FBS, 20%

DMEM, and 10% dimethyl sulfoxide. The analysis for contamination like was assessed frequently using mycoplasma PCR ELISA kit (Roche, Switzerland).

4.2.2 Total RNA isolation

Total RNA was isolated using trizol reagent (Invitrogen, USA). Cells suspended in trizol were added to an equal volume of chloroform, mixed gently, and centrifuged at 12,000 rpm for 10 min at 4°C. The aqueous phase was pipetted out without disturbing the contents and transferred to a fresh 1.5 mL tube and placed on ice. Then, isopropanol was added to the aqueous phase and centrifuged at 12,000 rpm for 10 min at 4°C. The white pellet thus formed was washed twice with 70% ethanol. Then the pellet was air-dried and dissolved in 20 µL of sterile nuclease-free water. The RNA concentration was quantified using Nanodrop ND 2000 spectrophotometer (Thermo Scientific, USA). Sterile nuclease-free water was used as blank and 1 µL of the sample was used to estimate the concentration of RNA in ng/µL. Absorbance at 260 nm/280 nm ratio between 1.8 and 2 indicated the purity of the isolated RNA.

4.2.3 Agarose Gel Electrophoresis

RNA integrity was assessed using agarose gel electrophoresis. 100 mL of 1x Tris/Borate/EDTA (TBE) buffer (0.13 M Tris, 5 mM boric acid, and 2.5 mM EDTA, pH 7.6) was produced with 0.8% agarose (Cytvia, USA). After the agarose was well dissolved in the oven, it was put into the gel tray and the comb was added. After giving the agarose 30 minutes to firm, the gel tray was submerged in the gel tank that contained one TBE buffer. After combining 1 µg of total RNA with 6x clear safe dye (Green Biosciences, India), the agarose gel was loaded, and it was run for 30 minutes at 100 V. The ChemiDoc™ MP imaging equipment (Biorad, USA) was used to see and record the gel.

4.2.4 cDNA Synthesis

Using the PrimeScript™ RT reagent kit with gDNA eraser (DSS Takara Clontech, Japan), 1 µg of total RNA was transformed into cDNA in two steps, following the manufacturer's instructions. To make 6 µL of sterile nuclease-free

water, add 2 μ L of 5X gDNA eraser buffer (2.5 μ M), 1 μ g of total RNA, and 1 μ L of gDNA eraser (10 mM). The tube was put in a heat cycler for two minutes at 42°C and Next, 4 μ L of 5X PrimeScript buffer 2, 1 μ L PrimerScript RT enzyme mix 1(40 U/ μ L), 1 μ L RT Primer mix, 10 μ L of reaction solution from previous step 1 and 4 μ L of sterile nuclease-free water were added to the reaction mixture mentioned above. A thermal cycler was used to place the tube under cycling conditions are 37°C at 15 min and 85°C for 5 sec. Finally store at 4°C.

4.2.5 Quantitative Real-Time- Polymerase Chain Reaction (QRT-PCR)

qRT-PCR was performed with gene-specific primers for β 4GALT5, β 4GALT6 genes and GAPDH as the normalizing control gene, using SensiFAST™ SYBR® Lo-ROX Kit (Bioline, UK) according to the manufacturer's protocol in triplicates. 10 μ L of master mix, 1 μ L of forward primer (10 μ M), 1 μ L of reverse primer (10 μ M), 1 μ L of cDNA template (50 ng) was added and the volume was made to 20 μ L with nuclease-free water. In the case of expression analysis of β 4GALT5, β 4GALT6 genes in cancer cell lines, the assay was performed as a 10 μ L (half-reaction) in 96 wells using EP motion 3000 (Eppendorf, Germany). In EP motion 3000, the workstation was set up by switching on UV light for 30 mins, the tubes containing master mix, template DNA, and primers were added to 1.5 ml microfuge tubes and fixed in the workstation. Then Eppendorf micro tips with a volume of 1 mL and 200 μ L were fixed on the workstation and the program was set up such that 1 μ L of template, 1 μ L of forward primer, 1 μ L of reverse primer, and 7 μ L of master mix (SensiFAST™ SYBR Lo-ROX kit, Meridian Biosciences, USA) was added to the 96 well plates and in the corresponding template GAPDH primers were added accordingly. After the addition of reaction components in the workstation, the 96 well plates were centrifuged at 1000 rpm to ensure the content is at the bottom of the well. Then the 96 well plates containing the reaction mix were analyzed using Quantstudio³ Realtime PCR (Thermo Fischer Scientific, USA). The program was set up as fast, 96 well plate, to calculate cycle threshold value (CT-value) and 40 cycles. The relative expression levels between target and control were measured with $2^{-\Delta\Delta CT}$ method. In this method, the relative gene expression is assumed to be 1 for the reference sample so that 2^0 will provide a

value of 1. It assumes 100% efficiency and uniform PCR amplification across all the samples. For β 4GALT5, β 4GALT6 genes expression analysis in cancer cell lines, HEK293 were used as control cell lines.

STUDY IN TISSUE SAMPLES

4.3 BC SAMPLE COLLECTION

4.3.1 FROZEN SECTION

Frozen section is an intra-operative diagnostic technique where a small piece of tissue is removed during surgery, immediately frozen, sectioned and stained with H&E to visualize under a microscope - all within minutes. This helps the pathologist provide a provisional diagnosis so the surgeon can take quick decisions related to surgical margins or cancer staging while the patient is still anesthetized. The frozen tissue is embedded in an optimal cutting temperature compound, cryosectioned at 4-6 microns thickness, briefly fixed and stained before conveying the diagnosis. Additional tissue may be submitted for permanent section later. The rapid turnaround is made possible by coordinated workflow between the surgical team, pathologists and technologists to transport, process and evaluate frozen sections through specialized equipment like cryostats. When combined with imprint cytology, frozen sections can analysis surgical specimens effectively with reasonable accuracy thereby guiding ongoing surgical procedures.

4.3.2 HISTOPATHOLOGY

Histopathology refers to the microscopic examination of tissue specimens to study the manifestations of disease. The microscopic scrutiny of tumors in histopathology starts from thorough gross examination of the tissue sample to note physical qualities before systematic processing via fixation, infiltration and embedding steps to support thin sectioning. The 2–5-micron sections when stained with hematoxylin-eosin facilitate differential uptake of basic dyes by nuclear and cytoplasmic elements, revealing microscopic manifestations of neoplasia to the pathologist. Careful inspection under the microscope delineates

architectural attributes based on cellularity patterns and matrix organization. The differentiation status is interpreted from features like nuclear pleomorphism, mitotic frequency, presence of abnormal forms. Additionally, invasion into neighboring tissues, vascularization profiles, necrosis depict the aggressive disease status.

4.3.3 SAMPLE TYPES & SAMPLE COUNT

TUMOR

Tumor samples are analyzed for various characteristics that provide crucial information about the nature and behavior of cancer. These include cellular morphology (appearance and structure of tumor cells), histological type (classification based on cell origin), grade (degree of abnormality), biomarkers (molecules indicating tumor behavior or treatment response), molecular profile (genetic and molecular characteristics), tumor heterogeneity (variations within the tumor), and tumor microenvironment (surrounding cells and structures). By examining these characteristics, doctors can determine the tumor type, stage, aggressiveness, and potential response to specific treatments, ultimately guiding the most appropriate therapy for each patient.

PAIR NORMAL

A pair normal sample is a sample of healthy tissue taken from the same individual as the tumor sample, serving as a control or reference for comparison. Its key characteristics include normal cellular morphology, gene expression profiles, genetic integrity without mutations, typical protein expression levels, normal epigenetic patterns, and an intact tissue microenvironment. By contrasting the pair normal sample with the tumor sample, researchers can identify the specific molecular and cellular changes associated with the cancer, enabling a deeper understanding of the disease mechanisms, identification of therapeutic targets, and development of personalized treatment strategies.

Here, I have collected around 10 sets of samples, which includes tumor and pair normal. All the samples were collected from straight forward surgery

process (MRM) technique, and it undergone microscopy examination by histopathology department. All the samples are negative viral markers that were reported by the serology department. Most of the samples are invasive ductal carcinoma (IDC) and few samples are ductal carcinoma in-situ (DCIS).



Fig. 4.1: Tumor & Paired Normal Breast Sample Collected from FS

4.4 BC SAMPLE PREPARATION

Tissue samples were collected after obtaining informed consent from breast cancer patients who underwent modified radical mastectomy without any prior treatment. The tissues adjacent to the tumor were obtained and analyzed to be histologically normal cells and were used as Paired-Normal (PN) samples. In general, patients were diagnosed with Invasive Ductal Carcinoma (IDC) and also patient was diagnosed with Ductal Carcinoma in Situ (DCIS). Tissues were snap-frozen in liquid nitrogen and stored at -80°C until further use.

Then the sample was ground with mortar and pestle, under the frozen state using liquid nitrogen to preserve the Rnase. The grinded tissue was added into the 1.5 μL tube containing trizol reagent and processed with RNA isolation process by standard protocol.

4.4.1 TOTAL RNA ISOLATION

Total RNA was isolated using trizol reagent (Invitrogen, USA). Cells suspended in trizol were added to an equal volume of chloroform, mixed gently, and centrifuged at 12,000 rpm for 10 min at 4°C . The aqueous phase was

pipetted out without disturbing the contents and transferred to a fresh 1.5 mL tube and placed on ice. Then, isopropanol was added to the aqueous phase and centrifuged at 12,000 rpm for 10 min at 4°C. The white pellet thus formed was washed twice with 70% ethanol. Then the pellet was air-dried and dissolved in 20 µL of sterile nuclease-free water. The RNA concentration was quantified using Nanodrop ND 2000 spectrophotometer (Thermo Scientific, USA). Sterile nuclease-free water was used as blank and 1 µL of the sample was used to estimate the concentration of RNA in ng/µL. Absorbance at 260 nm/280 nm ratio between 1.8 and 2 indicated the purity of the isolated RNA.

4.4.2 Agarose Gel Electrophoresis

RNA integrity was assessed using agarose gel electrophoresis. 100 mL of 1x Tris/Borate/EDTA (TBE) buffer (0.13 M Tris, 5 mM boric acid, and 2.5 mM EDTA, pH 7.6) was produced with 0.8% agarose (Cytvia, USA). After the agarose was well dissolved in the oven, it was put into the gel tray and the comb was added. After giving the agarose 30 minutes to firm, the gel tray was submerged in the gel tank that contained one TBE buffer. After combining 1 µg of total RNA with 6x clear safe dye (Green Biosciences, India), the agarose gel was loaded, and it was run for 30 minutes at 100 V. The ChemiDoc™ MP imaging equipment (Biorad, USA) was used to see and record the gel.

4.4.3 cDNA Synthesis

Using the PrimeScript™ RT-PCR Kit (DSS Takara Clontech, Japan), 1 µg of total RNA was transformed into cDNA in two steps, following the manufacturer's instructions. To make 10 µL of sterile nuclease-free water, add 1 µL of oligoDT primer (2.5 µM), 1 µg of total RNA, and 1 µL of dNTP mixture (10 mM). The tube was put in a heat cycler to be denaturated and annealed for five minutes at 65°C and ten minutes at 4°C. Next, 4 µL of 5X PrimeScript buffer, 0.5 µL RNase inhibitor (40 U/µL), 0.5 µL reverse transcriptase, and 20 µL of sterile nuclease-free water were added to the reaction mixture mentioned above. A thermal cycler was used to place the tube under cycling conditions are 30°C at 5 min, 42°C for 30 min, 95°C for 5 min, and 4°C.

4.4.4 Designing Primers

The primers for β 4GALT5, β 4GALT6 genes and housekeeping gene (GAPDH) are given in Table, which were designed using the Integrated DNA Technologies (Coralville, Iowa, USA) primer designing tool. The primers were validated by primer BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast>) and then given for synthesizing to Eurogentec, Belgium. The primers were dissolved in sterile nuclease-free water and diluted to a working concentration of 10 μ M for further use.

TABLE 4.2: List of Primers used for QRT-PCR Analysis

Gene	Primers for qRT-PCR
β4GALT5	FP: GAAGATGACGACCTCTGGAACAG RP: GCCGTTCTTTTGACTTCCTCAGC
β4GALT6	FP: CTGTGCTCAGGCGGATGAT RP: AGAGATATGTGTTGGCGATGO
GAPDH	FP: GTCTCCTCTGACTTCAACAGCG RP: ACCACCCTGTTGCTGTAGCCAA

4.4.5 Quantitative Real-Time- Polymerase Chain Reaction (QRT-PCR)

QRT-PCR was performed with gene-specific primers for β 4GALT5, β 4GALT6 genes and GAPDH as the normalizing control gene, using SensiFAST™ SYBR® Lo-ROX Kit (Bioline, UK) according to the manufacturer's protocol in triplicates. 10 μ L of master mix, 1 μ L of forward primer (10 μ M), 1 μ L of reverse primer (10 μ M), 1 μ L of cDNA template (50 ng) was added and the volume was made to 20 μ L with nuclease-free water. In the case of expression analysis of β 4GALT5, β 4GALT6 genes in breast cancer tissue (tumor & pair normal), the assay was performed as a 10 μ L (half-reaction) in 96 wells using EP motion 3000 (Eppendorf, Germany). In EP motion 3000, the workstation was set up by switching on UV light for 30 mins, the tubes containing master mix, template DNA, and primers were added to 1.5 ml microfuge tubes and fixed in

the workstation. Then Eppendorf micro tips with a volume of 1 mL and 200 μ L were fixed on the workstation and the program was set up such that 1 μ L of template, 1 μ L of forward primer, 1 μ L of reverse primer, and 7 μ L of master mix (SensiFASTTMSYBR Lo-ROX kit, Meridian Biosciences, USA) was added to the 96 well plates and in the corresponding template GAPDH primers were added accordingly. After the addition of reaction components in the workstation, the 96 well plates were centrifuged at 1000 rpm to ensure the content is at the bottom of the well. Then the 96 well plates containing the reaction mix were analyzed using Quantstudio³ Realtime PCR (Thermo Fischer Scientific, USA). The program was set up as fast, 96 well plate, to calculate cycle threshold value (CT-value) and 40 cycles. The relative expression levels between target and control were measured with $2^{-\Delta\Delta CT}$ method. In this method, the relative gene expression is assumed to be 1 for the reference sample so that 2^0 will provide a value of 1. It assumes 100% efficiency and uniform PCR amplification across all the samples [177]. For β 4GALT5, β 4GALT6 genes expression analysis in cancer cell lines, HEK293 were used as control cell lines.

BIOINFORMATICS ANALYSIS

A gene set is defined as comprising either a single or multiple genes. By default, if a gene set comprises multiple genes an average expression is calculated for the gene set. This average expression is next used for investigation across clinical and molecular groups, as well as for stratification of samples into groups based on gene expression quantiles. If a gene set is specified in file format, gene weights may be associated with each gene. A gene weight is a numerical value (positive or negative) that is multiplied with a gene's expression value prior to summing up the expression values of the genes in the gene set (the weighted average is normalized by the sum of the absolute values of the weights). Consequently, this allows gene sets to comprise genes with supposedly both positive and negative gene expression patterns.

4.5 EXPRESSION ANALYSIS OF GENES

4.5.1 *Gepia Database*

GEPIA (Gene Expression Profiling Interactive Analysis) is a web-based tool that integrates gene expression data from The Cancer Genome Atlas (TCGA) project, comprising over 9,000 tumor samples and 8,000 normal samples across 33 different cancer types, and the Genotype-Tissue Expression (GTEx) project, with data from over 8,000 normal tissue samples across 53 tissue types. It allows users to analyze and visualize gene expression patterns, compare expression levels between tumor and normal samples, perform survival analysis to assess the prognostic value of genes, identify genes with similar expression patterns, analyze correlations between gene expressions, and conduct pathway enrichment analysis. GEPIA provides interactive data visualization tools and enables researchers to explore the role of gene expression in cancer development, identify potential biomarkers, investigate biological processes and signaling pathways involved in diseases, and ultimately facilitate the understanding of molecular mechanisms underlying various conditions and the development of therapeutic strategies.

4.5.2 *GOBO Database*

The GOBO (Gene Expression-Based Outcome for Breast Cancer Online) database is a web-based tool that integrates gene expression data from multiple breast cancer cohorts, comprising over 1,800 patients with different breast cancer subtypes, enabling researchers to analyze the association between gene expression levels and clinical outcomes such as survival, relapse, and metastasis. It facilitates the identification of potential prognostic biomarkers, co-expressed gene networks, and multi-gene signatures with predictive value. Additionally, GOBO allows stratified analysis based on clinical parameters, interactive data visualization through Kaplan-Meier plots, box plots, and heat maps, and integration of gene expression data with other molecular and clinical data. By providing a comprehensive platform for exploring the prognostic and predictive value of gene expression patterns in breast cancer, GOBO contributes to the development of personalized treatment strategies, improved patient

outcomes, and a deeper understanding of the molecular mechanisms underlying breast cancer development, progression, and treatment response.

4.5.3 C-BIOPORTAL DATABASE

The cBioPortal for Cancer Genomics is a widely utilized web-based platform offering access to multidimensional cancer genomics data across various cancer types, including breast cancer. Users can explore gene expression profiles, mutations, copy number variations, and protein expression patterns, enabling comprehensive analysis of genomic alterations in breast cancer cohorts. Researchers leverage this platform to investigate the roles of B4GALT5 and B4GALT6 genes, exploring their expression levels across different breast cancer subtypes and correlating them with clinical parameters such as tumor stage, grade, and patient survival.

Additionally, cBioPortal integrates genomic data with clinical information, facilitating the identification of genetic alterations in B4GALT5 and B4GALT6 genes and their associations with patient outcomes and treatment responses. Comparative analysis across multiple breast cancer studies aids in identifying commonalities and differences in gene expression patterns, contributing to a deeper understanding of breast cancer heterogeneity. The platform's visualization tools enhance data interpretation, enabling researchers to identify patterns and associations related to B4GALT5 and B4GALT6 in breast cancer, thereby advancing biomarker discovery, prognostic assessment, and the identification of potential therapeutic targets.

4.5.4 COSMIC Database

COSMIC (Catalogue of Somatic Mutations in Cancer) is a comprehensive database providing curated information on somatic mutations in various cancer types, including breast cancer. Researchers utilize COSMIC to explore somatic mutations in the B4GALT5 and B4GALT6 genes within breast cancer samples, assessing their frequency and distribution across different subtypes. Additionally, while COSMIC primarily focuses on somatic mutations, it also includes limited data on gene expression and copy number variations, enabling analysis of

expression levels and copy number alterations of B4GALT5 and B4GALT6 in breast cancer.

Although COSMIC does not directly integrate clinical data, researchers can correlate genetic alterations in B4GALT5 and B4GALT6 with clinical parameters using additional datasets or platforms. This association helps assess the clinical relevance of these genes in breast cancer, including their potential impact on patient outcomes and treatment responses. Furthermore, COSMIC facilitates comparative analysis of genetic alterations across different cancer types, offering insights into the broader role of B4GALT5 and B4GALT6 in cancer biology beyond breast cancer.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 IN VITRO STUDY USING CELL LINE

5.1.1 CULTURING OF CANCER CELL LINES

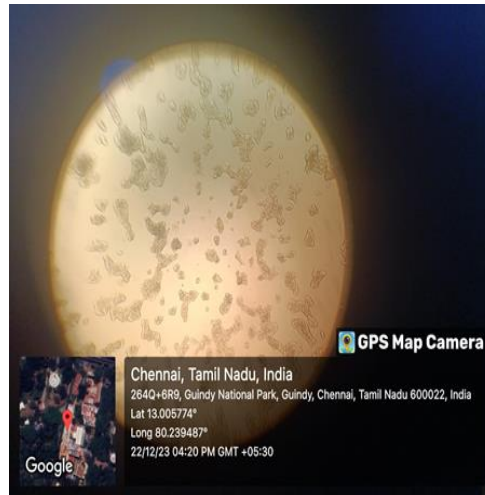


Fig. 5.1: HEK 293 cell lines, at 80% confluency growth attached to surface in T75 flask. Showing slightly elongated shape and slightly granular in appearance.



Fig:5.2: ZR751 cell lines, at 85% confluency growth attached to surface in T75 flask. Showing a slightly elongated rod shape.



Fig:5.3: MCF7 cell lines, at 90% confluency growth attached to surface in T75 flask. shows smaller, more circular, less irregular in shape.



Fig:5.4: Cell Culture Technique

5.1.2 RNA Isolation



Fig:5.5: RNA isolation

During the removing of the upper aqueous phase after adding chloroform step.

TABLE 5.1: Quantification Value for Cell Lines

CELL LINES TYPES	ng/ul	A260/A280
ZR751	2141.4	1.87
HEK293	2180.1	1.88
MCF7	1454.6	2.01

5.1.3 Agarose Gel Electrophoresis

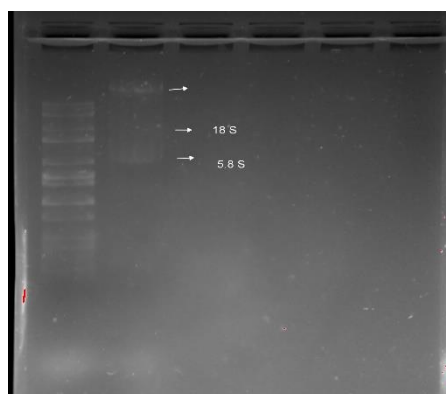


Fig:5.6: HEK 293 RNA Band

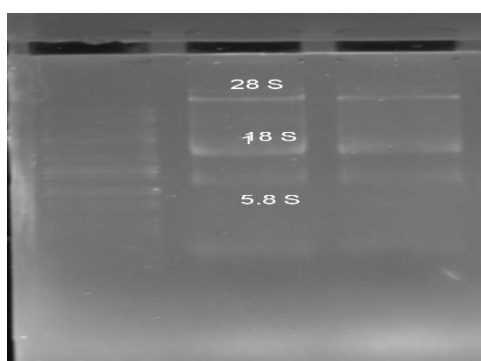


Fig:5.7: ZR751 RNA Band

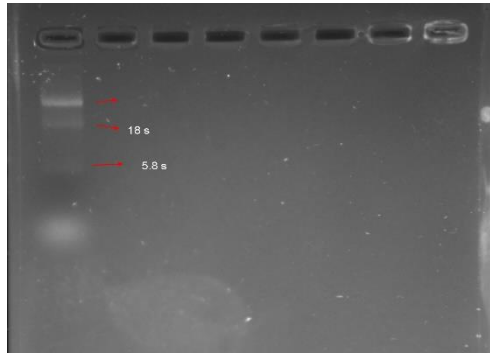


Fig:5.8: MCF7 RNA Band

Fig. 2.6, 2.7&2.8. Gel Doc analysis of RNA isolated were done for above three cell lines. All the rna band were seen (5.8s, 18s, 28).

5.1.4 QRT-PCR

E		F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
Target	Bar	Task	Reporter	Quencher	Quantity	Quantity	k	Quantity	S/RQ	RQ Min	RQ Max	CT	Ct Mean	Ct SD	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct	Delta Ct
41	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	27.651	29.947	1.974	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
42	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	30.443	29.947	1.974	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
43	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
44	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
45	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
46	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
47	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
48	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	35.346	34.664	0.965	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
49	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	33.982	34.664	0.965	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
50	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	35.772	35.772	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
51	G1	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
52	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	33.369	33.022	0.490	3.976	1.438	None	None	None	None	None	None	None	None	None	None	None	None	None	None
53	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	32.676	33.022	0.490	3.976	1.438	None	None	None	None	None	None	None	None	None	None	None	None	None	None
54	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
55	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
56	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
57	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
58	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
59	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
60	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
61	G2	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
62	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	31.579	31.963	0.458	2.817	1.433	None	None	None	None	None	None	None	None	None	None	None	None	None	None
63	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	32.187	31.963	0.458	2.817	1.433	None	None	None	None	None	None	None	None	None	None	None	None	None	None
64	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
65	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
66	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	30.228	38.228	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
67	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
68	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
69	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
70	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
71	G3	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
72	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	33.983	34.172	0.268	5.126	1.409	None	None	None	None	None	None	None	None	None	None	None	None	None	None
73	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	34.362	34.172	0.268	5.126	1.409	None	None	None	None	None	None	None	None	None	None	None	None	None	None
74	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
75	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
76	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
77	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
78	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
79	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
80	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
81	G4	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None	None
82	G5	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	34.270	34.614	0.487	5.568	1.438	None	None	None	None	None	None	None	None	None	None	None	None	None	None
83	G5	UNKNOWN	SYBR	None	None	None	None	None	None	None	None	34.959	34.614	0.487	5.568	1.438	None	None	None	None	None	None	None	None	None	None	None	None	None	None

Fig.5.9: Raw Data Extracted from the QRT-PCR Instrument

TC	HE	HC	d CT (TE-HE)	d CT (TC-HC)	dd CT (d CT TE- d CT TC)	2(ddCT)fold change	std dev	Average															
12	31.641	33.373	34.487	17.318	16.497	17.711	25.749	26.920	27.554	4.197	6.358	4.132	5.892	6.453	6.933	-1.695	-0.095	-2.801	3.23823	1.06812	6.96783	2.98401	3.75806
13	34.781	32.819	33.759	17.318	16.497	17.711	25.749	26.920	27.554	5.646	6.981	6.519	9.032	5.899	6.206	-3.387	1.082	0.313	10.459	0.47223	0.8049	5.67229	3.91206
14	34.207	35.313	35.456	17.318	16.497	17.711	25.749	26.920	27.554	6.651	7.343	7.374	8.459	8.393	7.902	-1.808	-1.049	-0.528	3.50169	2.06938	1.44232	1.0556	2.3378
15	37.582	37.787		17.318	16.497	17.711	25.749	26.920	27.554	9.229	9.472		11.834	10.867		-2.605	-1.395		0.68292	2.62946		4.39175	4.35619
16	33.175	34.362	34.929	17.318	16.497	17.711	25.749	26.920	27.554	4.260	6.465	7.987	7.426	7.462	7.375	-3.146	-0.907	0.611	8.85224	1.99597	0.65456	0.16091	0.23531
17	33.595	35.763	34.492	17.318	16.497	17.711	25.749	26.920	27.554	9.267	10.711	11.031	7.846	8.843	6.938	1.421	1.868	4.093	0.37338	0.27395	0.05861	0.16091	0.23531
18	33.265	34.799	35.953	17.318	16.497	17.711	25.749	26.920	27.554	8.649	10.861	10.200	7.516	7.879	8.399	1.133	2.981	1.801	0.45591	0.12061	0.28699	0.16467	0.28984
19																							
20																							
21																							
22																							
23																							
24																							

Fig:5.10: Calculation of Fold Change

TE denotes test experiment

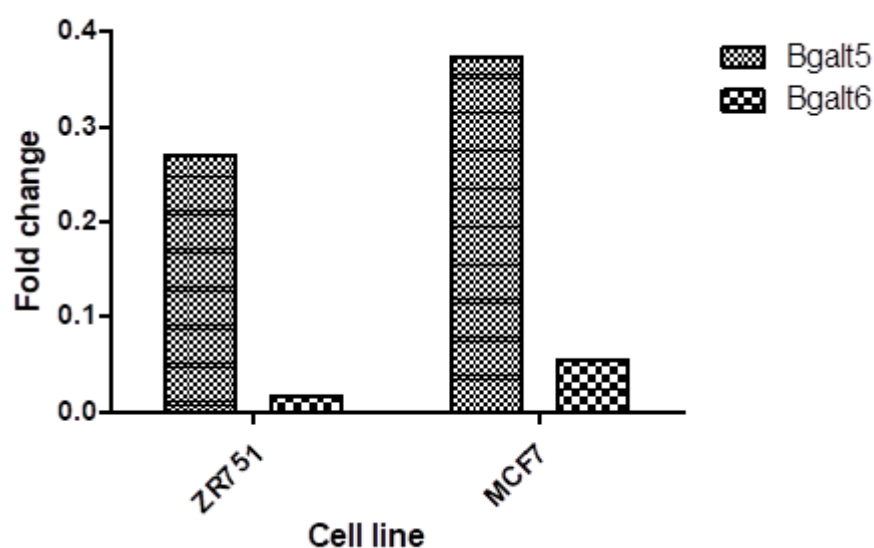
TC denotes test control

HE denotes house experiment

HC denotes house control

**TABLE 5.2: Fold Change Value for B4GALT5 & B4GALT6
GENES in Cell Lines**

CELL LINES	B4GALT5 FOLD CHANGE	B4GALT6 FOLD CHANGE
ZR751	0.270430342	0.017137934
MCF7	0.373727115	0.055229424



**Fig.5.11: Fold Change Comparison between the B4GALT5
& B4GALT 6 Genes**

Table 2.3 and Fig 2.11, HEK293 cell line is used as a house control to calculate the fold change value, which is compared with other two cell lines for predicting the expression level. B4GALT5 indicates the less expression gene or

less significant value (p value= 0.3 >0.05) . Which denotes the downregulation function in the breast cancer cell line. On the other hand, B4GALT6 gene shows lite upregualtion indication on breast cancer cell lines (P value =0.03<0.05). Which result in the little high expression of the these gene .

5.2 STUDY IN TISSUE SAMPLES

5.2.1 RNA ISOLATION

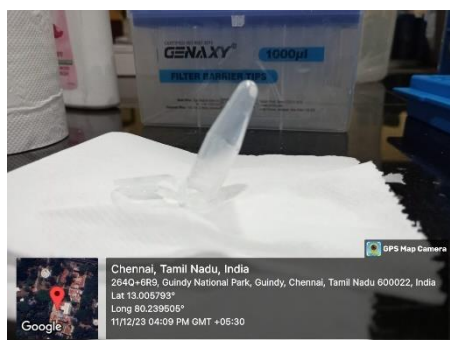


Fig.5.12: RNA isolation, at during the air-dry step for 10-20 mins.

TABLE 5.3: Quantification Value for Breast Cancer Tissue Samples (Tumor & Paired Normal)

SAMPLE NO	T, ng/ul	A260/A280	PN,ng/ul	A260/A280
1.	248.8	2.03	578.5	2.2
2.	551.8	2.04	102.8	2.1
3.	613.5	2.02	70	2.0
4.	947.4	2.01	198.9	2.3
5.	214.3	2.11	198.1	2.2

6.	155.7	2.0	52.9	2.3
7.	1252.1	2.01	72.9	2.4
8.	279.6	1.99	65.0	1.88
9.	181.7	1.8	43.3	2.3
10.	1610.4	2.1	155.5	2.2

5.2.2 CDNA SYNTHESIS



Fig.5.13: CDNA SYNTHESIS

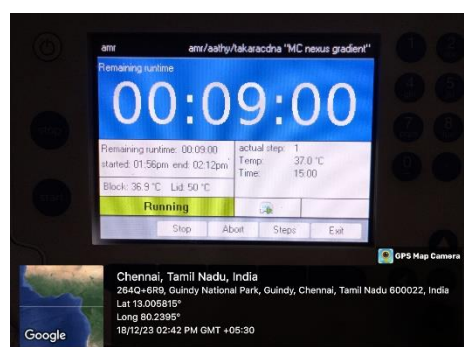


Fig.5.14: Thermal Cycle Instrument

5.2.3 Standardization of Primers

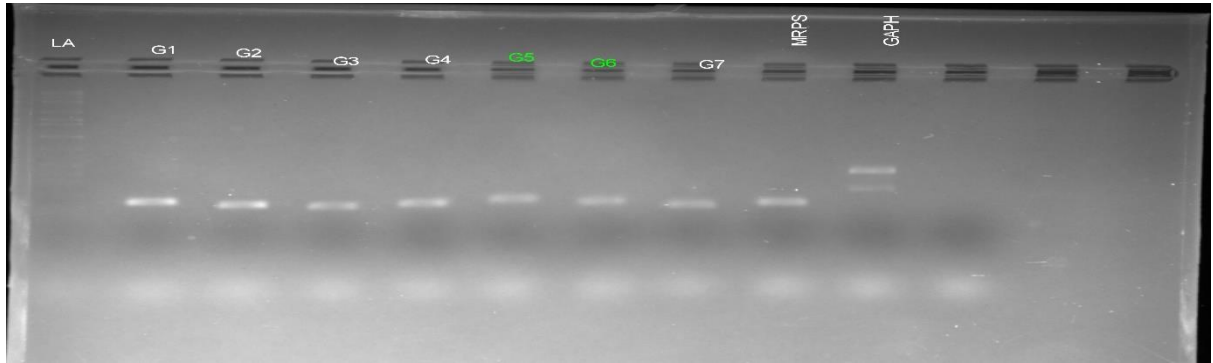


Fig.5.15 : B4GALT5 & B4GALT6 Bands Shows in the Range between 100-250 bp

5.2.4 QRT-PCR



Fig.5.16: Quantstudio 12th k Realtime PCR

TABLE 5.4: Fold Change Value for B4GALT5 & B4GALT6 GENES

SAMPLE NO	B4GALT5 -FOLD CHANGE	B4GALT6- FOLD CHANGE
1.	0.064415745	0.147684574
2.	0.009751639	0.018768036
3.	0.175827694	0.028881617
4.	0.36788586	0.106309832
5.	6.29035734	0.167732416
6.	0.488809195	0.084842859

7.	10.84592498	2.314509863
8.	1.423931816	1.217201004
9.	0.181206369	0.053086758
10.	0.772629988	0.312814962

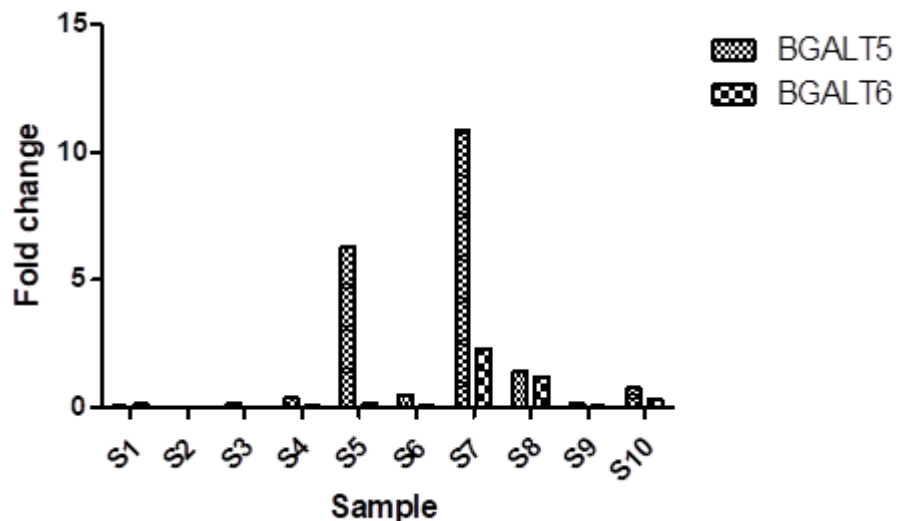


Fig.5.17: Fold Change Comparison between the B4GALT5 & B4GALT 6 Genes

Table 3.3 and Fig 3.4 paired normal tissue is used to compare with tumor sample to calculate the fold change value, which predicts the expression level of these genes. This final output of fold change from the breast tumor samples shows that both B4GALT5 and B4GALT6 are not significant or no expression in the breast cancer cell. In some cases, sample 5 and 7 shows high under expression of these genes as compared to other sample. These genes result in the downregulation function in the breast cancer cell.

Avg of B4GALT5 - **2.062074063**

Avg of B4GALT 6 - **0.445183192**

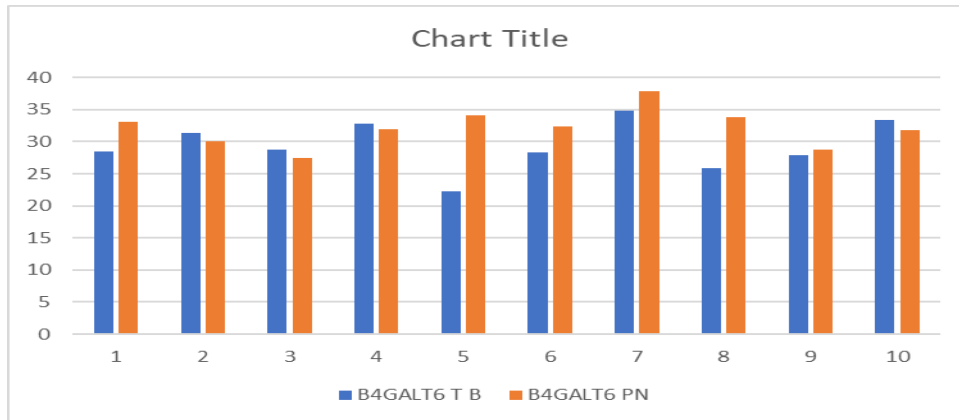


Fig.5.18: Comparison of B4GALT5 between Tumor & Paired Normal

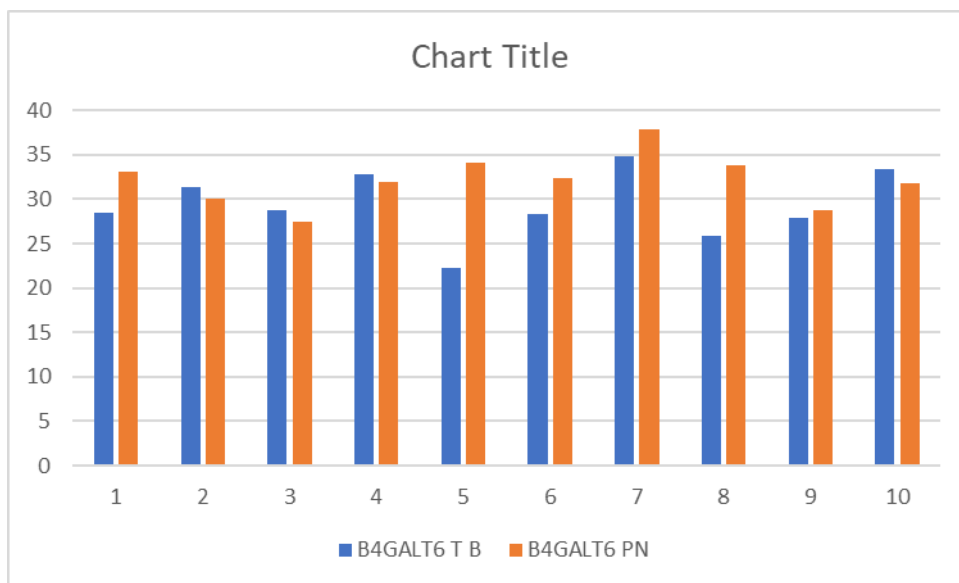


Fig.5.19: Comparison of B4GALT6 between Tumor & Paired Normal

In both figures, the expression of paired normal is higher than the tumor in both genes. It clearly shows that there is no expression of these genes in the breast tumor sample.

5.3 BIOINFORMATICS ANALYSIS

5.3.1 GEPIA DATABASE

B4GALT5

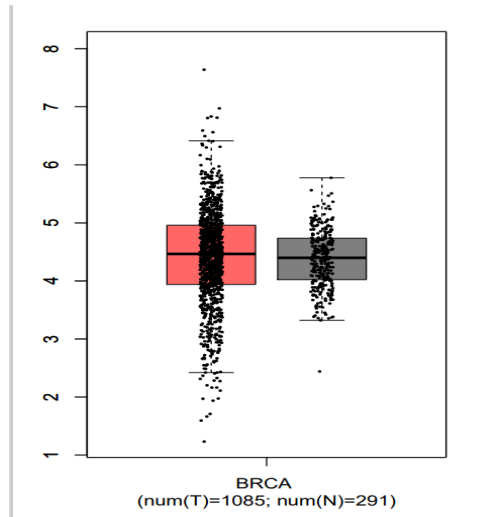


Fig.5.20: Box Plot of Breast Cancer Comparison between Tumor & Paired Normal with B4GALT5 Gene

Analysis of this box plot indicates that there is no significant of these gene in breast cancer when compared to the number of tumor vs paired normal samples. (P value is 0.39)

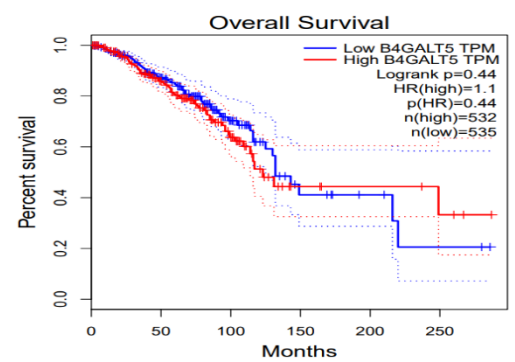


Fig.5.21: Survival Analysis of Breast Cancer Comparison between Tumor & Paired Normal with B4GALT5 Gene

The analysis of this plot indicates that the relationship between gene expression levels and patient survival outcomes, hazard ratio is 1.1 and P value

is 0.44. Which indicates that it is not significant value or there is no expression of these genes in BC.

B4GALT6

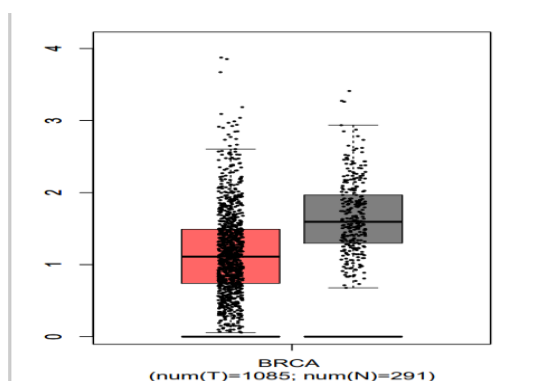


Fig.5.22: Box Plot of Breast Cancer Comparison between Tumor & Paired Normal with B4GALT6 Gene

Analysis of this box plot indicates that there is no significant of these gene in breast cancer when compared to the number of tumor vs paired normal samples.

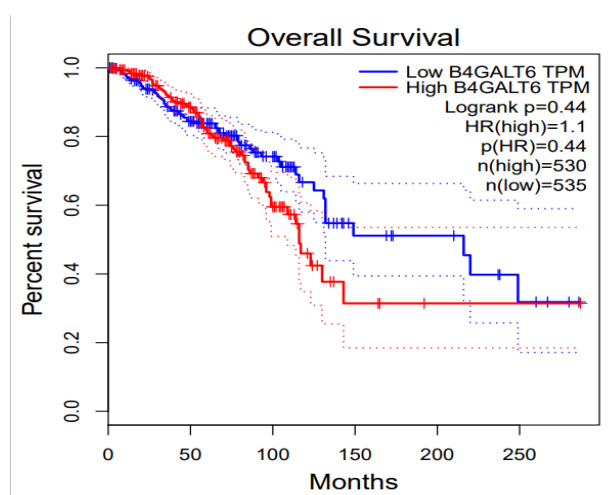


Fig.5.23: Survival Analysis of Breast Cancer Comparison between Tumor & Paired Normal with B4GALT6 Gene

The analysis of this plot indicates that the relationship between gene expression levels and patient survival outcomes, hazard ratio is 1.1 and P value

is 0.44. Which indicates that it is not significant value or there is no expression of these genes in BC

5.3.2 GOBO Database

B4GALT 5

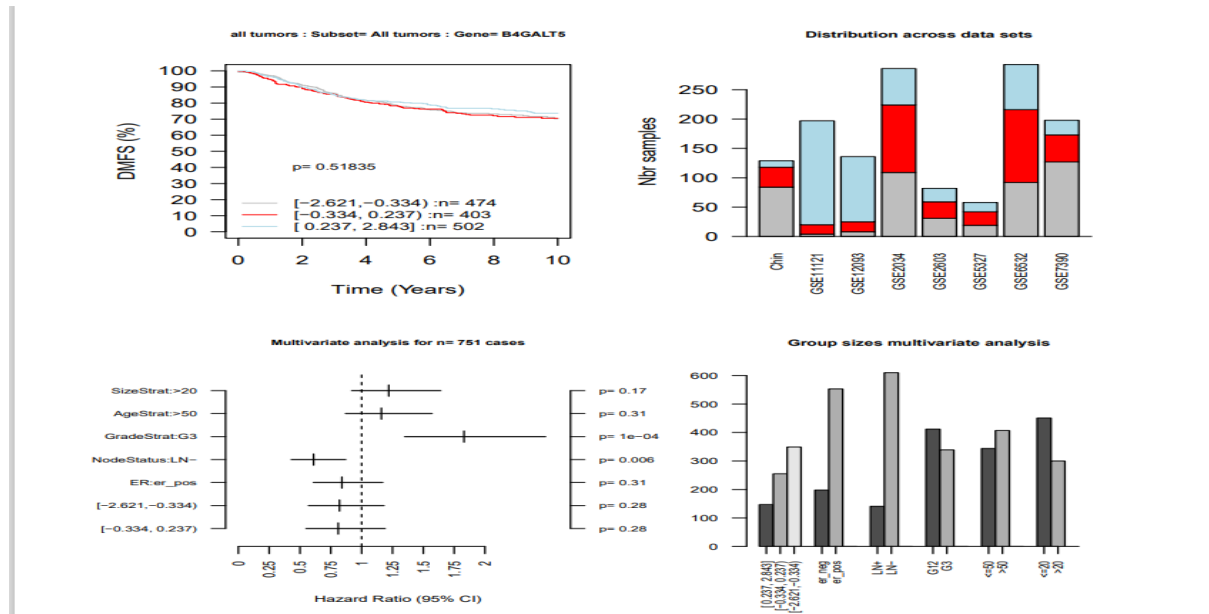


Fig.5.24: Gene set Analysis of B4GALT 5 Gene in all Tumors

GOBO analysis or Gene set analysis of B4GALT5 gene in all subsets of tumor also indicates that there is no significant expression of this gene in the BC, because the P value is great than the actual significant value ($0.5 > 0.05$).

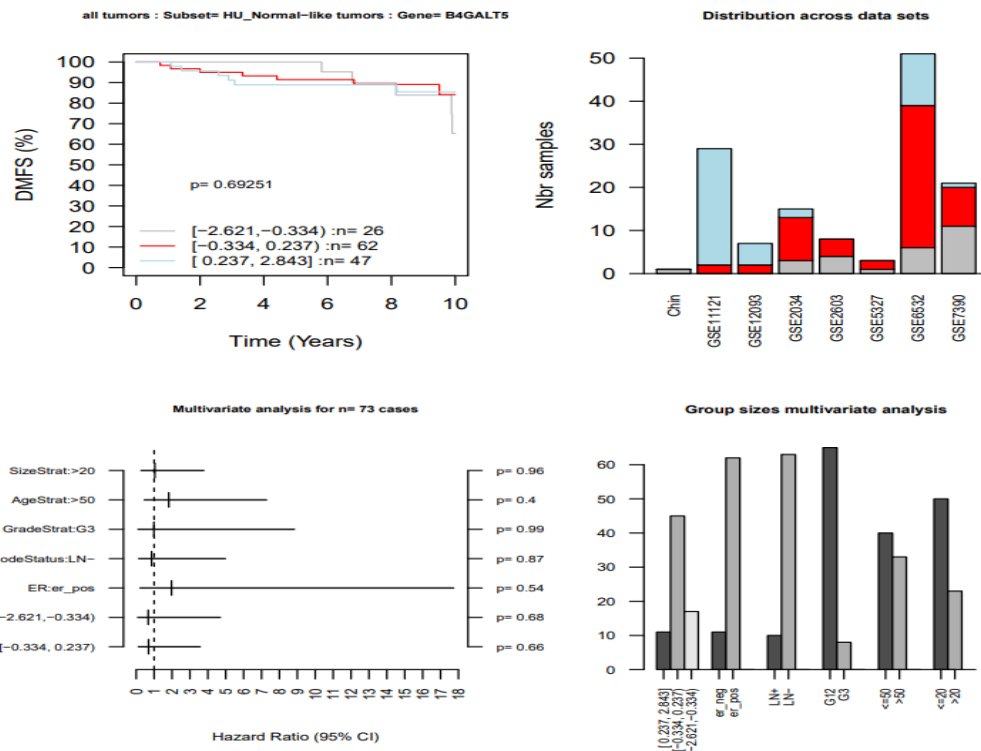


Fig.5.25: Gene set Analysis of B4GALT 5 Gene in Normal with Tumor

GOBO analysis or Gene set analysis of B4GALT5 gene in all subset of normal is compared with all the subset of the tumors, also indicates that there is no significant expression of this gene in the BC, because the P value is great than the actual significant value ($0.6 > 0.05$).

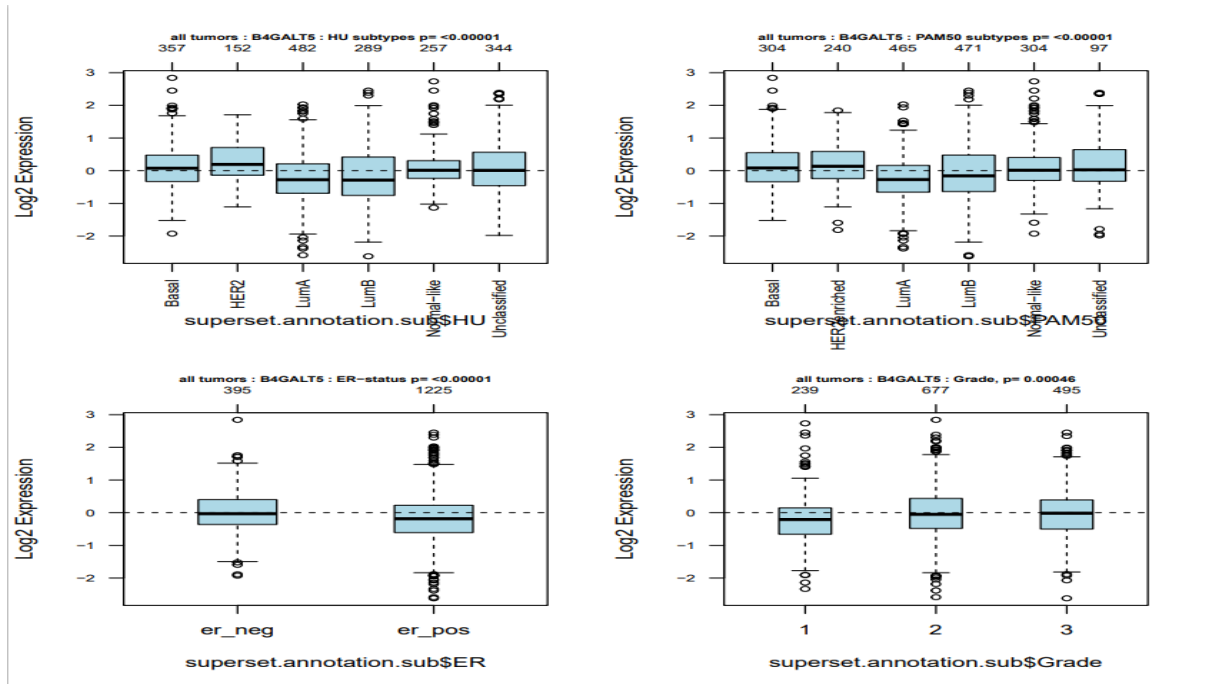


Fig.5.26: Comparison of B4GALT 5 GENE in Across all Types of Breast Cancer

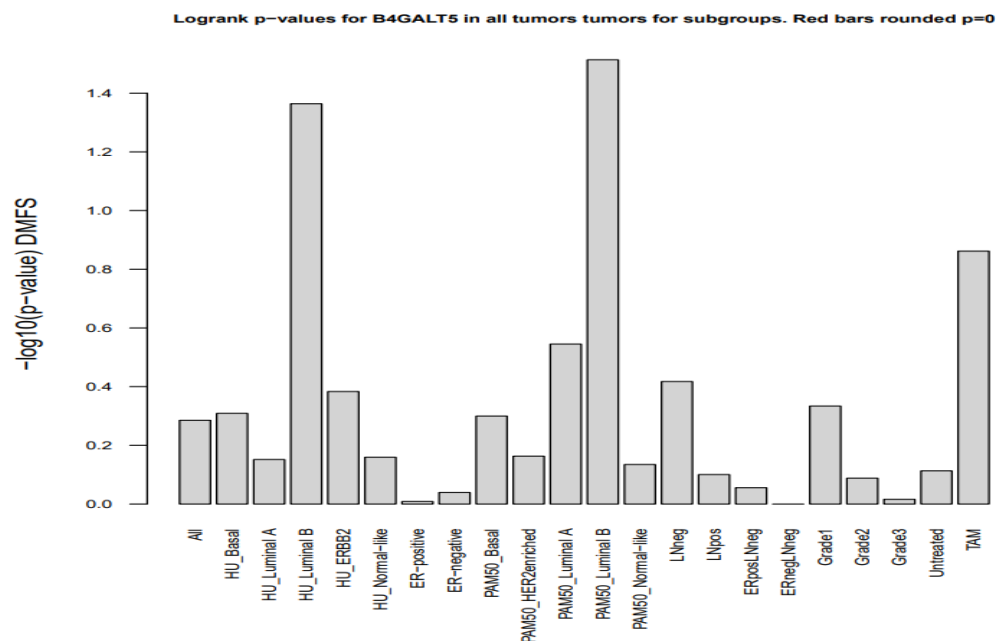


Fig.5.27: Comparison of B4GALT 5 Gene in Across all Subtypes of Breast Cancer

Fig5.26 and Fig5.27 also indicate the expression level of these gene when compared to the all the subsets of tumors, result in the no significant of these genes in the BC.

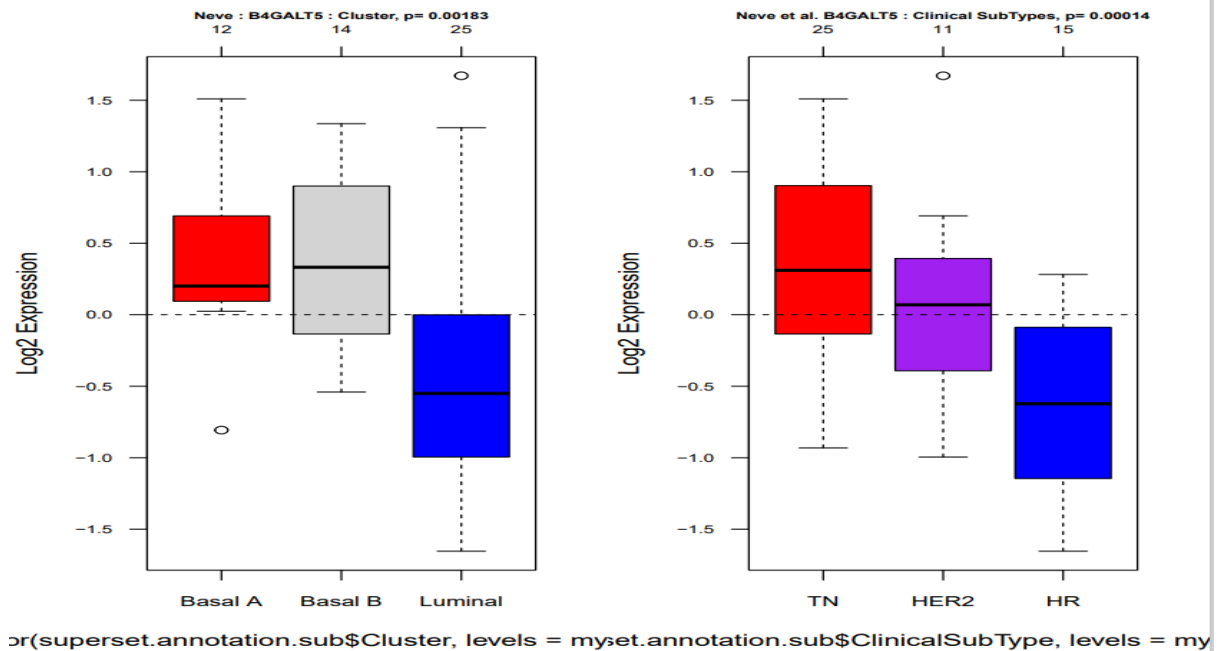


Fig.5.28: Box plot of B4GALT 5 Gene in Across All Subtypes of Breast Cancer

Analysis of these plot is showing some different result when the comparison between the all the molecular subsets of tumors ($P=0.001$) and all the subsets of the tumors ($P=0.0018$), it indicates that there small upregulated of these gene. It means that there is some expression of these genes in the BC and it also indicates that there is a significant of these gene.

GOBO analysis or Gene set analysis of B4GALT6 gene in all subsets of tumor also indicates that there is no significant expression of this gene in the BC, because the P value is great than the actual significant value ($0.6 > 0.05$).

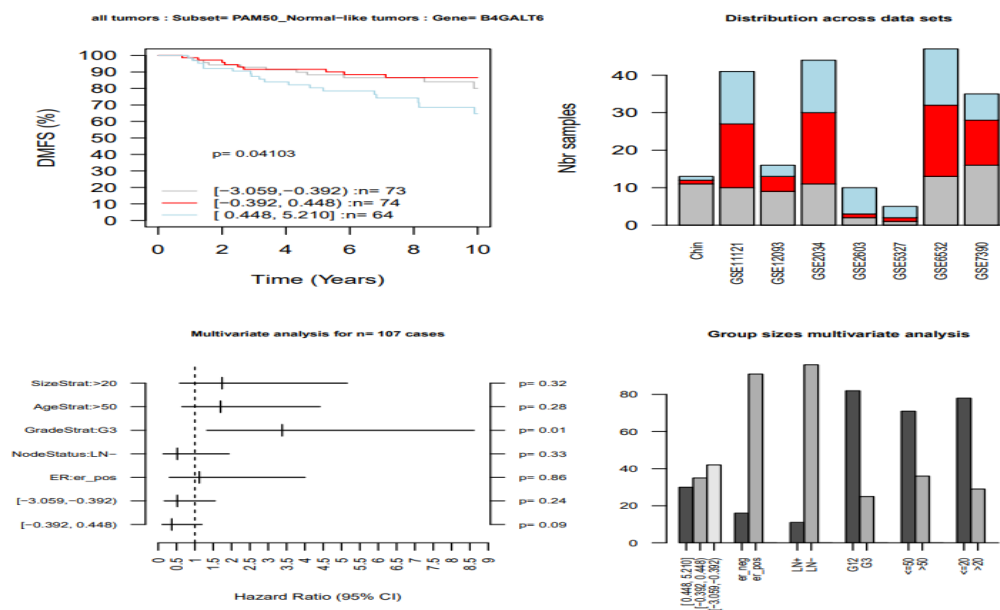


Fig.5.31: Gene set Analysis of B4GALT 6 Gene in Normal with Tumor

GOBO analysis or Gene set analysis of B4GALT6 gene in all subset of normal is compared with all the subset of the tumors, also indicates that there is some significant expression of this gene in the BC, because the P value is lesser than the actual significant value ($0.04 > 0.05$).

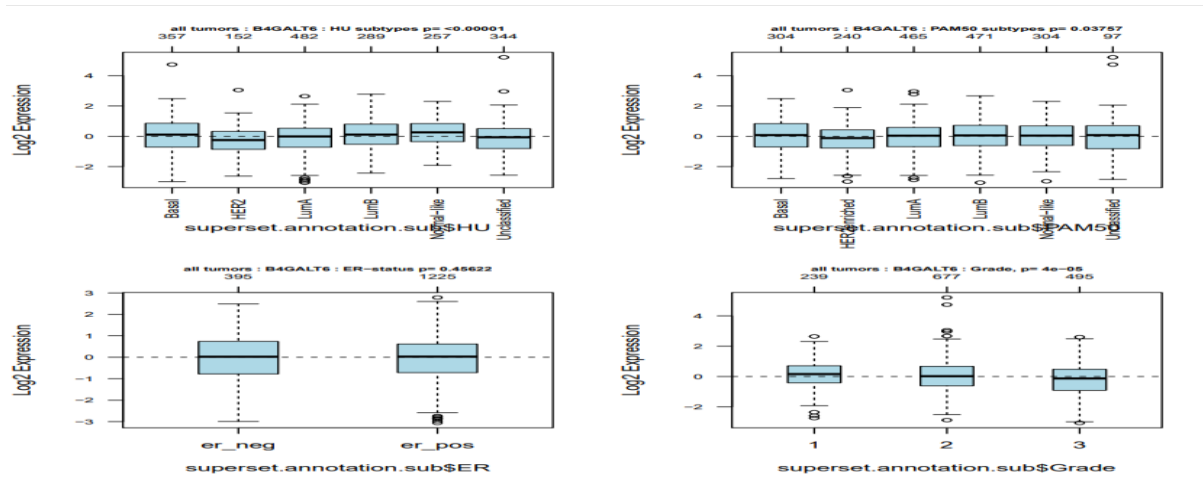


Fig.5.32: Comparison of B4GALT 6 GENE in Across All Types of Breast Cancer

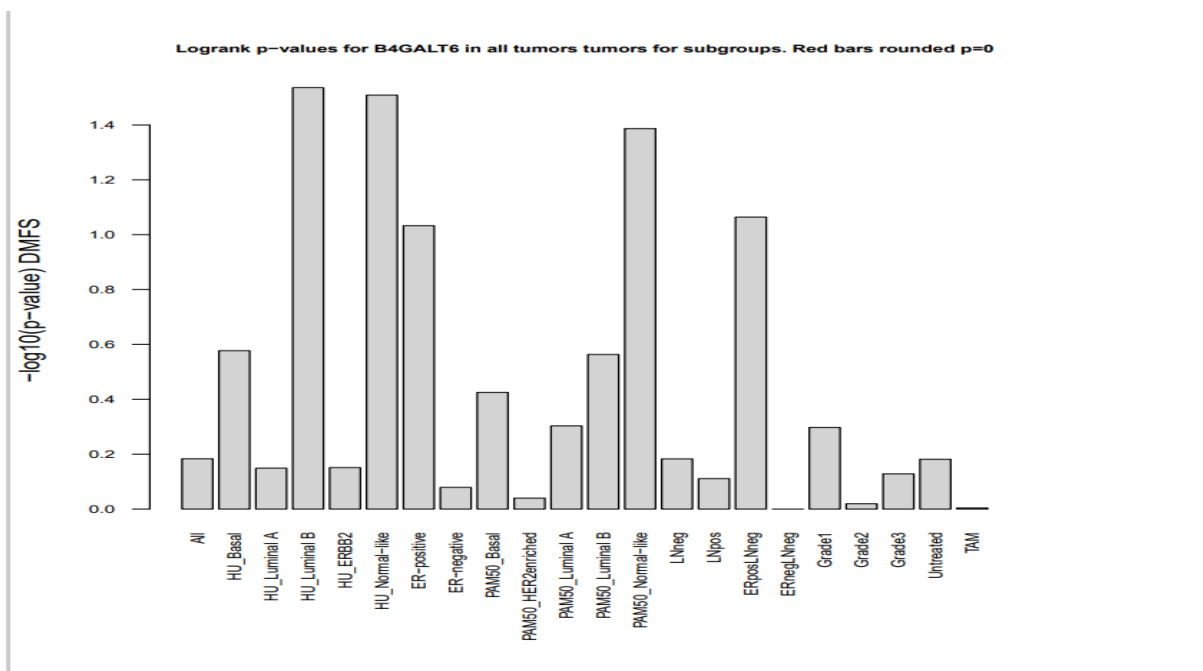


Fig.5.33: Comparison of B4GALT 6 Gene in Across All Subtypes of Breast Cancer

Fig5.31 and Fig5.32 also indicate the expression level of these gene when compared to the all the subsets of tumors, result in the no significant of these genes in the BC.

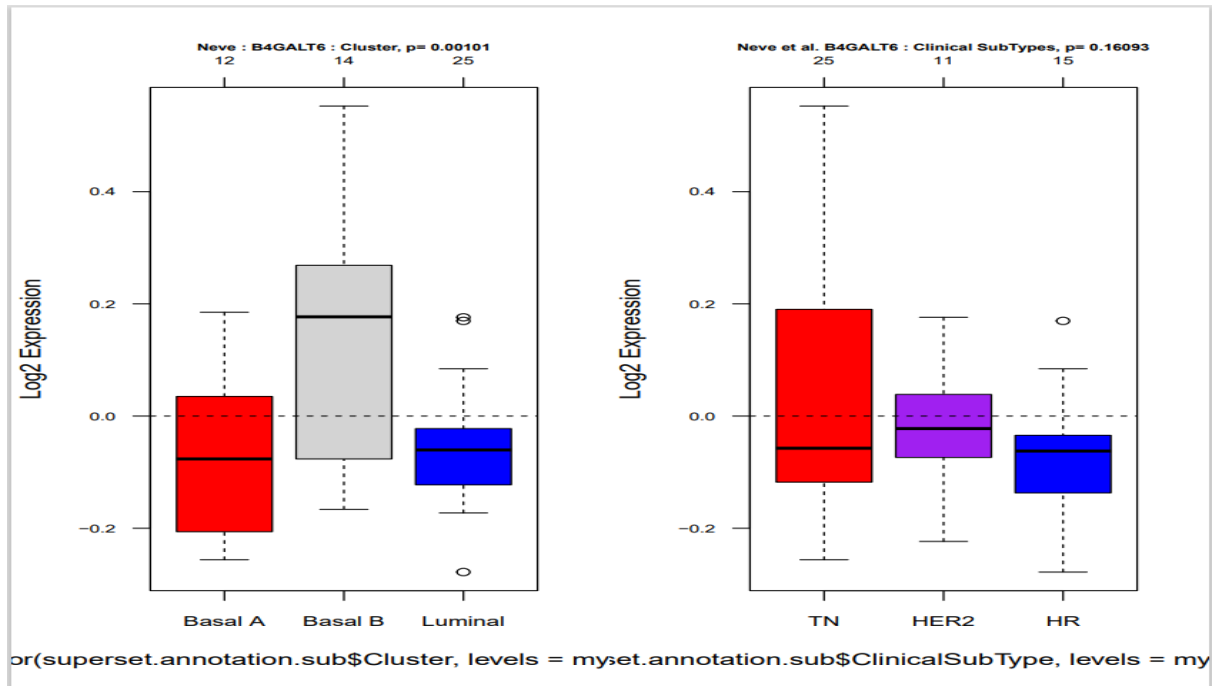


Fig.5.34: Box Plot of B4GALT 6 Gene in Across All Subtypes of Breast Cancer

Analysis of these plot is showing some different result when the comparison between the all the molecular subsets of tumors ($P=0.16$) and all the subsets of the tumors ($P=0.0018$), it indicates that there small upregulated of these gene shown in molecular types. It means that there is some expression of these genes in the BC, and it also indicates that there is a significant of these gene. In some cases, in subtypes of tumors, it shows downregulated of these genes in BC.

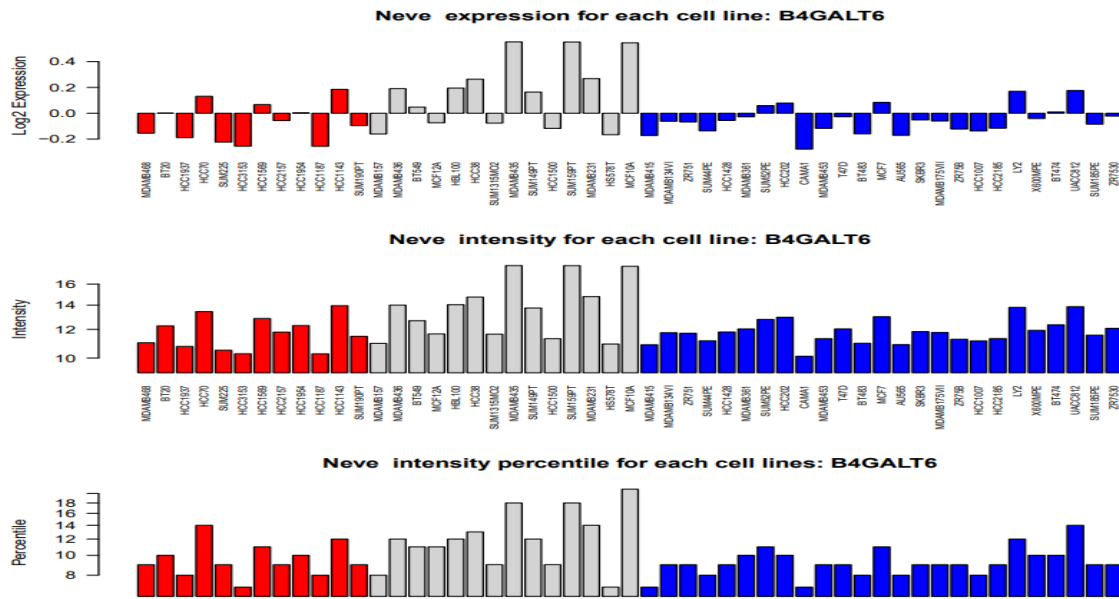


Fig.5.35: KaplanMeier Summary of B4GALT 6 Gene for each Cell Lines

Kaplanmeier summary indicates that expression level of these genes over a period and shows us the survival rate of the patient over the period of times as result there is also no significant level or no expression of these genes in the breast cancer cell when compared to the other cancer cell lines like MDAMB435, SUM159PT and MCF10A. It means that there is some upregulated function of these genes in these types of cell lines.

5.3.3 C-BIOPORTAL DATABASE

B4GALT 5

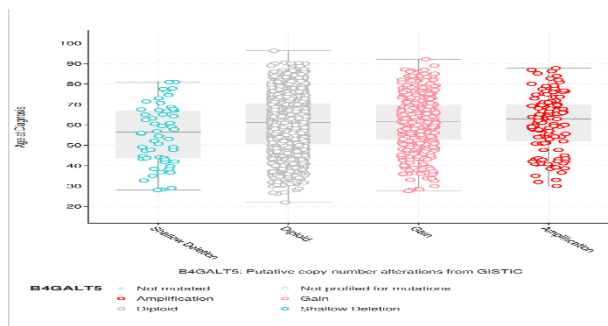


Fig.5.36: Alteration of B4GALT 5 in Breast Cancer Cell

This graphical representation shows us that alteration of B4GALT5 gene in the tumor sample genome occurs some amplification and others like not mutated, shallow deletion of these genes in the breast cancer cellular function. All of these expression range between the 50-60, which shows the people who are in the age higher than 50 have high chance of getting malignant.

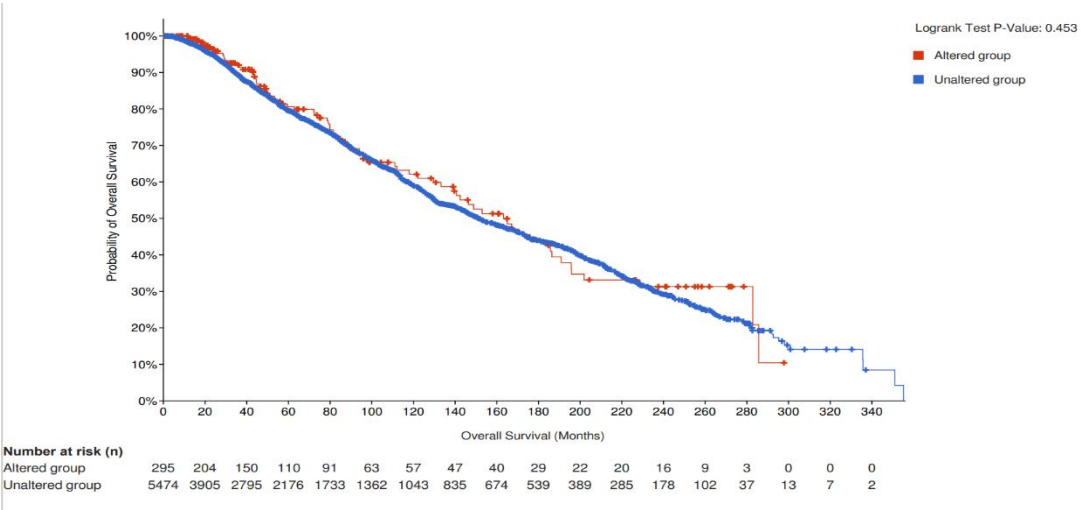


Fig.5.37: Graph between the Altered & Unaltered Group – Overall Survival Rate

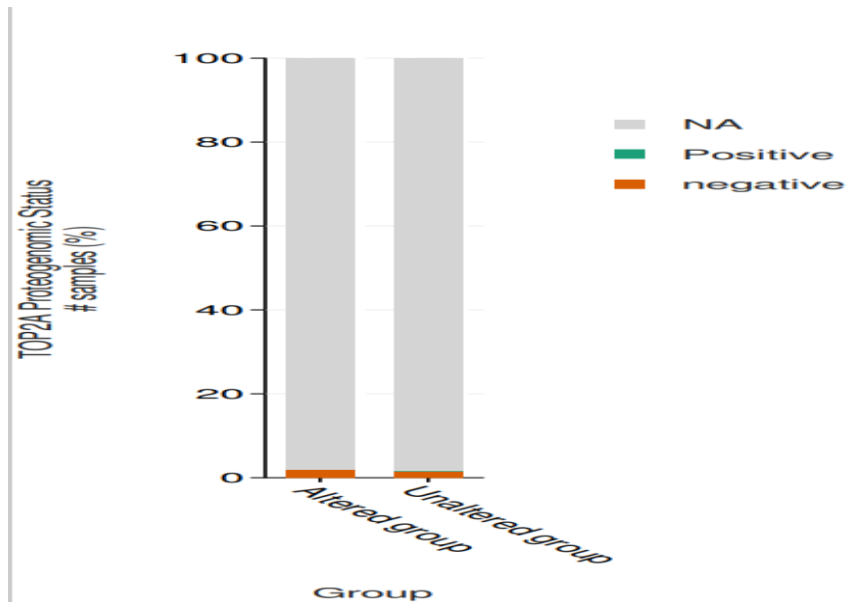


Fig.5.38: Proteogenomic Status of Samples

Fig 5.37 & 5.38, the survival rate and proteogenomic status indicates the alteration or mutation of these genes over a period interval. Comparison is

between the expression level vs the no expression of these gene in the tumor sample. This also shows us that there is no significant level (P value $=0.4 > 0.05$) of these genes. It ends in the downregulation function.

B4GALT 6

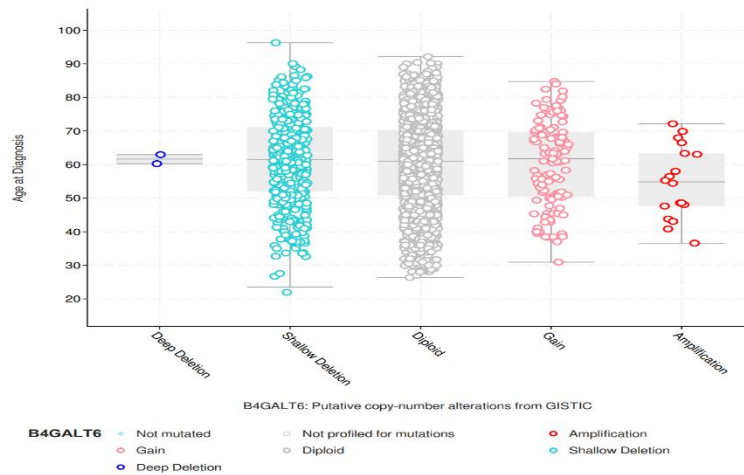


Fig.5.39: Alteration of B4GALT 6 in Breast Cancer Cell

This graphical representation shows us that alteration of B4GALT6 gene in the tumor sample genome occurs some amplification and others like not mutated, shallow deletion of these genes in the breast cancer cellular function. All of these expression range between the 60-70, which shows the people who are in the age higher than 60 have high chance of getting malignant.

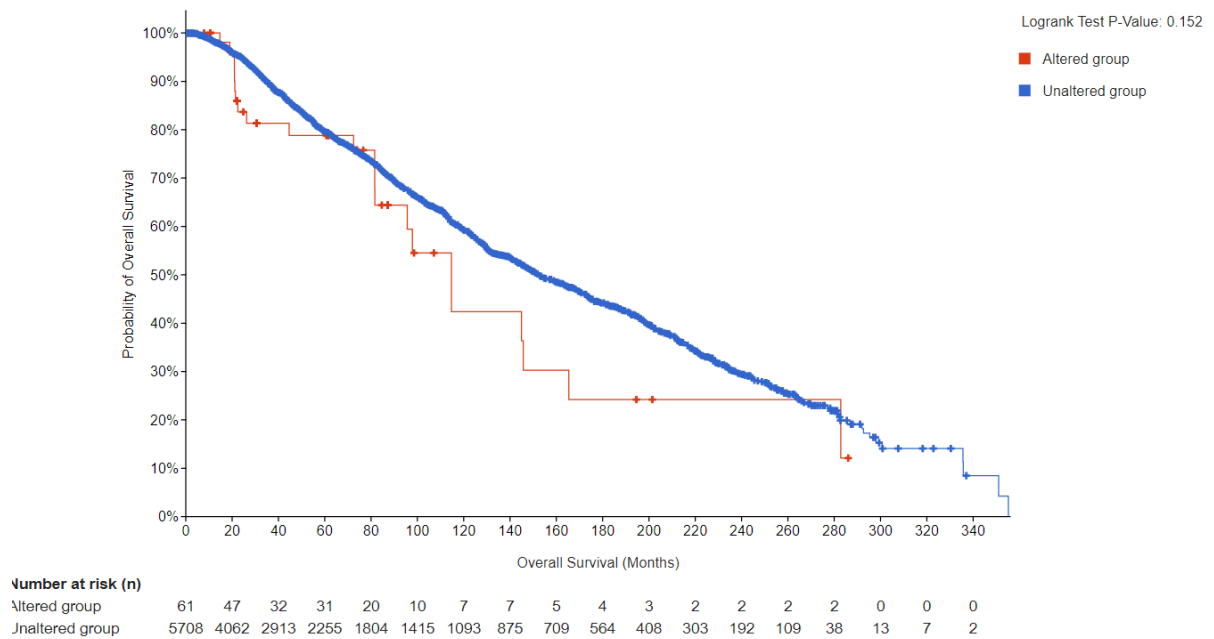


Fig.5.40: Graph between the Altered & Unaltered Group – Overall Survival Rate

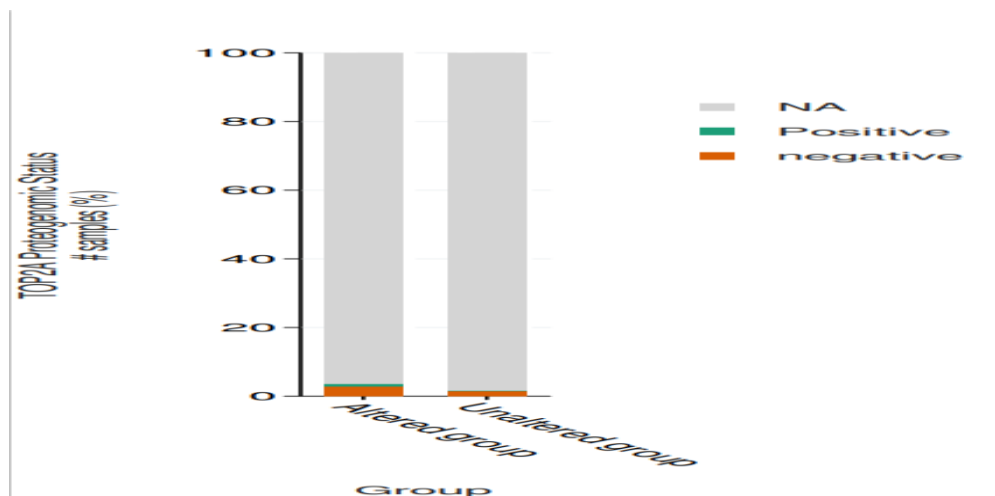


Fig.5.41: Proteogenomic Status of Samples

Fig5.40 & 5.41, the survival rate and proteogenomic status indicates the alteration or mutation of these genes over a period interval. Comparison is between the expression level vs the no expression of these gene in the tumor sample. This also shows us that there is no significant level ($P \text{ value} = 0.1 > 0.05$) of these genes. It ends in the downregulation function.

5.3.4 COSMIC DATABASE

B4GALT 5

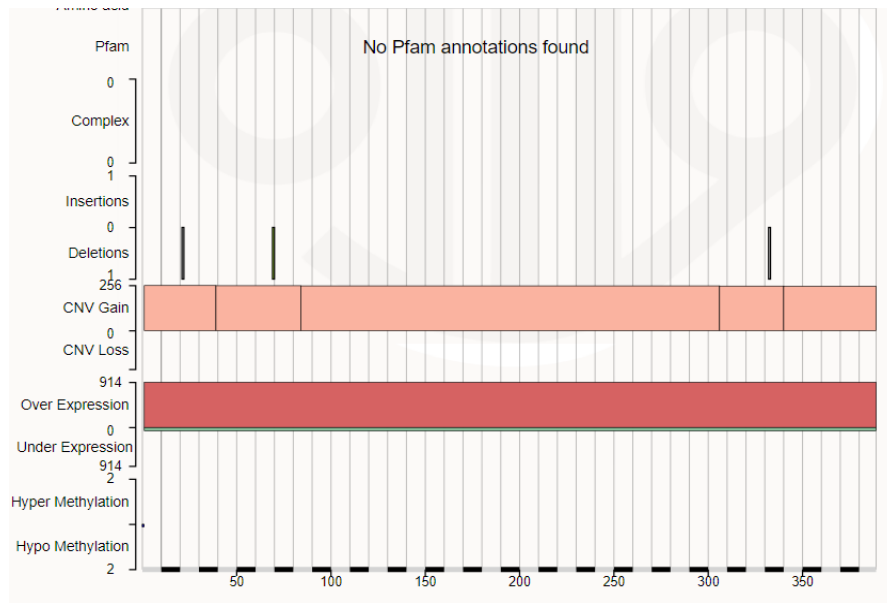


Fig.5.42: Gene View for *B4GALT 5*

The gene view histogram is a graphical view of mutations across B4GALT5. These mutations are displayed at the amino acid level across the full length of the gene. It shows us the different types of expression level like over and under expression of gene. It also indicates the hyper methylation and hypo methylation level and other results like deletion and insertion levels. It indicates the 914 patients were over expressed and 78 are under expressed. There is no sign of both hyper and hypo methylation.

Tissue	Point Mutations		Copy Number Variation		Gene Expression		Methylation	
	% Mutated	Tested	Variant %	Tested	% Regulated	Tested	% Diff. Methylated	Tested
Adrenal gland		624		267		72		-
Autonomic ganglia		1152		-		-		-
Biliary tract		861		-		-		-
Bone		589		-		-		-
Breast		2773		1492		1104		707
Central nervous system		2770		-		697		-
Cervix		360		299		307		-
Endometrium		786		586		602		-
Eye		203		80		-		-
Fallopian tube		2		-		-		-
Gastrointestinal tract (site indeterminate)		1		-		-		-
Genital tract		70		-		-		-
Haematopoietic and lymphoid		7278		-		221		-
Kidney		2355		995		600		513
Large intestine		2847		718		610		281
Liver		2288		663		373		244
Lung		3054		1006		1019		717
Meninges		255		-		-		-
NS		162		-		-		-
Oesophagus		2190		510		125		-
Ovary		941		684		266		-
Pancreas		2011		898		179		-
Parathyroid		35		-		-		-
Penis		10		-		-		-
Peritoneum		20		-		-		-
Pituitary		83		-		-		-
Placenta		31		-		-		-
Pleura		262		-		-		-
Prostate		2079		-		498		358
Salivary gland		120		-		-		-
Skin		2046		587		473		-
Small intestine		226		-		-		-
Soft tissue		818		264		263		-
Stomach		1582		472		285		-
Testis		233		-		-		-
Thymus		162		-		-		-
Thyroid		1137		-		513		-
Upper aerodigestive tract		1427		520		522		-
Urinary tract		805		399		408		-
Vagina		1		-		-		-
Vulva		6		-		-		-

Fig.5.43: Tissue Distribution of Mutations Across the Primary Tissue Types

The table shows the distribution of mutations across the primary tissue types that are curated by COSMIC. Histograms show the percentage of mutated samples for point mutations, CNV data and gene expression data. This table indicates that less mutation and variant among all types of tissue and also it shows that high expression rate of these genes seen in the adrenal gland and NS region when compare to others tissue's part. No sign for methylation.

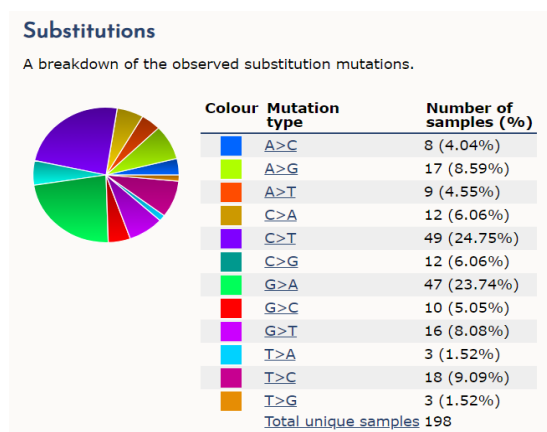


Fig.5.44: Substitution Mutations

This graphical representation shows us the alteration of the gene in amino acid sequence during protein synthesis like Silent Mutation, Missense Mutation, Nonsense Mutation etc. The amino acid sequence (C>T) shows high mutation and least one is (T>G) sequence.

B4GALT 6

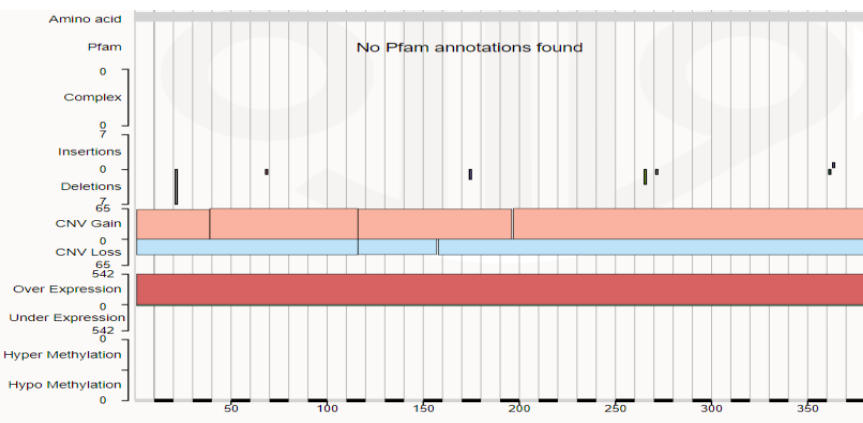


Fig.5.45: Gene view for B4GALT 6

The gene view histogram is a graphical view of mutations across B4GALT5. These mutations are displayed at the amino acid level across the full length of the gene. It shows us the different types of expression level like over and under expression of gene. It also indicates the hyper methylation and hypo methylation level and other results like deletion and insertion levels. It indicates the 552 patients were over expressed and 26 are under expressed. There is no sign of both hyper and hypo methylation.

Tissue	Point Mutations		Copy Number Variation		Gene Expression		Methylation	
	% Mutated	Tested	Variant %	Tested	% Regulated	Tested	% Diff. Methylated	Tested
Adrenal gland	62.4	62.4	-	-	-	79	-	-
Autonomic ganglia	11.82	11.82	-	-	-	-	-	-
Biliary tract	66.1	66.1	-	-	-	-	-	-
Bone	66.9	66.9	-	-	-	-	-	-
Breast	277.3	277.3	1.492	1.492	-	110.4	-	-
Central nervous system	277.0	277.0	1.025	1.025	-	69.7	-	-
Cervix	360	360	-	-	-	302	-	-
Endometrium	78.6	78.6	-	-	-	60.2	-	-
Eye	20.3	20.3	-	-	-	-	-	-
Fallopian tube	2	2	-	-	-	-	-	-
Gastrointestinal (digestive)	1	1	-	-	-	-	-	-
Genital tract	70	70	-	-	-	-	-	-
Heart (cardiac)	277.8	277.8	1.881	1.881	-	22.1	-	-
Kidney	23.84	23.84	9.95	9.95	-	600	-	-
Large intestine	26.7	26.7	7.18	7.18	-	610	-	-
Liver	26.7	26.7	-	-	-	37.3	-	-
Lung	20.83	20.83	1.006	1.006	-	10.19	-	-
Meninges	2.55	2.55	-	-	-	-	-	-
MS	1.62	1.62	-	-	-	-	-	-
Oesophagus	21.90	21.90	51.0	51.0	-	12.5	-	-
Stomach	20.1	20.1	8.83	8.83	-	2.88	-	-
Testes	20.1	20.1	1.49	1.49	-	-	-	-
Thyroid	2.5	2.5	-	-	-	-	-	-
Uterus	10	10	-	-	-	-	-	-
Peritoneum	20	20	-	-	-	-	-	-
Bladder	6.3	6.3	-	-	-	-	-	-
Placenta	3.1	3.1	-	-	-	-	-	-
Esophagus	2.83	2.83	-	-	-	-	-	-
Prostate	207.9	207.9	9.49	9.49	-	49.9	-	38.8
Salivary gland	1.0	1.0	-	-	-	-	-	-
Skin	20.45	20.45	-	-	-	47.3	-	-
Small intestine	2.28	2.28	-	-	-	-	-	-
Soft tissue	6.0	6.0	2.64	2.64	-	2.63	-	-
Stomach	1.882	1.882	4.72	4.72	-	2.88	-	-
Testis	2.23	2.23	1.49	1.49	-	-	-	-
Thymus	1.62	1.62	-	-	-	-	-	-
Thyroid	11.37	11.37	-	-	-	51.3	-	-
Upper aerodigestive tract	14.27	14.27	6.20	6.20	-	62.3	-	-
Uterus	20.5	20.5	3.32	3.32	-	40.8	-	-
Vagina	1	1	-	-	-	-	-	-
Vulva	6	6	-	-	-	-	-	-

Fig.5.46: Tissue Distribution of Mutations Across the Primary Tissue Types

The table shows the distribution of mutations across the primary tissue types that are curated by COSMIC. Histograms show the percentage of mutated samples for point mutations, CNV data and gene expression data. This table indicates that less mutation and variant among all types of tissue and also it also shows us the least gene expression rate of these gene in the patient rate. No sign for methylation.

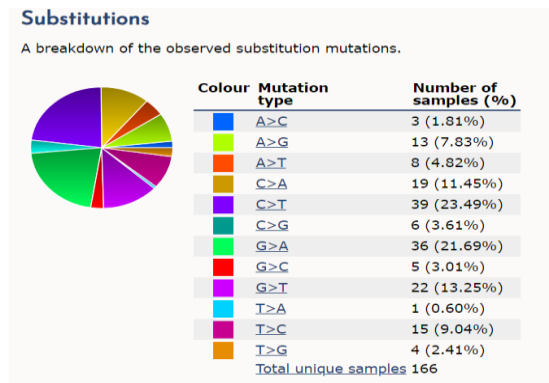


Fig.5.47: Substitution Mutations

This graphical representation shows us the alteration of the gene in amino acid sequence during protein synthesis like Silent Mutation, Missense Mutation, Nonsense Mutation etc. The amino acid sequence (C>T) shows high mutation and least one is (A>C) sequence when compared to other genomic sequence.

CHAPTER 6

STATISTICAL ANALYSIS

B4GALT5

TABLE 6.1: CT Value of B4GALT5, Tumors Vs Paired Normal

SAMPLE N	B4GALT5 T	B4GALT5 P
1	31.93845	35.40021
2	30.34079	28.09514
3	24.53109	25.826
4	32.85775	31.64695
5	22.28008	34.65547
6	28.34873	31.10994
7	34.795	40
8	25.939	33.717
9	27.854	37.932
10	33.345	36.673

>>PEARSON COEFFICIENT = **0.132093282**

>N = **10**

>T STATISTIC = **0.376919053**

>DEGREE OF FREEDOM = **8**

>>P- VALUE = **0.641982382**

>>> This statistical analysis of Pearson coefficient also indicates that the B4GALT 5 gene are not significant, or no expression was seen in the breast cancer cell is proved by this method (P-value = 0.6>0.05). So, it also verifies that these genes are downregulated cellular function.

> MEAN = **31.46624631**

> STANDARD DEVIATION = **4.698368128**

>> COEFFICIENT OF VARIANCE (%) = **14.93145411**

>> REGRESSION LINE

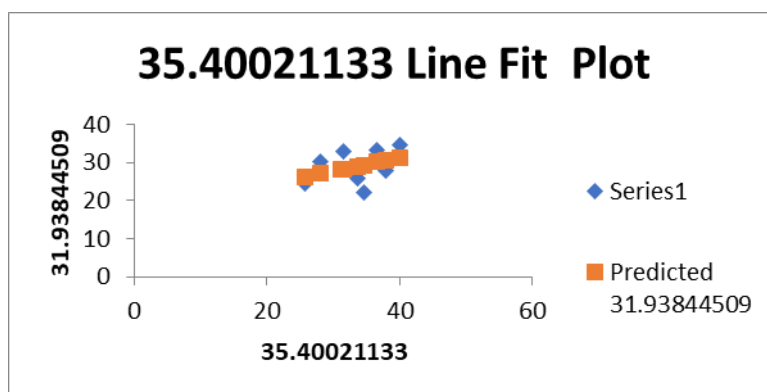


Fig.6.1: Linear Regression Line, R Square = 0.141955214.

The linear regression analysis reveals a slope coefficient of 35.400, indicating that for every unit increase in the independent variable. However, the modest R-squared value of 0.1419 suggests that only approximately 14.20% of the variance in the dependent variable is explained by the independent variable included in our model.

TABLE6.2- ANOVA test

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance</i>
Regression	1	20.62905	20.62905	1.158082	0.317549
Residual	7	124.6918	17.81311		
Total	8	145.3208			

>>> This test explains the significant level of the B4GALT 5 gene is 0.317, which determines the downregulated function, because the value great than 0.05 p-value.

B4GALT6

TABLE 6.3: CT Value of B4GALT6, Tumors Vs Paired Normal

SAMPLE N	B4GALT6 T	B4GALT6 P
1	28.4106	33.069
2	31.35	30.04896
3	28.734	27.423
4	32.85775	32.00222
5	22.28008	34.04316
6	28.34873	32.32834
7	34.795	37.928
8	25.939	33.829
9	27.854	28.833
10	33.345	31.844

>PEARSON COEFFICIENT = **0.397313062**

>N = **10**

>T STATISTIC = **1.224574149**

>DEGREE OF FREEDOM = **8**

>P- VALUE = **0.872215456**

>> This statistical analysis of Pearson coefficient also indicates that the B4GALT 6 gene are not significant, or no expression was seen in the breast cancer cell is proved by this method (P-value = $0.8 > 0.05$). So, it also verifies that these genes are downregulated cellular function.

> MEAN = **2.899690048**

> STANDARD DEVIATION = **3.584970575**

> COEFFICIENT OF VARIANCE (%) = **123.6328889**

>> REGRESSION LINE

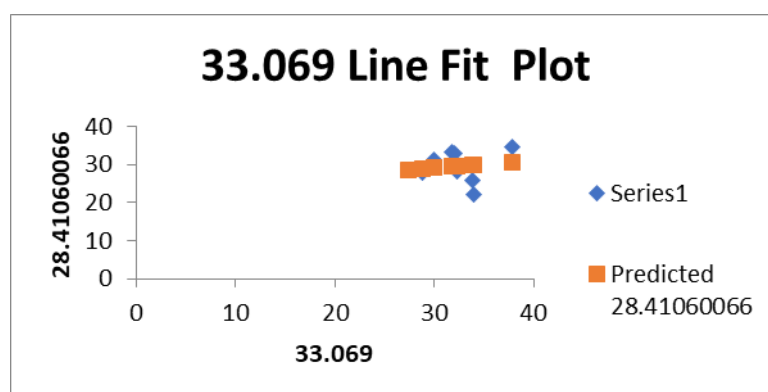


Fig.6.2: Linear Regression Line, $R^2 = 0.20625498$.

The linear regression analysis yielded a slope coefficient of 33.069, suggesting that for every unit increase in the independent variable, there is an average increase of 33.069 units in the dependent variable, assuming other factors remain constant. However, the low R^2 value of 0.0206 indicates that only approximately 2.06% of the variability in the dependent variable is explained by the independent variable in our model.

TABLE 6.4: ANOVA test

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance</i>
Regression	1	2.622644	2.622644	0.147666	0.712183
Residual	7	124.3248	17.76068		
Total	8	126.9474			

>>> This test explains the significant level of the B4GALT 6 gene is 0.712, which determines the downregulated function, because the value greater than 0.05 p-value.

CHAPTER 7

SUMMARY AND CONCLUSION

Here, the expression of B4GALT5 and B4GALT6 genes in breast cancer cells and breast cancer sample (tumor and paired normal). The analysis revealed selective lower expression of these genes in breast cancer cells. Gene expression results in the breast cancer cell lines shown to be downregulated in breast cancer cell as compared to the HEK 293 normal human cell lines (control cell). Similarly, expression analysis result of these genes in the breast tumor tissue samples when compared to the paired normal samples shows significantly downregulated.

In summary, B4GALT5 and B4GALT6 genes showed downregulated expression in breast tissue samples when compared to the paired normal samples. These genes were confirmed in breast cancer cells when compared to the control cell. The expression of these genes in breast cancer did not show significant concordance with expression levels of COX2 levels and requires further analysis. In the light of the emerging role of these genes in breast cancer, additional studies are required to assess their as biomarker and therapeutic targets by assessing their putative functions in the regulation of cellular processes.

These including functional studies to elucidate the mechanistic roles of these genes in breast cancer development, progression, metastasis, and therapeutic response; correlating their expression levels with clinical outcomes and evaluating their potential as prognostic or predictive biomarkers; exploring therapeutic strategies targeting B4GALT5 and B4GALT6, such as inhibitors or modulators; investigating their glycosylation patterns and substrates; identifying interaction partners and regulatory networks; analyzing subtype-specific expression and roles; evaluating combinatorial approaches with other therapies; utilizing in vivo models for more physiologically relevant studies; improve patient outcomes in breast cancer.

REFERENCES

- Borsig, L., Berger, E. G., & Malissard, M. (1997). Expression and Purification of His-Tagged β -1,4-Galactosyltransferase in Yeast and in COS Cells.
- Chatterjee, S., Hou, J. C., Bandaru, V. V. R., Pezhouh, M. K., Mannan, A. a. S. R., & Sharma, R. (2019). Lactosylceramide synthase β -1,4-GalT-V: A novel target for the diagnosis and therapy of human colorectal cancer.
- Chen, X., Jiang, J., Yang, J., Chen, C., Sun, M., Wei, Y., Guang, X., & Gu, J. (2006). Down-regulation of the expression of β 1,4-galactosyltransferase V promotes integrin β 1 maturation.
- Cooke, S. V., & Shur, B. D. (1994). Cell Surface beta1,4-Galactosyltransferase: Expression and Function. (cell surface galactosyltransferase/cell interactions/fertilization/migration/adhesion).
- Cui, S., Zhang, X., Li, Y., Huang, S., Wu, B., Zhao, F., Gao, J., Li, M. O., Hao, W., Tang, B., Xia, H., & Lin, X. (2023). UGCG modulates heart hypertrophy through B4GalT5-mediated mitochondrial oxidative stress and the ERK signaling pathway
- Ishii, T., Miyauchi, K., Nitta, Y., Kaneko, K., Maruyama, T., & Satô, T. (2018). Mechanism for Decreased Gene Expression of β 4-Galactosyltransferase 5 upon Differentiation of 3T3-L1 Mouse Preadipocytes to Adipocytes.
- Echeverri, D., & Orozco, J. (2022). β -1,4-Galactosyltransferase-V colorectal cancer biomarker immunosensor with label-free electrochemical detection.
- Guo, S., Satô, T., Shirane, K., & Furukawa, K. (2001). Galactosylation of N-linked oligosaccharides by human -1,4-galactosyltransferases I, II, III, IV, V, and VI expressed in Sf-9 cells.
- Hennet, T. (2002). The galactosyltransferase family. *Cellular and Molecular Life Sciences*, 59(7), 1081–1095.
- Kremer, J., Brendel, C., Mack, E. K. M., & Mack, H. I. D. (2020).
- Kumagai, T., Satô, T., Natsuka, S., Kobayashi, Y., Zhou, D., Senoo, T., Hayakawa, S., & Furukawa, K. (2010). Involvement of murine β -1,4-galactosyltransferase V in lactosylceramide biosynthesis.

- Kumagai, T., Tanaka, M., Yokoyama, M., Satô, T., Senoo, T., & Furukawa, K. (2009). Early lethality of β -1,4-galactosyltransferase V-mutant mice by growth retardation.
- Lee, J. H., Sundaram, S., Shaper, N. L., Raju, T. S., & Stanley, P. (2001). Chinese hamster ovary (CHO) cells may express six B4-Galactosyltransferases (B4GALTs).
- Lo, N. W., Shaper, J. H., Pevsner, J., & Shaper, N. L. (1998). The expanding beta 4-galactosyltransferase gene family: messages from the databanks.
- Ma, J., Kubushiro, K., Tashima, Y., Tsukazaki, K., Udagawa, Y., Nozawa, S., & Fukuda, M. N. (1997). Expression of human beta 1,4-galactosyltransferase in gynecological cancer cell lines.
- Ma, R., Decker, N. M., Anilus, V., Moskal, J. R., Burgdorf, J., Johnson, J. R., Basu, M., Banerjee, S., & Basu, S. (2009). Post-translational and transcriptional regulation of glycolipid glycosyltransferase genes in apoptotic breast carcinoma cells: VII. Studied by DNA-microarray after treatment with I-PPMP.
- Masri, K. A., Appert, H. E., & Fukuda, M. N. (1988). Identification of the full-length coding sequence for human galactosyltransferase (β -N-acetylglucosaminide: β 1,4-galactosyltransferase).
- Mayo, L., Trauger, S. A., Blain, M., Nadeau, M., Patel, B., Alvarez, J. I., Mascanfroni, I., Yeste, A., Kivisäkk, P., Kallas, K., Ellezam, B., Bakshi, R., Prat, A., Antel, J. P., Weiner, H. L., & Quintana, F. J. (2014).
- Miyazaki, T., Miyashita, R., Nakamura, S., Ikegaya, M., Kato, T., & Park, E. Y. (2019). Biochemical characterization and mutational analysis of silkworm *Bombyx mori* β -1,4-N-acetylgalactosaminyltransferase and insight into the substrate specificity of β -1,4-galactosyltransferase family enzymes.
- Nihei, M., Kawaguchi, S. I., Mineyama, A., & Satô, T. (2023). Transcriptional mechanism of the mouse B4-Galactosyltransferase 6 gene in mouse neuroblastoma cell line Neuro-2A.

Nilsson, T., Hoe, M. H., Slusarewicz, P., Rabouille, C., Watson, R., Hunte, F., Watzel, G., Berger, E., & Warren, G. (1994). Kin recognition between medial Golgi enzymes in HeLa cells.

Nishie, T., Hikimochi, Y., Zama, K., Fukusumi, Y., Ito, M., Yokoyama, H., Naruse, C., Ito, M., & Asano, M. (2010). β 4-Galactosyltransferase-5 is a lactosylceramide synthase essential for mouse extra-embryonic development.

Odunsi, K., Ghamande, S., Chandrasekaran, E. V., Ta, A., Moysich, K. B., Driscoll, D. L., Matta, K. L., & Lele, S. (2002). Evaluation of β 1,4-galactosyltransferase as a potential biomarker for the detection of subclinical disease after the completion of primary therapy for ovarian cancer.

Paone, J. F., Waalkes, T. P., Baker, R. R., & Shaper, J. H. (1980). Serum UDP-galactosyl transferase as a potential biomarker for breast carcinoma.

Pouncey, L., Easton, J., Heath, L. S., Grenet, J., & Kidd, V. J. (1991). β 1,4-galactosyltransferase gene expression is regulated during entry into the cell cycle and during the cell cycle.

Satô, T., & Furukawa, K. (2004). Transcriptional Regulation of the Human β -1,4-Galactosyltransferase V Gene in Cancer Cells.

Satô, T., & Furukawa, K. (2007). Sequential Action of Ets-1 and Sp1 in the Activation of the Human β -1,4-Galactosyltransferase V Gene Involved in Abnormal Glycosylation Characteristic of Cancer Cells.

Satô, T., Guo, S., & Furukawa, K. (2001). Occurrence of poly-N-acetyllactosamine synthesis in Sf-9 cells upon transfection of individual human β -1,4-galactosyltransferase I, II, III, IV, V and VI cDNAs.

Satô, T., Shirane, K., Kido, M., & Furukawa, K. (2000). Correlated Gene Expression between β -1,4-Galactosyltransferase V and N-Acetylglucosaminyltransferase V in Human Cancer Cell Lines.

Shirane, K., Satô, T., Segawa, K., & Furukawa, K. (1999). Involvement of β -1,4-Galactosyltransferase V in Malignant Transformation-Associated Changes in Glycosylation.

Taki, T., Nishiwaki, S., Handa, N., Hattori, N., & Handa, S. (1994). A New Method for Detecting β 1,4-Galactosyltransferase Activity in Sera of Cancer Patients.

Tokuda, N., Numata, S., Li, X. J., Nomura, T., Takizawa, M., Kondo, Y., Yamashita, Y., Hashimoto, N., Kiyono, T., Urano, T., Furukawa, K., & Furukawa, K. (2013). 4GalT6 is involved in the synthesis of lactosylceramide with less intensity than 4GalT5.

Ujita, M., Misra, A. K., McAuliffe, J. C., Hindsgaul, O., & Fukuda, M. (2000). Poly-N-acetyllactosamine Extension in N-Glycans and Core 2- and Core 4-branched O-Glycans Is Differentially Controlled by α -Extension Enzyme and Different Members of the β 1,4-Galactosyltransferase Gene Family.

Wang, X., Shi, N., Hui, M., Jin, H., Gao, S., Zhou, Q., Zhang, L., Yan, M., & Shen, H. (2021). The Impact of β -1,4-Galactosyltransferase V on Microglial Function.

Xu, S., Zhu, X., Zhang, S., Yin, S., Zhou, L., Chen, C., & Gu, J. (2001). Over-expression of β -1,4-galactosyltransferase I, II, and V in human astrocytoma.

Yoshihara, T., Satake, H., Nishie, T., Okino, N., Hatta, T., Otani, H., Naruse, C., Suzuki, H., Sugihara, K., Kamimura, E., Tokuda, N., Furukawa, K., Furukawa, K., Ito, M., & Asano, M. (2018). Lactosylceramide synthases encoded by B4galt5, and 6 genes are pivotal for neuronal generation and myelin formation in mice.

Youakim, A., & Shur, B. D. (1993). Effects of overexpression of β 1, 4-galactosyltransferase on glycoprotein biosynthesis in F9 embryonal carcinoma cells.

Zhang, L., Ren, J., Shi, P., Lu, D., Zhao, C., Su, Y., Zhang, L., & Huang, J. (2018). The Immunological Regulation Roles of Porcine β -1, 4 Galactosyltransferase V (B4GALT5) in PRRSV Infection.

Zhang, X., Cui, S., Ding, Y., Li, Y., Wu, B., Gao, J., Li, M., Xu, L., & Xia, H. (2024). Downregulation of B4GALT5 attenuates cardiac fibrosis through Lumican and Akt/GSK-3 β / β -catenin pathway.

Zhou, H., Ma, H., Wei, W., Ji, D., Song, X., Sun, J., Zhang, J., & Li, J. (2013). B4GALT family mediates the multidrug resistance of human leukemia cells by regulating the hedgehog pathway and the expression of p-glycoprotein and multidrug resistance-associated protein 1.

ANNEXURE

Review paper manuscript submitted to the journal

Manuscript Initial Submission letter [BMS-CCTR-2023-19]  Inbox x



cctr@benthamscience.net <admin@bentham.manuscriptpoint.com>

to me ▾

Referenced BMS-CCTR-2023-19

Submission Title: Salivary proteins as a prognostic and diagnostic marker in breast cancer.

Dear Dr Revathi Paramasivam Oviya,

We are pleased to inform you that your article entitled, "**Salivary proteins as a prognostic and diagnostic marker in breast cancer**" has been initially approved by the Editorial Advisory Board of the journal "**Current Cancer Therapy Reviews**" to proceed further.

This is to inform you that currently your article is under editorial evaluation by Bentham Author Services Support (BASS) editors.

The Bentham Authors Services Support (BASS) editors consist of competent professionals who perform detailed, initial scrutiny of all submitted manuscripts to assist authors in fine-tuning their manuscripts. This will help reduce the errors that may delay the review and publication process later. The core aim of the BASS editors is to ensure the complete, correct, and concise incorporation of all details needed at the initial stage, to ensure smooth and timely processing of the article to the peer-review stage.

Your article will be evaluated by the BASS editors for the following things:

*Language editing.

*Figure quality checking enhancement.

*Reference duplication-checking formatting.